MIT-CTP/5998
KEK-Cosmo-0408

Lyman-α Forest and its Cross-Correlation with High-Redshift Galaxies in Effective Field Theory at the Field Level

Roger de Belsunce,[1, 2, *] Mikhail M. Ivanov,[1, 2, †] James M. Sullivan,[1, 2, ‡] Shi-Fan Chen,[3, 4] and Kazuyuki Akitsu[5]

[1] *Center for Theoretical Physics – a Leinweber Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2] *The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, Cambridge, MA 02139, USA*
[3] *Department of Physics, Columbia University, New York, NY 10027, USA*
[4] *NASA Hubble Fellowship Program, Einstein Fellow*
[5] *Theory Center, Institute of Particle and Nuclear Studies,*
High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan

We present a field-level perturbative forward model for the Lyman-α (Ly-α) forest flux decrement. We validate it on two simulation suites: large-volume AbacusSummit N-body simulations with the Ly-α forest painted onto the dark matter field, and the Sherwood hydrodynamic simulations. Across the redshift range of the simulations ($z = 2.0$–3.2), the 3D and 1D power spectra of the model match the simulated Ly-α fields at the 1% (5%) level up to $k \approx 0.3\,(1.0)\,h\,\mathrm{Mpc}^{-1}$, with similar performance for the cross-correlation with massive dark matter halos. The counts-in-cells statistic shows excellent agreement down to cell radii of $2\,h^{-1}\,\mathrm{Mpc}$. Leveraging cosmic variance cancellation, the model enables precision measurements of Ly-α bias parameters and robustly detects the full set of quadratic line-of-sight bias operators, consistent with the notion of naturalness in effective field theory (EFT). We quantify the stochasticity of the Ly-α forest (the analog to the one-halo term), and find it to be white (scale- and orientation-independent) on large scales, matching EFT predictions. We further find that phenomenological flux power spectrum models, based on modulations of the linear-theory power spectrum, fail at the field level even on quasi-linear scales. For the currently observing Dark Energy Spectroscopic Instrument (DESI), we generate large-scale clustering mocks of the Ly-α forest to validate cosmological parameter inference pipelines. Looking ahead to its successor, DESI-II, we produce large-volume mocks of representative samples of Lyman-break galaxies (LBGs) and Ly-α emitters (LAEs), calibrated on Astrid hydrodynamic simulations and matched to observations at $z = 3$, enabling joint analyses of Ly-α forest and high-redshift galaxy data.

CONTENTS

---

[*] belsunce@mit.edu
[†] ivanov99@mit.edu
[‡] jms3@mit.edu; Brinson Prize Fellow

MIT-CTP/5998 KEK-
Cosmo-0408

在场级有效场论中 Lyman-α 森与高红移星系的交叉相关

罗杰・德・贝伦斯,1, 2, ⋆米哈伊尔・M・伊万诺夫,1, 2, †詹姆斯・M・沙利文,1, 2, ‡陈士凡,3, 4 和 秋津和之5 1麻省理工学院莱因韦伯理论物理中心，美国马萨诸塞州剑桥市，邮编 02139 2美国国家科学基金会人工智能与基础相互作用AI研究所，马萨诸塞州剑桥市，邮编 02139 3哥伦比亚大学物理系，美国纽约州纽约市，邮编 10027 4NASA 哈勃奖学金计划，爱因斯坦研究员 5高能加速器研究机构（KEK）粒子与核研究所理论中心，日本茨城县筑波市，邮编 305-0801

我们提出了一个用于Lyman-α（Ly-α）森林通量减量的场级微扰前向模型。我们在两个模拟套件上验证了该模型：大体积的AbacusSummit N体模拟（在暗物质场上绘制Ly-α森林）以及Sherwood流体动力学模拟。在模拟的红移范围内（$z = 2.0 - 3.2$），模型的3D和1D功率谱与模拟的Ly-α场在$k \approx 0.3$（1.0）h Mpc−1以内分别达到1%（5%）的匹配精度，同时与大质量暗物质晕的交叉相关也表现出类似的性能。计数-单元统计在单元半径降至2 h−1Mpc时仍显示出极佳的一致性。利用宇宙方差抵消，该模型能够精确测量Ly-α偏差参数，并稳健地检测到完整的一组二次视线偏差算子，这与有效场论（EFT）中的自然性概念一致。我们量化了Ly-α森林的随机性（类似于单晕项），并发现其在大尺度上是白噪声（尺度和方向无关的），与有效场理论（EFT）的预测一致。我们进一步发现，基于线性理论功率谱调制的表象通量功率谱模型，即使在准线性尺度上也无法在场层面上有效描述。针对当前观测的暗能量光谱仪（DESI），我们生成了Ly-α森林的大尺度聚类模拟，以验证宇宙学参数推断管线。展望其继任者DESI-II，我们生成了代表性样本的大体积莱曼断裂星系（LBGs）和Ly-α发射星系（LAEs）模拟，这些模拟经过Astrid流体动力学模拟校准，并与$z = 3$的观测数据匹配，从而能够对Ly-α森林和高红移星系数据进行联合分析。

目录

---

*belsunce@mit.edu † ivanov99@mit.edu ‡
jms3@mit.edu；布林森奖学金研究员

I. INTRODUCTION AND KEY RESULTS

The Lyman-α (Ly-α) forest consists of a series of absorption features in the spectra of distant quasars, produced by intervening neutral hydrogen along the line of sight. Since the 1990s, high-resolution observations with instruments such as the High Resolution Echelle Spectrometer [HIRES; 1, 2] and the Ultraviolet and Visual Echelle Spectrograph [UVES; 3, 4] have enabled precise measurements of small-scale fluctuations in the neutral hydrogen density. These observations established that the absorbing gas resides in the low-density, highly ionized intergalactic medium (IGM), providing a transparent link between the neutral hydrogen distribution and the underlying dark matter field. This connection, in turn, enables precision simulations of the Ly-α forest. Because the neutral gas is in photoionization equilibrium with an approximately uniform ultraviolet background, the Ly-α forest probes density fluctuations from cosmological down to Mpc scales and below over a wide redshift range ($2 \leq z \leq 5$) using ground-based observations.

High-resolution spectra have enabled analyses deep into the small-scale regime ($k_{\max} \lesssim 10\,h\,\mathrm{Mpc}^{-1}$) through measurements of the line-of-sight (or one-dimensional) power spectrum [5–12]. These measurements are sensitive to a broad range of fundamental physics, including neutrino properties [5, 8, 13–17], primordial black holes [18–20], dark matter models [21–31], the thermal history of the ionized IGM [32–43], non-minimal cosmological models [44–46], and the running of the spectral index [15, 47].

Over the past two decades, Ly-α forest surveys have expanded dramatically in both spectral resolution and the total cosmological volume probed. Large samples of medium-resolution quasar spectra from the extended Baryon Oscillation Spectroscopic Survey [eBOSS; 48] and, in particular, from the ongoing Dark Energy Spectroscopic Instrument [DESI; 49–52] now enable cross-correlation analyses across many independent lines of sight. These multi-skewer maps trace the large-scale structure of our Universe, containing information similar to that provided by spectroscopic galaxy and quasar surveys at high redshift, but with a significantly wider dynamic range [49, 53, 54].

Observations of the Ly-α forest constrain the expansion history of the Universe through measurements of the baryon acoustic oscillation (BAO) feature [51, 55–58] and through the broadband shape of the three-dimensional correlation function [53, 56, 59–61]. The mapping between neutral hydrogen and underlying dark matter has motivated the development of high-fidelity numerical simulations of the Ly-α forest [62–72]. These simulations are commonly used to calibrate linear-theory-based models augmented by phenomenological fitting functions that account for nonlinear growth, pressure smoothing, and line-of-sight velocity broadening [73–75].

At the current precision set by the finite sampling of quasar sightlines (i.e., the shot-noise limited regime for galaxies), such modeling approaches have yielded robust cosmological constraints [50, 51, 61, 76]. However, DESI is expected to observe up to one million quasar spectra over its lifetime, with forecasts indicating a cumulative precision below the 0.2% level when combining all tracers and redshift bins [49]. Achieving this level of precision requires exquisite control over theoretical systematics. Indeed, recent studies have demonstrated that phenomenological Ly-α forest models bias the inferred BAO scaling parameters – the primary observables for constraining the cosmic expansion history with DESI – at the 0.3% level [77, 78]. These findings indicate that existing modeling frameworks are approaching their limits, motivating the development of more accurate and robust theoretical descriptions for current DESI data and forthcoming surveys such as DESI-II, the WEAVE-QSO survey [79], the Prime Focus Spectrograph [PFS; 80] and 4MOST [81].

One can remove the bias on the BAO inference within the framework of the effective field theory (EFT) of large-scale structure, which has recently been extended to the Ly-α forest [77, 85–87]. The EFT formalism provides a perturbative description of large-scale dynamics by incorporating only the symmetries relevant to the tracer [88–91]. In the case of the Ly-α forest, these include the equivalence principle and rotational invariance around the line-of-sight direction $\hat{z}$, corresponding to the $SO(2)$ group [15, 85, 86, 92–95]. Whilst this paves the way to directly constrain cosmological parameters through the one-loop power spectrum (in the context of low-redshift galaxy surveys, see, e.g., [96–101]) a key challenge for cosmological analyses is that the large range of scales involved require large-volume, high-resolution simulations to validate inference pipelines.

High-resolution hydrodynamic simulations accurately capture the involved physics on small to intermediate scales but running a set of simulations covering a wide range of cosmological and astrophysical parameters, whilst capturing

---

一. 简介与主要结果

Lyman-α（Ly-α）森林由一系列出现在遥远类星体光谱中的吸收特征组成，这些特征是由视线方向上经过的中性氢产生的。自20世纪90年代以来，使用高分辨率仪器如高分辨率伊歇尔光谱仪（HIRES；1, 2）和紫外-可见伊歇尔光谱仪（UVES；3, 4）进行的高分辨率观测，使人们能够精确测量中性氢密度的小尺度波动。这些观测结果表明，产生吸收的气体存在于低密度、高度电离的星际介质（IGM）中，为中性氢分布与底层暗物质场之间提供了透明的联系。这一联系反过来又使得对Ly-α森林的精确模拟成为可能。由于中性气体与近似均匀的紫外背景处于光电离平衡，Ly-α森林可以利用地基观测在广泛的红移范围（2 ⩽ z ⩽ 5）内探测从宇宙学尺度到Mpc尺度及以下的密度波动。高分辨率光谱使得通过测量视线方向（或一维）功率谱[5 – 12]能够深入分析小尺度区域（k_max < ~10 h Mpc^{-1}）。这些测量对广泛的基础物理非常敏感，包括中微子性质[5, 8, 13 – 17]、原初黑洞[18 – 20]、暗物质模型[21 – 31]、电离IGM的热历史[32 – 43]、非最小宇宙学模型[44 – 46]以及谱指数的漂移[15, 47]。

在过去的二十年里，Ly-α森林调查在光谱分辨率和所探测的总体宇宙体积方面都显著扩展。来自扩展的重子声波振荡光谱调查（eBOSS；48）的大量中分辨率类星体光谱样本，尤其是来自正在进行的暗能量光谱仪（DESI；49 – 52）的样本，现在能够在许多独立视线之间进行交叉相关分析。这些多光游标图描绘了我们宇宙的大尺度结构，包含的信息类似于高红移光谱星系和类星体调查提供的信息，但具有显著更宽的动态范围[49, 53, 54]。

通过观测 Ly-α 森林，可以通过测量重子声学振荡（BAO）特征 [51, 55 – 58] 以及三维相关函数的宽带形状 [53, 56, 59 – 61] 来约束宇宙的膨胀历史。中性氢与潜在暗物质之间的映射促使了高保真 Ly-α 森林数值模拟的发展 [62 – 72]。这些模拟通常用于校准基于线性理论的模型，并辅以考虑非线性增长、压力平滑以及视线方向速度展宽的现象学拟合函数 [73 – 75]。

在由类星体视线的有限采样所设定的当前精度下（即在银河的射击噪声限制区域），此类建模方法已经产生了稳健的宇宙学约束 [50, 51, 61, 76]。然而，预计 DESI 在其运行期间将观测高达一百万条类星体光谱，预测显示在结合所有示踪器和红移区间时，累积精度将低于 0.2% 的水平 [49]。达到这一精度水平需要对理论系统误差进行极其精确的控制。事实上，近期研究表明，现象学 Ly-α 林模型会使推断的 BAO 缩放参数产生偏差——这些是利用 DESI 约束宇宙膨胀历史的主要观测量——偏差水平为 0.3% [77, 78]。这些发现表明，现有的建模框架正接近其极限，这推动了为当前的DESI数据以及即将到来的调查（如DESI-II、WEAVE-QSO调查[79]、主焦点光谱仪[PFS; 80]和4MOST[81]）开发更准确和更稳健的理论描述。

通过大尺度结构有效场论（EFT）的框架，可以消除对BAO推断的偏差，最近该框架已扩展至Ly-α森林 [77, 85 – 87]。EFT形式主义通过仅纳入与示踪物相关的对称性，对大尺度动力学提供微扰描述 [88 – 91]。在Ly-α森林的情况下，这些对称性包括等效原理和沿视线方向ˆz的旋转不变性，对应于SO(2)群 [15, 85, 86, 92 – 95]。虽然这为通过单环功率谱直接约束宇宙学参数铺平了道路（在低红移星系调查的背景下，参见例如 [96 – 101]），但对宇宙学分析来说，一个关键挑战是所涉及的大范围尺度需要大体积、高分辨率的模拟来验证推断流程。

高分辨率的流体动力学模拟能够在小到中等尺度上准确捕捉所涉及的物理过程，但在覆盖广泛的宇宙学和天体物理参数范围的同时进行一系列模拟

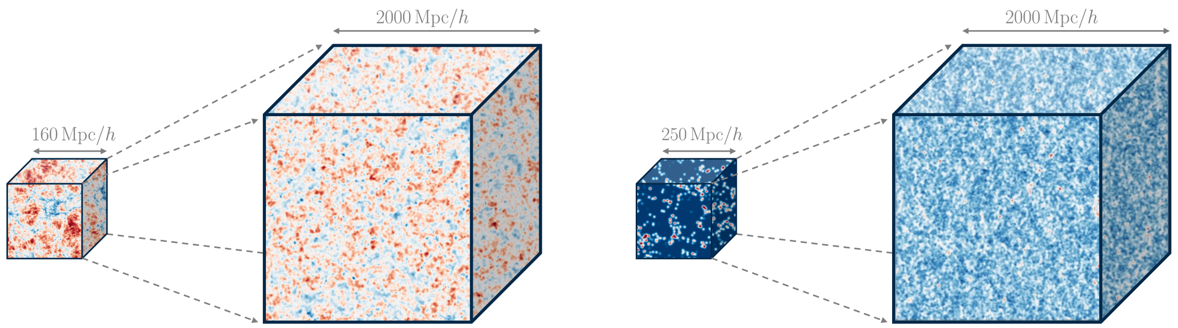

FIG. 1. **Summary of Results:** We fit an analytic, perturbative forward model at the field level to high-resolution hydrodynamic simulations, and use it to generate large-volume realizations. In the left panel, we fit small-volume ($V = (160\,h^{-1}\mathrm{Mpc})^3$) Sherwood hydrodynamic simulations of the Ly-α forest [82]. In the right panel, a similar framework is applied to high-redshift galaxy simulations (such as Lyman-α emitters and Lyman-break galaxies) from Astrid in a $(250\,h^{-1}\mathrm{Mpc})^3$ volume [83, 84]. These calibrated models enable the generation of large-volume simulations (here: $V = (2000\,h^{-1}\mathrm{Mpc})^3$) here depicted as larger 3D boxes. Red (blue) regions correspond to over- (under-)dense regions. The field-level fits benefit from cosmic variance cancellation when using the same set of initial conditions (ICs) for the forward model as for the input hydrodynamic simulation. Coincidentally, two simulations calibrated on different simulations using the same ICs can be used for cross-correlation analyses of the Ly-α forest and high-redshift galaxies – a key science driver for DESI-II.

cosmological volumes is prohibitively expensive. To mitigate this, several approaches have been developed to connect the observed flux to the underlying matter field for large-scale clustering simulations [78, 102–109]. In particular, deep learning-based reconstruction methods [110–114] yield promising results at intermediate to small scales but do not fully capture the long-wavelength quasi-linear modes. An alternative avenue are emulators that directly predict a summary statistic such as the power spectrum [115–118]. Whilst these offer per cent level accuracy for a wide range of scales they depend on (i) fits to power spectra of small-volume hydrodynamic simulations suffering from cosmic variance; and (ii) are restricted to the power spectrum. Whilst paired-fixed simulations reduce cosmic variance limitations [119], the resulting emulators cannot directly be generalized to other summary statistics. A forward model at the field level, however, needs to match *all* the amplitudes and phases of *all* Fourier modes – a more stringent test of the theoretical framework than comparing summary statistics – all whilst yielding higher-order moments of the field.

To enable cosmological analyses of the Ly-α forest from DESI and DESI-II that incorporate two- and, eventually, three-point statistics, we present a computationally efficient framework for generating large-volume simulations calibrated on high-fidelity hydrodynamic simulations. This work provides the theoretical background for Ref. [87], where this technique was first introduced for the Ly-α forest, and extends earlier developments in perturbative, field-level modeling [120–125]. The same framework generalizes to simulations of high-redshift galaxies, enabling validation of cross-correlation measurements between the Ly-α forest and galaxy positions. Such joint analyses help break the degeneracy between the growth rate f and the otherwise poorly constrained velocity-gradient bias [126, 127]. More broadly, the simulations developed here provide a controlled environment for validating end-to-end full-shape inference pipelines and for quantifying key physical systematics, including shifts of the BAO feature [77, 128].

We validate our perturbative forward model using two complementary sets of simulations. First, we employ the large-volume N-body simulations from the AbacusSummit suite, onto which the Ly-α forest is painted, spanning a volume of $V = 2^3\,(h^{-1}\,\mathrm{Gpc})^3$ [102]. Second, we use two sets of the Sherwood hydrodynamic simulations, which cover a volume of $V = 160^3\,(h^{-1}\,\mathrm{Mpc})^3$ and $V = 80^3\,(h^{-1}\,\mathrm{Mpc})^3$ [82]. A key advantage of performing field-level fits is the resulting cancellation of cosmic variance, which enables tight constraints on both cosmological and nuisance parameters even for small-volume simulations.

In this work, we present two main results. First, we fit and *predict* the shapes of the bias transfer functions using a perturbative bias expansion directly at the field level. Second, using the resulting field-level fits, we generate large-scale clustering simulations. Examples of these simulations are shown in Fig. 1, with the Ly-α forest displayed in the left panel and high-redshift galaxies in the right panel. The methodology developed here is critical for forthcoming cosmological analyses, as existing large-scale clustering mocks [129, 130] are approaching their limits when targeting scales an order of magnitude smaller than the current baseline for analyses of the Ly-α broadband shape of $r \approx 25\ h^{-1}\,\mathrm{Mpc}$ [61]. Especially, since the Ly-α forest is one of the only ways of probing the high-redshift Universe ($2 \lesssim z \lesssim 5$) ahead of next-generation surveys such as DESI-II and Spec-S5 [131].

This paper provides the theoretical foundation of the Ly-α forward model presented in a *Letter* in Ref. [87] and is organized as follows: We review the perturbative forward model of the Ly-α forest and halos (as proxies for high-redshift galaxies) in redshift space in Sec. II. In Sec. III we present the used synthetic data. We assess the performance of our perturbative forward model in Sec. IV and investigate the obtained transfer functions from our field-level fits in

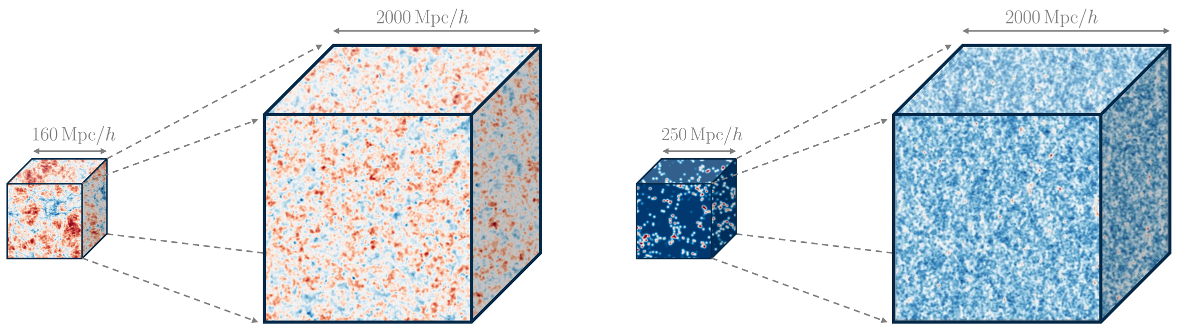

图 1. 结果摘要：我们在场域层面将解析的、微扰的前向模型拟合到高分辨率流体动力学模拟，并使用它生成大体积实现。在左图中，我们拟合了小体积 ($V = (160\ \mathrm{h-1Mpc})^3$) 的 Sherwood 氢森林流体动力学模拟 [82]。在右图中，类似的框架被应用于高红移星系模拟（如 Lyman-α 发射体和 Lyman 跃迁断裂星系），来自 Astrid 的 $(250\ \mathrm{h-1Mpc})^3$ 体积 [83, 84]。这些校准后的模型能够生成大体积模拟（此处：$V = (2000\ \mathrm{h-1Mpc})^3$），在图中显示为更大的三维盒子。红色（蓝色）区域对应过密（欠密）区域。当前向模型与输入的流体动力学模拟使用相同的初始条件 (ICs) 时，场域层面的拟合能够受益于宇宙方差的抵消。巧合的是，两个使用相同初始条件在不同模拟上校准的模拟可以用于Ly-α森林和高红移星系的交叉相关分析——这是DESI-II的一个关键科学驱动因素。

宇宙学体积的计算成本高得令人望而却步。为减轻这一问题，已经开发了几种方法，将观测到的通量与大规模聚类模拟的基础物质场联系起来 [78, 102 – 109]。特别是，基于深度学习的重建方法 [110 – 114] 在中到小尺度上取得了良好效果，但无法完全捕捉长波长的准线性模式。另一种途径是直接预测如功率谱等统计摘要的仿真器 [115 – 118]。虽然这些方法在广泛的尺度范围内能够达到百分比级的准确度，但它们依赖于 (i) 小体积流体动力学模拟的功率谱拟合，而这些模拟受到宇宙方差的影响；(ii) 并且仅限于功率谱。虽然配对固定（paired-fixed）模拟可以降低宇宙方差的限制 [119]，但由此得到的仿真器不能直接推广到其他统计摘要。然而，场级的前向模型需要匹配所有傅里叶模态的所有振幅和相位——这比比较摘要统计量对理论框架的检验更严格——同时还要产生场的高阶矩。

为了实现对来自 DESI 和 DESI-II 的 Ly-α 森林的宇宙学分析，并纳入二点及最终的三点统计，我们提出了一个计算上高效的框架，用于生成基于高保真流体力学模拟校准的大体积模拟。这项工作为文献 [87] 提供了理论背景，该文献首次将该技术引入 Ly-α 森林，并扩展了早期在微扰、场级建模方面的发展 [120 – 125]。相同的框架还可以推广到高红移星系的模拟，使得验证 Ly-α 森林与星系位置之间的交叉相关测量成为可能。这种联合分析有助于打破增长率 f 与其他很难约束的速度梯度偏差之间的简并性 [126, 127]。更广泛地说，这里开发的模拟为验证端到端的全形状推断流程以及量化关键物理系统效应提供了一个受控环境，包括BAO特征的位移 [77, 128]。

我们使用两组互补的模拟来验证我们的微扰前向模型。首先，我们采用 AbacusSummit 系列的大体积 N 体模拟，并在其上绘制了 Ly-α 森林，覆盖体积为 $V = 23(\ \mathrm{h-1Gpc})^3$[102]。其次，我们使用两组 Sherwood 流体动力学模拟，覆盖体积分别为 $V = 160^3(\ \mathrm{h-1Mpc})^3$ 和 $V = 80^3(\ \mathrm{h-1Mpc})^3$[82]。进行场级拟合的一个关键优势是能够抵消宇宙方差，即使对于小体积模拟，也可以对宇宙学参数和干扰参数进行严格约束。

在这项工作中，我们提出了两个主要结果。首先，我们直接在场层次上使用微扰偏差展开来拟合和预测偏差传递函数的形状。其次，利用得到的场层次拟合结果，我们生成大规模聚类模拟。这些模拟的示例如图 1 所示，左面板显示了 Ly-α 森林，右面板显示了高红移星系。这里开发的方法对于即将到来的宇宙学分析至关重要，因为现有的大规模聚类模拟 [129, 130] 在针对比当前 Ly-α 宽带形状分析基准 $r \approx 25\ \mathrm{h-1Mpc}$ 小一个数量级的尺度时已经接近其极限 [61]。特别是由于 Ly-α 森林是探测高红移宇宙（$2 \lesssim z \lesssim 5$）的唯一途径之一，为未来的下一代测量如 DESI-II 和 Spec-S5 [131] 做准备。

本文提供了在参考文献 [87] 的一篇快报中提出的 Ly-α 前向模型的理论基础，组织结构如下：我们在第 II 节回顾了 Ly-α 森林和晕（作为高红移星系的代理）在红移空间中的微扰前向模型。在第 III 节中，我们介绍了所使用的合成数据。我们在第 IV 节评估了微扰前向模型的性能，并研究了从我们的场级拟合中获得的传递函数。

Sec. V. In Sec. VI we compare the measured transfer functions to theoretically expected ones and use these transfer functions to create large-scale clustering mocks encompassing cosmological volumes in Sec. VII. We conclude and discuss future work in Sec. VIII.

II. BUILDING THE FIELD-LEVEL MODEL

The description of the cosmological Ly-α forest correlations begins with the linear theory model [73, 92],

$$\delta_F^{\rm lin}(\boldsymbol{k},z) = (b_1 - b_\eta f\mu^2)\delta_1(\boldsymbol{k},z) \ , \tag{1}$$

where b_1, b_η are linear bias parameters, $f = d\ln D_+/d\ln a$ is the logarithmic growth factor, $\mu \equiv k_\parallel/k = \hat{\mathbf{k}}\cdot\hat{\boldsymbol{z}}$ is the cosine of the angle to the line-of-sight, $\hat{\boldsymbol{z}}$, and δ_1 the linear density field, which can be rewritten as

$$\delta_1(\boldsymbol{k},z) = \delta_{1,0}(\boldsymbol{k})D_+(z)\,, \tag{2}$$

where $\delta_{1,0}(\boldsymbol{k})$ is the initial condition field extrapolated to redshift zero and $D_+(z)$ is the linear growth factor. In simulations, $\delta_{1,0}(\boldsymbol{k})$ is the random scalar field generated from the linear matter power spectrum of initial conditions. For brevity, we will suppress the time dependence of δ_1, implicitly assuming that this quantity is always evaluated at the simulation redshift z. The model in Eq. (1) reproduces the well-known linear theory model for the flux power spectrum,

$$\begin{aligned}\langle\delta_F(\boldsymbol{k})\delta_F(\boldsymbol{k}')\rangle &= (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}+\boldsymbol{k}')\langle\delta_F(\boldsymbol{k})\delta_F(\boldsymbol{k}')\rangle' = (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}+\boldsymbol{k}')P(k,\mu) \ , \\ P_{11}(k) &\underbrace{=}_{\rm linear}(b_1 - fb_\eta\mu^2)^2P_{\rm lin}(k)\,,\end{aligned} \tag{3}$$

where $P_{\rm lin}(k)$ is the linear matter power spectrum (evaluated at redshift z) where we use $\int_{\boldsymbol{k}} = \int d^3k/(2\pi)^3$. The above coefficient $(b_1 - fb_\eta\mu^2)^2$ is the well-known generalization of the Kaiser factor for galaxies [73, 92, 132], which can be recovered by setting $b_\eta = -1$.

From this discussion it is clear that all phases of the Ly-α field are captured by the field $\delta_1(\boldsymbol{k})$ in the linear approximation. It is thus convenient to split the theory model from Eq. (1) into parts that make the amplitude and phase dependence manifest. Absorbing the Kaiser factor into Eq. (1) into a momentum-dependent transfer function yields

$$\delta_F^{\rm lin}(\boldsymbol{k},z) = \beta_1(k,\mu)\delta_1(\boldsymbol{k},z)\,. \tag{4}$$

The advantage of this approach will become evident at the non-linear level, where the transfer function β_1 will also account for the one-loop corrections.

The success of the linear theory model in describing the simulated Ly-α field $\delta_F^{\rm truth}$ can be estimated using the error power spectrum,

$$P_{\rm err}(\boldsymbol{k}) = \langle|\delta_F^{\rm truth}(\boldsymbol{k}) - \delta_F^{\rm lin}(\boldsymbol{k})|^2\rangle' \ . \tag{5}$$

On very large scales (i.e. in the limit $k \to 0$), the non-linear corrections to our model are expected to be small, so the only expected source of error in the model should be the stochastic field ϵ,

$$\delta_F(\boldsymbol{k}) = \delta_F^{\rm lin}(\boldsymbol{k}) + \epsilon(\boldsymbol{k}) \ , \tag{6}$$

which by definition does not correlate with $\delta_F^{\rm lin}$ generated by cosmological fluctuations. In general, $\epsilon(\boldsymbol{k})$ is the Ly-α flux decrement component produced by small-scale processes unrelated to large-scale initial conditions. In the context of galaxies and halos, the field $\epsilon(\boldsymbol{k})$ captures the shot noise which arises due to the discreteness of tracers. Using this analogy, the stochastic field of the Ly-α forest could be thought to originate from the discreteness of absorption lines. Due to their large numbers it is expected to be extremely small. This logic, however, is not fully correct because it assumes that the discreteness is generated at the level of absorption lines. One can consider that Ly-α absorption is produced by neutral hydrogen clouds whose distribution also has a stochastic component. This discreteness of the Ly-α clouds can be used as a first proxy to understand the stochasticity of the forest.

Within the EFT for LSS framework [89–91, 94, 133–135], one uses perturbative Taylor expansions to parameterize unknown functions. In this approach, the error power spectrum assumes the following expansion valid in the $k \to 0$ limit:

$$P_{\rm err}(k,\mu) = n_0(1 + \alpha_1k^2 + \alpha_2k^2\mu^2 + ...)\,, \tag{7}$$

第五节。在第六节中，我们将测量的传递函数与理论上预期的传递函数进行比较，并在第七节中使用这些传递函数创建涵盖宇宙卷的大片聚类模拟。我们在第八节总结并讨论未来的工作。

二.　建立田间级模型

宇宙学 Ly-α 森林相关性的描述从线性理论模型 [73, 92] 开始，

$$\delta_F^{\rm lin}(\boldsymbol{k},z) = (b_1 - b_\eta f\mu^2)\delta_1(\boldsymbol{k},z) \ , \tag{1}$$

其中 b1, bη 是线性偏置参数，f = d ln D+/ 1 是线性密度场，可以重写为

$$\delta_1(\boldsymbol{k},z) = \delta_{1,0}(\boldsymbol{k})D_+(z)\,, \tag{2}$$

其中 δ1,0(k) 是外推到红移为零的初始条件场，D+(z) 是线性增长因子。在模拟中，δ1,0(k) 是从初始条件的线性物质功率谱生成的随机标量场。为了简洁起见，我们将省略 δ1 的时间依赖性，隐含假设该量总是在模拟红移 z 下进行评估。方程 (1) 中的模型重现了众所周知的流量功率谱线性理论模型，

$$\begin{aligned}\langle\delta_F(\boldsymbol{k})\delta_F(\boldsymbol{k}')\rangle &= (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}+\boldsymbol{k}')\langle\delta_F(\boldsymbol{k})\delta_F(\boldsymbol{k}')\rangle' = (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}+\boldsymbol{k}')P(k,\mu) \ , \\ P_{11}(k) &\underbrace{=}_{\rm linear}(b_1 - fb_\eta\mu^2)^2P_{\rm lin}(k)\,,\end{aligned} \tag{3}$$

其中 Plin(k) 是线性物质功率谱（在红移 z 处计算），我们使用 k= d3k/(2π)³。该上述系数 (b1 − fbημ2)² 是众所周知的星系 Kaiser 因子的推广 [73, 92, 132]，通过设定 bη = −1 可以恢复。由此讨论可以清楚地看出，在线性近似下，Ly-α 场的所有相位都被场 δ1(k) 捕获。因此，将方程 (1) 的理论模型拆分为显式显示振幅和相位依赖的部分是很方便的。将 Kaiser 因子吸收到方程 (1) 中作为动量依赖的传递函数会得到

$$\delta_F^{\rm lin}(\boldsymbol{k},z) = \beta_1(k,\mu)\delta_1(\boldsymbol{k},z)\,. \tag{4}$$

这种方法的优势将在非线性层面显现出来，此时转移函数β1也将考虑一环修正。

线性理论模型在描述模拟 Ly-α 场 δtruth F 中的成功 可以使用误差来估计功率谱

$$P_{\rm err}(\boldsymbol{k}) = \langle|\delta_F^{\rm truth}(\boldsymbol{k}) - \delta_F^{\rm lin}(\boldsymbol{k})|^2\rangle' \ . \tag{5}$$

在非常大尺度上（即在 k →0 的极限下），我们模型的非线性修正预计很小，因此模型中唯一预期的误差来源应当是随机场 ϵ,

$$\delta_F(\boldsymbol{k}) = \delta_F^{\rm lin}(\boldsymbol{k}) + \epsilon(\boldsymbol{k}) \ , \tag{6}$$

从定义上讲，这与 δlin F 无关 由宇宙学涨落产生。一般来说，ϵ(k) 是 Ly-α 由与大尺度初始条件无关的小尺度过程产生的通量减量分量。在星系和晕的背景下，场ϵ(k)捕捉到由追踪体离散性引起的射击噪声。使用这个类比，Ly-α森林的随机场可以被认为源自吸收线的离散性。由于吸收线的数量众多，预计其非常小。然而，这一逻辑并不完全正确，因为它假设离散性是在吸收线层面上产生的。可以认为Ly-α吸收是由中性氢云产生的，而这些氢云的分布也具有随机成分。Ly-α云的这种离散性可以作为理解森林随机性的第一个代理。

在 LSS 的 EFT 框架内 [89 − 91, 94, 133 − 135]，人们使用微扰泰勒展开来参数化未知函数。在这种方法中，误差功率谱在 k →0 极限下具有以下展开式：

$$P_{\rm err}(k,\mu) = n_0(1 + \alpha_1k^2 + \alpha_2k^2\mu^2 + ...)\,, \tag{7}$$

where $n_0, \alpha_{1,2}$ are dimensional constants whose values are set by the Ly-α physics. For galaxies, n_0 can be estimated at leading order as $1/\bar{n}$, where $\bar{n} = N/V$ is the galaxy number density (N is the number of galaxies in the comoving volume V). The relevant distance scale in this case is the mean separation $R = (V/N)^{1/3}$ between the individual galaxies. Applying the same argument to the Ly-α forest and assuming for simplicity that the individual tracer is a neutral hydrogen cloud whose separation to its neighboring cloud is about 0.5 h^{-1}Mpc, we get an estimate

$$n_0 \sim R^3 \sim 0.1\ [h^{-1}\mathrm{Mpc}]^3 \quad \alpha_{1,2} \sim R^2 \sim 0.3\ [h^{-1}\mathrm{Mpc}]^2\ . \tag{8}$$

As we shall see, this naive estimate will turn out to be quite accurate for the actual Ly-α simulations. Importantly, the stochasticity expansion features scales that are in general different from those that appear in the perturbative bias expansion. The latter can be estimated as a non-linear scale, where matter density fluctuations become of order one [136],

$$\frac{k_{\mathrm{NL}}^3}{2\pi^2} P_{\mathrm{lin}}(k_{\mathrm{NL}}, z) = 1\,, \quad \Rightarrow \quad k_{\mathrm{NL}} \approx 5\ h\mathrm{Mpc}^{-1} \quad \text{at} \quad z = 2.8\,. \tag{9}$$

If the theory model is accurate, we expect to recover the scale-dependence suggested by Eq. (7) on large scales. If there is a significant correction to the naive linear model (4), this will generate a noticeable scale and orientation dependence not captured by Eq. (7). What corrections do we expect?

First, the displacements of dark matter particles in our simulations are large, and have to be treated non-perturbatively. If unaccounted for, they produce large distortions of baryon acoustic oscillations, which show up as a mismatch between the linear model and actual phases of the density field. This effect is well understood, and can be corrected for by using the linear density field $\tilde{\delta}_1$ shifted by the Zel'dovich displacement in lieu of δ_1 [137]. In what follows $\tilde{\delta}_1$ will be referred to as the shifted linear field.

$$\tilde{\delta}_1(\boldsymbol{k}) = \int d^3\boldsymbol{q}\ \delta_1(\boldsymbol{q}) e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}_1(\boldsymbol{q})+f\hat{\boldsymbol{z}}(\boldsymbol{\psi}_1(\boldsymbol{q})\cdot\hat{\boldsymbol{z}}))}\ , \tag{10}$$

where $\boldsymbol{q}$ denotes Lagrangian space (initial) coordinates, and $\boldsymbol{\psi}_1$ is the Zel'dovich displacement

$$\boldsymbol{\psi}_1(\boldsymbol{q}) = \int d^3k\, e^{i\boldsymbol{q}\cdot\boldsymbol{k}} \frac{i\boldsymbol{k}}{k^2} \delta_1(\boldsymbol{k}). \tag{11}$$

Note that $\tilde{\delta}_1$ above has an infinite Taylor expansion in the linear field δ_1. In perturbation theory one can write this as:

$$\tilde{\delta}_1 = \sum_{n=1}^{3} \left(\prod_{i=1}^{n} \int_{\boldsymbol{k}_i} \delta_1(\boldsymbol{k}_i) \right) (2\pi)^3 \delta_D^{(3)}(\boldsymbol{k} - \boldsymbol{k}_{1\ldots n}) \tilde{K}_n(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n)\,, \tag{12}$$

where $\tilde{K}_1(\boldsymbol{k}) = 1$ and

$$\tilde{K}_2(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{\boldsymbol{k}\cdot\boldsymbol{k}_1}{2k_1^2} + \frac{\boldsymbol{k}\cdot\boldsymbol{k}_2}{2k_2^2} + \frac{(f\mu k)}{2}\left(\frac{k_{1z}}{k_1^2} + \frac{k_{2z}}{k_2^2}\right)\,, \quad \text{etc.} \tag{13}$$

where $k_{iz} = (\boldsymbol{k}_i \cdot \hat{\boldsymbol{z}})$, and $\boldsymbol{k} \equiv \boldsymbol{k}_1 + \ldots + \boldsymbol{k}_n$ for the n'th kernel. This expansion is very similar to that of the Zel'dovich matter density field [137–140]

$$\delta_{\mathrm{Z}}(\boldsymbol{k}) = \int d^3\boldsymbol{q}\ e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}_1(\boldsymbol{q})+f\hat{\boldsymbol{z}}(\boldsymbol{\psi}_1(\boldsymbol{q})\cdot\hat{\boldsymbol{z}}))} = \sum_{n=1} \left(\prod_{i=1}^{n} \int_{\boldsymbol{k}_i} \delta_1(\boldsymbol{k}_i) \right) (2\pi)^3 \delta_D^{(3)}(\boldsymbol{k} - \boldsymbol{k}_{1\ldots n}) F_n^{\mathrm{ZA}}(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n)\,, \tag{14}$$

with $F_1^{\mathrm{ZA}} = 1 + f(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{z}})^2$, and

$$F_2^{\mathrm{ZA}} = \frac{1}{2}\frac{(\boldsymbol{k}\cdot\boldsymbol{k}_1)(\boldsymbol{k}\cdot\boldsymbol{k}_2)}{k_1^2k_2^2} + \frac{f}{2}\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k})k_{2z}(k\mu)}{k_1^2k_2^2} + \frac{f}{2}\frac{(\boldsymbol{k}_2\cdot\boldsymbol{k})k_{1z}(k\mu)}{k_1^2k_2^2} + \frac{f^2(\mu k)^2}{2k_1^2k_2^2}k_{1z}k_{2z}\,, \quad \text{etc.} \tag{15}$$

These building blocks will be useful in our future discussion. The shifts implemented by the Zel'dovich displacement introduce higher order non-linear effects. In the following, we will discuss these effects more systematically.

Nonlinearities in the bias expansion are a second important source of corrections. In the perturbative Eulerian bias formulation these are given by

$$\delta_F(\boldsymbol{k}) = \sum_{n=1} \delta_F^{(n)} = \sum_{n=1} \Big[\prod_{j=1}^{n} \int \frac{d^3\mathbf{k}_j}{(2\pi)^3} \delta_1(\boldsymbol{k}_j) \Big] K_n(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n) (2\pi)^3 \delta_D^{(3)}(\boldsymbol{k} - \boldsymbol{k}_1 - \ldots - \boldsymbol{k}_n)\,, \tag{16}$$

其中 n0、α1,2 是由 Ly-α 物理决定的维度常数。对于星系，n0 可以在领先阶近似为 1/¯ n，其中 ¯ n = N/V 是星系的数密度（N 是共动体积 V 中星系的数量）。在这种情况下，相关的距离尺度是各个星系之间的平均间距 R = (V/N)1/3。将相同的推理应用到 Ly-α 林，并为了简化假设单个示踪物是一个中性氢云，其与相邻云的间隔约为 0.5 h−1Mpc，我们得到一个估计

$$n_0 \sim R^3 \sim 0.1\ [h^{-1}\mathrm{Mpc}]^3 \quad \alpha_{1,2} \sim R^2 \sim 0.3\ [h^{-1}\mathrm{Mpc}]^2\ . \tag{8}$$

正如我们将看到的，这一天

不同于微扰偏差展开中出现的尺度。后者可以被估计为非线性尺度，在该尺度上物质密度波动达到数量级一 [136]。

$$\frac{k_{\mathrm{NL}}^3}{2\pi^2} P_{\mathrm{lin}}(k_{\mathrm{NL}}, z) = 1\,, \quad \Rightarrow \quad k_{\mathrm{NL}} \approx 5\ h\mathrm{Mpc}^{-1} \quad \text{at} \quad z = 2.8\,. \tag{9}$$

如果理论模型是准确的，我们期望在大尺度上恢复由方程 (7) 提示的尺度依赖性。如果对天真的线性模型 (4) 存在显著修正，这将产生方程 (7) 没有捕捉到的明显尺度和方向依赖性。我们预计会有哪些修正?

首先，我们模拟中的暗物质粒子位移很大，必须以非微扰方式处理。如果不加以考虑，它们会对重子声学振荡产生较大扭曲，从而导致线性模型与密度场的实际相位不匹配。这种效应是可以理解的，并且可以通过使用由泽尔多维奇位移移动的线性密度场 ˜ δ1 代替 δ1 来加以修正 [137]。在下文中，˜ δ1 将被称为移位线性场。

$$\tilde{\delta}_1(\boldsymbol{k}) = \int d^3\boldsymbol{q}\ \delta_1(\boldsymbol{q}) e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}_1(\boldsymbol{q})+f\hat{\boldsymbol{z}}(\boldsymbol{\psi}_1(\boldsymbol{q})\cdot\hat{\boldsymbol{z}}))}\ , \tag{10}$$

其中 q 表示拉格朗日空间（初始）坐标，ψ1 是泽尔多维奇位移

$$\boldsymbol{\psi}_1(\boldsymbol{q}) = \int d^3k\, e^{i\boldsymbol{q}\cdot\boldsymbol{k}} \frac{i\boldsymbol{k}}{k^2} \delta_1(\boldsymbol{k}). \tag{11}$$

请注意，上述˜ δ1在线性场δ1中具有无限的泰勒展开。在微扰理论中，可以将其写为:

$$\tilde{\delta}_1 = \sum_{n=1}^{3} \left(\prod_{i=1}^{n} \int_{\boldsymbol{k}_i} \delta_1(\boldsymbol{k}_i)\ (2\pi)^3 \delta_D^{(3)}(\boldsymbol{k} - \boldsymbol{k}_{1\ldots n}) \tilde{K}_n(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n)\,, \right. \tag{12}$$

其中 ˜ K1(k) = 1 并且

˜ K2(k1, k2) = k · k1 2k2 1 + $\frac{\boldsymbol{k}\cdot\boldsymbol{k}_2}{2k_2^2} + \frac{(f\mu k)}{2}\left(\frac{k_{1z}}{k_1^2} + \frac{k_{2z}}{k_2^2}\right)$, etc. (13)

其中 kiz = (ki · ˆ z)，且 k ≡k1 + ... + kn 对于第 n 个核。这一展开与 Zel’dovich 物质密度场的展开非常相似 [137 − 140]

$\delta_{\mathrm{Z}}(\boldsymbol{k}) = \int$ d3q e−ik · (q+ψ1(q)+f ˆ z(ψ1(q) · ˆ z)) = $\sum_{n=1} \left(\prod_{i=1}^{n} \int_{\boldsymbol{k}_i} \delta_1(\boldsymbol{k}_i)\quad (2\pi)^3 \delta_D^{(3)}(\boldsymbol{k} - \boldsymbol{k}_{1\ldots n}) F_n^{\mathrm{ZA}}(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n)\right.$, (14)

FZA 1 = 1 + f(ˆ k · ˆ z)²，FZA 2 = 1/2

(k · k1)(k · k2) / (k1² k2²)

+ f/2 (k1 · k)k2z(kμ) k2 1k2 2 + f/2 (k2 · k)k1z(kμ) k2 1k2 2 + f 2(μk)² / 2k² 1k² 2 $k_{1z}k_{2z}$, etc. (15)

这些构建模块将在我们未来的讨论中非常有用。Zel'dovich 位移引入的变化会产生更高阶的非线性效应。在接下来的内容中，我们将更系统地讨论这些效应。

偏置展开中的非线性是第二个重要的修正来源。在微扰欧拉偏置中

这些由……给出

$$\delta_F(\boldsymbol{k}) = \sum_{n=1} \delta_F^{(n)} = \sum_{n=1} \Big[\prod_{j=1}^{n} \int \frac{d^3\mathbf{k}_j}{(2\pi)^3} \delta_1(\boldsymbol{k}_j) \Big] K_n(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_n) (2\pi)^3 \delta_D^{(3)}(\boldsymbol{k} - \boldsymbol{k}_1 - \ldots - \boldsymbol{k}_n)\,, \tag{16}$$

where K_n are non-linear kernels. General perturbative corrections to the power spectrum from individual powers in this series have the form

$$P_{nm} = s_{nm} \langle \delta_F^{(n)}(\boldsymbol{k}) \delta_F^{(m)}(-\boldsymbol{k}) \rangle' \,, \tag{17}$$

where s_{nm} is the combinatorial factor and $n = m = 1$ at the linear level, $n + m = 4, 6, 8$ etc. at the one, two, three-loop orders, respectively. At the quadratic order one has the following general set of operators consistent with the line-of-sight rotations [85, 94]:

$$K_2(\boldsymbol{k}_1, \boldsymbol{k}_2) \equiv b_1 F_\delta + b_2 F_{\delta^2} + b_{\mathcal{G}_2} F_{\mathcal{G}_2} + b_\eta F_\eta + b_{\delta\eta} F_{\delta\eta} + b_{\eta^2} F_{\eta^2} + b_{\Pi^{[2]}_\parallel} F_{\Pi^{[2]}_\parallel} + b_{(KK)_\parallel} F_{(KK)_\parallel} \;, \tag{18}$$

where the momentum-dependent quadratic kernels are given by

$$\begin{aligned} &F_{\delta^2} = \frac{1}{2} \,, \quad F_{\mathcal{G}_2} = \left(\frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2}{k_1^2 k_2^2} - 1 \right) \,, \quad F_\delta = 1 + \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{2k_1^2} + \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{2k_2^2} + \frac{2}{7} F_{\mathcal{G}_2} + f \frac{\mu_1 \mu_2}{2} \left(\frac{k_2}{k_1} + \frac{k_1}{k_2} \right) \,, \\ &F_\eta = -f\mu^2 \left(1 + \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{2k_1^2} + \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{2k_2^2} + \frac{4}{7} F_{\mathcal{G}_2} \right) - f^2 \frac{\mu_1 \mu_2}{2} \left(\frac{k_2}{k_1} \mu_2^2 + \frac{k_1}{k_2} \mu_1^2 \right) \,, \quad F_{\delta\eta} = -f \frac{\mu_2^2 + \mu_1^2}{2} \,, \\ &F_{\eta^2} = f^2 \mu_1^2 \mu_2^2 \,, \quad F_{(KK)_\parallel} = \mu_1 \mu_2 \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{k_1 k_2} - \frac{\mu_1^2 + \mu_2^2}{3} + \frac{1}{9} \,, \quad F_{\Pi^{[2]}_\parallel} = \mu_1 \mu_2 \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{k_1 k_2} - \frac{5}{7} \mu^2 F_{\mathcal{G}_2} \,, \end{aligned} \tag{19}$$

where $\mu_i \equiv \frac{\hat{\boldsymbol{z}} \cdot \boldsymbol{k}_i}{k_i}$ are the cosines between the line-of-sight and momentum vectors. One can see that the kernels above can be related to the shifted density and Zel'dovich kernels that we have introduced before. For instance, it is easy to see that

$$F_\delta = \tilde{K}_2 + F_{\delta\eta} + \frac{2}{7} F_{\mathcal{G}_2} \;. \tag{20}$$

Likewise, the kernel F_η can be rewritten as

$$F_\eta = F_\delta - \left(F_2^{\rm ZA} - \frac{3}{7} f\mu^2 F_{\mathcal{G}_2} \right) + \frac{3}{14} F_{\mathcal{G}_2} + F_{\eta^2} - F_{\delta\eta} = \tilde{K}_2 - \left(F_2^{\rm ZA} - \frac{3}{7} f\mu^2 F_{\mathcal{G}_2} \right) + \frac{1}{2} F_{\mathcal{G}_2} + F_{\eta^2} \;. \tag{21}$$

Finally, we have

$$F_{(KK)_\parallel} = F_{\Pi^{[2]}_\parallel} + \frac{5}{7} \mu^2 F_{\mathcal{G}_2} - \frac{\mu_1^2 + \mu_2^2}{3} + \frac{1}{9} = F_{\Pi^{[2]}_\parallel} + \frac{5}{7} \mu^2 F_{\mathcal{G}_2} + \frac{2}{3f} F_{\delta\eta} + \frac{2}{9} F_{\delta^2} \;. \tag{22}$$

Therefore, using the new line-of-sight velocity divergence field

$$\eta_{\rm new} \equiv -\left(\delta_Z - \frac{3}{7} f\mu^2 \mathcal{G}_2 \right) \,, \tag{23}$$

the quadratic kernels can be rewritten as

$$\begin{aligned} K_2(\boldsymbol{k}_1, \boldsymbol{k}_2) \equiv &(b_1 + b_\eta) \tilde{K}_2 + b_\eta F_{\eta_{\rm new}} + b_2 F_{\delta^2} + \left(\frac{2}{7} b_1 + \frac{1}{2} b_\eta + b_{\mathcal{G}_2} \right) F_{\mathcal{G}_2} \\ &+ (b_1 + b_{\delta\eta}) F_{\delta\eta} + (b_{\eta^2} + b_\eta) F_{\eta^2} + b_{\Pi^{[2]}_\parallel} \left(F_{(KK)_\parallel} - \frac{5}{7} \mu^2 F_{\mathcal{G}_2} - \frac{2}{3f} F_{\delta\eta} - \frac{2}{9} F_{\delta^2} \right) + b_{(KK)_\parallel} F_{(KK)_\parallel} \\ = &(b_1 + b_\eta) \tilde{K}_2 - b_\eta \left(F_2^{\rm ZA} - \frac{3}{7} f\mu^2 F_{\mathcal{G}_2} \right) + \left(b_2 - \frac{2}{9} b_{\Pi^{[2]}_\parallel} \right) F_{\delta^2} + \left(\frac{2}{7} b_1 + \frac{1}{2} b_\eta + b_{\mathcal{G}_2} - \frac{5}{7} \mu^2 b_{\Pi^{(2)}} \right) F_{\mathcal{G}_2} \\ &+ \left(b_1 + b_{\delta\eta} - \frac{2}{3f} b_{\Pi^{[2]}_\parallel} \right) F_{\delta\eta} + (b_{\eta^2} + b_\eta) F_{\eta^2} + (b_{(KK)_\parallel} + b_{\Pi^{[2]}_\parallel}) F_{(KK)_\parallel} \end{aligned} \tag{24}$$

If we use the k and μ dependent transfer functions, there are only seven independent quadratic operators: the effect of $\Pi^{(2)}_\parallel$ is fully absorbed into transfer functions. In addition, the velocity bias terms can be fully captured by the Zel'dovich field. This implies that at the quadratic level the full EFT can be described by the following forward model:

$$\begin{aligned} \delta_F^{\rm model}(\mathbf{k}) = &\beta_1^F(k,\mu) \tilde{\delta}_1(\mathbf{k}) + \beta_\eta^F(k,\mu) \left(\delta_Z(\mathbf{k}) - \frac{3}{7} f\mu^2 \tilde{\mathcal{G}}_2 \right) + \beta_2^F(k,\mu) (\tilde{\delta_1^2}) \\ &+ \beta_{\mathcal{G}_2}^F(k,\mu) \tilde{\mathcal{G}}_2(\mathbf{k}) + \beta_{\delta\eta}^F(k,\mu) [\tilde{\delta\eta}](\mathbf{k}) + \beta_{\eta^2}^F(k,\mu) (\tilde{\eta^2})(\mathbf{k}) + \beta_{KK_\parallel}^F(k,\mu) (\tilde{KK})_\parallel \,, \end{aligned} \tag{25}$$

其中 Kn 是非线性核。来自该级数中各个幂的功率谱的一般微扰修正具有以下形式

$$P_{nm} = s_{nm} \langle \delta_F^{(n)}(\boldsymbol{k}) \delta_F^{(m)}(-\boldsymbol{k}) \rangle' \,, \tag{17}$$

其中 snm 是组合因子，并且在线性阶段 n = m = 1，一、二、三环阶分别为 n + m = 4、6、8 等。在二次阶上，有与视线方向旋转一致的一般算符集合 [85, 94]：

K2(k1, k2) ≡ b1Fδ + b2Fδ2 + bG2FG2 + bηFη + bδηFδη + bη2Fη2 + bΠ[2] // FΠ[2] // + b(KK) // F(KK) // , (18)

其中动量依赖的二次核由以下公式给出

$F_{\delta^2} = \frac{1}{2}$, $F_{\mathcal{G}_2} = \left(\frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2}{k_1^2 k_2^2} - 1 \right)$, $F_\delta = 1 + \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{2k_1^2}$+ (k1 · k2) 2k2 2 + 2 7FG2 + f μ1μ2 2 $\left(\frac{k_2}{k_1} + \frac{k_1}{k_2} \right)$,

$F_\eta = -f\mu^2 \Big(1 + \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{2k_1^2}$+ (k1 · k2) 2k2 2 + $\frac{4}{7} F_{\mathcal{G}_2} \Big) - f^2 \frac{\mu_1 \mu_2}{2}$ $\frac{k_2}{k_1} \mu_2^2 + \frac{k_1}{k_2} \mu_1^2 \Big)$, $F_{\delta\eta} = -f \frac{\mu_2^2 + \mu_1^2}{2}$, (19)

Fη2 = f2μ2 1μ2 2, F(KK) // = μ1μ2 (k1 · k2) k1k2 $- \frac{\mu_1^2 + \mu_2^2}{3} + \frac{1}{9}$, FΠ[2] // = μ1μ2 (k1 · k2) k1k2 $- \frac{5}{7} \mu^2 F_{\mathcal{G}_2}$,

其中 μi ≡ ẑ · ki ki 是在视线方向与动量矢量之间的余弦。可以看到，上述核可以与我们之前引入的移位密度和泽尔多维奇核相关。例如，很容易看到

Fδ = ˜ K2 + Fδη + 2 7FG2 . (20)

同样，核 Fη 可以重写为

$$F_\eta = F_\delta - \left(F_2^{\rm ZA} - \frac{3}{7} f\mu^2 F_{\mathcal{G}_2} \right) + \frac{3}{14} F_{\mathcal{G}_2} + F_{\eta^2} - F_{\delta\eta} = \tilde{K}_2 - \left(F_2^{\rm ZA} - \frac{3}{7} f\mu^2 F_{\mathcal{G}_2} \right) + \frac{1}{2} F_{\mathcal{G}_2} + F_{\eta^2} \;. \tag{21}$$

最后，我们有

F(KK) // = FΠ[2] // + 5 7μ2FG2 −μ2 1+ μ2 2 3 + 1 9 = FΠ[2] // + 5 7μ2FG2 + 2 3f Fδη + 2 9Fδ2 . (22)

因此，使用新的视线速度散度场

$$\eta_{\rm new} \equiv -\left(\delta_Z - \frac{3}{7} f\mu^2 \mathcal{G}_2 \right) \,, \tag{23}$$

二次核可以重写为

K2(k1, k2) ≡(b1 + bη) ˜ K2 + bηFηnew + b2Fδ2 + 2 7b1 + 1 2bη + bG2) $F_{\mathcal{G}_2}$

+ (b1 + bδη)Fδη + (bη2 + bη)Fη2 + bΠ[2] // F(KK) // −5 7μ2FG2 −2 3f Fδη −2 9Fδ2 + $b_{(KK)}$ $F_{(KK)_\parallel}$ (24)

=(b1 + bη) ˜ K2 −bη $F_2^{\rm ZA} - \frac{3}{7} f\mu^2 F_{\mathcal{G}_2}$ + $\Big(b_2 - \frac{2}{9} b_{\Pi^{[2]}_\parallel}$ F_{δ^2} + 2 7b1 + 1 2bη + bG2 −5 7μ2bΠ(2)) $F_{\mathcal{G}_2}$

+ $b_1 + b_{\delta\eta} - \frac{2}{3f} b_{\Pi^{[2]}_\parallel} \Big)$ Fδη + (bη2 + bη)Fη2 + (b(KK) // + bΠ[2] //)F(KK) //

如果我们使用依赖于 k 和 μ 的传递函数，只有七个独立的二次算符：Π(2) // 的效应完全被吸收到传递函数中。此外，速度偏差项可以完全通过 Zel’dovich 场来描述。这意味着在二次水平上，完整的 EFT 可以通过以下前向模型来描述：

δmodel F (k) = βF 1(k, μ)˜ δ1(k) + βF η(k, μ) (δZ(k) −3 7fμ2 ˜ G2) + $\beta_2^F(k,\mu)(\tilde{\delta_1^2})$ (25)

+ βF G2(k, μ) ˜ G2(k) + βF δη(k, μ) ˜ [δη](k) + βF η2(k, μ) ˜ (η2)(k) + βF KK // (k, μ) ˜ (KK) // , where we promoted the bias parameters in Eq. (24) to momentum-dependent transfer functions, and also shifted each operator with the Zel'dovich displacement,

$$\tilde{\mathcal{O}}(\boldsymbol{k}) = \int d^3\boldsymbol{q}\ \mathcal{O}(\boldsymbol{q}) e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}(\boldsymbol{q})+f\hat{\boldsymbol{z}}(\boldsymbol{\psi}(\boldsymbol{q})\cdot\hat{\boldsymbol{z}}))}\ , \tag{26}$$

which accounts for IR-resummation of bulk flows that affect the phases from these operators at higher orders [141–147]. It is easy to see that the model (25) reproduces the linear Ly-α forest bias on large scales,

$$\delta_F^{\rm model}(\boldsymbol{k})|_{k\to 0} = ((\beta_1^F + \beta_\eta^F) + \beta_\eta^F f\mu^2)\delta_1(\boldsymbol{k})\,, \tag{27}$$

so that the transfer functions β_1 and β_η are constant in this limit, and the relationship

$$b_1 = \beta_1^F + \beta_\eta^F\,, \quad b_\eta = -\beta_\eta^F\,, \tag{28}$$

consistently holds through the quadratic order in our forward model. Beyond the $k \to 0$ limit the linear transfer functions β_1 and β_η absorb any physical corrections correlated with the linear density field. For instance, they automatically account for all cubic operators contributing to the Ly-α power spectrum at the one-loop order and all the relevant counterterms. Indeed, the cross-correlation between these terms and the linear matter field in EFT can be cast as a non-linear correction $\Delta\beta_1$ to the β_1 transfer function,

$$\begin{aligned}\langle\delta_1\delta_F^{(3)}\rangle' + \langle\delta_1\delta_F^{\rm ctr}\rangle' &= 3P_{\rm lin}(k)\int_{\boldsymbol{p}} K_3(\boldsymbol{p},-\boldsymbol{p},\boldsymbol{k})P_{11}(p) + k^2 P_{11}(k)\sum_{m=0}^{2} c_m\mu^{2m} \\ &\equiv \Delta\beta_1(k,\mu)P_{\rm lin}(k) = \langle\Delta\beta_1(k,\mu)\delta_1\delta_1\rangle'\ .\end{aligned} \tag{29}$$

Thus, by construction, our forward model absorbs small-scale effects such as baryonic feedback and gas smoothing into the transfer functions. In particular, the linear forward model with the transfer function β_1 produces the power spectrum

$$P_{\rm F}(k, k_\parallel) = F_{\rm NL}(k, k_\parallel)P_{11}(k)\,, \tag{30}$$

equivalent to expressions which appear in many popular phenomenological models [73, 74, 148]. Likewise, the transfer functions of the quadratic operators automatically encapsulate all the bispectrum higher order corrections that correlate with the linear and quadratic fields, including the deterministic bispectrum counterterms.

It is important to stress that fitting the density field is a much more difficult task than fitting the power spectrum. For instance, a forward model

$$\delta_F^{\rm model}(\boldsymbol{k})\Big|_{\rm pheno} = \sqrt{\frac{P_{\rm F,NL}(\boldsymbol{k})}{P_{\rm lin}(k)}}\delta_1(\boldsymbol{k})\,, \tag{31}$$

where $P_{\rm F,NL}$ is the simulated non-linear flux power spectrum, by construction reproduces the Ly-α power spectrum perfectly. However, this model is expected to fail at the field level because it misses higher order Ly-α bias operators and proper IR resummation. We will demonstrate the breakdown of this model below.

Our full forward model is equivalent to the full EFT at the quadratic order, and includes some additional terms beyond that order. Thanks to the transfer functions, at the power spectrum level our model is equivalent to the full 1-loop EFT including the cubic operators. To reflect this fact, with some abuse of terminology, we will refer to our model as "effectively cubic" in what follows. However, we stress that the model does not currently contain all the necessary cubic terms, like Γ_3 or η^3. The cubic terms that cannot be absorbed into the transfer functions will then contribute to the model error, and in particular to the noise power spectrum. The effect of these terms, however, vanishes in the $k \to 0$ limit, except those proportional to the δ^3 operator, entering through P_{33} in the EFT nomenclature. This is the only independent cubic operator whose power spectrum is constant in the $k \to 0$ limit. Hence, in principle, it could contribute to n_0. In order to reduce the noise power spectrum in the $k \to 0$ limit, we explicitly include the δ^3 operator in our model. At fourth order, there are two power spectrum contributions constant on large scales, P_{24} (two-loop power spectrum order) and P_{44} (three-loop power spectrum order). Our quadratic transfer functions, however, account for all higher order corrections that correlate with the quadratic fields. Hence, the contributions from P_{24} terms are already included in our model, and the first missing perturbative contribution to the noise on large scales is the constant piece from P_{44} at the three-loop order. This can be estimated to be negligibly small on the redshifts of interest, and we will ignore this term in what follows.

我们将公式 (24) 中的偏置参数提升为动量依赖的传递函数，并且还将每个算符用 Zel'dovich 位移进行了平移，

$$\tilde{\mathcal{O}}(\boldsymbol{k}) = \int d^3\boldsymbol{q}\ \mathcal{O}(\boldsymbol{q}) e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}(\boldsymbol{q})+f\hat{\boldsymbol{z}}(\boldsymbol{\psi}(\boldsymbol{q})\cdot\hat{\boldsymbol{z}}))}\ , \tag{26}$$

这解释了红外重求和（IR-resummation）对这些算子在高阶中相位的影响。很容易看出，模型（25）在大尺度上重现了线性Ly-α森林偏差，

$$\delta_F^{\rm model}(\boldsymbol{k})|_{k\to 0} = ((\beta_1^F + \beta_\eta^F) + \beta_\eta^F f\mu^2)\delta_1(\boldsymbol{k})\,, \tag{27}$$

以便在这个极限下传递函数 β1 和 βη 是常数，并且关系

$$b_1 = \beta_1^F + \beta_\eta^F\,, \quad b_\eta = -\beta_\eta^F\,, \tag{28}$$

在我们的前向模型中，这在二次方阶数上始终成立。超出 k →0 极限时，线性传递函数 β1 和 βη 会吸收与线性密度场相关的任何物理修正。例如，它们会自动考虑所有对 Ly-α 功率谱在一环阶贡献的三次算符以及所有相关的对策项。实际上，这些项与 EFT 中线性物质场的交叉关联可以被表示为 β1 传递函数的非线性修正 Δβ1。

$$\begin{aligned}\langle\delta_1\delta_F^{(3)}\rangle' + \langle\delta_1\delta_F^{\rm ctr}\rangle' &= 3P_{\rm lin}(k)\int_{\boldsymbol{p}} K_3(\boldsymbol{p},-1,\text{ k)P11(p) + k2P11(k)}\quad \sum_{m=0}^{2} c_m\mu^{2m} \\ &\equiv \Delta\beta_1(k,\mu)P_{\rm lin}(k) = \langle\Delta\beta_1(k,\mu)\delta_1\delta_1\rangle'\ .\end{aligned} \tag{29}$$

因此，通过构建，我们的前向模型将小尺度效应，例如重子反馈和气体平滑，吸收到传递函数中。特别是，具有传递函数 β1 的线性前向模型产生了功率谱

$$P_{\rm F}(k, k_\parallel) = F_{\rm NL}(k, k_\parallel)P_{11}(k)\,, \tag{30}$$

等同于出现在许多流行现象学模型中的表达式 [73, 74, 148]。同样，二次算子的传递函数会自动包含与线性和二次场相关的所有双谱高阶修正，包括确定性双谱反项。

重要的是要强调，拟合密度场比拟合功率谱要困难得多。
例如，一个前向模型

$$\delta\text{模型 F (k) 表型} = \sqrt{\text{PF,NL(k) Plin(k)}}\,\text{δ1(k)}\,, \tag{31}$$

其中 PF,NL 是模拟的非线性通量功率谱，按构造完美重现 Ly-α 功率谱。然而，该模型预计在场层面会失效，因为它缺失了更高阶的 Ly-α 偏差算子和适当的红外重求和。我们将在下面展示该模型的失效。

我们的完整前向模型在二次阶上等价于完整的有效场论（EFT），并且还包括一些高于该阶数的额外项。得益于传输函数，在功率谱层面上，我们的模型等价于包含三次算符的完整一环（1-loop）EFT。为了反映这一事实，在某种程度上滥用术语，我们将在下面将我们的模型称为"有效三次"。然而，我们强调，这个模型目前并不包含所有必要的三次项，例如 Γ3 或 η3。那些不能被吸收到传输函数中的三次项将会对模型误差做出贡献，特别是对噪声功率谱的贡献。然而，这些项的效应在 k → 0 极限下会消失，除了那些与 δ3 算符成比例、通过 EFT 命名中的 P33 进入的项。这是唯一在 k → 0 极限下其功率谱为常数的独立三次算符。因此，原则上，它可能会对 n0 做出贡献。为了在 k →0 极限下降低噪声功率谱，我们在模型中明确包含了 δ3 算子。在四阶时，有两个在大尺度上恒定的功率谱贡献，P24（两环功率谱阶）和 P44（三环功率谱阶）。然而，我们的二次传递函数已经考虑了与二次场相关的所有高阶修正。因此，P24 项的贡献已经包含在我们的模型中，而在大尺度上噪声的第一个缺失扰动贡献是来自三环阶的 P44 的常数项。可以估计，在感兴趣的红移下这个值可以忽略不计，因此在接下来的分析中我们将忽略该项。

Finally, in order to reduce the numerical noise due to degeneracies in our bias expansion, we orthogonalize all the relevant operators using the Gram-Schmidt algorithm [122] for each (k, μ) bin,

$$\tilde{\mathcal{O}}_a^{\perp} = \sum_b M_{ab} \tilde{\mathcal{O}}_b \ , \tag{32}$$

where $\tilde{\mathcal{O}}_b = \{\tilde{\delta}_1, \tilde{\delta}_1^2, \tilde{\mathcal{G}}_2, \tilde{\delta}_1^3, (K\tilde{K})_{\parallel}, \tilde{\eta}_{\rm new}, \tilde{\eta^2}, [\tilde{\delta\eta}]\}$, and M_{ab} is the rotation matrix [122] constructed from

$$O_{ab} \equiv \langle \tilde{\mathcal{O}}_a(\boldsymbol{k}) \tilde{\mathcal{O}}_b^*(\boldsymbol{k}) \rangle' \, ,$$

using the Cholesky decomposition procedure. One can check that the operator $\Pi_{\parallel}^{[2]}$ does not contribute to our model because $\langle |\tilde{\mathcal{O}}^{\perp}_{\Pi_{\parallel}^{[2]}}|^2 \rangle \approx 0$, which is equivalent to the statement that its contribution is absorbed by the transfer functions. This yields the final forward model

$$\begin{aligned} \delta_F^{\rm model}(\mathbf{k}) = {} & \beta_1^F(k,\mu)\tilde{\delta}_1(\mathbf{k}) + \beta_\eta^F(k,\mu)\eta^{\perp}_{\rm new} + \beta_2^F(k,\mu)(\tilde{\delta}_1^2)^{\perp}(\mathbf{k}) + \beta_3^F(k,\mu)(\tilde{\delta}_1^3)^{\perp}(\mathbf{k}) \\ & + \beta_{\mathcal{G}_2}^F(k,\mu)\tilde{\mathcal{G}_2}^{\perp}(\mathbf{k}) + \beta_{\delta\eta}^F(k,\mu)[\tilde{\delta\eta}]^{\perp}(\mathbf{k}) + \beta_{\eta^2}^F(k,\mu)\tilde{\eta}^{2,\perp}(\mathbf{k}) + \beta_{KK_\parallel}^F(k,\mu)(K\tilde{K})_{\parallel}^{\perp}(\mathbf{k}) \, . \end{aligned} \tag{33}$$

Note that the above model appears as a simple generalization of the forward model for redshift-space galaxies,

$$\delta_g^{\rm model}(\mathbf{k}) = \beta_1(k,\mu)\tilde{\delta}_1(\boldsymbol{k}) + \eta_{\rm new}(\boldsymbol{k}) + \beta_2(k,\mu)(\tilde{\delta}_1^2)^{\perp}(\mathbf{k}) + \beta_3(k,\mu)(\tilde{\delta}_1^3)^{\perp}(\boldsymbol{k}) + \beta_{\mathcal{G}_2}(k,\mu)\tilde{\mathcal{G}_2}^{\perp}(\boldsymbol{k}) \, . \tag{34}$$

Note that the above model *does not have* a transfer function in front of $\eta_{\rm new}$, defined in Eq. (23), which reflects that halos conserve denisty when switching from real space to redshift space.

In what follows we will explore the implications of our Ly-α forward model.

III. SIMULATION DATA

We fit our field-level perturbative model to two synthetic Ly-α forest data sets: First, large Ly-α forest mocks constructed from the N-body simulation suite ABACUSSUMMIT (hereinafter Abacus) [149]. Second, hydrodynamic simulations of the intergalactic medium (IGM) from the Sherwood suite [75, 82]. The Abacus simulation will be the primary data set of this work. We now introduce both simulations, briefly summarizing [75, 82, 102], to which the reader is referred for a fuller presentation.

The Abacus simulations, centered at redshift $z = 2.5$, are large N-body simulations with the Ly-α forest painted on top of them using a simplistic fluctuating Gunn-Peterson approximation (FGPA) [102]. Note that Ref. [102] provides four different FGPA implementations resulting in four different realizations of the Abacus Ly-α simulations. The main difference arises for the value of the bias parameter associated with the gradient of the peculiar velocity, b_η, which is lower for models I and II and larger for models III and IV than current constraints obtained from DESI data [51]. The simulations encompass a comoving volume of $V = 2000^3(h^{-1}\,{\rm Mpc})^3$ with 6912^3 particles, each of mass $M_{\rm part} = 2.1 \times 10^9$ [78, 102, 150]. The simulation is based on a fiducial *Planck* 2018 cosmology with $\Omega_b h^2 = 0.02237$, $\Omega_c h^2 = 0.12$, $h = 0.6736$, $A_s = 2.0830 \times 10^{-9}$, $n_s = 0.9649$, $w_0 = -1$, $w_a = 0$. The quasi-stellar objects (QSO) use a simplistic halo occupation distribution (HOD) model given in equations (1) and (2) in [78] with as mass range approximately yielding a linear bias that matches observations $b_q \approx 3.3$ and a number density of $\sim 1.75 \times 10^{-4}$ $(h^{-1}\,{\rm Mpc})^{-3}$. The parameters describing the HOD are $\log_{10}(M_{\rm cut}) = 13.2$ which characterizes the minimum halo mass to host a central galaxy, $\log_{10}(M_1) = 13.8$ the typical halo mass that hosts one satellite galaxy, $\sigma = 0.65$ the steepness of the transition from 0 to 1 in the number of central galaxies, $\alpha = 0.8$ the power law index on the number of satellite galaxies, ic the incompleteness parameter, and $\kappa = 1.11$ multiplied by $M_{\rm cut}$ gives the minimum halo mass to host a satellite galaxy.

We use two sets of Sherwood hydrodynamic simulations: (i) one snapshot at redshift $z = 2.8$ encompassing a comoving volume $V = 160^3(h^{-1}\,{\rm Mpc})^3$ where the number of CDM particles and gas particles in the simulation is 2048^3; and (ii) a series of snapshots at redshifts $z = 2.0, 2.4, 2.8, 3.2$ in a $V = 80^3(h^{-1}\,{\rm Mpc})^3$ volume with 1024^3 particles keeping the resolution fixed across simulations. The initial conditions are generated using the `N-GenIC` code [151] with the corresponding number of either 2048^3 or 1024^3 particles. The fiducial cosmology is based on the best-fit *Planck* 2013 cosmology with cosmological parameters set to $\Omega_m = 0.308$, $\Omega_b = 0.0482$, $h = 0.678$, $\sigma_8 = 0.829$, $n_s = 0.961$ [152]. Additionally, the Sherwood simulations come with a halo catalog built using a friends-of-friends algorithm and with an available mass range of $10^9 \leq M_\odot \lesssim 10^{14}$. For all simulations, redshift-space distortions have been applied along the $\hat{\boldsymbol{z}}$-axis and particles have been assigned using a triangular-shaped cloud (TSC) algorithm.

最后，为了减少由于我们偏差展开中的简并性而产生的数值噪声，我们对所有的进行正交化
对于每个 (k, μ) 区间，使用 Gram-Schmidt 算法 [122] 的相关算子，

$$\tilde{\mathcal{O}}_a^{\perp} = \sum_b M_{ab} \tilde{\mathcal{O}}_b \ , \tag{32}$$

其中 ˜ Ob = {˜ δ1,˜ δ2 1, ˜ G2,˜ (δ3KK) // , ˜ ηnew,˜ η2,˜ [δη]}, 而 Mab 是由 [122] 构造的旋转矩阵
1,

Oab ≡⟨˜ Oa(k) ˜ O∗ b(k)⟩′ ,

使用Cholesky分解过程。可以检查算子Π[2] // 不对我们的模型有贡献
因为⟨|˜ O⊥ Π[2]|2⟩≈0，这等同于其贡献被传递函数吸收的说法。

这产生了最终的前向模型

δmodel F (k) = βF 1(k, μ)˜ δ1(k) + βF η(k, μ)η⊥ new+ βF 2(k, μ) ˜ (δ2 1)⊥(k) + βF 3(k, μ) ˜ (δ3 1) ⊥(k)

$$+ \beta_{\mathcal{G}_2}^F(k,\mu)\tilde{\mathcal{G}_2}^{\perp}(\mathbf{k}) + \beta_{\delta\eta}^F(k,\mu)[\tilde{\delta\eta}]^{\perp}(\mathbf{k}) + \beta_{\eta^2}^F(k,\mu)\tilde{\eta}^{2,\perp}(\mathbf{k}) + \beta_{KK_\parallel}^F(k,\mu)(K\tilde{K})_{\parallel}^{\perp}(\mathbf{k}) \, . \tag{33}$$

请注意，上述模型似乎是红移空间星系前向模型的一个简单推广，

δ模型 g (k) = β1(k, μ)˜ δ1(k) + η新(k) + β2(k, μ) ˜ (δ2 1)⊥(k) + β3(k, μ) ˜ (δ3 1)⊥(k) + βG2(k, μ) ˜ G2⊥(*k*) . (34)

请注意，上述模型在 ηnew 前没有传递函数，如公式 (23) 定义，这反映了在从实空间切换到红移空间时，晕保持密度守恒。

在下文中，我们将探讨我们的 Ly-α 前向模型的意义。

三. 模拟数据

我们将我们的场级微扰模型拟合到两个合成的Ly-α森林数据集：首先，是由N体模拟套件AbacusSummit（以下简称Abacus）[149]构建的大型Ly-α森林模拟；其次，是来自Sherwood套件[75, 82]的星际介质（IGM）流体动力学模拟。Abacus模拟将是本文的主要数据集。现在我们介绍这两种模拟，简要总结[75, 82, 102]，详细内容请读者参阅原文。

阿巴库斯模拟，以红移 z = 2.5 为中心，是大型 N 体模拟，并在其上使用简单的涨落冈-彼得森近似（FGPA）[102]构建 Ly-α 森林。请注意，参考文献 [102] 提供了四种不同的 FGPA 实现，从而产生四种不同的阿巴库斯 Ly-α 模拟实现。主要区别在于与奇异速度梯度相关的偏差参数 bη 的取值，对于模型 I 和 II，该值较低；而对于模型 III 和 IV，该值较当前从 DESI 数据获得的约束值 [51] 较高。模拟的共动体积为 $V = 2000^3(h^{-1}Mpc)^3$，包含 69123 个粒子，每个粒子的质量为 Mpart = 2.1 × 10^9 [78, 102, 150]。该模拟基于 Planck 2018 的标准宇宙学模型，参数为 Ωbh^2 = 0.02237，Ωch^2 = 0.12，h = 0.6736，As = 2.0830 × 10^{-9}，ns = 0.9649，w0 = −1，wa = 0。类星体（QSO）使用了一个简化的晕占据分布（HOD）模型，该模型在文献[78]的方程（1）和（2）中给出，其质量范围大致产生与观测相符的线性偏差 bq ≈ 3.3 和约为 ~1.75 × 10−4(h−1Mpc)−3 的数密度。描述 HOD 的参数包括 log10 (Mcut) = 13.2，表征承载中心星系的最小晕质量；log10 (M1) = 13.8，为承载一个卫星星系的典型晕质量；σ = 0.65，为中心星系数量从 0 到 1 过渡的陡度；α = 0.8，为卫星星系数量的幂律指数；ic 为不完整性参数；以及 ϰ = 1.11 乘以 Mcut，得到承载卫星星系的最小晕质量。

我们使用了两套 Sherwood 流体动力学模拟：(i) 在红移 z = 2.8 的一个快照，覆盖一个本征体积 V = 1603(h−1Mpc)^3，其中模拟中的 CDM 粒子和气体粒子数量为 2048^3；以及 (ii) 在红移 z = 2.0、2.4、2.8、3.2 的一系列快照，体积为 V = 803(h−1Mpc)^3，粒子数量为 1024^3，在各模拟中保持分辨率不变。初始条件使用 N-GenIC 代码[151] 生成，粒子数量分别为 2048^3 或 1024^3。标定宇宙学模型基于 Planck 2013 最优拟合宇宙学模型，宇宙学参数设置为 Ωm = 0.308，Ωb = 0.0482，h = 0.678，σ8 = 0.829，ns = 0.961 [152]。此外，Sherwood 模拟还附带了一个使用朋友-朋友算法构建的晕目录，质量范围在 10 ≤M⊙≲10^{14}。对于所有模拟，红移空间畸变已沿 ˆ z 轴应用，并且粒子已使用三角形云 (TSC) 算法分配。

To explore cross-correlations of the Ly-α forest with high-redshift galaxies we use hydrodynamic simulations of two types of star-forming galaxies in the state-of-the-art Astrid hydrodynamical simulations (see, e.g., [83, 84, 124, 153, 154]). In particular, we use the samples consistent with the observed (angular) clustering and number density of Ly-α emitters (LAEs) and Lyman-break galaxies (LBGs) at $z = 3$. The samples are designed to approximately match the linear bias b_1 and number density of existing and future LSS surveys for LAEs ("ODIN" [155] and S5 [131, 155, 156]) and similarly for LBGs ("CARS" and S5; [131, 157]).

IV. RESULTS

In this section, we use the perturbative forward model introduced in Sec. II on the synthetic data presented in Sec. III. We compare the Ly-α forest auto-correlation as well as its cross-correlation with dark matter halos (used as proxies for high-redshift galaxies and quasars). We compare four metrics to assess the quality of the forward model: (i) the two-dimensional flux power spectrum, $P(k,\mu)$ which is the average of the squared norm of the Fourier transform of the flux decrement, δ_F, in bins of the Fourier wavenumber, k, with the cosine of the angle to the line-of-sight, $\mu = k_\parallel/k$. The corresponding error bars are computed from the square root of the diagonal of the Gaussian covariance based on the number of expected Fourier modes per bin $P(k,\mu)\sqrt{2/N(k,\mu)}$. (ii) The one-point probability density function which is a more stringent test of the model since we are effectively capturing information beyond the two-point function. (iii) The error power spectrum (or mean-squared model error), given in Eq. (5), is a measure of the success of the model. (iv) To investigate the agreement at the level of the phases and amplitudes, we compute the cross-correlation coefficient:

$$r_{cc}(\delta_F^{\rm truth}, \delta_F^{\rm model}) = \frac{\langle \delta_F^{\rm model}(\boldsymbol{k})[\delta_F^{\rm truth}(\boldsymbol{k})]^* \rangle}{\left(\langle|\delta_F^{\rm truth}(\boldsymbol{k})|^2\rangle\langle|\delta_F^{\rm model}(\boldsymbol{k})|^2\rangle\right)^{1/2}}\,, \tag{35}$$

for the simulated ('truth') and forward modeled ('model') fields. The cross-correlation coefficient allows us to investigate how spatially coherent the densities are in Fourier space and is a measure of the amount of cosmological information that can be extracted when using the model to describe Ly-α forest measurements [122].

We show our results as a function of Fourier wavenumber k and in three angular bins centered at $\mu = 0.17, 0.50, 0.83$. The maximum wavenumber shown is given by the Nyquist frequency which is $k_{\rm Ny} = 1.61\,h\,{\rm Mpc}^{-1}$ for Abacus with a resolution of $N_{\rm cell} = 1024$ in a $2^3(\,h^{-1}\,{\rm Gpc})^3$ volume and $k_{\rm Ny} = 5.03,\ 10.05\,h\,{\rm Mpc}^{-1}$ for Sherwood in $160^3(\,h^{-1}\,{\rm Mpc})^3$ and $80^3(\,h^{-1}\,{\rm Mpc})^3$ volumes with $N_{\rm cell} = 256$, respectively. Throughout this work, we will discuss model I from Abacus and compare these results to the ones obtained from the Sherwood hydrodynamic simulations (if not already discussed in Ref. [87]). In App. B we confirm the robustness of the perturbative forward model to different realizations (models II-IV) of the Abacus simulations stemming from different approaches to calibrating the FGPA procedure to plant a Ly-α forest on top of an N-body simulation (see Ref. [102] for more details).

A. Ly-α forest forward model

In this section, we present the results of the Ly-α forest forward model. In Fig. 2 we compare the linear (left panel) and cubic (right panel) forward models at the power spectrum level for the Abacus simulation. Note that our linear theory model matches the phenomenological fitting functions employed in current Ly-α forest analyses (see, e.g., [51, 73, 158]). We find that linear theory does not *exactly* capture the largest scales $k \lesssim 10^{-2}\,h\,{\rm Mpc}^{-1}$ and recovers the input power spectrum at the 5% level only down to $k_{\rm max} \approx 0.1\,h\,{\rm Mpc}^{-1}$. Similarly, the linear error power spectrum, given in Eq. (5), exhibits an orientation dependence and amplitude at the level of a few percent of the amplitude of the total power spectrum indicating poor model performance. For the cubic model, we find excellent agreement at the largest scales between the simulation and model power spectra. It captures the power spectrum down to $k_{\rm max} \approx 0.8\,h\,{\rm Mpc}^{-1}$ at the 5% level.[1] We note that the error power spectrum exhibits a scale dependence above $k \gtrsim 0.6\,h\,{\rm Mpc}^{-1}$ leading to increased stochasticity.[2]

In Fig. 3 we compare both Ly-α forest fields at the field level using the one-point probability density function (PDF). The break down of the linear model is striking, even on large scales.[3] We find excellent agreement between

---

[1] Current DESI Ly-α full-shape analyses use the correlation function down to a minimum scale of $r = 25\,h^{-1}$ Mpc [61] which corresponds to $k \sim \pi/25 = 0.13\,h\,{\rm Mpc}^{-1}$ which is the range of validity where the present forward model is accurate at the sub-percent level.

[2] We refer the reader to Ref. [159] to further reduce the Ly-α forest stochasticity by computing non-linear particle displacements using simulations in the context of Hybrid effective field theory (HEFT). This extends the range of validity of the forward model to $k \approx 1\,h\,{\rm Mpc}^{-1}$.

[3] This is qualitatively similar, yet less pronounced than on the field-level methodology applied to the Sherwood hydrodynamic simulations, see [87], illustrating that the Abacus mocks can access larger scales and thus more quasi-linear modes.

为了探索Ly-α森林与高红移星系的交叉相关，我们使用最先进的Astrid流体力学模拟中两类恒星形成星系的流体力学模拟（参见，例如，[83, 84, 124, 153, 154]）。特别地，我们使用与观测到的Ly-α发射体（LAEs）和Lyman断裂星系（LBGs）在z = 3时的（角度）聚类和数密度一致的样本。这些样本旨在大致匹配现有和未来大型结构（LSS）调查中LAEs的线性偏差b1和数密度（"ODIN" [155] 和 S5 [131, 155, 156]）以及LBGs（"CARS" 和 S5 [131, 157]）

四. 结果

在本节中，我们使用第 II 节中介绍的微扰前向模型，对第 III 节中展示的合成数据进行分析。我们比较了 Ly-α 森林的自相关以及其与暗物质晕的交叉相关（暗物质晕被用作高红移星系和类星体的代理）。我们比较了四个指标来评估前向模型的质量：（i）二维通量功率谱 P(k, μ)，它是在傅里叶波数 k 和与视线方向夹角余弦 μ = k // /k 的分箱中，通量减量 δF 的傅里叶变换平方模的平均值。相应的误差条是基于每个分箱的傅里叶模式数量计算的高斯协方差对角线平方根，即 P(k, μ) ± √(2/N(k, μ))。（ii）单点概率密度函数，这对模型是更严格的测试，因为我们实际上捕捉到了超出二点函数的信息。(iii) 错误功率谱（或均方模型误差），如公式(5)所示，是衡量模型成功程度的一个指标。 (iv) 为了研究相位和振幅层面的匹配情况，我们计算交叉相关系数：

$$r_{cc}(\delta_F^{\rm truth}, \delta_F^{\rm model}) = \frac{\langle \delta_F^{\rm model}(\boldsymbol{k})[\delta_F^{\rm truth}(\boldsymbol{k})]^* \rangle}{\left(\langle|\delta_F^{\rm truth}(\boldsymbol{k})|^2\rangle\langle|\delta_F^{\rm model}(\boldsymbol{k})|^2\rangle\right)^{1/2}}\,, \tag{35}$$

对于模拟的（"真实"）和前 forward modeled ('model') fields. The cross-correlation coefficient allows us to in-

性，并且是使用模型描述Ly-α森林测量时可以提取的宇宙学信息量的一个度量 [122]。

我们将结果展示为傅里叶波数 k 的函数，并在三个以 μ = 0.17、0.50、0.83 为中心的角度区间中。所显示的最大波数由奈奎斯特频率给出，对于分辨率为 Ncell = 1024、体积为 23(h−1Gpc)³ 的 Abacus，kNy = 1.61 h Mpc−1；对于 Sherwood，在体积为 160³(h−1Mpc)³ 和 80³(h−1Mpc)³ 且 Ncell = 256 的情况下，kNy 分别为 5.03 和 10.05 h Mpc−1。在本文中，我们将讨论来自 Abacus 的模型 I，并将这些结果与 Sherwood 流体动力学模拟得到的结果进行比较（如果在文献 [87] 中尚未讨论）。在附录 B 中，我们确认了微扰前向模型对于 Abacus 模拟的不同实现（模型 II-IV）的稳健性，这些模拟源自不同的方法来校准 FGPA 程序，以在 N 体模拟中植入 Ly-α 森林（更多细节见文献 [102]）。

A. Ly-α森林前向模型

在本节中，我们展示了Ly-α森林前向模型的结果。在图2中，我们比较了Abacus模拟的功率谱水平上的线性（左图）和三次（右图）前向模型。请注意，我们的线性理论模型与当前Ly-α森林分析中使用的经验拟合函数一致（参见例如，[51, 73, 158]）。我们发现，线性理论并不能完全捕捉最大的尺度 k < ~10−2 h Mpc−1，且仅在 kmax ≈0.1 h Mpc−1 处能够以约5%的精度恢复输入功率谱。类似地，给定在公式（5）中的线性能误差功率谱表现出方向依赖性，其幅度约为总功率谱幅度的几百分点，表明模型表现不佳。对于三次模型，我们发现模拟与模型功率谱在最大尺度上具有极好的一致性。它捕捉了功率谱直到 kmax ≈0.8 h Mpc−1，在 5% 水平上。我们注意到误差功率谱在 k > ~0.6 h Mpc−1 以上显示出尺度依赖性，导致随机性增加。在图 3 中，我们使用单点概率密度函数 (PDF) 在场的层面比较了两种 Ly-α 林场。即使在大尺度上，线性模型的崩溃也令人震惊。我们发现之间有极好的一致性。

---

1 当前的 DESI Ly-α 全形分析使用相关函数的最小尺度为 r = 25 h−1 Mpc [61]，对应 k ~ π/25 = 0.13 h Mpc−1，这是当前前向模型在亚百分比水平上精确的有效范围。2 我们建议读者参考文献 [159]，通过在混合有效场论（HEFT）框架下使用模拟计算非线性粒子位移，以进一步减少 Ly-α 森林的随机性。这将前向模型的有效范围扩展到 k ≈ 1 h Mpc−1。3 这在定性上与应用于 Sherwood 流体动力学模拟的场层面方法类似，但效果不那么显著。

参见[87]，说明Abacus模拟可以访问更大的尺度，从而更多准线性模式。

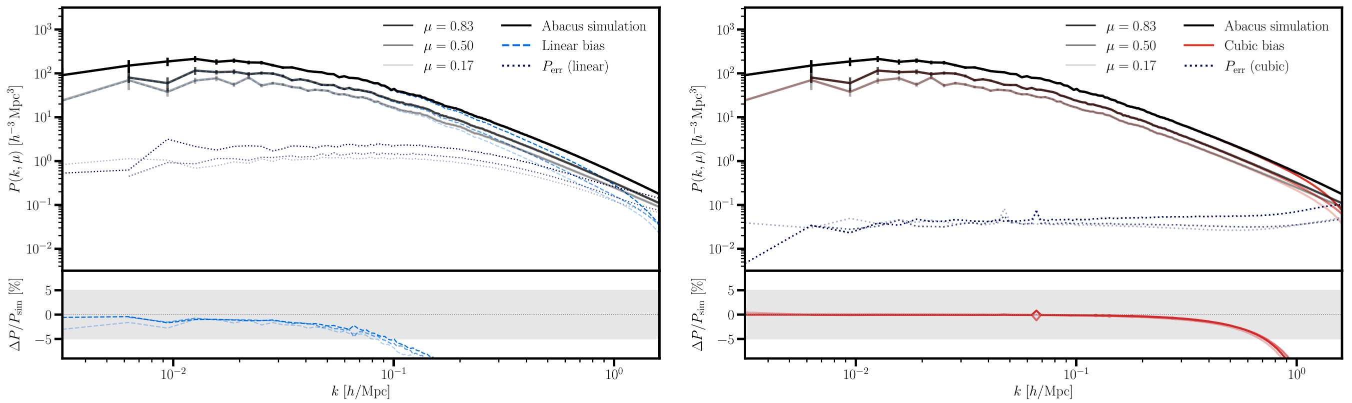

FIG. 2. **Best-Fit Abacus Power Spectra:** Comparison between the measured two-dimensional power spectrum from the Abacus simulation (black) and the best-fit forward model obtained from a linear and a cubic bias expansion in three angular wedges. *Left*: Linear theory prediction (blue dashed line), which shows that the forward model breaks down beyond $k_{\rm max} \sim 0.1\,h\,{\rm Mpc}^{-1}$ and a large error power spectrum $P_{\rm err}(\mathbf{k},\hat{\mathbf{z}}) \equiv \langle|\tilde{\delta}_{\rm sim} - \tilde{\delta}_{\rm EFT}|^2\rangle$ (blue dashed lines), indicating poor model performance. *Right*: Cubic bias model prediction (red solid line), which matches the simulation much more closely, with substantially reduced residuals and a suppressed $P_{\rm err}$. In each panel, the power spectrum is shown in bins of Fourier wavenumber k and angle to the line of sight, parametrized by $\mu = k_\parallel/k$, with darker (lighter) lines for $\mu = 0.83$ (0.17). The bottom panel displays the percent difference between the simulation and model power spectra. A gray band highlights the ±5% region in the bottom panel.

	Variance					Skewness					Kurtosis				
R	$\delta_{\rm truth}$	$\delta_{\rm lin.}$	$\delta_{\rm best-fit}$	$\Delta\delta_{\rm lin.}$	$\Delta\delta$	$\delta_{\rm truth}$	$\delta_{\rm lin.}$	$\delta_{\rm best-fit}$	$\Delta\delta_{\rm lin.}$	$\Delta\delta$	$\delta_{\rm truth}$	$\delta_{\rm lin.}$	$\delta_{\rm best-fit}$	$\Delta\delta_{\rm lin.}$	$\Delta\delta$
1	0.0159	0.0109	0.0149	0.0051	0.0010	-1.0635	-3.7654	-1.2825	0.8168	-0.0412	1.4303	26.5017	2.5290	11.4647	4.0812
2	0.0075	0.0060	0.0074	0.0015	0.0001	-0.7810	-2.7355	-0.8410	1.1211	0.0031	0.8954	13.3336	1.0496	8.4781	7.0318
5	0.0022	0.0020	0.0022	0.0002	0.0000	-0.3764	-1.4703	-0.3882	1.2552	0.0624	0.2078	3.6621	0.1712	4.7769	4.4458
10	0.0007	0.0007	0.0007	0.0000	0.0000	-0.1635	-0.8150	-0.1646	1.1150	0.0772	0.0275	1.1027	0.0022	3.2269	1.9103
30	0.0001	0.0001	0.0001	0.0000	0.0000	-0.0161	-0.2487	-0.0124	0.4271	0.0299	0.0081	0.1016	0.0037	0.6529	0.2548

TABLE I. **Statistical Moments Abacus:** Statistical moments (variance, skewness, and kurtosis) of the Abacus simulation ($\delta_{\rm sim}$) and the perturbative forward model ($\delta_{\rm EFT}$) as well as their residuals for different isotropic smoothing scales R (in h^{-1} Mpc). Following baseline expectation, we find increasing agreement between the one-point probability density functions when removing small-scale modes. The mean is consistent with zero for all fields. Note that skewness and kurtosis vanish for Gaussian fields indicating that our forward model captures higher-order moments, i.e., non-Gaussian information, of the field.

the EFT model and the simulations down to cell sizes of $\approx 2\,h^{-1}$ Mpc with the linear model breaking down for very large cell sizes of $10 - 30\,h^{-1}$ Mpc. Note that the number of pixels from the EFT forward model (blue dashed line) that are visibly deviating from the distribution obtained from the simulation (black line) is less than 0.01%. In Tab. I we quantify the agreement between the modeled and true field using the moments of the field with the same Gaussian kernel smoothing radii ($R = 1\,,2\,,5\,,10\,,30\ h^{-1}$ Mpc) applied to them. Whilst this is not applicable to observed Ly-α forest data [86], it allows us to isolate larger scales. The residuals (denoted by $\Delta\delta$ in Tab. I) for the best-fit EFT model are consistently smaller than the residuals from the linear model, highlighting the importance of higher order bias parameters.

1. *Impact of IR resummation*

Standard Eulerian perturbation theory does not fully capture the non-linear damping of the baryon acoustic oscillation (BAO) feature. This stems from linear matter displacements, which are sensitive to long-wavelength (infrared, IR) modes which are in turn responsible for the slow convergence of the perturbative expansion [141]. While individual diagrams in standard perturbation theory exhibit IR-enhanced terms, these enhancements cancel out when all diagrams are summed, as required by the equivalence principle [144, 161]. However, the presence of the BAO feature prevents a complete cancellation, leaving residual IR sensitivity [142, 145]. IR resummation has been introduced to

---

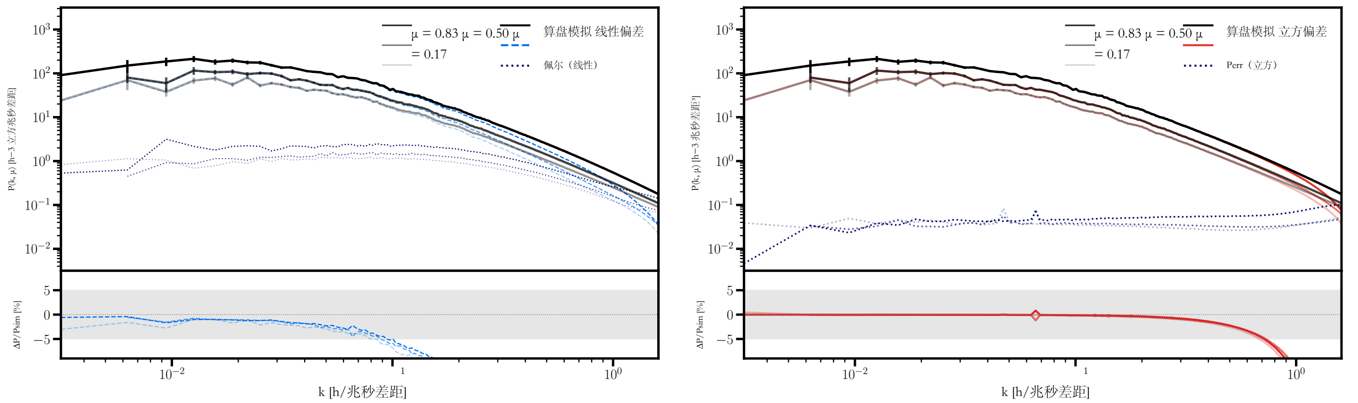

图 2. 最佳拟合 Abacus 功率谱：比较来自 Abacus 仿真的二维功率谱测量值（黑色）与在三个角度楔形中通过线性和三次偏差展开获得的最佳拟合前向模型。左图：线性理论预测（蓝色虚线），显示前向模型在 kmax ~0.1 h Mpc−1 之外失效，并出现较大的误差功率谱 Perr(k, ˆ z) ≡⟨|˜ δsim −˜ δEFT|2⟩（蓝色虚线），表明模型性能较差。右图：三次偏差模型预测（红色实线），与仿真匹配得更紧密，残差显著减少，Perr 也被抑制。在每个面板中，功率谱按照傅里叶波数 k 和视线方向夹角参数 μ = k∥ /k 的不同分箱显示，其中 μ = 0.83 (0.17) 用深（浅）色线表示。底部面板显示仿真与模型功率谱之间的百分比差异。底部面板的灰色带突出了 ±5% 的区域。

	方差					偏度					峰度				
R	$\delta_{\rm truth}$	$\delta_{\rm lin.}$	$\delta_{\rm best-fit}$	$\Delta\delta_{\rm lin.}$	$\Delta\delta$	$\delta_{\rm truth}$	$\delta_{\rm lin.}$	$\delta_{\rm best-fit}$	$\Delta\delta_{\rm lin.}$	$\Delta\delta$	$\delta_{\rm truth}$	$\delta_{\rm lin.}$	$\delta_{\rm best-fit}$	$\Delta\delta_{\rm lin.}$	$\Delta\delta$

1 0.0159 0.0109 0.0149 0.0051 0.0010 -1.0635 -3.7654 -1.2825 0.8168 -0.0412 1.4303 26.5017 2.5290 11.4647 4.0812 2 0.0075 0.0060 0.0074
0.0015 0.0001 -0.7810 -2.7355 -0.8410 1.1211 0.0031 0.8954 13.3336 1.0496 8.4781 7.0318 5 0.0022 0.0020 0.0022 0.0002 0.0000 -0.3764 -1.4703
-0.3882 1.2552 0.0624 0.2078 3.6621 0.1712 4.7769 4.4458 10 0.0007 0.0007 0.0007 0.0000 0.0000 -0.1635 -0.8150 -0.1646 1.1150 0.0772 0.0275
1.1027 0.0022 3.2269 1.9103 30 0.0001 0.0001 0.0001 0.0000 0.0000 -0.0161 -0.2487 -0.0124 0.4271 0.0299 0.0081 0.1016 0.0037 0.6529 0.2548

表 I. 统计矩算表：Abacus 模拟 (δsim) 和微扰前向模型 (δEFT) 的统计矩（方差、偏度和峰度）以及它们在不同各向同性平滑尺度 R（单位 h−1Mpc）下的残差。根据基线预期，我们发现在去除小尺度模时，一点概率密度函数之间的匹配度增加。所有场的均值均与零一致。请注意，对于高斯场，偏度和峰度为零，这表明我们的前向模型捕捉到了场的高阶矩，即非高斯信息。

EFT 模型和模拟可以达到约 2 h−1Mpc 的单元格大小，而对于非常大的单元格大小 10−30 h−1Mpc，线性模型会失效。请注意，来自 EFT 前向模型（蓝色虚线）的像素数明显偏离从模拟（黑线）得到的分布的比例少于 0.01%。在表 I 中，我们使用应用相同高斯核平滑半径（R = 1、2、5、10、30 h−1Mpc）的场的矩来量化模型场与真实场之间的一致性。虽然这不适用于观测的 Ly-α 森林数据 [86]，但它允许我们孤立出更大的尺度。对于最佳拟合的 EFT 模型，残差（在表 I 中记为 Δδ）始终小于线性模型的残差，凸显了高阶偏差参数的重要性。

1. 红外重求和的影响

标准欧拉扰动理论无法完全捕捉重子声学振荡（BAO）特征的非线性阻尼。这源于线性物质位移，它们对长波长（红外，IR）模敏感，而这些模又导致扰动展开的收敛性较慢 [141]。尽管标准扰动理论中的各个图在红外增强项上有所体现，但当所有图相加时，这些增强会被抵消，这是等效原理所要求的 [144, 161]。然而，BAO 特征的存在阻止了完全抵消，留下了残余的红外敏感性 [142, 145]。红外重求和已被引入以

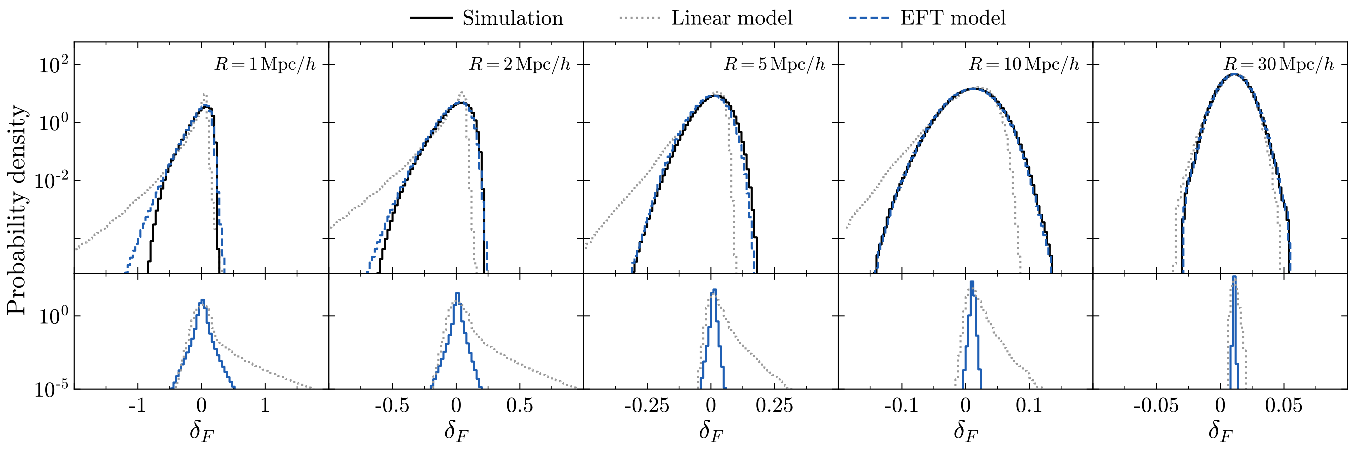

FIG. 3. **One-point PDF Abacus:** Comparison of one-point probability density function from Abacus simulation (black) to perturbative forward model (dashed blue) and linear model (gray lines) with increasing degree of isotropic Gaussian smoothing applied to the fields $R = 1 - 30\,h^{-1}\,\mathrm{Mpc}$ from left to right. The bottom panel shows the residuals. The differences between the linear forward model and the simulation are visible for *all* smoothing scales. This emphasizes the importance of the higher order terms in our bias expansion.

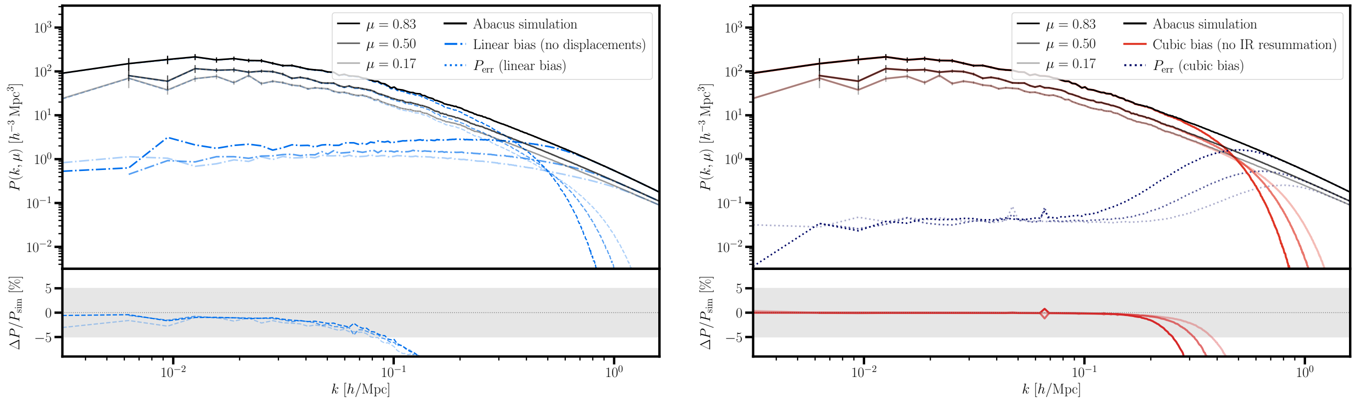

FIG. 4. **Best-Fit Abacus Power Spectrum without IR Resummation:** Same as Fig. 2 but with the bulk flows left unresummed for the linear (*left panel*) and cubic (*right panel*) forward model, to isolate the effect of IR resummation. Rather than discarding the displacements (which would also remove the displacement mode-coupling kernels and no longer constitute a consistent Eulerian PT), we Taylor-expand the shift operator to consistent order, retaining the kernels $\tilde{K}_2$, $\tilde{K}_3$ while omitting the exponential resummation ("EPT"; see text). The power spectrum is shown in three angular wedges $\mu = 0.17, 0.50, 0.83$ for the AbacusSummit simulation, with the dotted lines showing the error power spectrum $P_{\rm err}$. The broadband shape is preserved: the cubic residuals remain within 5% down to $k \approx 0.1\,h\,\mathrm{Mpc}^{-1}$ and the cubic $P_{\rm err}$ stays roughly an order of magnitude below the linear one. This confirms that the IR resummation does not affect the broadband shape (cf. the shifted/IR-resummed operator equivalence [160]). Near the BAO scale ($k \approx 0.1\,h\,\mathrm{Mpc}^{-1}$) the unresummed model fails to damp the acoustic feature, imprinting oscillatory, μ-dependent structure on $P_{\rm err}$. The bottom panel shows the percent difference between model and simulation, with a gray band marking the ±5% region.

rigorously account for these large soft modes in the BAO signal [142–147, 160, 162–167].

To isolate the effect of IR resummation we regenerate the field-level forward model *without* resumming the bulk flows. We stress that this is distinct from discarding the displacement altogether: the shift operator $e^{-i\mathbf{k}\cdot\mathbf{\Psi}}$ in Eq. (26) generates both the perturbative displacement mode-coupling and the all-orders resummation of bulk flows. Setting $\mathbf{\Psi} \to 0$ would remove the former as well, yielding a model that is no longer a consistent Eulerian PT. Instead, we Taylor-expand the shift to consistent order in the linear field, retaining the displacement kernels ($\tilde{K}_2$, $\tilde{K}_3$) while omitting the exponential resummation. Since shifted and IR-resummed Eulerian operators are equivalent [160], this expanded model agrees with the fully resummed one on the smooth (no-wiggle) component of the power spectrum and differs only in the treatment of the oscillatory (wiggle) part.

The result is shown in Fig. 4 for the linear (*left panel*) and cubic (*right panel*) expansions. In contrast to discarding

---

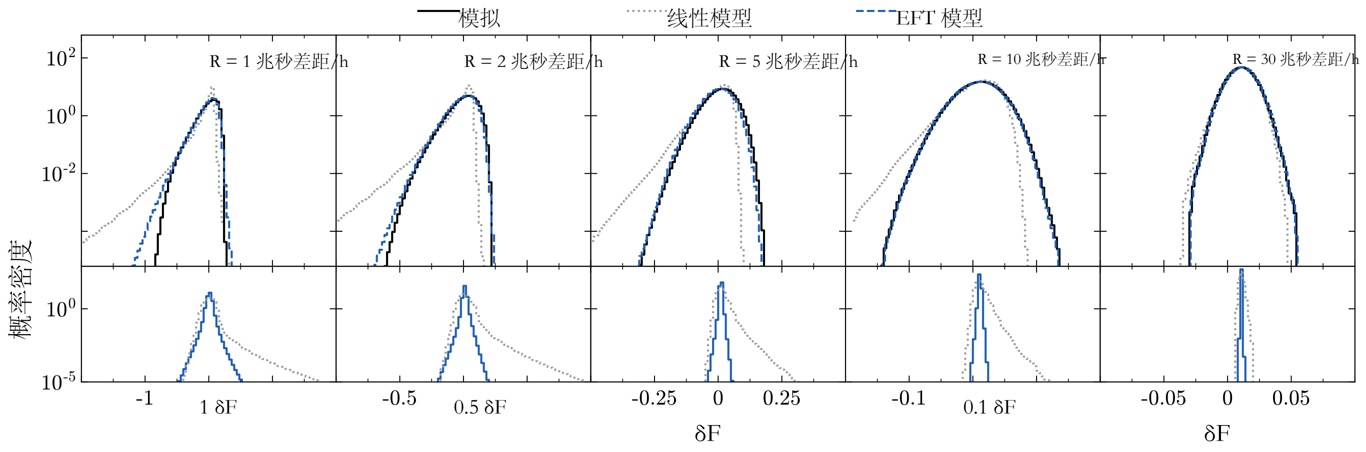

图 3. 单点 PDF 算盘图：比较来自 Abacus 模拟（黑色）的单点概率密度函数与微扰前向模型（蓝色虚线）和线性模型（灰色线）在不同各向同性高斯平滑程度下的表现，从左到右平滑半径 R = 1 − 30 h−1Mpc。底部面板显示残差。线性前向模型与模拟之间的差异在所有平滑尺度下都可见。这强调了在我们的偏差展开中高阶项的重要性。

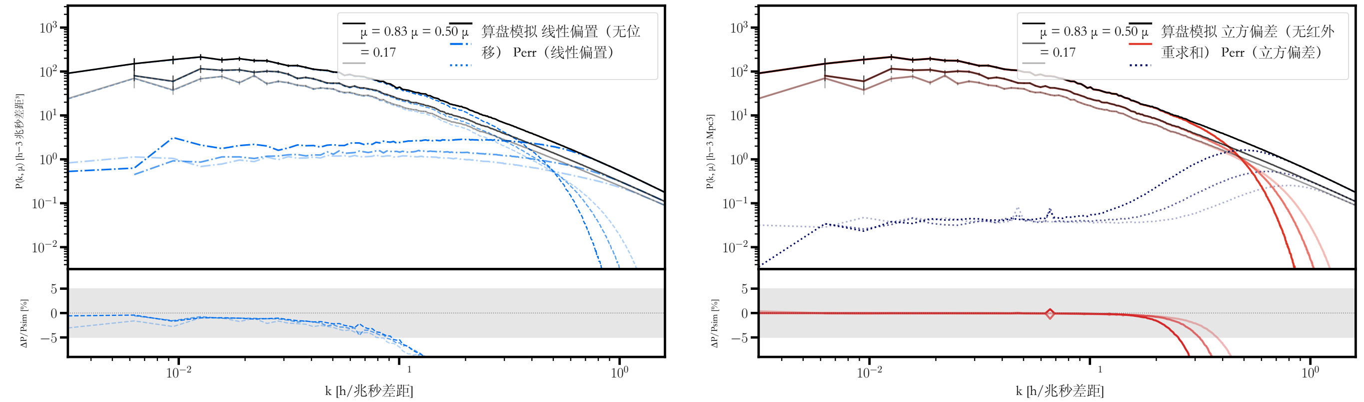

图4. 未进行红外重求和的最佳拟合Abacus功率谱：与图2相同，但散流未对线性（左图）和三次（右图）前向模型进行重求和，以隔离红外重求和的影响。我们不是丢弃位移（这也会移除位移模式耦合核，并不再构成一致的欧拉扰动理论），而是对位移算符进行一致阶次的泰勒展开，保留核 ˜ K2, ˜ K3，同时省略指数重求和（“EPT”；见正文）。功率谱在AbacusSummit模拟中以三个角向楔 μ = 0.17、0.50、0.83 显示，虚线表示误差功率谱 Perr。整体形状保持不变：三次残差在 k≈0.1 h Mpc−1 时仍在5%以内，三次 Perr 大约比线性 Perr 低一个数量级。这确认了红外重求和不会影响整体形状（参见移位/IR重求和算符等效 [160]）。在BAO尺度附近（k ≈0.1 h Mpc−1），未重求和模型未能削弱声学特征，在Perr上留下振荡的、依赖于μ的结构。底部面板显示模型与模拟之间的百分比差异，灰色带标示了±5%的区域。

严格考虑这些在 BAO 信号中的大软模 [142 − 147, 160, 162 − 167]。

为了孤立红移重求和（IR resummation）的影响，我们在不对大尺度流进行重求和的情况下重新生成场级前向模型。我们强调，这与完全丢弃位移不同：位移算子 e−ik · Ψ 在公式(26)中既生成微扰位移的模态耦合，也生成大尺度流的全阶重求和。如果将 Ψ →0，将同时去掉前者，从而产生不再符合一致的欧拉微扰理论（Eulerian PT）的模型。相反，我们将位移算子在线性场中进行一致阶的泰勒展开，保留位移核（˜K2, ˜K3），但省略指数重求和。由于移位且经红移重求和的欧拉算子是等价的 [160]，该展开模型在功率谱的平滑（无波动）部分与完全重求和模型一致，仅在振荡（波动）部分的处理上有所不同。

结果如图4所示，左图为线性展开，右图为三次展开。与舍弃相反

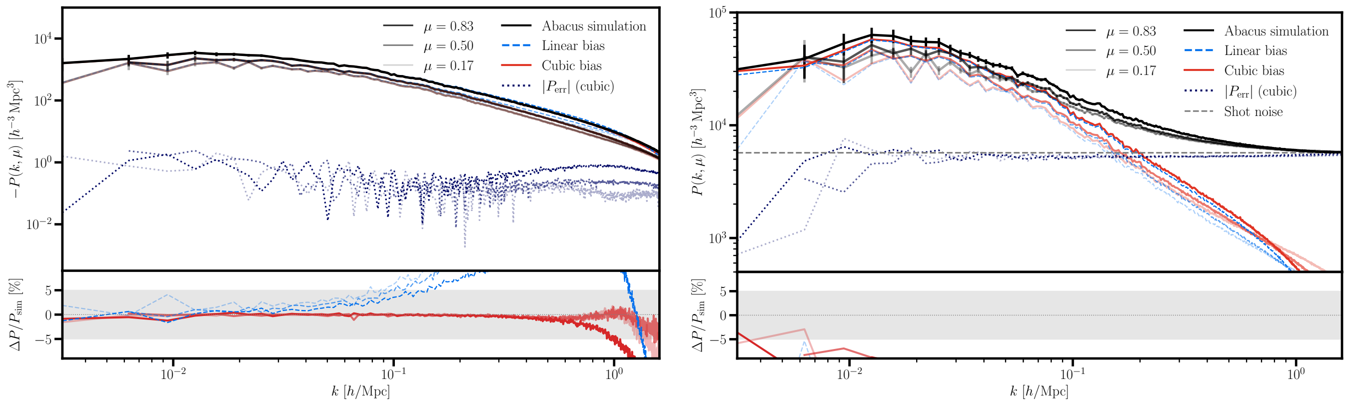

FIG. 5. **Best-Fit Abacus Power Spectra incl. Halos:** Same as Fig. 2 for the cross-correlation of the Ly-α forest with halos (*left plot*) and the halo auto-correlation using all available halo masses (*right plot*). The halo auto-correlation error power spectrum matches the shot noise ($\overline{n} = 1.75 \cdot 10^{-4}$). The model dependence of the cross spectrum is shown in Fig. 30.

the displacements altogether, the broadband shape is preserved: the percent-level residuals remain within 5% down to $k \approx 0.1\,h\,{\rm Mpc}^{-1}$ for the cubic model, which continues to outperform linear theory, with the two error power spectra separated by roughly an order of magnitude. The signature of the omitted resummation is instead localized around the acoustic scale, where the unresummed model fails to damp the BAO feature: this produces a residual mismatch between the modeled (undamped) and simulated (damped) wiggles that imprints oscillatory, μ-dependent structure on the error power spectrum in this regime. This demonstrates that IR resummation is essential specifically for recovering the BAO feature at the field level, while having a negligible effect on the broadband shape.

B. Cross-correlation of the Ly-α forest with massive halos

In this section, we extend the results from Sec. A to perform the field-level modeling in redshift space for galaxies [see, e.g., 122–125]. This allows to compute the cross correlation of the Ly-α forest with massive halos (a proxy for high-redshift galaxies and quasars). Cross-correlations are a key source of cosmological information for Ly-α analyses and help break the degeneracy of the growth rate f with the (unknown) velocity gradient bias [51, 126, 127]. In the left panel of Fig. 5 we show the cross spectrum of both tracers which matches remarkably well for the cubic model up to $k_{\rm max} \approx 1\,h\,{\rm Mpc}^{-1}$ at the five per cent level. The linear theory model fails already beyond $k \gtrsim 4 \cdot 10^{-2}\,h\,{\rm Mpc}^{-1}$. It is interesting to note that the mean-squared-error is reduced in the cross-correlation compared to the halo auto-correlation. This indicates that cross-correlation measurements, even with highly-biased tracers (as proposed by DESI-II), are a fruitful avenue to extract cosmological parameters.

The halo auto spectrum, shown in the right panel, shows a significantly larger error power spectrum which stems from the fact that the field-level model is only valid up to the shot noise. (Following baseline expectation, subtraction of the shot noise yields agreement at the per cent level down to $k \lesssim 0.3\,h\,{\rm Mpc}^{-1}$.) The halo sample for the Abacus simulation consists of 1,403,076 objects which, in turn, yields a shot noise level of $1/\overline{n} = 5.7 \cdot 10^3$ $[h^{-1}{\rm Mpc}]^3$ with masses in the range $10.8 \leq \log_{10}(M/(h^{-1}M_\odot)) \leq 14.2$.[4]

C. The error power spectrum

To quantify the performance of our forward model we use the error (or noise) power spectrum, defined in Eq. (5), which reflects the agreement at the level of the phases. In particular, a successful forward model should produce an error spectrum with a small amplitude and weak scale and orientation-dependence, i.e., EFT predicts the shape of the error power spectrum in the large scale limit given in Eq. (7). In this regime, we find the following values for the leading order stochasticity parameters at $k_{\rm max} = 0.6\ h{\rm Mpc}^{-1}$:

$$\begin{aligned} \text{Sherwood } (L = 160\,h^{-1}\,{\rm Mpc}){:} \quad & n_0 = 0.16\ [h^{-1}{\rm Mpc}]^3\,, \quad \alpha_1 = -0.22\ [h^{-1}{\rm Mpc}]^2\,, \quad \alpha_2 = 0.75\ [h^{-1}{\rm Mpc}]^2\,, \\ \text{Abacus } (L = 2\,h^{-1}\,{\rm Gpc}){:} \quad & n_0 = 0.04\ [h^{-1}{\rm Mpc}]^3\,, \quad \alpha_1 = -0.06\ [h^{-1}{\rm Mpc}]^2\,, \quad \alpha_2 = 0.17\ [h^{-1}{\rm Mpc}]^2\,. \end{aligned} \tag{36}$$

---

[4] We verified that splitting the mass range to select massive or light halos does not affect our conclusion and we observer the same behavior. Further we verified that including additional cubic operators, such as $\mathcal{S}_3$, does not improve the error power spectrum.

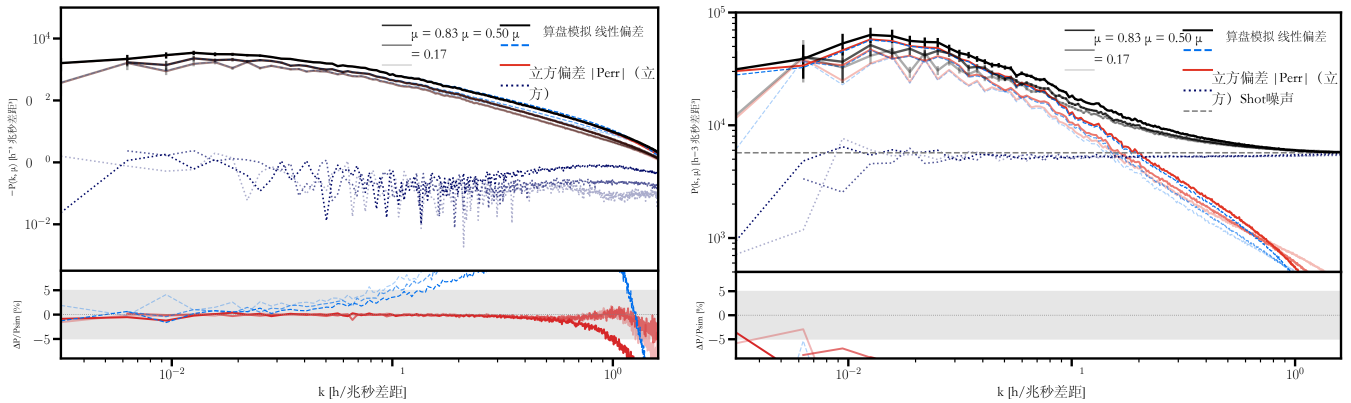

图5. 包含晕的最优拟合Abacus功率谱：与图2相同，但用于Ly-α森林与晕的交叉相关（左图）以及使用所有可用晕质量的晕自相关（右图）。晕自相关误差功率谱与泊松噪声相符（n = 1.75 · 10−4）。交叉谱的模型依赖性如图30所示。

总的位移下，宽带形状得以保持：对于三次模型，百分比级的残差仍保持在 5% 以内，直到 k ≈0.1 h Mpc−1，并且继续优于线性理论，两者的误差功率谱大约相差一个数量级。被省略的重排的特征则集中在声学尺度附近，此时未重排模型未能抑制 BAO 特征：这导致建模（未抑制）和模拟（已抑制）波动之间的残差不匹配，在该区域的误差功率谱上留下随 μ 变化的振荡结构。这表明 IR 重排对于在场水平上恢复 BAO 特征是必不可少的，而对宽带形状几乎没有影响。

B. Ly-α森林与大质量晕的交叉相关

在本节中，我们将 A 节的结果扩展到在红移空间对星系进行场级建模 [参见，例如 122 – 125]。这使得计算 Ly-α 森林与大质量晕（作为高红移星系和类星体的代理）的交叉相关成为可能。交叉相关是 Ly-α 分析中宇宙学信息的关键来源，并有助于打破增长率 f 与（未知的）速度梯度偏置之间的退化 [51, 126, 127]。在图 5 的左侧面板中，我们展示了两种示踪物的交叉谱，对于立方模型来说，直到 kmax ≈1 h Mpc−1 的五个百分点水平匹配得非常好。线性理论模型在 k> ~4 · 10−2 h Mpc−1 就已经失效。有趣的是，与晕自相关相比，交叉相关的均方误差有所减少。这表明，即使使用高度偏置的示踪物（如 DESI-II 提出的），交叉相关测量也是提取宇宙学参数的重要途径。

右侧面板所示的晕自动功率谱显示了明显更大的误差功率谱，这是因为场级模型仅在散弹噪声范围内有效。(按照基线预期，减去散弹噪声后，在 k< ~0.3 h Mpc−1 时可实现百分比水平的吻合。) Abacus 模拟的晕样本包含 1,403,076 个天体，从而产生的散弹噪声水平为 1/n = 5.7 · 10^3 [h−1Mpc]^3，质量范围为 10.8 ⩽ log10(M/(h−1M⊙)) ⩽ 14.2。

C. 误差功率谱

为了量化我们的前向模型的性能，我们使用误差（或噪声）功率谱，如公式（5）所定义，它反映了相位层面的吻合程度。特别是，一个成功的前向模型应当生成一个振幅较小且尺度和方向依赖性较弱的误差谱，即 EFT 在大尺度极限下预测了误差功率谱的形状，如公式（7）所示。在此范围内，我们在 kmax = 0.6 hMpc−1 下得到以下高阶随机性参数的值：

舍伍德 (L = 160 h−1Mpc): n0 = 0.16 [h−1Mpc]³, α1 = −0.22 [h−1Mpc]², α2 = 0.75 [h−1Mpc]2,
算珠（L = 2 h−1Gpc）: n0 = 0.04 [h−1Mpc]³, α1 = −0.06 [h−1Mpc]2, α2 = 0.17 [h−1Mpc]2。 (36)

---

4 我们验证了将质量范围分开以选择大质量或轻质量晕不会影响我们的结论，我们观察到相同的行为。此外，我们还验证了包括额外的三次算符，例如 S3，并不能改善误差功率谱。

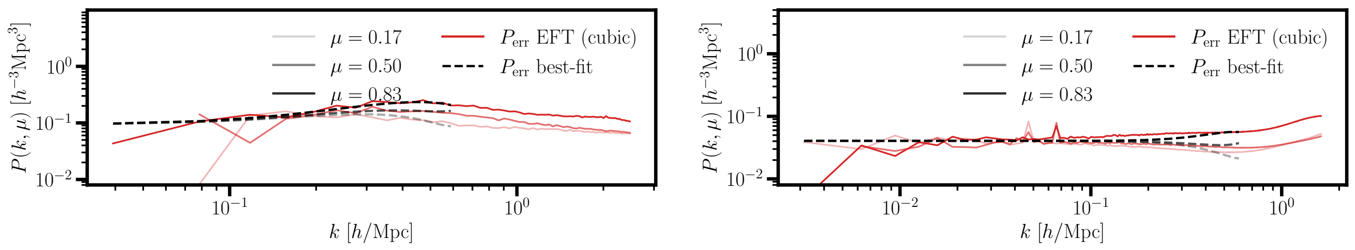

FIG. 6. **Redshift-space error power spectrum** $P_{\rm err}$ corresponding to the best-fit transfer functions shown in Fig. 15 for the Sherwood simulation (*left panel*) and Abacus simulation (*right panel*). $P_{\rm err}$ is computed in three angular bins with a maximum wavenumber of $k = 1\,h\,{\rm Mpc}^{-1}$ for the fits. We show the polynomial fits to $P_{\rm err}$ up to $k = 0.6\,h\,{\rm Mpc}^{-1}$, the wavenumber up to which we can accurately reconstruct the power spectrum. The scale and angular dependence sets in for both simulations at around $k \approx 0.4\,h\,{\rm Mpc}^{-1}$. The best-fit coefficients are tabulated in Tab. II. For Abacus the best-fit error power spectra on models two to four are shown in Fig. 29.

These measurements are consistent with the estimates given in Eq. (8), pointing to a characteristic stochasticity scale $R_{\rm stoch} \sim 0.5\ h^{-1}$Mpc for Sherwood and $R_{\rm stoch} \sim 0.3\ h^{-1}$Mpc for Abacus. In addition, these values are consistent with the Ly-α stochasticity measurements from the ACCEL^2 simulations [77], though the latter have significant errorbars. To the best of our knowledge, this is the first ever precise determination of the stochasticity of the Lyman-α forest field.

Interestingly, we see a high level of consistency between Sherwood and Abacus simulations in term of their stochasticity, even though they represent very different small-scale models. Since the Abacus mocks capture dark matter physics only, the similarity between the Sherwood and Abacus results suggest that the Lyman-α stochasticity may be driven by the stochasticity of the dark matter distribution.

In addition, we notice that the scale dependence of the stochastic power spectrum becomes more shallow around 1 hMpc^{-1}, where the error power spectrum contributes significantly (by more than 5%) to the total power. This signals the breakdown of the EFT gradient expansion for the stochastic effects. Theoretical modeling of the error power spectrum in this regime may represent a serious challenge. The connection to dark matter stochasticity mentioned above opens up the possibility to model the Ly-α stochasticity with the hybrid EFT technique where the Lagrangian bias expansion for the Lyman-α forest is supplemented with the non-linear displacements from an N-body simulation. This topic is further explored in Ref. [159], where it has been shown that the incorporation of the non-linear displacements that include the stochastic components in them indeed allows one to reduce the scale and orientation dependence of the error power spectrum extending the range of validity of EFT beyond $k \approx 1\,h\,{\rm Mpc}^{-1}$.

To investigate the error power spectrum beyond the $k \to 0$ limit, we fit the following functional form to it

$$P_{\rm err}(k,\mu) = a_0 + a_2k^2 + a_3k^3 + a_4k^4 + \sum_{i=2,4} a_{ii}(k\mu)^i\,, \tag{37}$$

where we have the k^3 term to account for the observed smooth scaling of the error power spectrum in the non-linear regime. Odd powers of k in the error power spectrum are forbidden in EFT; their inclusion here is purely phenomenological. The best-fit parameters for both sets of simulations are tabulated in Tab. II and a subset are shown in Fig. 6. Following baseline expectation, the error power spectrum has a an amplitude that is more than two orders of magnitude smaller than the signal and constant on large scales. In particular, the onset of the scale-and orientation-dependence occurs around $k \approx 0.4 - 0.6\,h\,{\rm Mpc}^{-1}$ ultimately setting an upper limit on the applicability of EFT as a theory model (see Ref. [159]). We perform the same fits to the error power spectrum obtained from the cross-correlation of the Ly-α forest with halos and find that the amplitude of $P_{\rm err}$ is not driven by the larger one obtained from the halos but rather by the one from the Ly-α forest.

D. Cross-correlation coefficient

A second stringent test of the field-level model is provided by the cross-correlation coefficient, r_{cc}, defined in Eq. (35). It relates the power spectrum of the simulated field to the mean-squared model error through $P_{\rm err} = P_{\rm truth}(1 - r_{cc}^2)$ and thus quantifies how faithfully the model reproduces the phases of the field. As such, r_{cc} is our key metric for assessing the range of validity of the forward model. However, the scale at which the error power spectrum begins to rise with r_{cc} correspondingly departing from unity does not by itself define the maximum scale of applicability. We

---

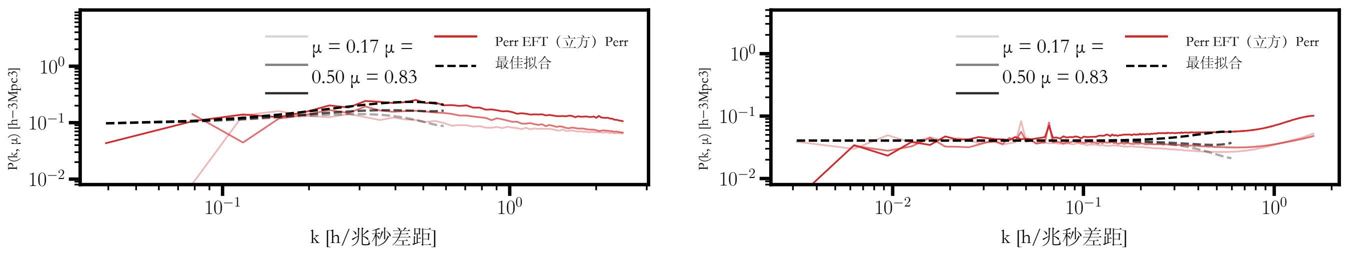

图6. 与图15中所示谢伍德模拟（左图）和Abacus模拟（右图）的最佳拟合传递函数对应的红移空间误差功率谱 Perr。Perr 在三个角度区间内计算，拟合的最大波数为 k = 1 h Mpc−1。我们展示了对 Perr 的多项式拟合，适用于 k = 0.6 h Mpc−1 以内，这是我们能够准确重建功率谱的波数范围。尺度和角度依赖性在两个模拟中均在大约 k ≈ 0.4 h Mpc−1 时出现。最佳拟合系数列于表 II 中。对于 Abacus，模型二到模型四的最佳拟合误差功率谱显示在图29中。

这些测量结果与方程(8)中给出的估计值一致，指出Sherwood的特征随机性尺度约为 Rstoch ~0.5 h−1Mpc，而Abacus约为 Rstoch ~0.3 h−1Mpc。此外，这些值与来自ACCEL2模拟[77]的Ly-α随机性测量结果一致，尽管后者的误差较大。据我们所知，这是迄今为止首次对莱曼-α森林场的随机性进行精确测定。

有趣的是，尽管 Sherwood 和 Abacus 模拟代表非常不同的小尺度模型，但在随机性方面，我们仍然看到它们之间具有高度一致性。由于 Abacus 模拟仅捕捉暗物质物理特性，Sherwood 和 Abacus 结果的相似性表明 Lyman-α 的随机性可能是由暗物质分布的随机性驱动的。

此外，我们注意到随机功率谱的尺度依赖在约 1 hMpc−1 附近变得更为平缓，此时误差功率谱对总功率的贡献显著（超过 5%）。这表明随机效应的有效场论梯度展开出现了失效。在这一范围内对误差功率谱进行理论建模可能是一项重大挑战。上述与暗物质随机性相关的联系为使用混合有效场论技术对 Lyman-α 随机性建模提供了可能，其中 Lyman-α 森林的拉格朗日偏差展开式与 N 体模拟中的非线性位移相结合。该主题在文献 Ref. 中有进一步探讨。[159]，其中已显示，包含随机分量的非线性位移的引入确实可以减少误差功率谱的尺度和方向依赖性，从而将EFT的有效范围扩展到 k ≈1 h Mpc−1 以上。为了研究 k →0 极限之外的误差功率谱，我们将以下函数形式拟合到其上

$$\text{Perr(k, μ) = a0 + a2k2 + a3k3 + a4k4 +} \quad \underset{i=2,4}{} a_{ii}(k\mu)^i\,, \tag{37}$$

在这里我们引入了 k3 项，以解释非线性区域中观测到的误差功率谱的平滑标度。误差功率谱中的奇数 k 次幂在有效场论（EFT）中是被禁止的；这里的引入完全是出于经验性的考虑。两组模拟的最佳拟合参数列在表 II 中，其中一部分展示在图 6 中。按照基线预期，误差功率谱的振幅比信号小两个数量级以上，并且在大尺度上是常数。特别地，尺度和方向依赖性的出现大约在 k ≈ 0.4 − 0.6 h Mpc−1，这最终决定了 EFT 作为理论模型的适用上限（参见参考文献 [159]）。我们对从 Ly-α 森与哈密顿量的交叉相关中得到的误差功率谱进行了相同的拟合，发现 Perr 的振幅并不是由来自哈密顿量的较大值驱动，而是由来自 Ly-α 森的值驱动。

D. 互相关系数

场级模型的第二项严格检验由交叉相关系数 rcc 提供，其定义见公式 (35)。它通过 Perr = Ptruth(1 − rcc²) 将模拟场的功率谱与均方模型误差联系起来，从而量化模型对场相位重现的忠实度。因此，rcc 是我们评估前向模型适用范围的关键指标。然而，误差功率谱开始随 rcc 上升并相应偏离 1 的尺度，并不能单独定义适用的最大尺度。我们

Data		a_0	a_2	a_3	a_4	a_{22}	a_{44}
Abacus	I	0.041	−0.097	0.016	0.069	0.230	−0.276
–	II	0.048	−0.195	0.433	−0.345	0.046	0.235
–	III	0.069	−0.531	1.149	−0.814	0.358	−0.570
–	IV	0.063	−0.549	1.163	−0.820	0.438	−0.668
–	I × halos	0.116	0.061	0.841	−1.486	2.368	−0.945
–	II × halos	0.121	−0.226	1.461	−1.798	2.251	−1.018
–	III × halos	0.254	1.897	−5.069	2.916	2.499	−2.814
–	IV × halos	0.268	1.694	−4.875	2.833	2.819	−2.958
Sherwood	$L = 160\,h^{-1}\,\mathrm{Mpc}$	0.154	−0.005	−0.412	0.316	0.404	−0.337
–	× halos	0.034	0.955	−3.074	2.096	1.441	−1.092

TABLE II. **Best-fit P_{err} fits:** Coefficients for the polynomial fit to P_{err} given in Eq. (37) for the Abacus simulations (*top eight rows*) and the Sherwood simulation (*bottom two row*), shown in Fig. 6. In addition we quote the best-fit parameters from fits to the error power spectrum of the cross-correlation with all available halos, denoted by "× halos". The corresponding transfer functions are tabulated in Tabs. III and IV and a_n, a_{nn} are in units of $[h^{-1}\mathrm{Mpc}]^{3+n}$. The best-fit plots for Abacus models two to four are shown in Fig. 29.

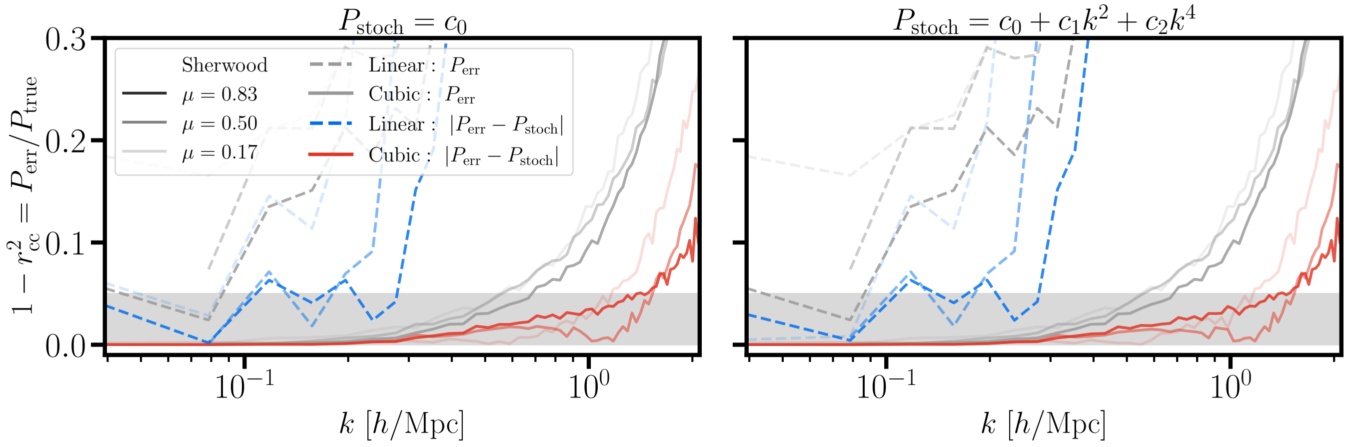

FIG. 7. **Cross-Correlation Coefficients Ly-α Sherwood:** Fractional mean-squared model error $1 - r_{cc}^2 = |P_{\mathrm{err}}/P_{\mathrm{true}}|$ where $r_{cc} = r_{cc}(\delta^{\mathrm{truth}}, \delta^{\mathrm{model}})$ is the cross-correlation coefficient between simulations and model. The results are shown for the Sherwood simulation using the linear (dashed lines) and cubic EFT model (solid lines) in three angular bins, μ. The gray lines denote the "raw" cross-correlation coefficient and the colored lines illustrate the subtraction of the stochastic counterterms, given in Eq. (38), from the error power spectrum. The gray shaded region indicates the 5% error band.

therefore examine how well the error power spectrum can be captured by EFT stochastic counterterms [136],

$$P_{\mathrm{1D}}^{\mathrm{stoch.}}(k_{\parallel}) = c_0 + c_1 k_{\parallel}^2 + c_2 k_{\parallel}^4\,, \tag{38}$$

where the coefficient c_n has units of $[h^{-1}\mathrm{Mpc}]^{(2n+1)}$ for $n = 0, 1, 2$. As discussed in Ref. [87], the small-scale modeling is affected by non-Poisson stochasticity (the analog of the one-halo term in galaxy clustering), which renders the error power spectrum scale- and direction-dependent in a way that EFT cannot capture beyond the gradient expansion. This stochastic noise sets the limit of applicability of EFT [122, 123, 168, 169], which we discuss in the following for the 3D power spectrum and in Sec. E for the 1D power spectrum.

In Fig. 7 we show the cross-correlation coefficients obtained for the linear (dashed lines) and cubic EFT model (solid lines) for the large Ly-α Sherwood simulation at $z = 2.8$. The gray lines use the "raw" error power spectrum, while the colored lines are obtained after subtracting the stochastic counterterms [see Eq. (38)] from it. Before subtraction, the cubic EFT model reproduces the phases in Fourier space to better than 1% up to $k \lesssim 0.3\,h\,\mathrm{Mpc}^{-1}$ ($k \lesssim 0.24\,h\,\mathrm{Mpc}^{-1}$) for Abacus (Sherwood), with the 5%-level reached at $k \lesssim 0.7\,h\,\mathrm{Mpc}^{-1}$ ($k \lesssim 0.6\,h\,\mathrm{Mpc}^{-1}$). Since Abacus covers a larger volume, it contains more quasi-linear modes and the linear model captures a larger range of scales than for Sherwood for which it captures the large-scale modes only at the 3% on large scales.

Data		a_0	a_2	a_3	a_4	a_{22}	a_{44}
算盘	我	0.041	−0.097	0.016	0.069	0.230	−0.276
–	二	0.048	−0.195	0.433	−0.345	0.046	0.235
–	三	0.069	−0.531	1.149	−0.814	0.358	−0.570
–	IV	0.063	−0.549	1.163	−0.820	0.438	−0.668
–	I × halos	0.116	0.061	0.841	−1.486	2.368	−0.945
–	II halos	0.121	−0.226	1.461	−1.798	2.251	−1.018
–	III × 光环	0.254	1.897	−5.069	2.916	2.499	−2.814
–		0.268	1.694	4.875	2.833	2.819	−2.958
舍伍德 L = 160 h−1Mpc		0.154	−0.005	−0.412	0.316	0.404	−0.337
–	× halos				2.096	1.441	−1.092

表 II. 最佳拟合 Perr 拟合：用于 Abacus 模拟（前八行）和 Sherwood 模拟（后两行）的 Perr 多项式拟合系数，如方程 (37) 所示，见图 6。此外，我们还引用了对与所有可用晕的交叉相关误差功率谱拟合的最佳拟合参数，用“× halos”表示。相应的传递函数列在表 III 和 IV 中，an、ann 的单位为 [h−1Mpc]^(3+n)。Abacus 模型二到四的最佳拟合图见图 29。

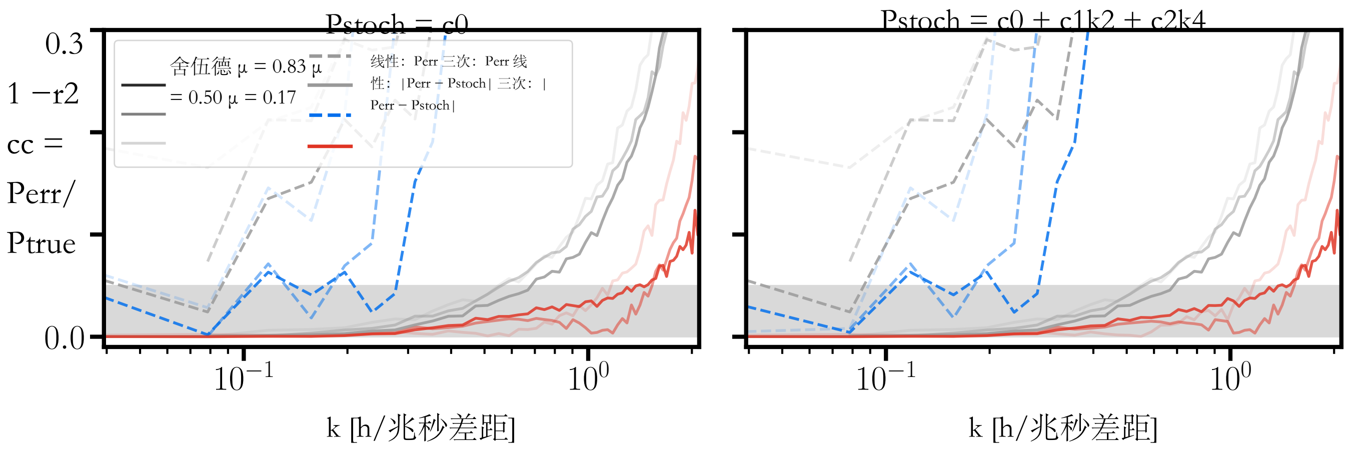

图 7. 交叉相关系数 Ly-α Sherwood：分数均方模型误差 1 − r^2_cc = |P_err / P_true|，其中 r_cc = r_cc(δ_truth, δ_model) 是模拟与模型之间的交叉相关系数。结果显示了 Sherwood 模拟中使用线性（虚线）和三次 EFT 模型（实线）在三个角度区间 μ 下的情况。灰色线表示“原始”交叉相关系数，彩色线表示从误差功率谱中减去随机反项（Eq. (38) 所给）。灰色阴影区域表示 5% 误差带。

因此，研究误差功率谱能在多大程度上被 EFT 随机对项所捕捉 [136]，

$$P_{\mathrm{1D}}^{\mathrm{stoch.}}\,(k\,/\!/\,) = c0 + c1k^2\,/\!/\; + c2k^4\,/\!/\,, \tag{38}$$

其中系数 cn 的单位为 [h−1Mpc](2n+1)，n = 0, 1, 2。如文献 [87] 所述，小尺度建模受到非泊松随机性的影响（类似于星系聚类中的单晕项），这使得误差功率谱在尺度和方向上依赖的方式超过了 EFT 在梯度展开之外所能捕捉的范围。这种随机噪声决定了 EFT 的适用范围 [122, 123, 168, 169]，我们将在下文讨论 3D 功率谱，并在附录 E 中讨论 1D 功率谱。

在图7中，我们显示了在 z = 2.8 的大型 Ly-α Sherwood 模拟中获得的线性（虚线）和三次 EFT 模型（实线）的互相关系数。灰色曲线使用“原始”误差功率谱，而彩色曲线是在从中减去随机对偶项后获得的 [参见公式 (38)]。在减去之前，对于 Abacus（Sherwood），三次 EFT 模型在傅里叶空间中再现相位的精度优于 1%，直到 k< ~0.3 h Mpc−1（k < ~0.24 h Mpc−1），达到 5% 误差水平为 k< ~0.7 h Mpc−1（k < ~0.6 h Mpc−1）。由于 Abacus 覆盖了更大的体积，它包含更多准线性模，因此线性模型能够捕捉到比 Sherwood 更大范围的尺度，而 Sherwood 只能在大尺度上以 3% 的水平捕捉大尺度模。

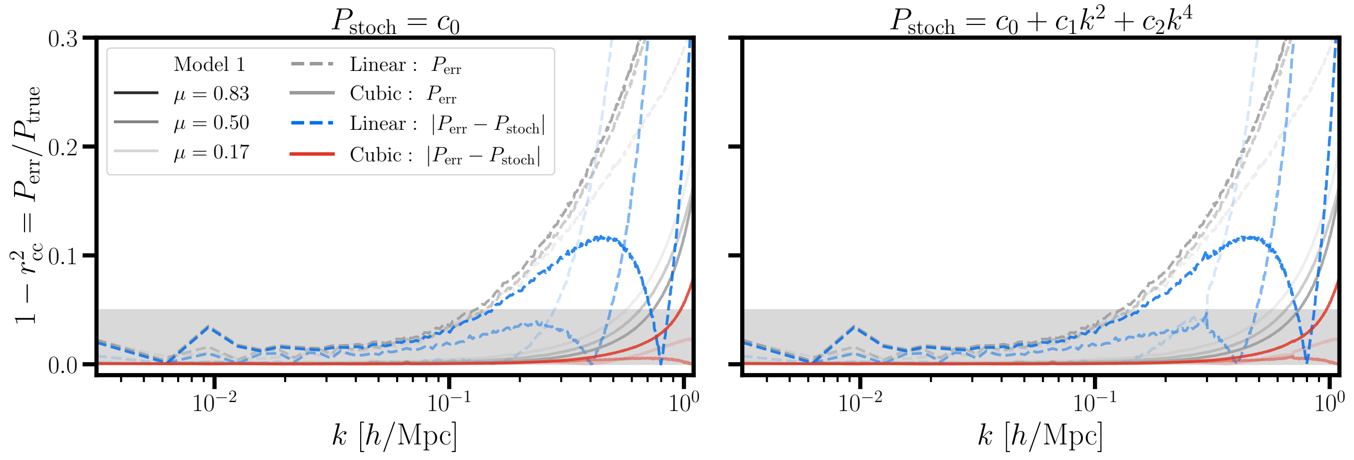

FIG. 8. **Cross-Correlation Coefficients Ly-α Abacus:** Same as Fig. 7 for the Abacus simulation model I. The turnover in the linear model is an artifact from an over-subtraction at high k and the result of plotting the norm.

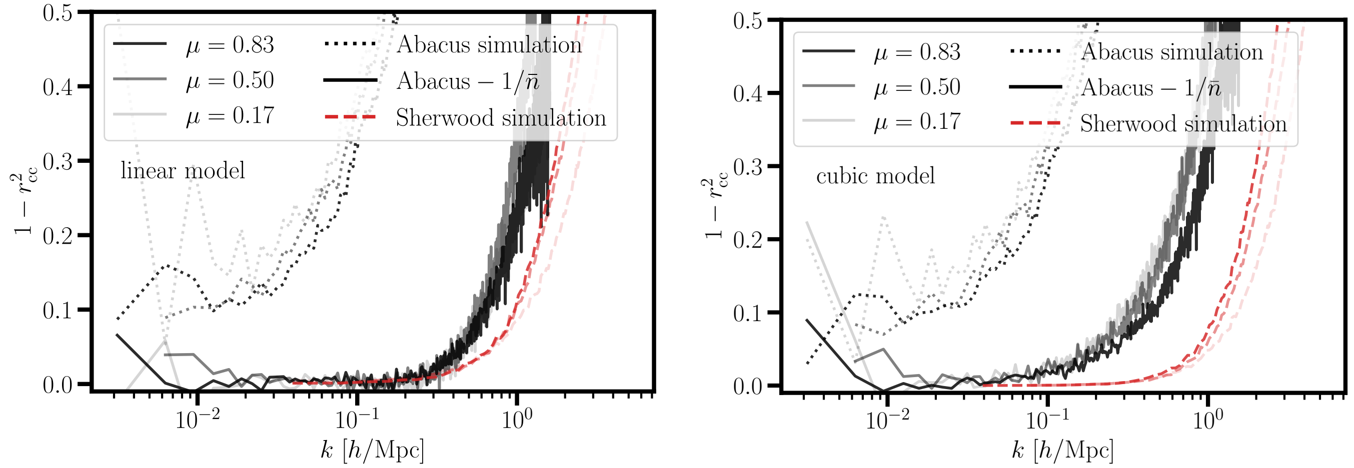

FIG. 9. **Cross-Correlation Coefficients Halos:** Same as Fig. 7 for the halo auto-correlation using all available halo masses for both simulations. As discussed in Fig. 5, the error power spectrum for Abacus converges to the shot noise, setting a clear upper bound on the cross-correlation coefficient. Therefore, we show the shot noise subtracted version (solid black line) and r_{cc} including the shot noise (dotted black line). This shows that for quasar (or galaxy) samples with lower shot noise, the field-level methodology extends its reach – similar to the Ly-α forest. For Sherwood the cross-correlation coefficient is very close to one since the provided halo catalogs are (unrealistically) light.

After subtracting the stochastic contribution from the error power spectrum[5], we extend the reach of both models: for Sherwood, the linear model reaches $k \lesssim 0.3\,h\,{\rm Mpc}^{-1}$ at the 5% level and the cubic model reaches $k \lesssim 0.5\,(1.3)\,h\,{\rm Mpc}^{-1}$ at the 1% (5%) level, with a similar performance for Abacus as shown in Fig. 8. From this analysis, the majority of the improvement stems from subtracting a constant offset with little weight stemming from the c_1 and c_2 terms. We note that going to higher redshifts, e.g., $z = 3.2$ using the smaller Sherwood boxes, does not significantly improve the performance of the cubic model for the 3D power spectrum. We emphasize that, while both simulations use entirely different prescriptions to model the Ly-α forest, they yield consistent results, demonstrating the robustness of the field-level methodology and setting the maximum scale for linear theory at $k \lesssim 0.1 - 0.2\,h\,{\rm Mpc}^{-1}$.[6]

---

[5] This procedure is equivalent to the joint modeling of the deterministic and stochastic parts of the Ly-α power spectrum.

[6] It is interesting to compare the obtained cross-correlation coefficient to the one obtained from other approaches, e.g. using the promising deep learning reconstruction presented in figure 6 in Ref. [114]. Whilst their cross-correlation coefficient significantly improves using their reconstruction approach, their model has a floor of approximately five per cent on large scales ($k \lesssim 1\,h\,{\rm Mpc}^{-1}$). We leave the exploration of a hybrid approach combining EFT on large scales with a deep-learning-based reconstruction on small scales to future work.

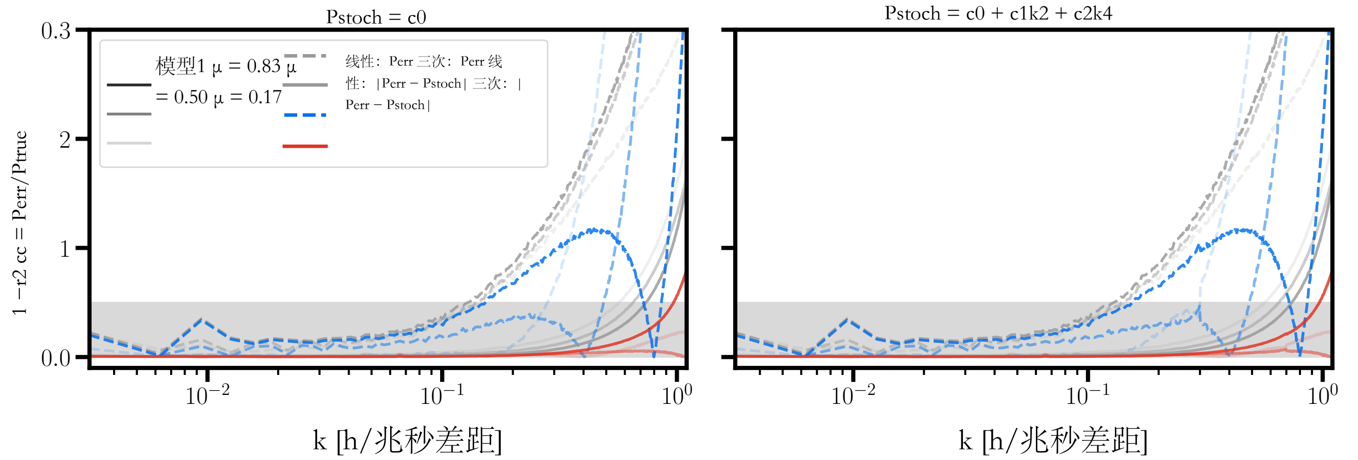

图8. 交叉相关系数 Ly-α Abacus：与图7相同，但针对 Abacus 模拟模型 I。线性模型的拐点是由于在高 k 下过度减法引起的人为现象，同时也是绘制归一化结果的结果。

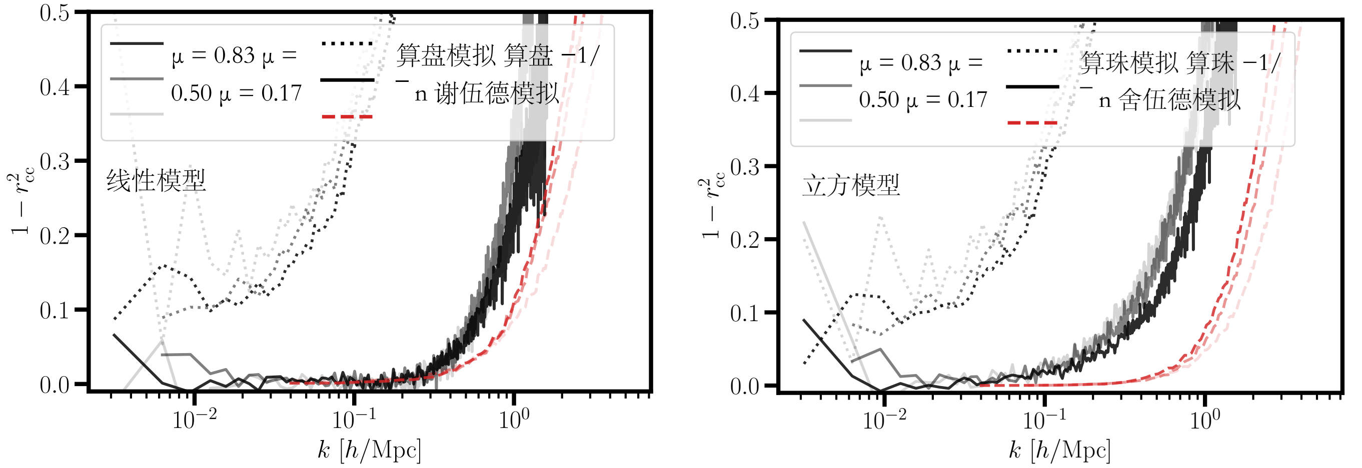

图9. 交叉相关系数晕: 与图7相同，用于使用两个模拟的所有可用晕质量进行晕自相关。如图5中讨论的那样，Abacus的误差功率谱收敛到散粒噪声，为交叉相关系数设定了明确的上限。因此，我们展示了减去散粒噪声的版本（实黑线）以及包含散粒噪声的rcc（虚黑线）。这表明，对于散粒噪声较低的类类星体（或星系）样本，场层方法可以扩展其适用范围——类似于Ly-α森林。对于Sherwood，由于提供的晕目录（不现实地）较轻，交叉相关系数非常接近1。

在从误差功率谱中减去随机贡献之后，我们扩展了两种模型的适用范围：对于 Sherwood，线性模型在 5% 水平下可达到 k< ~0.3 h Mpc−1，而三次模型在 1%（5%）水平下可分别达到 k< ~0.5（1.3）h Mpc−1，Abacus 的表现与图 8 所示类似。通过这一分析，大部分改进来自于减去一个常数偏移，c1 和 c2 项的权重很小。我们注意到，在更高红移（例如使用较小的 Sherwood 盒子在 z = 3.2）时，三次模型对于三维功率谱的性能并没有显著提高。我们强调，尽管两种模拟都完全使用不同的方法来模拟 Ly-α 林，它们仍产生一致的结果，证明了现场级方法的稳健性，并将线性理论的最大尺度设定在 k< ~0.1−0.2 h Mpc−1。

---

5 这一过程相当于对 Ly-α 功率谱的确定性部分和随机部分的联合建模。6 将得到的互相关系数与其他方法得到的系数进行比较是很有趣的，例如使用参考文献 [114] 图 6 中展示的有前景的深度学习重建方法。虽然他们使用重建方法显著提高了互相关系数，但他们的模型在大尺度（k< ~1 h Mpc−1）上仍有大约五个百分点的下限。我们将结合大尺度上的 EFT 与小尺度上基于深度学习的重建的混合方法的探索留作未来研究。

工作。

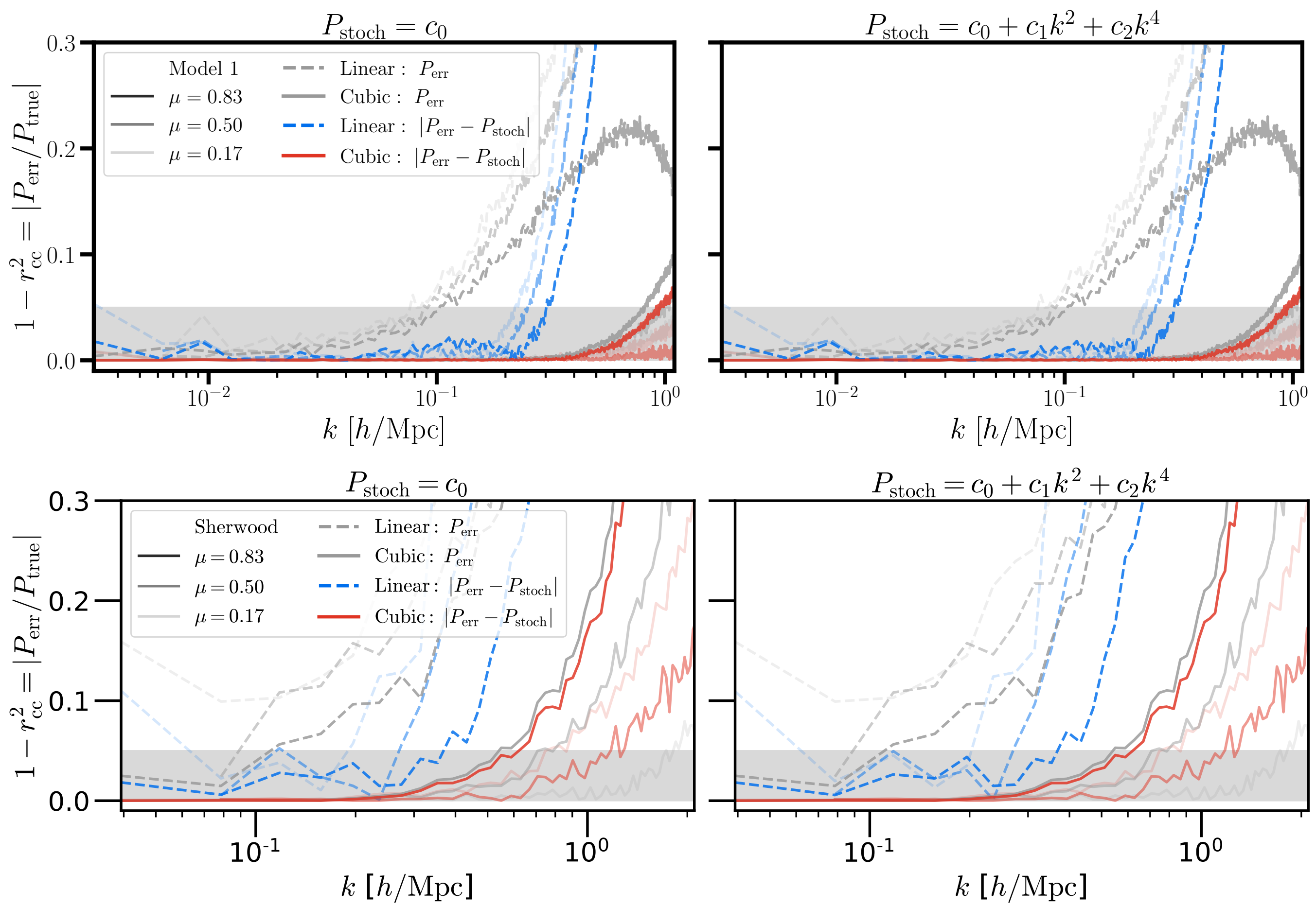

FIG. 10. **Cross-correlation coefficient Ly-α × Halos:** Same as Figs. 7 for the cross-correlation of the Ly-α forest with halos for Abacus Model I (*top row*) and Sherwood (*bottom row*) using all available halo masses. The gray lines denote the “raw” cross-correlation coefficient and the colored lines illustrate the subtraction of the stochastic counterterms, given in Eq. (38), from the error power spectrum. The gray shaded region indicates the 5% error band.

In addition to the Ly-α forest, we show the fractional mean-squared model error for the halo auto-correlation in Fig. 9, using the linear model (left panel) and the cubic model (right panel). Here the cubic model offers almost no improvement over the linear theory model, indicating that we are in the shot-noise limited regime for halos. Massive and light halos serve as proxies for quasars and high-redshift galaxies, respectively. In contrast to those from Abacus, the Sherwood halos are light and carry a very low shot noise; we therefore subtract the shot noise only from Abacus in Fig. 9, moving from the dotted to the solid black lines for both models.

Finally, in Fig. 10 we show the 3D cross-correlation coefficient for the cross-correlation of the Ly-α forest with halos, using all available halo masses for Abacus (top row) and Sherwood (bottom row). As before, we model the stochastic contributions to the error power spectrum: this yields a noticeable improvement for the linear model (gray to blue dashed lines), extending its reach to $k \lesssim 0.1\,h\,{\rm Mpc}^{-1}$, and a negligible improvement for the cubic model (gray to red solid lines), which reaches the 5%-level at $k \lesssim 1\,h\,{\rm Mpc}^{-1}$. Overall, the cubic forward model provides an excellent fit to the cross-correlation data.

E. One-dimensional power spectrum

In Fig. 11 we additionally compare the one-dimensional power spectrum – a key summary statistic in the context of Ly-α forest analyses (see, e.g., [5–12]). The left panel shows the model performance without any smoothing applied to the simulation and the cubic forward model recovers the shape of the true power spectrum well while showing a residual offset of ∼ 15%. In contrast, the linear theory model fails at recovering the P1D of the input simulation,

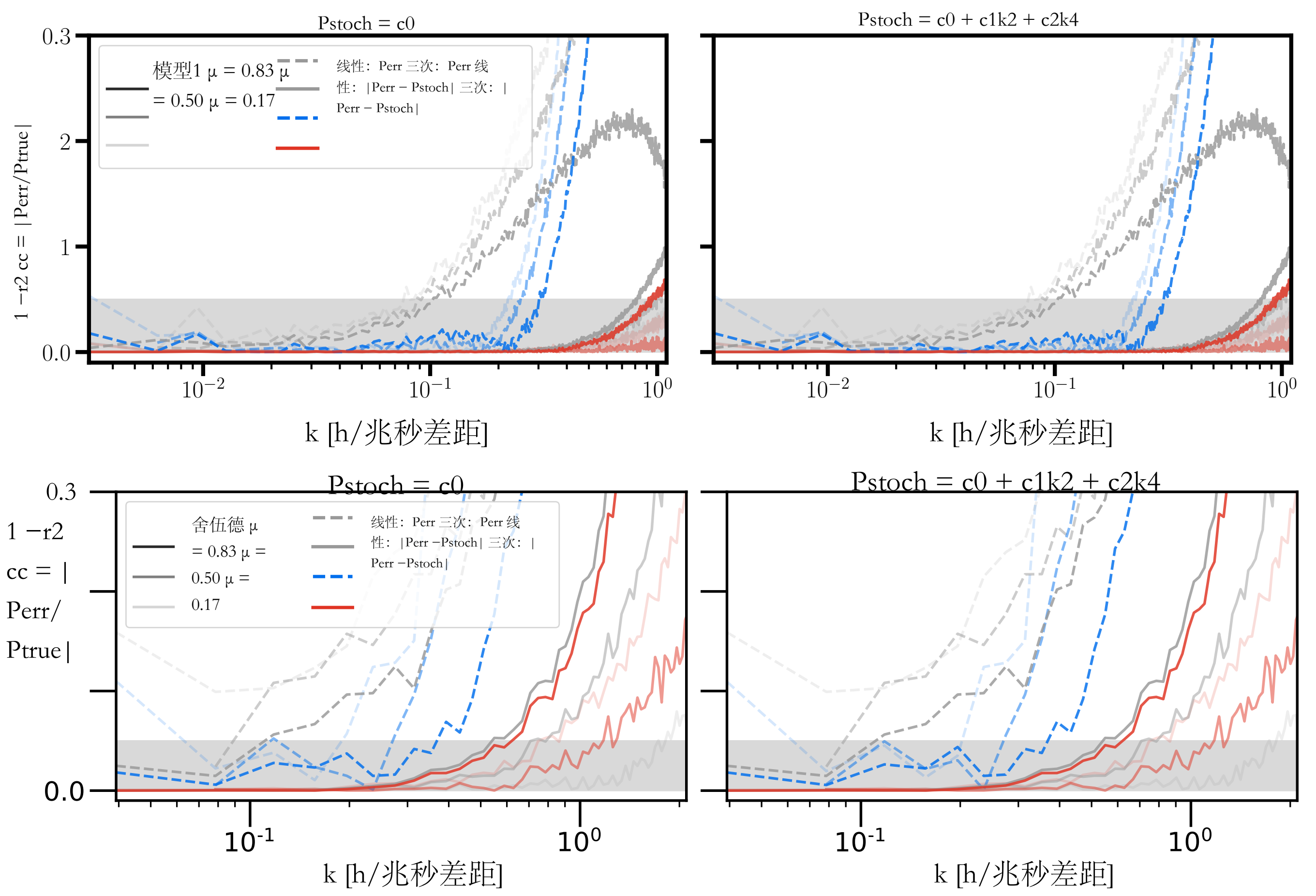

图 10. Ly-α × 暗晕的互相关系数：与图 7 相同，用于 Abacus 模型 I（上排）和 Sherwood（下排）中 Ly-α 森林与暗晕的互相关，使用所有可用的暗晕质量。灰色曲线表示“原始”互相关系数，彩色曲线显示从误差功率谱中减去随机补偿项（见公式 (38)）。灰色阴影区域表示 5% 的误差范围。

除了Ly-α森林之外，我们在图9中展示了使用线性模型（左图）和立方模型（右图）时，晕自相关的平均平方模型误差分数。在这里，立方模型几乎没有比线性理论模型提供改进，这表明我们处于晕的散粒噪声限制区域。大质量和低质量晕分别作为类星体和高红移星系的代理。与Abacus的结果相比，Sherwood的晕质量较低且具有非常低的散粒噪声；因此在图9中，我们仅从Abacus中减去散粒噪声，使两种模型从虚线变为黑色实线。

最后，在图10中，我们展示了使用所有可用的Abacus（上行）和Sherwood（下行）晕质量的Ly-α森林与晕的三维互相关系数。和以前一样，我们对误差功率谱的随机贡献进行建模：这对线性模型（灰色到蓝色虚线）带来了显著的改进，将其适用范围扩展到 k < ~0.1 h Mpc−1，而对立方模型（灰色到红色实线）的改进可以忽略不计，该模型在 k < ~1 h Mpc−1 时达到5%水平。总体而言，立方前向模型对互相关数据的拟合效果非常好。

E. 一维功率谱

在图 11 中，我们还比较了一维功率谱——这是 Ly-α 森林分析中的一个关键总结统计量（参见例如 [5 – 12]）。左侧面板显示在未对模拟应用任何平滑的情况下模型的表现，三次前向模型能够很好地恢复真实功率谱的形状，同时显示出约 15% 的残余偏差。相比之下，线性理论模型未能恢复输入模拟的 P1D。

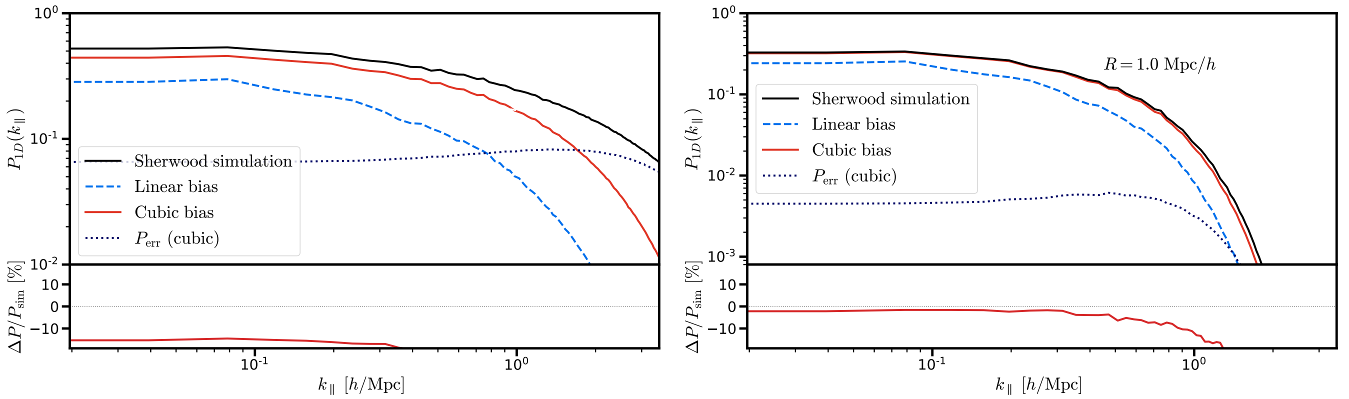

FIG. 11. **Best-Fit 1D Power Spectra Sherwood:** Comparison between the measured one-dimensional power spectrum from the Sherwood simulations with two different smoothing scales using a Gaussian kernel: left no smoothing, right $R = 1\,h^{-1}$ Mpc. The measured P1D is shown in black together with the best-fit forward model obtained from a linear (blue dashed) and a cubic bias (red solid) expansion, respectively. For the cubic bias expansion we show the error power spectrum (blue dotted line). For the baseline without smoothing, we find a 15% offset between the cubic model and the true power spectrum and an even larger off set for the linear model. In both panels, the P1D is shown in bins of Fourier wavenumber $k_\parallel$ in units of $h\,\mathrm{Mpc}^{-1}$. The bottom panel displays the percent difference between the simulation and model power spectra. A gray band highlights the ±5% region in the bottom panel.

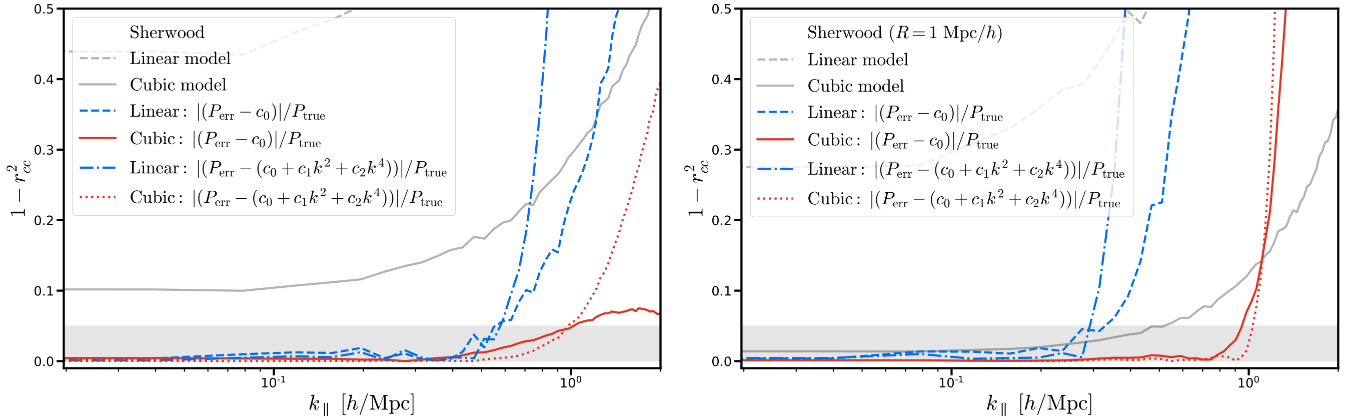

FIG. 12. **Cross-correlation coefficient P1D Sherwood:** Corresponding cross-correlation coefficient $(1 - r_{\rm cc}^2)$ to the one-dimensional power spectra shown in Fig. 11. The colored lines have, first, only the constant low-k limit subtracted from the P1D denoted by c_0 and, second, a higher-order polynomial $c_0 + c_1k_\parallel^2 + c_2k_\parallel^4$. This plot illustrates the importance of using a higher order biasing model compared to purely linear theory as the range of five-percent-level validity of the model is extended up to $k \approx 1\,h\,\mathrm{Mpc}^{-1}$ for the baseline case without smoothing.

even on large scales, to better than 40% (here: outside of the residual plot range). The right panel includes a Gaussian smoothing kernel of $R = 1\,h^{-1}$ Mpc and removes small-scale information. This results in recovery of the input simulation power spectrum to better than 5% (1%) down to $k_\parallel \lesssim 0.6\,(0.3)\,h\,\mathrm{Mpc}^{-1}$ for the cubic model and the linear theory results are off by more than ≈ 20%. Our results suggest that there is more perturbative information in the P1D than discussed in Ref. [170]. In fact, up to $k_\parallel \approx 1\ h\mathrm{Mpc}^{-1}$ the P1D in simulations is dominated by perturbative modes. While we agree with [170] on the significant role of stochastic contributions in P1D, we point out that this noise can be well modeled within EFT.

The error power spectra of the cubic model (and for the linear model on large scales) are constant and do not show any scale dependence. The error power spectrum of the cubic model (dotted blue line in the left panel) of the Sherwood simulation is flat with an amplitude of $\approx 0.07\,[h^{-1}\,\mathrm{Mpc}]^3$. It can be modeled, as discussed in Sec. D. In Fig. 12 we show the corresponding cross-correlation coefficients where we first only remove the constant low-$k_\parallel$ limit going from gray to solid colored lines (blue: linear theory, red: cubic model). The removal of the residual noise using Eq. (38) from the linear (cubic) model extends the range of validity (when investigating $1 - r_{\rm cc}^2$) to $k_\parallel \approx 0.3\,(1.0)\,h\,\mathrm{Mpc}^{-1}$ resulting in the dash-dotted blue lines for the linear theory and dotted red lines for the cubic model, respectively. Note that the smoothing of the field is non-physical and only serves to illustrate the importance of using a higher-order bias

---

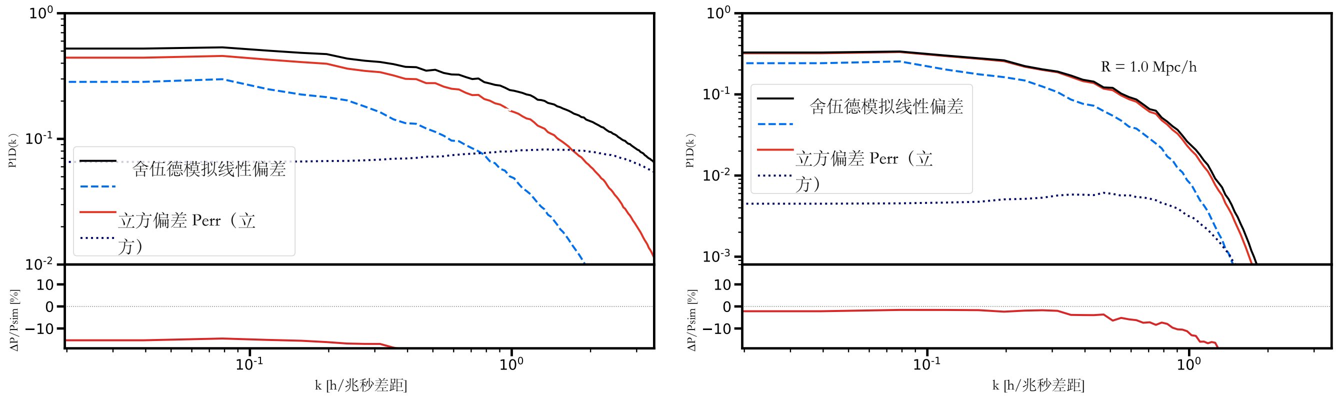

图11. Sherwood 最佳拟合一维功率谱：比较 Sherwood 模拟中测得的一维功率谱，使用两种不同的高斯核平滑尺度：左图无平滑，右图 R = 1 h−1Mpc。测得的 P1D 显示为黑色，同时显示了通过线性（蓝色虚线）和三次偏差（红色实线）展开得到的最佳拟合前向模型。对于三次偏差展开，我们显示了误差功率谱（蓝色虚线）。对于不进行平滑的基线，我们发现三次模型与真实功率谱之间存在 15% 的偏差，而线性模型的偏差更大。在两个面板中，P1D 以傅里叶波数 k // （单位 h Mpc−1）分箱显示。底部面板显示了模拟与模型功率谱之间的百分比差异。灰色带突出显示底部面板中的 ±5% 区域。

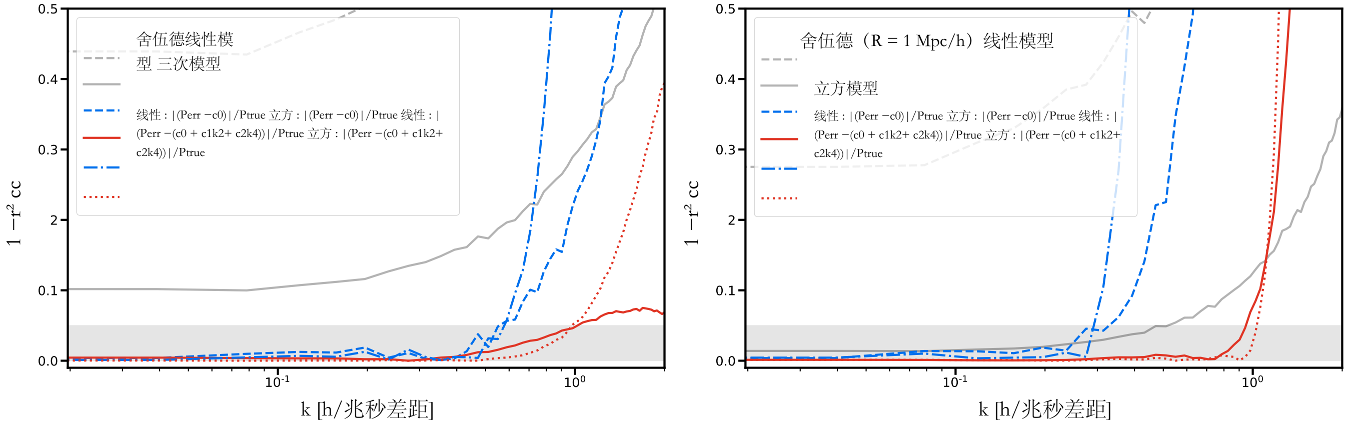

图12. 交叉相关系数 P1D Sherwood：对应于图11所示一维功率谱的交叉相关系数 (1 − r^2_cc)。彩色曲线首先仅从 P1D 中减去常数低 k 极限，用 c0 表示，其次减去高阶多项式 c0 + c1 k^2_ // + c2 k^4_ //。该图展示了相比纯线性理论使用高阶偏差模型的重要性，因为在基准情况下（未进行平滑处理）模型的五个百分点有效范围延伸至 k ≈ 1 h Mpc^{-1}

即使在大尺度上，也能达到优于40%的精度（这里指残差图范围之外）。右侧面板包括一个 R = 1 h−1Mpc 的高斯平滑核，并移除了小尺度信息。这导致在立方模型中，输入模拟功率谱的恢复精度优于5%（1%），直到 k // < ~0.6 (0.3) h Mpc−1，而线性理论结果偏离超过≈20%。我们的结果表明，P1D 中的摄动信息比参考文献 [170] 中讨论的更多。事实上，高达 k // ≈1 hMpc−1 时，模拟中的 P1D 仍主要由摄动模式主导。虽然我们同意 [170] 关于 P1D 中随机贡献重要性的观点，但我们指出，这种噪声可以在有效场论（EFT）框架内很好地建模。

立方模型的误差功率谱（以及大尺度下的线性模型）是常数，并且不显示任何尺度依赖性。Sherwood 模拟中立方模型的误差功率谱（左面板中的蓝色虚线）是平的，幅度约为 ≈0.07 [h−1Mpc]^3。正如第 D 节所讨论的，可以对其进行建模。在图 12 中，我们展示了相应的交叉相关系数，其中我们首先仅去除了从灰色到实线彩色线的常数低 k // 限制（蓝色：线性理论，红色：立方模型）。使用公式 (38) 从线性（立方）模型中去除残余噪声，将有效性范围（在研究 1 − r²_cc 时）扩展到 k // ≈ 0.3（1.0）h Mpc−1，分别得到线性理论的蓝色点划线和立方模型的红色虚线。注意，对场的平滑是非物理的，仅用于说明使用高阶偏置的重要性。

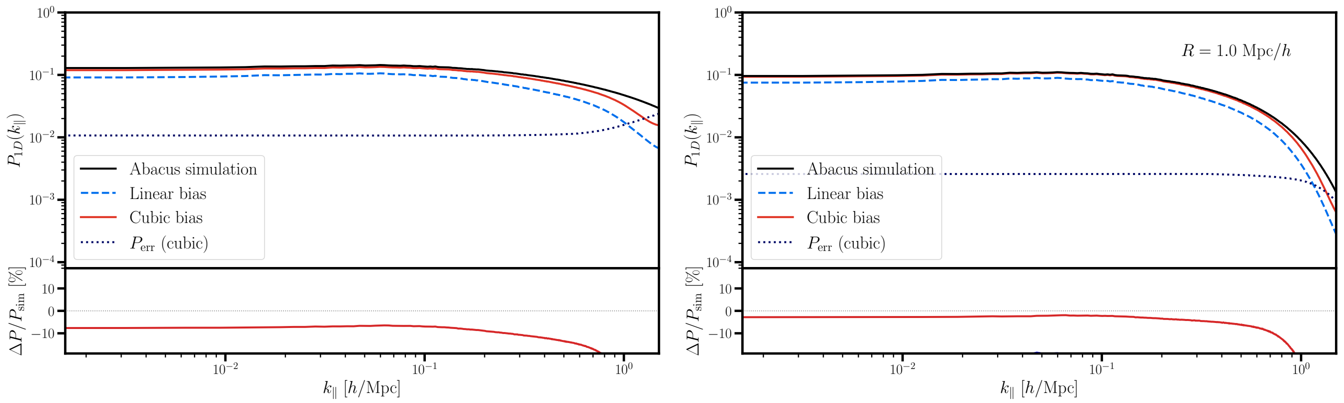

FIG. 13. **Best-Fit 1D Power Spectra Abacus:** Same as Fig. 11 for the Abacus simulation and model I.

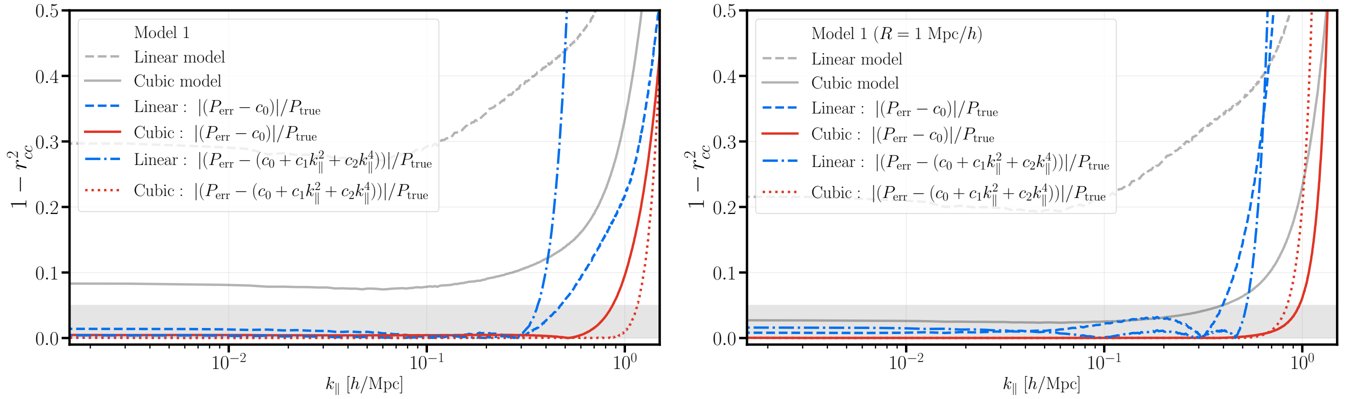

FIG. 14. **Cross-correlation coefficient P1D Abacus:** Cross-correlation coefficient $(1 - r_{\rm cc}^2)$ for the one-dimensional power spectrum of the Abacus simulation and model I where we model the residual noise floor using Eq. (38); same as Fig. 12.

expansion to better capture the small-scale clustering. We perform the same steps on the smoothed field and show the results in the right panel of the same figure. As visible from the cross-correlation coefficients, for the 3D power spectrum this sets in at $k \gtrsim 0.9 - 1.0\,h\,{\rm Mpc}^{-1}$ and for the 1D power spectrum at $k_\parallel \gtrsim 1.0\,h\,{\rm Mpc}^{-1}$, respectively.

In Fig. 13 we show the corresponding results on the P1D for the Abacus simulations and in Fig. 14 we show their cross-correlation coefficients. Qualitatively, the results agree with the ones obtained from Sherwood. Removing the stochastic terms we extend the range of validity of the cubic (linear) model to $k_\parallel \approx 1.0\,(0.4)\,h\,{\rm Mpc}^{-1}$ for the unsmoothed field. Analogously to the P3D results, the linear model has a bigger k-reach on Abacus compared to Sherwood given the larger box size and available number of quasi-linear modes. In Figs. 32-33 we show the model dependence of our forward model in the context of the P1D.

V. TRANSFER FUNCTION FITS

So far we have compared the best-fit perturbative forward model to a set of simulations and found an excellent agreement between the two. However, the best-fit transfer functions that minimize the error power spectrum (or mean-squared error) have been completely free in each k-bin. To build intuition for the transfer functions from Eq. (33) we fit these as polynomial expansions in k and μ, following Ref. [87], given by

$$\beta_i(k,\mu) = c_0 + c_{01}\mu^2 + \left(c_1 + c_{12}\mu^2 + c_{14}\mu^4\right)\cdot k + \left(c_4 k^2 + c_{22}\mu^2 + c_{44}k^2\mu^4\right)\cdot k^2\,, \tag{39}$$

which is different to the one used in Refs. [123, 124] to capture the angular dependence of the transfer functions at low-k. In particular, we allow for higher order corrections, i.e. the two-loop terms, by allowing for odd powers of k. Note that in the $k \to 0$ limit we recover the "Kaiser"-type expression. The fits are performed jointly for all μ-bins using non-linear least-squares minimization, weighting each data point by k to account for the different number of

---

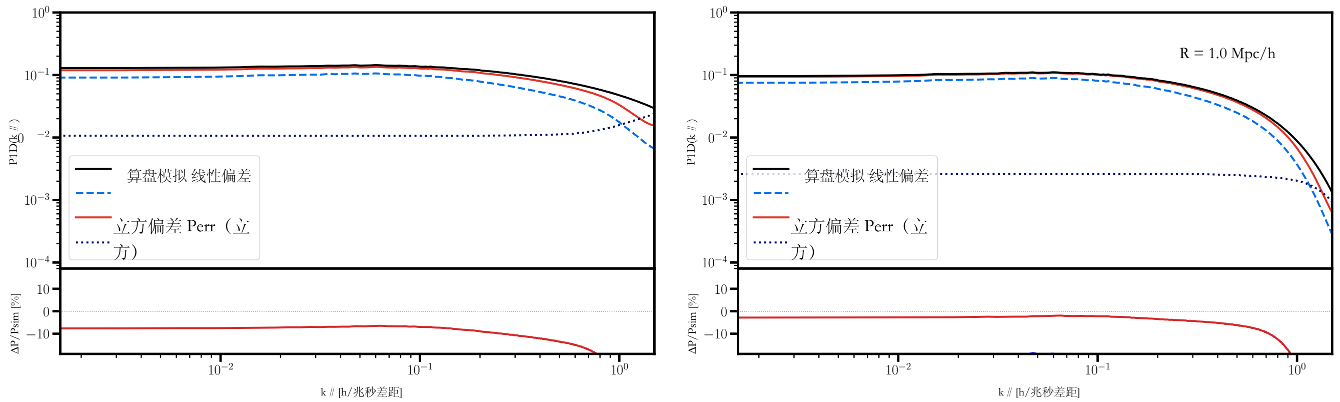

图13. 最佳拟合一维功率谱算盘：与图11相同，用于算盘模拟和模型I。

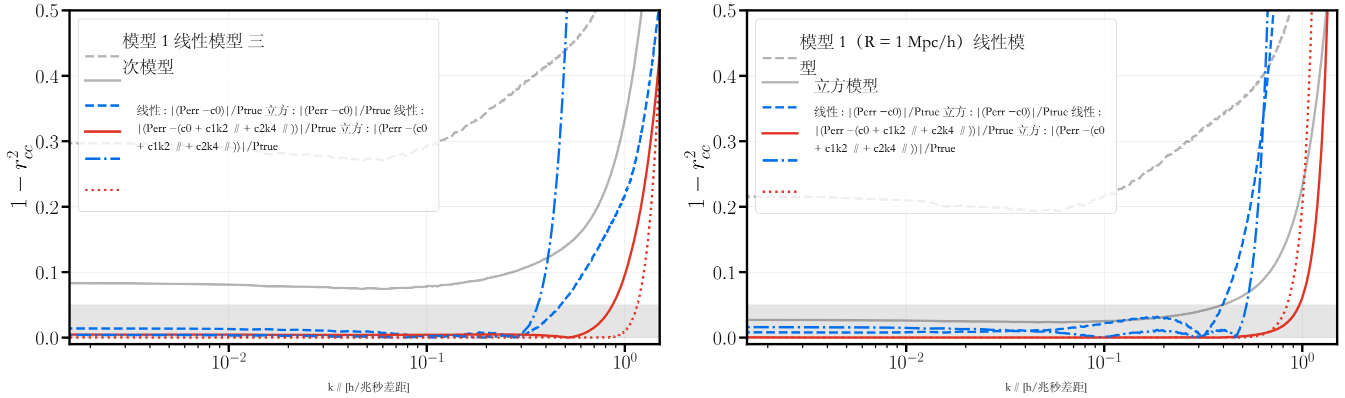

图14. 交叉相关系数 P1D Abacus：Abacus 模拟和模型 I 的一维功率谱的交叉相关系数 (1 − r2_cc)，其中我们使用公式 (38) 对残余噪声底进行建模；与图12相同。

扩展以更好地捕捉小尺度聚类。我们在平滑场上执行相同的步骤，并在同一图的右侧面板中显示结果。从交叉相关系数可以看出，对于三维功率谱，这在 k>~0.9−1.0 h Mpc−1 出现，对于一维功率谱则在 k // >~1.0 h Mpc−1 出现。图 13 显示了 Abacus 模拟中 P1D 的相应结果，图 14 显示了它们的交叉相关系数。从定性上看，这些结果与 Sherwood 得到的结果一致。去掉随机项后，我们将立方（线性）模型的有效范围扩展到未平滑场的 k // ≈1.0（0.4）h Mpc−1。与 P3D 的结果类似，由于 Abacus 的盒子更大且可用的准线性模式更多，线性模型在 Abacus 上的 k 范围比 Sherwood 更大。在图 32-33 中，我们展示了在 P1D 情境下前向模型的模型依赖性。

五　传递函数拟合

到目前为止，我们已经将最佳拟合扰动前向模型与一组模拟进行了比较，发现二者之间有极好的吻合。然而，为了最小化误差功率谱（或均方误差）的最佳拟合传递函数在每个 k-bin 中都是完全自由的。为了建立对公式 (33) 中传递函数的直觉，我们按照参考文献 [87] 的方法，将这些函数拟合为 k 和 μ 的多项式展开，具体为

$$\beta_i(k,\mu) = c_0 + c_{01}\mu^2 + \left(c_1 + c_{12}\mu^2 + c_{14}\mu^4\right)\cdot k + \left(c_4 k^2 + c_{22}\mu^2 + c_{44}k^2\mu^4\right)\cdot k^2\,, \tag{39}$$

这与参考文献[123, 124]中用于捕捉低-k传输函数角度依赖性的方式不同。特别地，我们通过允许k的奇数次幂来考虑高阶修正，即双环项。注意，在k →0极限下，我们会恢复“Kaiser”类型的表达式。拟合是对所有μ-bin联合进行的，使用非线性最小二乘最小化，并通过k对每个数据点加权，以考虑不同的

TF	c_0	c_{01}	c_1	c_{12}	c_{14}	c_4	c_{22}	c_{44}
β_1	−0.225	−0.384	0.146	0.661	−0.074	−0.017	−0.458	0.130
β_2	0.254	0.145	−0.090	0.158	−0.091	0.001	−0.121	0.041
β_{G2}	−0.140	−0.354	−0.006	−0.073	0.062	0.017	0.195	−0.133
β_3	−0.048	0.013	−0.021	−0.155	−0.014	0.001	0.134	−0.007
$\beta_{KK\|}$	−0.129	−0.554	−0.274	−0.263	1.177	0.053	−0.011	−0.473
β_η	−0.378	−0.319	−0.093	2.083	−1.546	−0.068	−0.599	0.297
β_{η^2}	0.217	0.018	−0.373	0.362	0.335	0.079	−0.062	−0.069
$\beta_{\delta\eta}$	−0.129	−0.120	0.193	−0.941	1.008	−0.001	0.326	−0.440

TABLE III. **Coefficients of Best-Fit Sherwood Transfer Functions:** Best-fit parameters for the transfer function (TF) model, $\beta(k,\mu)$, given in Eq. (39) and illustrated in Fig. 15 obtained from the Sherwood simulation. Each wavenumber bin is weighted by k to down weight small scale modes with a cut off at $k_{\max} = 1\,h\,\mathrm{Mpc}^{-1}$. The coefficients c_n, c_{nm} are given in units of $[h^{-1}\mathrm{Mpc}]^n$.

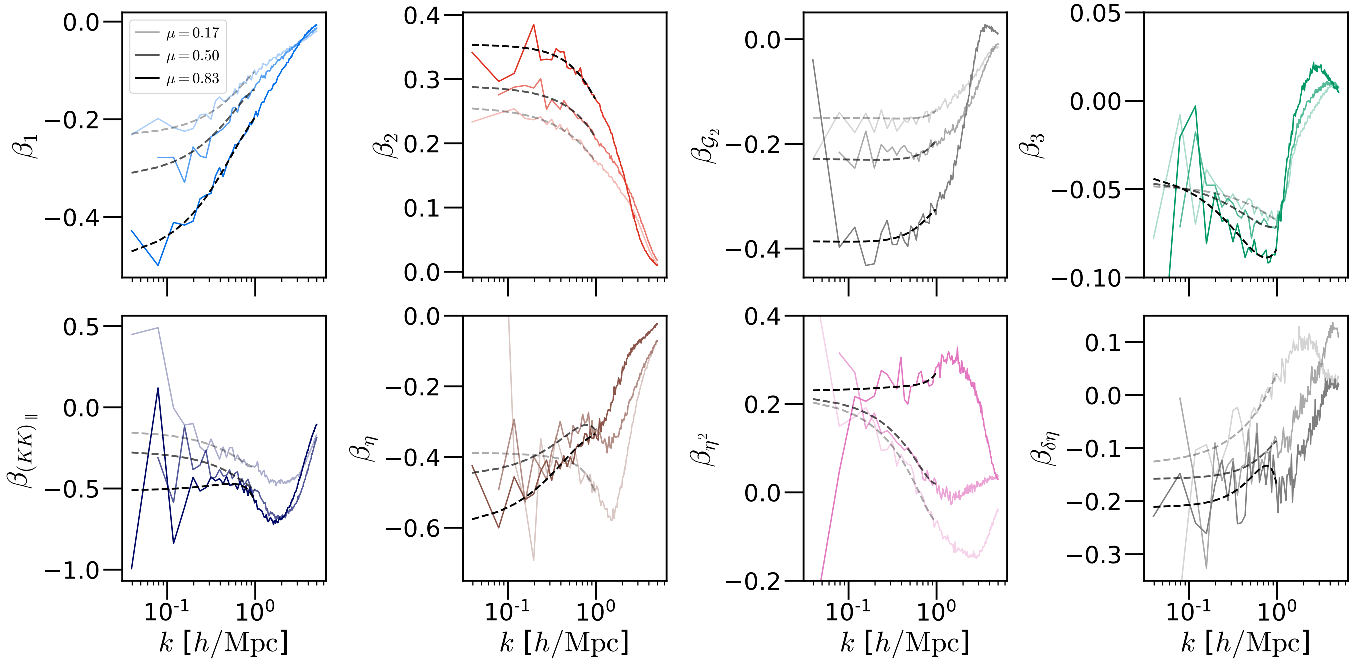

FIG. 15. **Coefficients of Best-Fit Sherwood Transfer Functions:** Best-fit transfer functions $\beta_i(k,\mu)$ for the cubic EFT model obtained from fits to the Sherwood simulations. The corresponding polynomial model for the transfer functions $\beta(k,\mu)$ is given in Eq. (39) which reduces to the Kaiser model in the low-k limit. The large fluctuations in the first two k-bins stem from the small number of available modes. Weighting each bin by its number of k modes down weights very noisy bins. The coefficients are tabulated in Tab. III.

modes with a cut-off at $k = 1\,h\,\mathrm{Mpc}^{-1}$.[7]

Sherwood simulation: The resulting best-fit coefficients obtained for each transfer function are tabulated in Tab. III and compared to the measured transfer functions from Sherwood in Fig. 15. This illustrates that the transfer functions can be approximated by smooth functional forms. We compare the coefficients of the polynomial fits in Tab. III to the bias parameters from fits of the one-loop power spectrum to the Sherwood simulation [66, 85] in the following. Note that the orthogonalization of the transfer functions mixes bias parameters and absorbs higher-order corrections, *i.e.* the present EFT model includes all contributions up to two-loop order (bar the P_{33} term stemming from the cubic field δ^3). From β_1 the parameters c_0 and c_{01} can directly be compared to b_1 and fb_η as well as c_0 from β_2 to b_2. We find very good agreement for b_1 and b_2, yet a larger difference for b_η and $c_{01}(\beta_1)$. We emphasize,

---

[7] We verified that reducing the maximum wavenumber chosen in the fit to $k = 0.6\,h\,\mathrm{Mpc}^{-1}$ does not change our conclusions.

TF	c_0	c_{01}	c_1	c_{12}	c_{14}	c_4	c_{22}	c_{44}
β_1	−0.225	−0.384	0.146	0.661	−0.074	−0.017	−0.458	0.130
β_2	0.254	0.145	0.090					0.041
β_{G2}	−0.140	−0.354	−0.006	−0.073	0.062	0.017	0.195	−0.133
β_3		0.013	−0.021	−0.155	−0.014	0.001	0.134	0.007
βKK //	−0.129	−0.554	−0.274	−0.263		0.053	−0.011	−0.473
η	−0.378	−0.319	−0.093	2.083	−1.546	−0.068	−0.599	
β_{η^2}						0.079	−0.062	−0.069
$\beta_{\delta\eta}$	−0.129	−0.120	0.193	−0.941	1.008	−		

表 III. 最佳拟合 Sherwood 传递函数的系数：根据 Sherwood 模拟获得的传递函数（TF）模型 β(k, μ) 的最佳拟合参数，如公式 (39) 所示，并在图 15 中进行了说明。每个波数区间按 k 加权，以降低小尺度模式的权重，截止值为 kmax = 1 h Mpc−1。系数 cn、cnm 的单位为 [h−1Mpc]n。

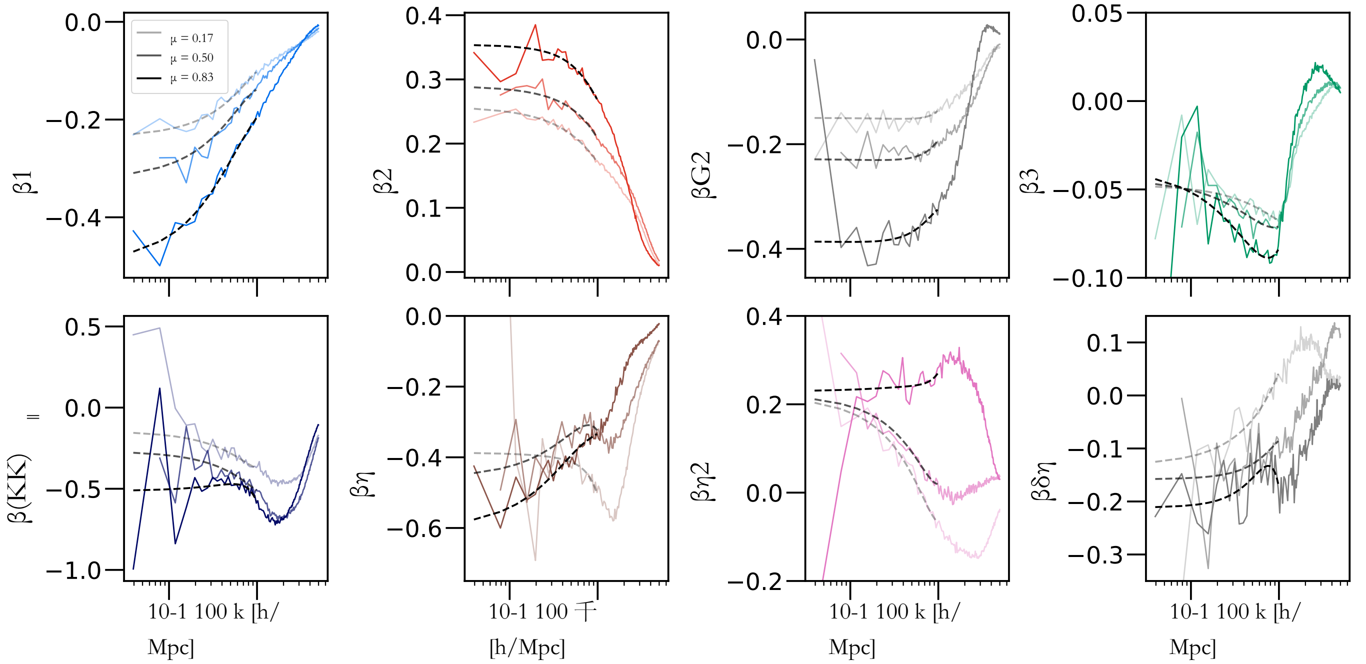

图15. 最佳拟合Sherwood传递函数的系数：从Sherwood模拟拟合中获得的立方EFT模型的最佳拟合传递函数 βi(k, μ)。传递函数 β(k, μ) 的对应多项式模型见公式(39)，在低k极限下退化为Kaiser模型。前两个k区间的大波动源于可用模态数量较少。通过按每个区间的k模态数量加权，可以降低噪声较大的区间的权重。系数列于表III。

在 k = 1 h Mpc−1 处有截止的模式

Sherwood 模拟：每个传递函数得到的最佳拟合系数列在表 III 中，并在图 15 中与 Sherwood 的测量传递函数进行比较。这表明传递函数可以用平滑的函数形式来近似。我们将表 III 中多项式拟合的系数与单回路功率谱拟合到 Sherwood 模拟的偏置参数 [66, 85] 进行比较。请注意，传递函数的正交化会混合偏置参数并吸收高阶修正，即目前的 EFT 模型包括所有到两回路阶的贡献（除了来自三次场 δ3 的 P33 项）。从 β1 可以直接将参数 c0 和 c01 与 b1 和 fbη 进行比较，以及从 β2 的 c0 与 b2 进行比较。我们发现 b1 和 b2 的一致性非常好，但 bη 和 c01(β1) 的差异较大。我们强调，

---

7 我们验证了将拟合中选择的最大波数降低到 k = 0.6 h Mpc−1 不会改变我们的结论。

TF	c_0	c_{01}	c_1	c_{12}	c_{14}	c_4	c_{22}	c_{44}
β_1	−0.152	−0.156	0.082	0.283	−0.059	−0.009	−0.176	0.062
β_2	0.108	0.067	−0.017	0.086	−0.089	−0.021	−0.066	−0.052
$\beta_{\mathcal{G}_2}$	−0.087	−0.167	0.021	0.044	0.086	0.020	0.005	0.110
β_3	−0.003	−0.012	−0.020	−0.117	0.095	0.017	0.097	0.012
$\beta_{KK_\parallel}$	−0.026	−0.154	−0.207	0.640	−0.755	−0.060	−0.997	0.922
β_η	−0.213	0.064	0.124	−0.275	0.121	−0.011	−0.036	0.187
β_{η^2}	0.073	−0.043	−0.111	0.104	0.092	0.018	−0.009	0.058
$\beta_{\delta\eta}$	−0.070	0.007	−0.016	−0.318	0.017	−0.011	0.581	−0.347

TABLE IV. **Coefficients of Best-Fit Abacus Transfer Functions:** Coefficients of the fitted transfer functions (TF) for Abacus simulation "one", line-of-sight z and Model I for the polynomial given in Eq. (39) and plotted in Fig. 16. The coefficients c_n, c_{nm} are given in units of $[h^{-1}\mathrm{Mpc}]^n$.

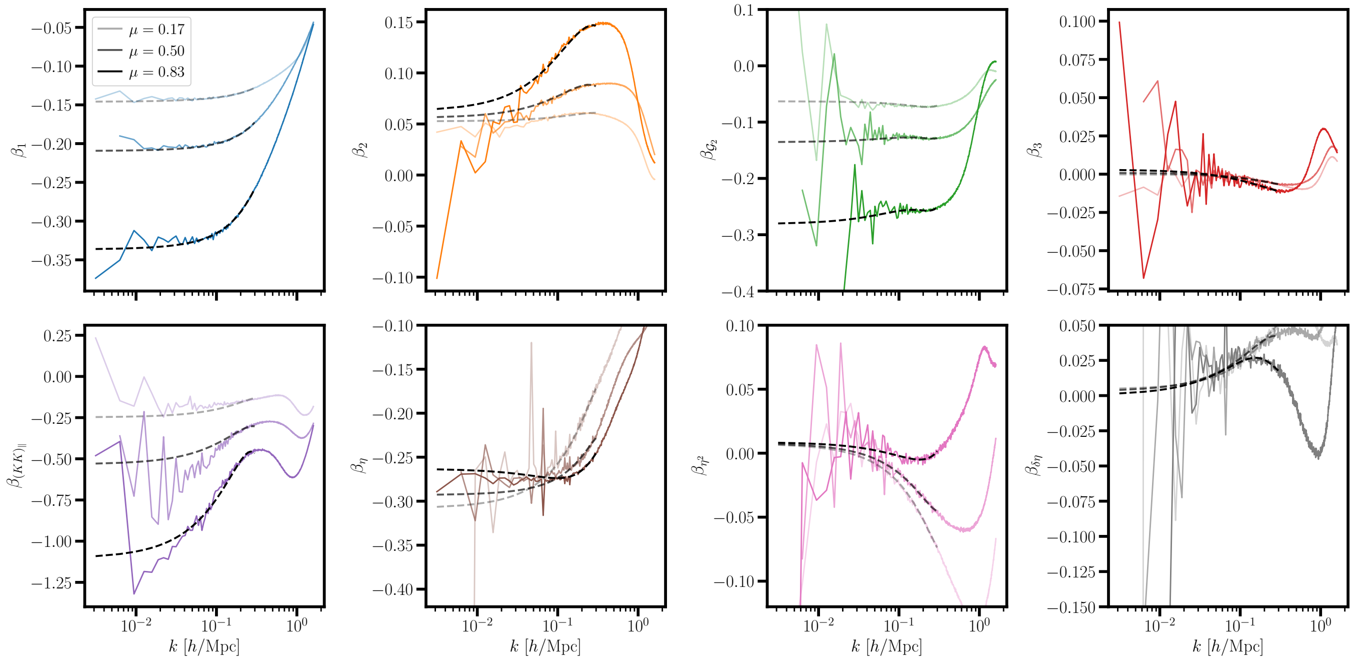

FIG. 16. **Best-Fit Abacus Transfer Functions:** Same as Fig. 15 but for Abacus model I. The coefficients are tabulated in Tab. IV.

however, that fits of the one-loop power spectrum in the small hydrodynamic simulations can be strongly affected by cosmic variance and since the present fits do not have an associated covariance matrix, care should be taken when comparing the two. Further, the equivalence principle imposes that $c_0(\beta_\eta)$ matches $c_{01}(\beta_1)$ which is an independent confirmation of our theoretical model as we fit for the transfer functions completely independently of each other.

Abacus simulation: Analogously to the analysis on Sherwood simulations, we fit the low-k limit of the transfer functions obtained from the field-level technique applied to the Abacus simulations (see Sec. A). In Tab. IV we tabulate the coefficients of the transfer functions. Here we only tabulate the values for simulation "one", Model I and line-of-sight z which can be directly compared to the one-loop power spectrum fits presented in table 4 in Ref. [78]. In particular, the coefficients c_0 and a_{01} for β_1 are directly related to b_1 and $-fb_\eta$ and c_0 of β_2 to b_2. For all three we find reasonable agreement. In contrast to Sherwood for which we find a strong discrepancy between c_{01} of β_1 indicating that the limited volume and the cosmic variance strongly affect constraints on the bias parameters of the one-loop power spectrum.

In Fig. 16 we show the fitted transfer functions and the resulting polynomial fits. Given the large box size, the transfer functions are computed out to large scales. Whilst the results agree qualitatively with Sherwood we note that the β_2 transfer functions exhibit a scale dependence already on large scales. The key takeaway from this section is that (i) we can measure each transfer function from the simulated data; and (ii) these can be approximated through

---

TF	c_0	c_{01}	c_1	c_{12}	c_{14}	c_4	c_{22}	c_{44}
β_1	−0.152	−0.156	0.082	0.283	−0.059	−0.009	−0.176	0.062
β_2	0.108	0.067	−0.017	0.086	−0.089	−0.021	−0.066	−0.052
$\beta_{\mathcal{G}_2}$	0.087	0.167	0.021					
β_3	−0.003	−0.012	−0.020	−0.117	0.095	0.017	0.097	0.012
βKK //	−0.026	−0.154	−0.207	0.640	−0.755	−0.060	−0.997	0.922
η				−	0.121	−0.011	−0.036	0.187
β_{η^2}	0.073	−0.043	−0.111	0.104				0.058
$\beta_{\delta\eta}$	−	0.007	−0.016	−0.318	0.017	−0.011	0.581	−0.347

表 IV. 最佳拟合算筹传递函数的系数：Abacus 模拟“one”的传递函数 (TF) 拟合系数，视线 z 和模型 I，对应于方程 (39) 给出的多项式，并绘制在图 16 中。系数 cn, cnm 的单位为 [h−1Mpc]ⁿ。

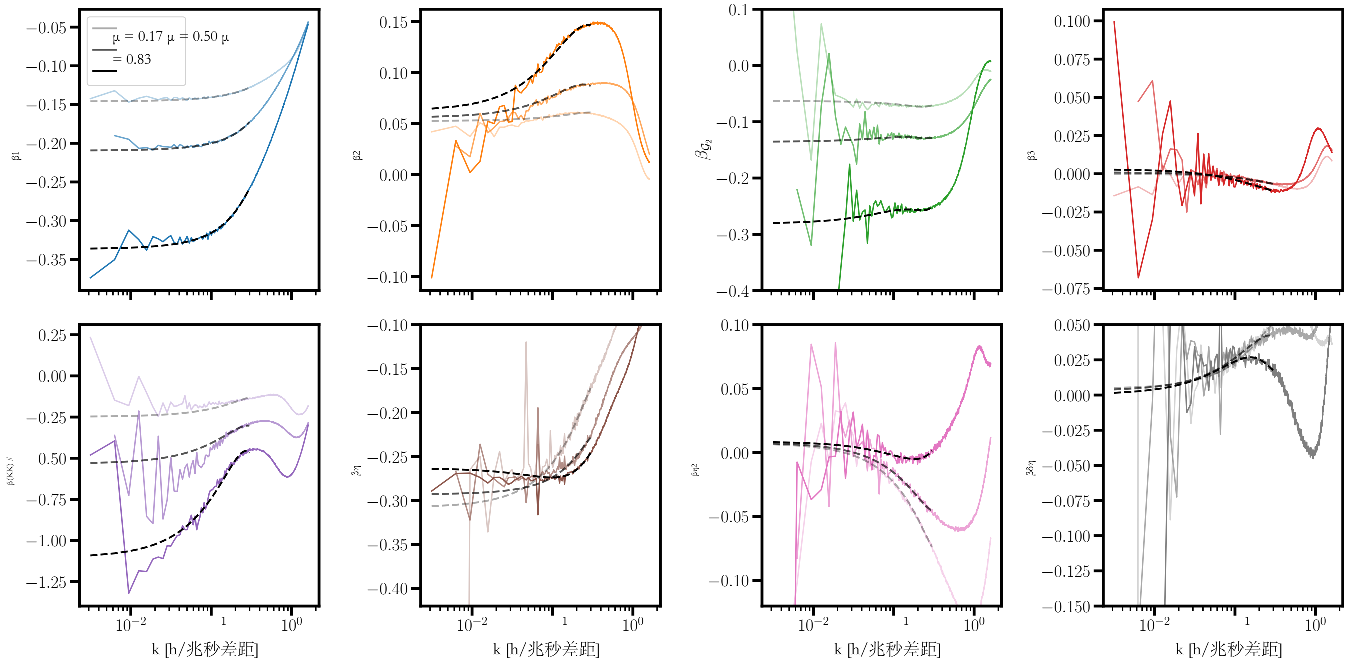

图16. 最佳拟合算盘传递函数：与图15相同，但适用于算盘模型I。系数列在表IV中。

然而，在小型流体动力学模拟中，对单环功率谱的拟合可能会受到宇宙方差的强烈影响，并且由于当前的拟合没有相关的协方差矩阵，因此在比较两者时应当谨慎。此外，等效原理规定 c0(βη) 必须与 c01(β1) 匹配，这也独立地验证了我们的理论模型，因为我们对传递函数的拟合是完全独立进行的。

Abacus 模拟：类似于对 Sherwood 模拟的分析，我们拟合了从应用于 Abacus 模拟的场级技术获得的传递函数的低-k 极限（见附录 A）。在表 IV 中，我们列出了传递函数的系数。这里我们只列出了模拟“one”、模型 I 和视线方向 z 的数值，这些数值可以直接与参考文献 [78] 表 4 中的一环功率谱拟合结果进行比较。尤其是，β1 的系数 c0 和 a01 与 b1 和 −fbη 直接相关，而 β2 的 c0 与 b2 相关。对于这三者，我们发现合理的一致性。与 Sherwood 不同，我们在 β1 的 c01 上发现了明显的差异，这表明有限的体积和宇宙方差对一环功率谱偏置参数的约束有强烈影响。

在图16中，我们展示了拟合的传递函数及其得到的多项式拟合结果。鉴于较大的盒子尺寸，传递函数被计算到了大尺度。虽然结果在定性上与Sherwood一致，但我们注意到β2的传递函数在大尺度上已经表现出尺度依赖性。本节的关键结论是：(i) 我们可以从模拟数据中测量每个传递函数；(ii) 这些可以通过……进行近似。

simple polynomials.

A. High-redshift galaxy transfer functions

In this work, we use a set of transfer functions matching the clustering properties of observations at $z = 3$ obtained from calibrations on state-of-the-art Astrid hydrodynamical simulations (see, e.g., [153, 154]), representing Ly-α emitters (LAEs) and Lyman-break galaxies (LBGs). Here, we follow Ref. [154] to which the reader is referred for a fuller presentation of the employed high-redshift galaxy samples. These are designed to approximately match the linear bias, b_1, and number density for LBGs ("CARS"; [157]) and LAEs ("ODIN" [155]) and a futuristic sample based on projections for Stage-V spectroscopy (denoted by S5; [131]), summarized as LBG and LAE S5 in table II of Ref. [154].

We fit the redshift space transfer functions for both galaxy samples as polynomial expansions in k and μ, using a functional form inspired by the one used in Ref. [124]

$$\beta_1(k,\mu) = c_0 + c_1 k + c_2 k^2 + c_4 k^4 + c_{22}(k\mu)^2 + c_{44}(k\mu)^4\,, \tag{40}$$

$$\beta_{i\neq 1}(k,\mu) = c_0 + c_2 k^2 + c_4 k^4 + c_{22}(k\mu)^2 + c_{44}(k\mu)^4\,, \tag{41}$$

$$P_{\rm err}(k,\mu) = a_0 + a_2 k^2\,. \tag{42}$$

The fits are performed jointly for all μ-bins using nonlinear least-squares minimization, weighting each data point by k with a cut off at $k = 0.4\,h\,\mathrm{Mpc}^{-1}$.[8] The resulting best-fit transfer functions are shown in Fig. 17.

VI. TRANSFER FUNCTIONS IN PERTURBATION THEORY

In this section we will develop the perturbation theory modeling for the field-level transfer functions and use them to measure EFT parameters at the field-level, connecting the fitted transfer functions to a theory prediction. In particular, addressing the question whether the large freedom in k and μ when minimizing the mean-squared-error is justified. The key idea of this method is that in the absence of orthogonalization the transfer functions should be constant on large-scales – by construction. These constants then can be matched to the EFT parameters. However, this constancy of transfer functions in the $k \to 0$ limit is violated by the Gram-Schmidt orthogonalization, which on large scales removes the operators which can be represented as other operators times a μ-dependent transfer function. Specifically, it removes the contributions of the $\eta_{\rm new}$, given in Eq. (23), operator and the $\Pi^{[2]}_{\parallel}$ operators at the linear and quadratic orders, respectively. As a result of the Gram-Schmidt procedure, we end up with μ-dependent, but k-independent transfer functions on large scales. We will deal with this by forward modeling the orthogonalization procedure.

The first step of our fitting procedure is to specify the EFT perturbation theory basis appropriate for the transfer function computations. It is natural to use the EFT formulation used to compute the one-loop predictions with the `CLASS-PT` code [85, 171]. However, it turns out that it is convenient to use an equivalent, but slightly different EFT model for the purpose of the matching, which we discuss here.

To start off, we introduce the field level EFT at the linear level, i.e. we ignore all quadratic operators and the orthogonalization procedure. Therefore, we take the present EFT model and assume all transfer functions to be constants, yielding

$$\delta_F = \beta_1 \tilde{\delta}_1 + \beta_\eta \delta_Z = (\beta_1 + \beta_\eta + \beta_\eta f \mu^2)\delta_1(\boldsymbol{k}) + \mathcal{O}(\delta_1^2)\,. \tag{43}$$

In the $k \to 0$ limit, where the linear theory prediction is dominant, our model simply reproduces the Kaiser formula provided that the transfer functions β_1 and β_η take constant values. These values then can be converted into the usual Ly-α EFT linear bias parameters from `CLASS-PT` via

$$b_1 = \beta_1 + \beta_\eta\,, \quad b_\eta = -\beta_\eta\,. \tag{44}$$

What we actually measure though are orthogonalized transfer functions, i.e. the model we use to fit the simulation snapshots takes the following form in the linear approximation

$$\delta_F^{\rm model} = \beta_1^F(k,\mu)\tilde{\delta}_1 + \beta_\eta(k,\mu)\delta_Z^{\perp} + \mathcal{O}(\delta_1^2)\,. \tag{45}$$

---

[8] We verified that varying the maximum used wavenumber by $\Delta k = 0.1\,h\,\mathrm{Mpc}^{-1}$ does not affect our conclusions.

简单多项式。

A. 高红移星系传递函数

在这项工作中，我们使用了一组传递函数，这些函数匹配了从最先进的Astrid流体动力学模拟校准中获得的 $z = 3$ 观测的聚类特性（参见，例如，[153, 154]），代表了Ly-α发射星系（LAEs）和Lyman断裂星系（LBGs）。在这里，我们遵循参考文献[154]，读者可查阅该文献以获取所使用的高红移星系样本的完整介绍。设计这些样本是为了大致匹配 LBGs（“CARS”；[157]）和LAEs（“ODIN” [155]）的线性偏差 b1 和数密度，以及基于第五阶段光谱学预测的未来样本（记为 S5；[131]），在参考文献[154]的表 II 中总结为 LBG 和 LAE S5。

我们将两个星系样本的红移空间传递函数拟合为 k 和 μ 的多项式展开，使用一个受参考文献 [124] 中使用的形

$$\beta_1(k,\mu) = c_0 + c_1 k + c_2 k^2 + c_4 k^4 + c_{22}(k\mu)^2 + c_{44}(k\mu)^4\,, \tag{40}$$

$$\beta_{i\neq 1}(k,\mu) = c_0 + c_2 k^2 + c_4 k^4 + c_{22}(k\mu)^2 + c_{44}(k\mu)^4\,, \tag{41}$$

$$P_{\rm err}(k,\mu) = a_0 + a_2 k^2\,. \tag{42}$$

这些拟合是对所有 μ-区间联合进行的，使用非线性最小二乘法最小化，每个数据点按 k 加权，并在 k = 0.4 h Mpc−1 处设定截止。得到的最佳拟合传递函数如图 17 所示。

六 微扰理论中的传递函数

在本节中，我们将发展用于场级传递函数的微扰理论建模，并利用它们在场级测量EFT参数，将拟合的传递函数与理论预测联系起来。特别是，讨论在最小化均方误差时 k 和 μ 的大自由度是否合理。该方法的关键思想是，在没有正交化的情况下，传递函数在大尺度上应该是恒定的——这是通过构造实现的。这些常数随后可以与EFT参数匹配。然而，在 k →0 极限下传递函数的这种恒定性会被Gram-Schmidt正交化破坏，因为在大尺度上，它会移除那些可以表示为其他算子乘以 μ 依赖传递函数的算子。具体来说，它会移除 η_new（如公式(23)所示）算子的贡献，以及在线性和二次阶分别的 Π[2] ∥ 算子。由于Gram-Schmidt过程的结果，我们最终在大尺度上得到依赖于μ但与k无关的传递函数。我们将通过正向建模正交化过程来处理这一点。

我们拟合程序的第一步是指定适用于传递函数计算的有效场理论（EFT）微扰理论基础。自然的选择是使用用于通过 CLASS-PT 代码 [85, 171] 计算一环预测的 EFT 公式。然而，事实证明，为了匹配的目的，使用一个等效但稍有不同的 EFT 模型更为方便，我们将在这里讨论该模型。

首先，我们在线性水平上介绍场论有效场论（EFT），即我们忽略所有二次算符和正交化过程。因此，我们采用当前的EFT模型，并假设所有传递函数为常数，从而得到

$$\delta_F = \beta_1 \tilde{\delta}_1 + \beta_\eta \delta_Z = (\beta_1 + \beta_\eta + \beta_\eta f \mu^2)\delta_1(\boldsymbol{k}) + \mathcal{O}(\delta_1^2)\,. \tag{43}$$

在 k →0 的极限下，当线性理论预测占主导时，我们的模型只要传递函数 β1 和 βη 取常数值，就可以简单地重现 Kaiser 公式。然后，这些值可以通过 CLASS-PT 转换为通常的 Ly-α EFT 线性偏差参数。

$$b_1 = \beta_1 + \beta_\eta\,, \quad b_\eta = -\beta_\eta\,. \tag{44}$$

不过，我们实际测量的是正交化的传递函数，即我们用于拟合模拟快照的模型在线性近似下采用以下形式

$$\delta_F^{\rm model} = \beta_1^F(k,\mu)\tilde{\delta}_1 + \beta_\eta(k,\mu)\delta_Z^{\perp} + \mathcal{O}(\delta_1^2)\,. \tag{45}$$

---

[8] We verified that varying the maximum used wavenumber by $\Delta k = 0.1\,h\,\mathrm{Mpc}^{-1}$ does not affect our conclusions.

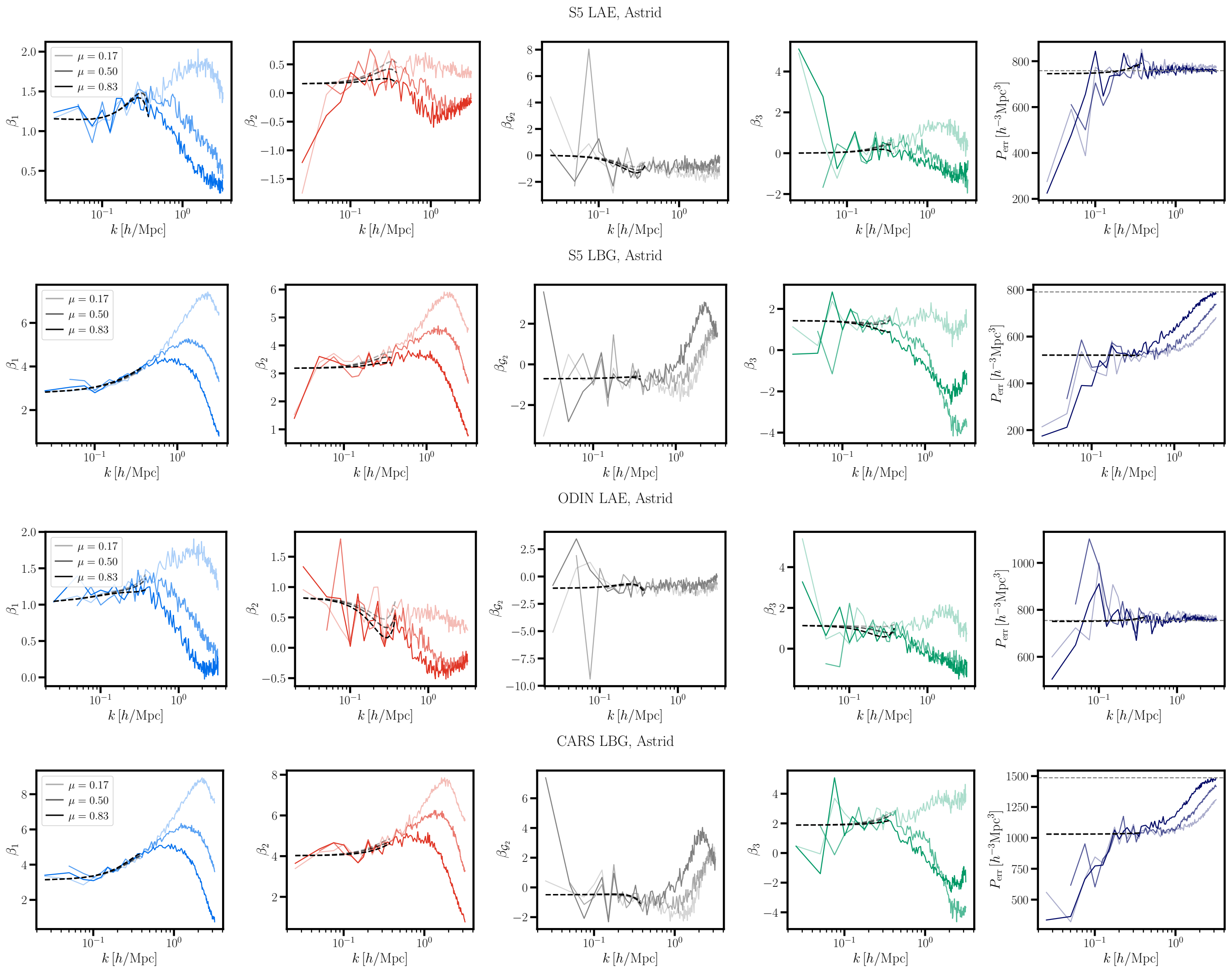

FIG. 17. Best-fit transfer functions $\beta_i(k,\mu)$ for the cubic EFT model obtained from fits to the Astrid simulation tuned to an "S5" and "ODIN" sample in the two top rows and both bottom rows show the same for an LBG-type sample calibrated on Astrid simulations. The first four columns show the redshift-space transfer functions for galaxies β_1, β_2, $\beta_{\mathcal{G}_2}$, β_3, and the last column shows the error power spectrum $P_{\rm err} \equiv \langle|\delta_F^{\rm truth}(\boldsymbol{k}) - \delta_F^{\rm model}(\boldsymbol{k})|^2\rangle$ a quantitative measure of the model performance. Note that for the error spectrum we use the constant large-scale theoretical limit as input for the field-level mocks. The intensity of the lines denotes the angular wedges μ given in three bins. The dashed black curves are fits to the measured transfer functions from Ref. [153] using the polynomial model given in Eqs. (40)-(42). The horizontal dashed gray lines in the last column indicates the effective shot noise of the sample $(1/\overline{n})$.

Since the Zel'dovich density field is 100% correlated with the shifted field $\tilde{\delta}_1$ in linear theory, we have

$$\delta_Z^{\perp} = \delta_Z - \frac{\langle\delta_Z\tilde{\delta}_1\rangle'}{\langle|\tilde{\delta}_1|^2\rangle'}\tilde{\delta}_1 = (1+f\mu^2)\delta_1 - (1+f\mu^2)\delta_1 + \mathcal{O}(\delta_1^2) = 0 + \mathcal{O}(\delta_1^2)\,, \tag{46}$$

i.e. at the linear order we simply have

$$\delta_F = \beta_1^F(k,\mu)\tilde{\delta}_1 + \mathcal{O}(\delta_1^2)\,, \tag{47}$$

so that both coefficients of our constant transfer function model can be matched from the μ-dependence of β_1. Using the definition of β_1^F and our new EFT model we get

$$\beta_1^F(k,\mu) = \frac{\langle\delta_F\tilde{\delta}_1\rangle'}{\langle|\tilde{\delta}_1|^2\rangle'}\bigg|_{k\to 0} = \beta_1 + \beta_\eta + \beta_\eta f\mu^2\,. \tag{48}$$

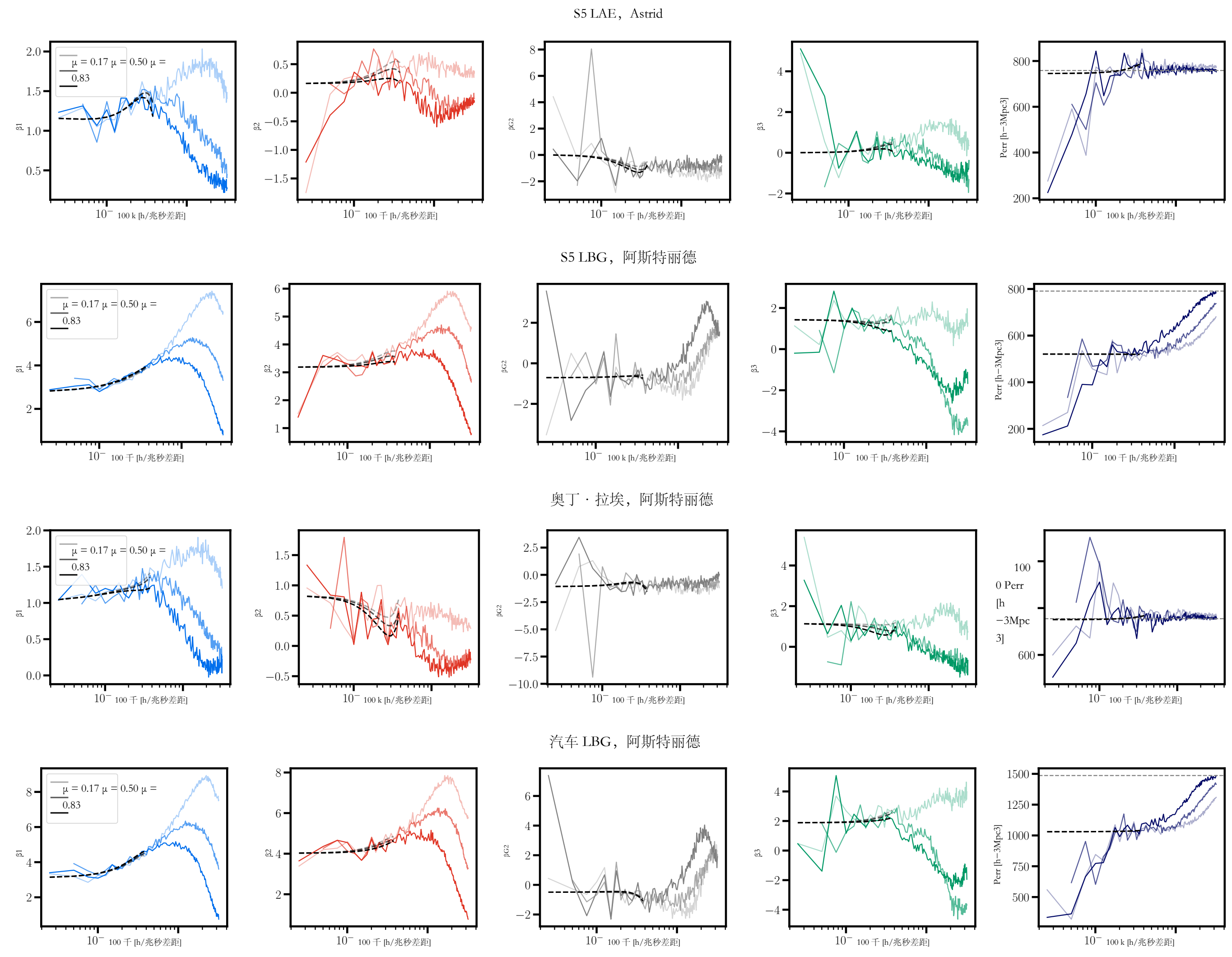

图17. 从拟合Astrid模拟获得的三次EFT模型的最佳拟合传递函数βi(k, μ)，前两行分别针对“S5”和“ODIN”样本进行调优，下两行显示针对LBG型样本（在Astrid模拟上校准）的相同结果。前四列展示了星系在红移空间的传递函数β1、β2、βG2、β3，最后一列显示误差功率谱Perr ≡⟨|δtruth F (k)−δmodel F (k)|2⟩，这是衡量模型性能的一个定量指标。注意，对于误差谱，我们使用常数大尺度理论极限作为场级模拟的输入。线条的强度表示角楔μ，分为三类。黑色虚线曲线是使用公式(40)-(42)中的多项式模型对参考文献[153]中测得的传递函数进行拟合的结果。最后一列的灰色虚线表示样本的有效测光噪声(1/n)。

由于泽尔多维奇密度场在线性理论中与平移场 ˜δ1 完全相关，我们有

δ⊥ Z = δ $\frac{\langle\delta_Z\delta_1\rangle'}{\langle|\tilde{\delta}_1|^2\rangle'}\tilde{\delta}_1 = (1+f\mu^2)\delta_1 - (1+f\mu^2)\delta_1 + \mathcal{O}(\delta_1^2) = 0 + \mathcal{O}(\delta_1^2)\,,$ (46)

O(δ1²) = 0

即，在一阶近似下我们仅有 have

$$\delta_F = \beta_1^F(k,\mu)\tilde{\delta}_1 + \mathcal{O}(\delta_1^2)\,, \tag{47}$$

这样我们就可以通过β1的μ依赖性匹配我们恒定传递函数模型的两个系数。使用βF 1的定义和我们新的有效场论模型，我们得到

βF 1(k, μ) = ⟨δF ˜ ′ δ1⟩′ ⟨|˜ δ1|2⟩′ ′ $\bigg|_{k\to 0} = \beta_1 + \beta_\eta + \beta_\eta f\mu^2\,.$ (48)

Once we have β_1, β_η measurements from fits to the $\beta_1^F(k,\mu)$ data, we can convert them into the usual EFT parameters using Eq. (44). This matching exercise shows that it is easier to match the transfer functions to our new EFT model which is directly equivalent to the forward model up to the Gram-Schmidt process. These coefficents then can be converted into the usual bias parameters. Let us now take a look at this constant EFT transfer function model at the quadratic order. The corresponding Ly-α kernel then reads

$$\begin{aligned}
K_2(\boldsymbol{k}_1,\boldsymbol{k}_2) =& \beta_1 \tilde{K}_2 + \beta_\eta \left(F_2^{\rm ZA} - \frac{3}{7} f\mu^2 F_{\mathcal{G}_2}\right) + \beta_2 2F_{\delta^2} + \beta_3 6F_{\delta^3} \\
&+ \beta_{\mathcal{G}_2} F_{\mathcal{G}_2} + \beta_{\delta\eta} F_{\delta\eta} + \beta_{\eta^2} F_{\eta^2} + \beta_{KK_\parallel} F_{(KK)_\parallel} + \beta_{\Pi^{(2)}_\parallel} F_{\Pi^{(2)}_\parallel} \\
\equiv& (\beta_1+\beta_\eta) F_\delta - \beta_1 F_{\delta\eta} + \beta_\eta\left(-F_\eta + F_{\eta^2} - F_{\delta\eta}\right) + 2\beta_2 F_{\delta^2} + \left(-\frac{2}{7}\beta_1 + \frac{3}{14}\beta_\eta + \beta_{\mathcal{G}_2}\right) F_{\mathcal{G}_2} \\
&+ \beta_{\delta\eta} F_{\delta\eta} + \beta_{\eta^2} F_{\eta^2} + \beta_{\Pi^{[2]}_\parallel} \Pi^{[2]}_\parallel + \beta_{(KK)_\parallel} F_{(KK)_\parallel} \\
=& (\beta_1+\beta_\eta) F_\delta - \beta_\eta F_\eta + 2\beta_2 F_{\delta^2} + \left(-\frac{2}{7}\beta_1 + \frac{3}{14}\beta_\eta + \beta_{\mathcal{G}_2}\right) F_{\mathcal{G}_2} \\
&+ (-\beta_1 - \beta_\eta + \beta_{\delta\eta}) F_{\delta\eta} + (\beta_{\eta^2} + \beta_\eta) F_{\eta^2} + \beta_{\Pi^{[2]}_\parallel} F_{\Pi^{[2]}_\parallel} + \beta_{(KK)_\parallel} F_{(KK)_\parallel}
\end{aligned} \tag{49}$$

Note that this model explicitly has the $\Pi^{[2]}_\parallel$ operator. The above can be compared with the usual EFT kernel in Eq. (18). Matching the coefficients we get:

$$\begin{aligned}
&b_1 = \beta_1 + \beta_\eta\,, \quad b_\eta = -\beta_\eta\,, \quad b_2 = 2\beta_2\,, \quad b_3 = 6\beta_3\,, \quad b_{\mathcal{G}_2} = -\frac{2}{7}\beta_1 + \frac{3}{14}\beta_\eta + \beta_{\mathcal{G}_2}\,, \\
&b_{\delta\eta} = -\beta_1 - \beta_\eta + \beta_{\delta\eta}\,, \quad b_{\eta^2} = \beta_{\eta^2} + \beta_\eta\,, \quad b_{\Pi^{[2]}_\parallel} = \beta_{\Pi^{[2]}_\parallel}\,, \quad b_{(KK)_\parallel} = \beta_{(KK)_\parallel}\;.
\end{aligned} \tag{50}$$

The above matching allows one to determine linear and quadratic EFT parameters from the large-scale limits of the EFT transfer functions. Importantly, the above matching is consistent with the linear theory matching. Let us also note that while the δ^3 operator is present in the forward model, its correlator with δ_1 is redundant at the one-loop order, while its correlations with the quadratic operators start only at the two-loop order which is beyond the scope of our work. Therefore, we will ignore this operator in our EFT transfer function model.

The correlators between relevant operators $\langle\mathcal{O}_A\mathcal{O}_B\rangle$ can be readily extracted from `CLASS-PT` [171], which computes all possible cross-correlations above as part of the P_{22} computation routine. In practice, however, this problem is complicated by four effects. First, as discussed before, we measure the transfer functions of the orthogonal operators, and thus we need to account for the Gram-Schmidt procedure in our forward model. Second, there are non-linear corrections that introduce scale-dependence of the transfer functions, making it hard to extract their constant parts from small simulation boxes. Third, large-scales are affected by residual cosmic variance, which is non-negligible for quadratic operators. And fourth, related to the previous points, we need to know the covariance between the transfer function data. In the following, we will address each point.

First, we forward model the Gram-Schmidt process. Using the non-orthogonalized bias expansion $\delta_F = \sum_A \beta_A \mathcal{O}_A$, the measured transfer function of an operator $\mathcal{O}_a$ follows the model

$$\beta_a^F(k,\mu) = \frac{\langle\delta_F \mathcal{O}_a^\perp\rangle'}{\langle|\mathcal{O}_a^\perp|^2\rangle'} = \frac{\langle\delta_F M_{aB}\mathcal{O}_B\rangle'}{\langle|\mathcal{O}_a^\perp|^2\rangle'} = \frac{1}{\langle|\mathcal{O}_a^\perp|^2\rangle'} M_{aB}\langle\delta_F\mathcal{O}_B\rangle' = \frac{1}{\langle|\mathcal{O}_a^\perp|^2\rangle'}\sum_A M_{aB}\beta_A\langle\mathcal{O}_A\mathcal{O}_B\rangle'\,, \tag{51}$$

where indices A, B run over all the eight operators including $\Pi^{[2]}_\parallel$, in contrast to the a index which runs only over seven operators. M_{aB} is the $\boldsymbol{k}$-dependent rotation matrix built from the cross spectra of the original, non-orthogonal operators. If we had access to only one bin, the inversion of the above equation would not be possible. However, using data from multiple bins, we can obtain the solution to the above problem as a least squares problem.

Second, to account for scale-dependent non-linearities, we introduce additional polynomials in our transfer function model with free fitting parameters for every μ-bin,

$$\beta_a^F(k,\mu)\Big|_{\rm model} = \frac{1}{\langle|\mathcal{O}_a^\perp|^2\rangle'}\sum_A M_{aB}\beta_A\langle\mathcal{O}_A(\boldsymbol{k})\mathcal{O}_B(-\boldsymbol{k})\rangle' + a_{a,\mu}k^2 + b_{a,\mu}k^4\,. \tag{52}$$

Third, to reduce the residual cosmic variance we use the correlators $\langle\mathcal{O}_A\mathcal{O}_B\rangle$ computed directly at the field level using the modes available in the simulation box. Fourth, we assume the transfer function errors $\propto 1/k$, as follows from

---

一旦我们从 βF 1(k, μ) 数据的拟合中获得了 β1、βη 测量值，我们就可以使用公式 (44) 将它们转换为常用的 EFT 参数。这个匹配实验表明，将传递函数匹配到我们新的 EFT 模型更为容易，该模型在 Gram-Schmidt 过程之上与前向模型完全等价。然后这些系数可以转换为常用的偏差参数。现在让我们来看一下二次阶的这个常数 EFT 传递函数模型。相应的 Ly-α 核函数为

K2(k1, k2) =β1˜ K2 + βη $\left(F_2^{\rm ZA} - \frac{3}{7} f\mu^2 F_{\mathcal{G}_2}\right) + \beta_2 2F_{\delta^2} + \beta_3 6F_{\delta^3}$

$+ \beta_{\mathcal{G}_2} F_{\mathcal{G}_2} + \beta_{\delta\eta} F_{\delta\eta} + \beta_{\eta^2} F_{\eta^2} + \beta_{KK_\parallel} F_{(KK)_\parallel} + \beta_{\Pi^{(2)}_\parallel} F_{\Pi^{(2)}_\parallel}$

≡(β1 + βη)Fδ − β1Fδη + βη −Fη + Fη2 −Fδη $+ 2\beta_2 F_{\delta^2} +$ −2 7β1 + 3 14βη + βG2$\Big)$ $F_{\mathcal{G}_2}$ (49)

+ βδηFδη + βη2Fη2 + βΠ[2] // Π[2] // +

β(KK) // F(KK) // =(β1 + βη)Fδ −βηFη + 2β2Fδ2 + � −2

7β1 + 3 14βη + βG2 $F_{\mathcal{G}_2}$

+ (−β1 −βη + βδη)Fδη + (βη2 + βη)Fη2 + βΠ[2] // FΠ[2] // + β(KK) // F(KK) //

请注意，这个模型明确具有 Π[2] // 操作符。上述内容可以与通常的 EFT 核进行比较

方程 (18)。匹配系数我们得到：

$$\begin{aligned}
&b_1 = \beta_1 + \beta_\eta\,, \quad b_\eta = -\beta_\eta\,, \quad b_2 = 2\beta_2\,, \quad b_3 = 6\beta_3\,, \quad b_{\mathcal{G}_2} = -\frac{2}{7}\beta_1 + \frac{3}{14}\beta_\eta + \beta_{\mathcal{G}_2}\,, \\
&b_{\delta\eta} = -\beta_1 - \beta_\eta + \beta_{\delta\eta}\,, \quad b_{\eta^2} = \beta_{\eta^2} + \beta_\eta\,, \quad b_{\Pi^{[2]}_\parallel} = \beta_{\Pi^{[2]}_\parallel}\,, \quad b_{(KK)_\parallel} = \beta_{(KK)_\parallel}\;.
\end{aligned} \tag{50}$$

上述匹配允许根据EFT传递函数的大尺度极限来确定线性和二次EFT参数。重要的是，上述匹配与线性理论匹配是一致的。我们还需要注意，虽然δ3算子存在于前向模型中，但它与δ1的关联在单圈阶次下是多余的，而它与二次算子的关联仅在双圈阶次才出现，这超出了我们工作的范围。因此，我们将在我们的EFT传递函数模型中忽略该算子。

相关算符之间的关联函数⟨OAOB⟩可以从 CLASS-PT [171] 中轻松提取，CLASS-PT 在 P22 计算例程中会计算上述所有可能的交叉关联。然而，在实践中，这个问题受到四个因素的影响。首先，如前所述，我们测量正交算符的传递函数，因此我们需要在前向模型中考虑 Gram-Schmidt 过程。第二，存在非线性修正，它引入了传递函数的尺度依赖性，使得从小尺寸模拟盒中提取其常数部分变得困难。第三，大尺度会受到残余宇宙方差的影响，对于二次算符来说这是不可忽略的。第四，与前述点相关，我们需要了解传递函数数据之间的协方差。在下面的内容中，我们将逐点讨论。首先，我们前向建模 Gram-Schmidt 过程。使用非正交化的偏差展开 δF = P AβAOA，算符 Oa 的测量传递函数遵循该模型

βF a(k, μ) = ⟨δF O⊥ a⟩′ ⟨|O⊥ a| = 1⟨|O⊥ a|2⟩′ 1⟨|O $\sum_A$ MaBβA⟨OAOB⟩′ , (51)

2⟩′ = ⟨δF MaBOB⟩′ ⟨|O⊥ a|2⟩′ MaB⟨δF OB⟩′ = a|2⟩′

其中指标 A、B 遍历包括 Π[2] // 在内的所有八个算符，相比之下，指标 a 仅遍历七个算符。MaB 是由原始的非正交算符的互谱构建的依赖于 k 的旋转矩阵。如果我们只能访问一个区间，上述方程的逆运算将不可行。然而，使用来自多个区间的数据，我们可以将上述问题作为最小二乘问题来求解。

其次，为了考虑尺度依赖的非线性，我们在传递函数中引入了额外的多项式

每个 μ-bin 都有自由拟合参数的模型，

$$\beta_a^F(k,\mu)\Big|_{\rm model} = \frac{1}{\langle|\mathcal{O}_a^\perp|^2\rangle'}\sum_A M_{aB}\beta_A\langle\mathcal{O}_A(\boldsymbol{k})\mathcal{O}_B(-\boldsymbol{k})\rangle' + a_{a,\mu}k^2 + b_{a,\mu}k^4\,. \tag{52}$$

第三，为了减少残余的宇宙方差，我们使用在场级别直接计算的相关函数⟨OAOB⟩，利用模拟盒中可用的模态。第四，我们假设传递函数误差 ∝1/k，如下所示

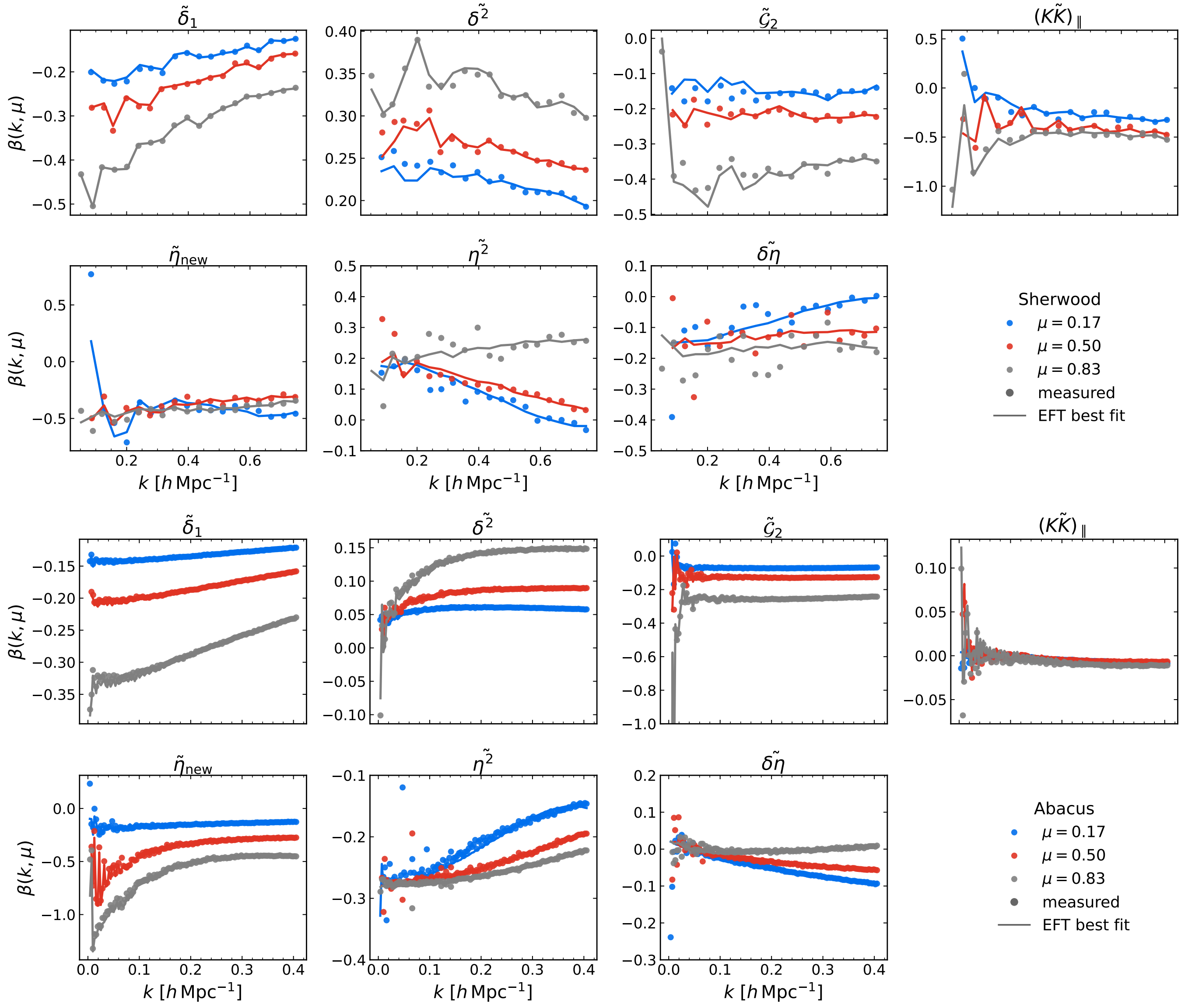

FIG. 18. **Transfer Functions:** Best-fitting EFT models for the seven transfer functions obtained from the Sherwood simulations (*top panel*) and the Abacus FGPA mocks (*bottom panel*). We compare the measured transfer functions (filled circles) to the EFT predictions (solid lines) in three angular bins μ (red where the angular bin is centered at $\mu = 0.83$ is close to the line-of-sight) and Fourier wavenumber k. Given the smaller statistical error bars of the Sherwood simulation, two-loop effects set in at a lower k compared to Abacus. Since the small-scale physics in the Sherwood simulations is much more accurate (by construction) than in Abacus, we measure and predict the transfer functions down to $k = 0.8\,h\,{\rm Mpc}^{-1}$ for Sherwood and to $k = 0.4\,h\,{\rm Mpc}^{-1}$ for Abacus. Note that we remove δ^3 from the fits to the measured transfer functions, as discussed in the main text.

mode counting. Since we do not have access to many simulations from which to extract the errors and the covariance, we apply the following procedure to account for the fact that β_1 must have the smallest errors of all transfer functions because it is a linear operator. We first fit β_1 only using the linear theory model corrected by the polynomials,

$$\beta_1^F(k,\mu) = \beta_1 + \beta_\eta(1+f\mu^2) + a_{1\mu}k^2 + b_{1\mu}k^4\,, \tag{53}$$

and then use the extracted values of β_1 and β_η as priors in the full fit including quadratic operators. We use $k_{\rm max} = 0.5\ (0.2)\ h{\rm Mpc}^{-1}$ at the first stage (fitting β_1) and $k_{\rm max} = 0.8\ (0.4)\ h{\rm Mpc}^{-1}$ at the second stage for Sherwood (Abacus). The results are shown in Fig. 18 and we highlight a key aspects: Given the larger number of quasi-linear modes in Abacus, our transfer function fits match better the measured ones. This is particularly noticeable for $\tilde{\eta}_{\rm new}$, $\tilde{\eta}^2$, and $\tilde{\delta\eta}$ where we get shape mismatches for Sherwood. The obtained best-fit values β_A then can be easily converted

---

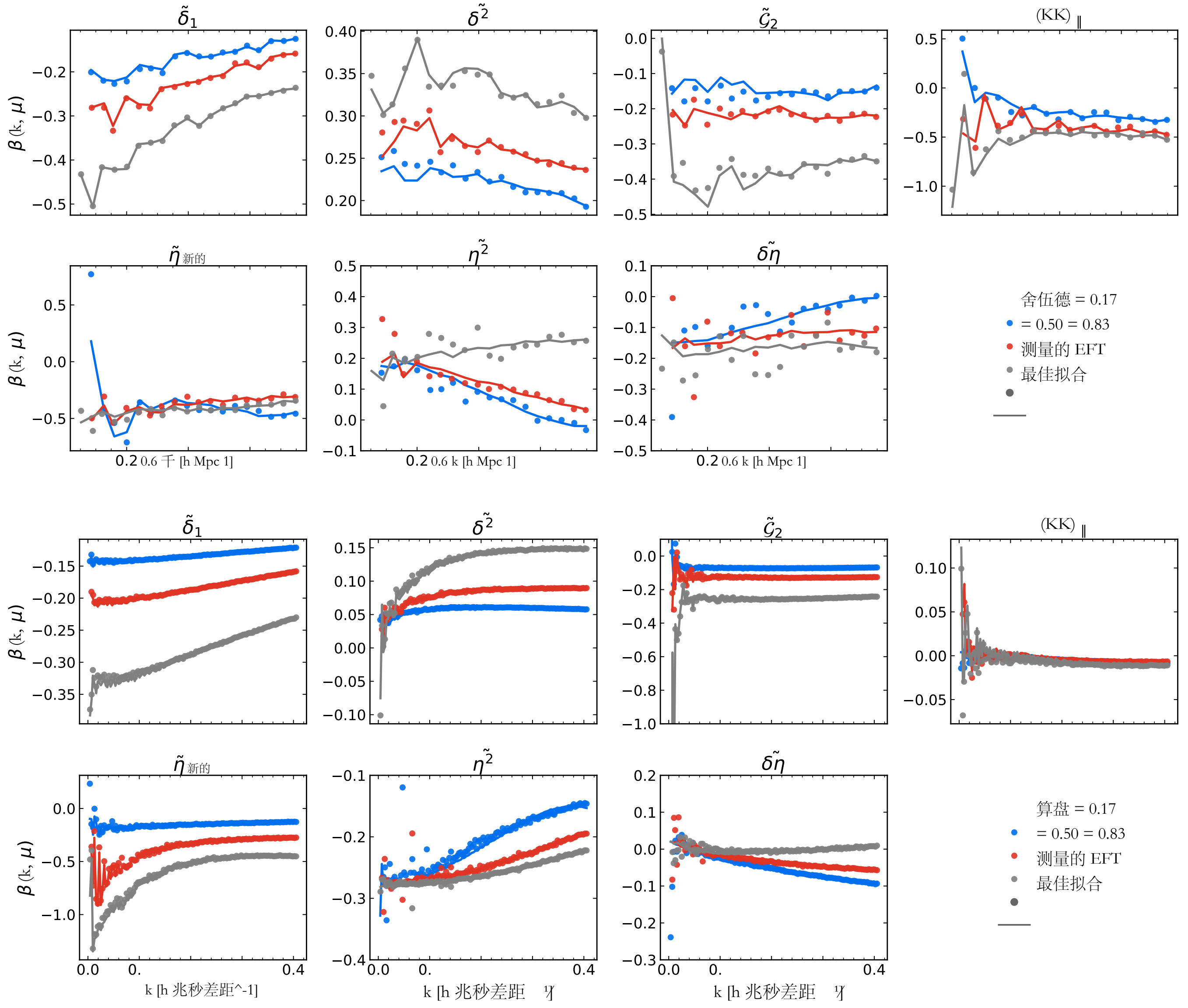

图18. 传递函数：来自Sherwood模拟（上图）和Abacus FGPA模拟（下图）的七个传递函数的最佳拟合EFT模型。我们将测量的传递函数（实心圆）与EFT预测（实线）在三个位于不同角度的角度区间 μ（红色表示角度区间中心在 μ = 0.83 接近视线方向）和傅里叶波数 k 上进行比较。由于Sherwood模拟的统计误差条较小，相比Abacus，两圈效应在较低的 k 范围内就开始显现。由于Sherwood模拟中的小尺度物理更加精确（根据设计），我们对Sherwood传递函数测量和预测的范围达到 k = 0.8 h Mpc−1，而对Abacus则达到 k = 0.4 h Mpc−1。注意，如正文所述，我们在对测量的传递函数进行拟合时移除了 δ3。

模式计数。由于我们无法获取大量模拟以提取误差和协方差，因此我们采用以下程序来考虑到 β1 必须具有所有传递函数中最小的误差这一事实，因为它是一个线性算子。我们首先仅使用由多项式修正的线性理论模型来拟合 β1，

$$\beta_1^F(k,\mu) = \beta_1 + \beta_\eta(1+f\mu^2) + a_{1\mu}k^2 + b_{1\mu}k^4\,, \tag{53}$$

然后使用提取的 β1 和 βη 的值作为带二次算符的完整拟合中的先验。我们在第一阶段（拟合 β1）使用 kmax = 0.5 (0.2) hMpc−1，在第二阶段对于 Sherwood (Abacus) 使用 kmax = 0.8 (0.4) hMpc−1。结果如图 18 所示，我们强调一个关键方面：由于 Abacus 中准线性模式数量更多，我们的传递函数拟合与测量值匹配得更好。这在 ˜ ηnew、˜ η2 和 ˜ δη 上特别明显，而在 Sherwood 上我们得到形状不匹配。得到的最佳拟合值 βA 然后可以很容易地转换。

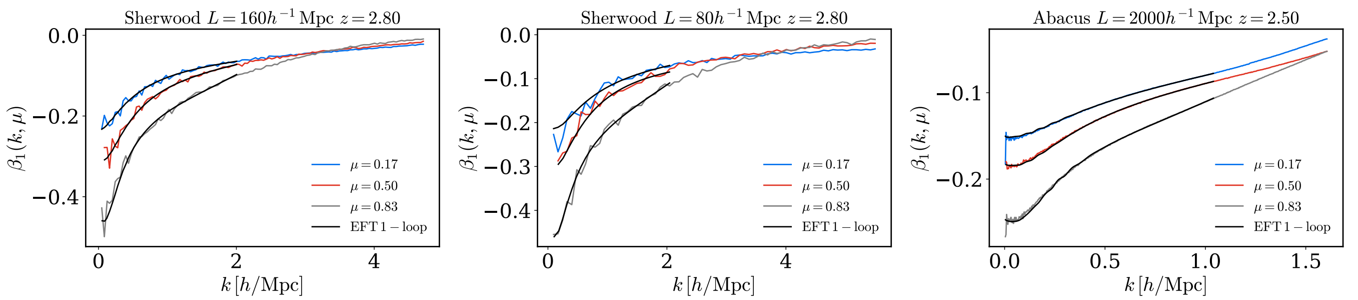

FIG. 19. **Best-fit EFT TF:** We show the best-fitting EFT models for the β_1 transfer function for the Sherwood (*left and center panel*) and the Abacus FGPA (*right panel*) mock. The transfer functions are shown in three angular bins μ (blue: $\mu = 0.17$, red: $\mu = 0.50$, gray: $\mu = 0.83$) and as a function of Fourier wavenumber k. The EFT 1-loop best-fit transfer functions are shown as solid black lines up to $k_{\max} = 2.0\,h\,\mathrm{Mpc}^{-1}$ ($k_{\max} = 1.0\,h\,\mathrm{Mpc}^{-1}$) for Sherwood (Abacus) which provide a good fit to the data. See Fig. 34 for the other Abacus models.

into the usual EFT parameters via Eq. (50), yielding the tabulated best-fit parameters in Tab. V. The obtained linear bias parameters are fully consistent with full-shape fits obtained from the one-loop power spectrum to Sherwood [136] and the Abacus simulations [78]. The cubic operators and counterterms can be extracted from the scale and orientation dependence of the β_1 transfer function. For that we use the full one-loop EFT model

$$\beta_1^F(k,\mu) = \frac{P_{F\tilde{1}}}{P_{\tilde{1}\tilde{1}}} = \frac{K_1^{\mathrm{lin+ctr}}(\boldsymbol{k})P_{11} + P_{22}^{F\tilde{1}} + P_{13}^{F\tilde{1}}}{P_{11} + \tilde{P}_{22} + 2\tilde{P}_{13}} \quad , \tag{54}$$

built from the non-linear kernels of the field-level shifted operators computed in [125, 172],

$$\begin{aligned} K_1^{\mathrm{lin+ctr}} &= b_1 - b_\eta f\mu^2 + k^2(c_0 + c_1\mu^2 + c_2\mu^4) \ , \\ \tilde{P}_{22} &= 2\int_{\boldsymbol{q}} [\tilde{K}_2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})]^2 P_{\mathrm{lin}}(q)P_{\mathrm{lin}}(|\boldsymbol{k}-\boldsymbol{q}|)\,, \quad \tilde{P}_{13} = 3P_{\mathrm{lin}}(k)\int_{\boldsymbol{q}} \tilde{K}_3(\boldsymbol{k},-\boldsymbol{q},\boldsymbol{q})P_{\mathrm{lin}}(q)\,, \\ P_{22}^{F\tilde{1}} &= 2\int_{\boldsymbol{q}} \tilde{K}_2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})K_2(-\boldsymbol{k}+\boldsymbol{q},-\boldsymbol{q})P_{\mathrm{lin}}(q)P_{\mathrm{lin}}(|\boldsymbol{k}-\boldsymbol{q}|)\,, \\ P_{13}^{F\tilde{1}} &= 3K_1(\boldsymbol{k})P_{\mathrm{lin}}(k)\int_{\boldsymbol{q}} \tilde{K}_3(\boldsymbol{k},-\boldsymbol{q},\boldsymbol{q})P_{\mathrm{lin}}(q) + 3P_{\mathrm{lin}}(k)\int_{\boldsymbol{q}} K_3(\boldsymbol{k},-\boldsymbol{q},\boldsymbol{q})P_{\mathrm{lin}}(q) \ , \end{aligned} \tag{55}$$

and impose the linear and quadratic bias parameter measurements as priors. This procedure is then applied at $k_{\max} = 2\ h\mathrm{Mpc}^{-1}$ ($1\ h\mathrm{Mpc}^{-1}$) for the Sherwood (Abacus) simulations yielding the bias parameters tabulated in Tab. V with b_1 and b_η being similar to the fitting value from the previous stage. The best-fit β_1 transfer functions from these fits are displayed in Fig. 19.

To illustrate the results further and compare both hydrodynamic simulations, we show in Fig. 20 the relation of the non-linear bias parameters $b_{\mathcal{O}}$ as a function of the linear bias b_1 and linear fits to the $b_1 - b_{\mathcal{O}}$ and $b_1 - c_n$ relations, tabulated in Tab. VI. It is interesting to note that within the $80\,h^{-1}$ Mpc Sherwood boxes the results follow the bias relation but for the larger volume we do see a difference in the linear bias parameters even though the resolution is kept constant for both Sherwood boxes. Note that we do not have error bars since we are performing best-fits and are not exploring the parameter space using e.g. MCMC chains. Note that the cubic parameters and counterterms have a significant degree of correlation between them. We believe that the unbroken degeneracies between these parameters are responsible for certain large values of cubic parameters, like $b_{\eta\Pi_\parallel^{[2]}}$. A more robust way to measure these parameters will be to explicitly add all necessary operators to the forward model, and fit their Wilson coefficients from the appropriate transfer functions.

We conclude that the EFT successfully predicts the shape of the transfer functions, from which the bias parameters can be measured. In future our fitting procedure can be improved with accurate covariances, two-loop computations, and the inclusion of cubic operators at the field level. We leave these for future investigation.

VII. LARGE-SCALE CLUSTERING MOCKS

One of the main bottlenecks in cosmological analyses of the Ly-α forest is the generation of large-volume high-resolution mocks. Therefore, approximate prescriptions such as log-normal mocks [129], or augmented LPT mocks

---

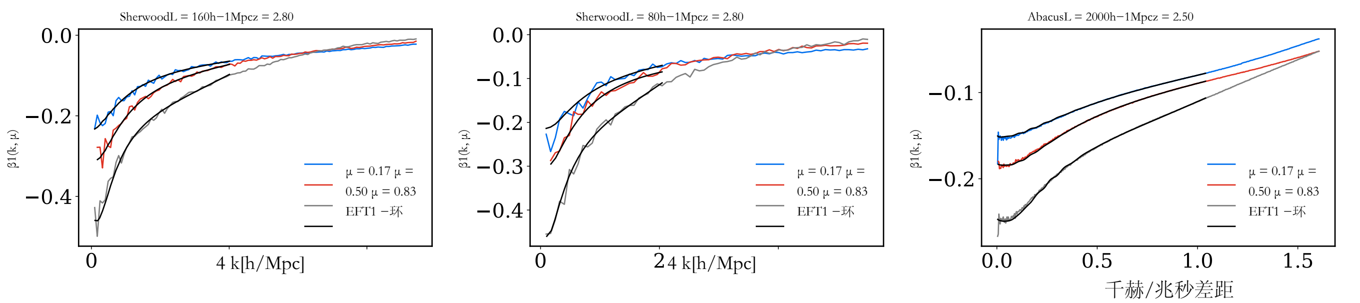

图19. 最佳拟合EFT传递函数: 我们展示了Sherwood（左和中面板）和Abacus FGPA（右面板）模拟中β1传递函数的最佳拟合EFT模型。传递函数以三个角度区间μ显示（蓝色: μ = 0.17，红色: μ = 0.50，灰色: μ = 0.83），并作为傅里叶波数k的函数。EFT一回路最佳拟合传递函数以黑色实线显示，Sherwood的kmax = 2.0 h Mpc−1（Abacus的kmax = 1.0 h Mpc−1）以内，该拟合能够很好地匹配数据。其他Abacus模型见图34。

通过公式（50）转化为常规的 EFT 参数，从而得到表 V 中列出的最佳拟合参数。获得的线性偏差参数与从 Sherwood [136] 和 Abacus 模拟 [78] 的一圈功率谱的全形状拟合结果完全一致。三次算子和对策项可以从 β1 传递函数的尺度和方向依赖中提取。为此，我们使用完整的一圈 EFT 模型

$$\text{βF 1(k, μ)} = \frac{P_{F\tilde{1}}}{P_{\tilde{1}\tilde{1}}} = \frac{K_1^{\mathrm{lin+ctr}}(\boldsymbol{k})P_{11} + P_{22}^{F\tilde{1}} + P_{13}^{F\tilde{1}}}{P_{11} + \tilde{P}_{22} + 2\tilde{P}_{13}} \quad , \tag{54}$$

1˜ 1

由在 [125, 172] 中计算的场级位移算子非线性核构建，

$$\begin{aligned} K_1^{\mathrm{lin+ctr}} &= \text{b1 − bηfμ2 + k2(c0 + c1μ2 + c2μ4) ,} \\ \tilde{P}_{22} &= 2 \ {}_{\boldsymbol{q}} \text{[˜ K2(k −q, q)]\^{}2 Plin(q) Plin(|k −q|) , ˜ P13 = 3Plin(k)} \int_{\boldsymbol{q}} \text{˜ K3(k, −q, q)Plin(q) ,} \\ P_{22}^{F\tilde{1}} &= 2\int_{\boldsymbol{q}} \text{˜ K2(k −q, q)K2(−k + q, −q)Plin(q)Plin(|k −q|) ,} \\ P_{13}^{F\tilde{1}} &= 3K_1(\boldsymbol{k})P_{\mathrm{lin}}(k)\int_{\boldsymbol{q}} \text{˜ K3(k, −q, q)Plin(q) + 3Plin(k)} \int_{\boldsymbol{q}} \text{K3(k, −q, q)Plin(q) ,} \end{aligned} \tag{55}$$

并将线性和二次偏差参数测量作为先验。然后将此过程应用于 kmax = 2 hMpc−1（Sherwood 模拟）和 1 hMpc−1（Abacus 模拟），得到表 V 中列出的偏差参数，其中 b1 和 bη 与前一阶段的拟合值相似。这些拟合得到的最佳拟合 β1 传递函数显示在图 19 中。

为了进一步说明结果并比较两种水动力学模拟，我们在图20中展示了非线性偏差参数 bO 与线性偏差 b1 的关系，以及 b1 − bO 和 b1 − cn 关系的线性拟合，具体列在表VI中。有趣的是，在 80 h−1Mpc 的 Sherwood 盒子中，结果遵循偏差关系，但对于更大体积，我们确实看到线性偏差参数有所不同，即使两个 Sherwood 盒子的分辨率保持不变。需要注意的是，我们没有误差棒，因为我们是在进行最优拟合，并没有使用例如 MCMC 链的方法探索参数空间。还需要注意的是，三次参数和反项之间存在显著的相关性。我们认为，这些参数之间未破坏的简并性导致了某些三次参数的较大值，比如 bηΠ[2] ∥。一种更健全的方法来测量这些参数是明确地将所有必要的算子添加到前向模型中，并从适当的传递函数中拟合它们的Wilson系数。

我们得出结论：有效场论成功预测了传递函数的形状，从中可以测量偏差参数。未来，我们的拟合程序可以通过精确的协方差、两环计算以及在场级别引入三次算符来改进。我们将这些留待未来研究。

七. 大规模集群模拟

在对Ly-α森林进行宇宙学分析时的主要瓶颈之一是生成大体积高分辨率的模拟。因此，诸如对数正态模拟 [129] 或增强型LPT模拟这样的近似方法

	Sherwood					**Abacus**			
L	$160\,h^{-1}$ Mpc	$80\,h^{-1}$ Mpc				$2\,h^{-1}$ Gpc			
z	2.8	2.0	2.4	2.8	3.2	2.5			
Models	–	–	–	–	–	I	II	III	IV
b_1	−0.210	−0.104	−0.148	−0.202	−0.270	−0.150	−0.133	−0.136	−0.132
b_η	0.323	0.199	0.262	0.336	0.410	0.149	0.136	0.287	0.321
b_2	0.392	0.170	0.265	0.415	0.619	0.128	0.127	0.100	0.107
b_{G_2}	−0.045	−0.075	−0.089	−0.087	−0.066	−0.016	−0.012	−0.007	−0.007
$b_{(KK)_\parallel}$	−0.423	−0.269	−0.415	−0.636	−0.852	−0.046	−0.061	−0.016	−0.030
b_{η^2}	−0.250	−0.055	−0.036	−0.001	−0.061	−0.203	−0.164	−0.268	−0.284
$b_{\delta\eta}$	0.079	−0.114	−0.130	−0.104	−0.022	0.073	0.064	0.106	0.106
$b_{\Pi^{[2]}_\parallel}$	0.347	0.207	0.288	0.386	0.590	0.024	0.008	−0.074	−0.066
b_{Γ_3}	−0.154	−0.754	−0.855	−0.627	0.736	0.814	0.599	1.253	1.223
$b_{\Pi^{[3]}_\parallel}$	1.544	0.043	0.109	0.214	0.512	1.603	1.349	3.795	3.908
$b_{\delta\Pi^{[2]}_\parallel}$	0.239	−4.141	−6.325	−9.348	−12.373	3.135	2.615	6.035	6.091
$b_{K\Pi^{[2]}_\parallel}$	0.841	1.279	1.009	−0.353	−3.930	−1.225	−0.814	−3.901	−3.939
$b_{\eta\Pi^{[2]}_\parallel}$	−2.157	−12.177	−17.132	−22.932	−26.187	8.093	6.490	15.917	15.954
c_0 $[h^{-1}\mathrm{Mpc}]^2$	0.002	−0.007	−0.013	−0.024	−0.039	0.015	0.014	0.019	0.018
c_2 $[h^{-1}\mathrm{Mpc}]^2$	−0.130	−0.011	−0.008	−0.006	−0.005	−0.020	−0.018	0.022	0.015
c_4 $[h^{-1}\mathrm{Mpc}]^2$	0.219	0.048	0.057	0.078	0.115	0.083	0.075	0.061	0.075

TABLE V. **Best-fit EFT bias parameters from TFs:** Best-fit EFT parameters obtained from the transfer functions, introduced in Sec. VI, via Eq. (50) for both sets of simulations analyzed in the present work: (i) from the $L = 160\,h^{-1}$ Mpc and $L = 80\,h^{-1}$ Mpc Sherwood simulations at four different redshifts using the same effective resolution; and (ii) the additional Abacus models I-IV obtained from a box of length $L = 2\,h^{-1}$ Gpc using a different resolution. The top block quotes the linear and quadratic bias operators, the middle block the cubic ones and the bottom block are the counterterms. We emphasize that these values should be interpreted with caution. The quoted bias parameters are best-fit values for which uncertainties cannot yet be reliably estimated, as the covariance of the transfer functions has not been measured (and is expected to be very small because of the Gram-Schmidt orthogonalization). Furthermore, the cubic bias parameters are significantly degenerate because they are inferred from a global fit to β_1 rather than measured directly at the field level. We leave a more robust determination of these parameters to future work.

$b_\mathcal{O}$	p_1	p_0	$b_\mathcal{O}$	p_1	p_0	$b_\mathcal{O}$	p_1	p_0
b_η	-1.273	0.071	$b_{(KK)_\parallel}$	3.562	0.102	$b_{K\Pi^{[2]}_\parallel}$	31.792	5.256
b_2	-2.729	-0.127	$b_{\Pi^{[2]}_\parallel}$	-2.287	-0.046	$b_{\eta\Pi^{[2]}_\parallel}$	85.125	-4.199
b_{G_2}	-0.066	-0.091	b_{Γ_3}	-9.012	-2.006	c_0	0.196	0.015
b_{η^2}	0.002	-0.038	$b_{\Pi^{[3]}_\parallel}$	-2.806	-0.288	c_2	-0.035	-0.014
$b_{\delta\eta}$	-0.580	-0.198	$b_{\delta\Pi^{[2]}_\parallel}$	50.060	1.014	c_4	-0.409	0.000

TABLE VI. **Linear relation between b_1 and $b_\mathcal{O}$:** Coefficients of the linear fits of the EFT bias parameters and counterterms to the linear bias b_1, $b_\mathcal{O}(b_1) = p_1\,b_1 + p_0$, obtained from the four Sherwood snapshots at $z = 2.0, 2.4, 2.8, 3.2$ ($L = 80\,h^{-1}$Mpc) and shown in Fig. 20. Counterterms c_n are in units of $(h^{-1}\mathrm{Mpc})^2$.

[173] as well as the present Abacus FPGA mocks [102] have been used in the past. We present and validate a new approach to generate (in principle, arbitrarily) large Ly-α mocks calibrated on small-scale hydrodynamic simulations which can be used for pipeline testing and covariance matrix estimation. As a first proof-of-principle we generate large-scale clustering mocks with a box length of $V = 2^3(\,h^{-1}\,\mathrm{Gpc})^3$ and a cell size of $\approx 1.95\,h^{-1}$ Mpc calibrated on

---

	舍伍德					**算盘**			
L	$160\,h^{-1}$ Mpc	$80\,h^{-1}$ Mpc				$2\,h^{-1}$ Gpc			
z	2.8	2.0	2.4	2.8	3.2	2.5			
模型	–	–	–	–	–	我	二	三	IV
b_1	−0.210	−0.104	−0.148	−0.202	−0.270	−0.150	−0.133	−0.136	−0.132
b_η	0.323								
b_2	0.392	0.170	0.265	0.415	0.619	0.128	0.127	0.100	0.107
b_{G_2}	−0.045	−0.075	−0.089	−0.087	−0.066	−0.016	−0.012	−0.007	−0.007
$b_{(KK)_\parallel}$	−0.423	−0.269	−0.415	−0.636	−0.852	−0.046	−0.061	−0.016	−0.030
b_{η^2}	−0.250	−0.055	−0.036	−0.001	−0.061	−0.203	−0.164	−0.268	−0.284
$b_{\delta\eta}$	0.079	−0.114	−0.130	−0.104	−0.022				
$b_{\Pi^{[2]}_\parallel}$	0.347					0.024	0.008	−0.074	−0.066
b_{Γ_3}	−0.154	−0.754	−0.855	−0.627	0.736	0.814	0.599	1.253	1.223
$b_{\Pi^{[3]}_\parallel}$	1.544				0.512	1.603	1.349	3.795	3.908
$b_{\delta\Pi^{[2]}_\parallel}$	0.239	−4.141	−6.325	−9.348	−12.373	3.135	2.615	6.035	6.091
$b_{K\Pi^{[2]}_\parallel}$	0.841	1.279	1.009	−0.353	−3.930	−1.225	−0.814	−3.901	−3.939
$b_{\eta\Pi^{[2]}_\parallel}$	−2.157	−12.177	−17.132	−22.932	−26.187	8.093	6.490	15.917	15.954
c0 [h−1兆秒差距]^2	0.002	−0.007	−0.013	−0.024	−0.039	0.015	0.014	0.019	0.018
c2 [h−1兆秒差距]2	−0.130	−0.011	−0.008	−0.006	−0.005	−0.020	−0.018	0.022	0.015
c4 [h−1兆秒差距]^2	0.219							0.061	0.075

表 V. 来自 TF 的最佳拟合 EFT 偏差参数：通过第 VI 节引入的传递函数，并使用公式 (50) 获得的最佳拟合 EFT 参数，针对本文分析的两组模拟：(i) 来自 L = 160 h−1Mpc 和 L = 80 h−1Mpc Sherwood 模拟，在四个不同红移使用相同有效分辨率；(ii) 额外的 Abacus I-IV 模型，使用边长 L = 2 h−1Gpc 的模拟盒，并采用不同分辨率。顶部区块引用线性和二次偏差算子，中部区块引用立方偏差算子，底部区块为反项。我们强调，这些数值应谨慎解释。所引用的偏差参数是最佳拟合值，其不确定性尚无法可靠估计，因为传递函数的协方差尚未测量（并且由于 Gram-Schmidt 正交化，预计其非常小）。此外，三次偏差参数显著退化，因为它们是通过对 β1 的全局拟合推断的，而不是在现场直接测量的。我们将这些参数的更稳健确定留待未来工作中进行。

$b_\mathcal{O}$	p_1	p_0	$b_\mathcal{O}$	p_1	p_0	$b_\mathcal{O}$	p_1	p_0
b_η	-1.273	0.071	$b_{(KK)_\parallel}$	3.562	0.102	$b_{K\Pi^{[2]}_\parallel}$	31.792	5.256
b_2	-2.729	-0.127	$b_{\Pi^{[2]}_\parallel}$	-2.287	-0.046	$b_{\eta\Pi^{[2]}_\parallel}$	85.125	-4.199
bG2	-0.066	-0.091	b_{Γ_3}	-9.012	-2.006	c_0	0.196	0.015
b_{η^2}	0.002	-0.038	$b_{\Pi^{[3]}_\parallel}$	-2.806	-0.288	c_2	-0.035	-0.014
bδη	-0.580	-0.198	$b_{\delta\Pi^{[2]}_\parallel}$	50.060	1.014	c_4	-0.409	0.000

表 VI. b1 与 bO 的线性关系：EFT 偏差参数和反项对线性偏差 b1 的线性拟合系数 bO(b1) = p1 b1 + p0，来自 z = 2.0、2.4、2.8、3.2 的四个 Sherwood 快照（L = 80 h−1Mpc），如图 20 所示。反项 cn 的单位为 (h−1Mpc)^2。

[173] 以及目前的 Abacus FPGA 模拟 [102] 在过去也被使用过。我们提出并验证了一种新的方法来生成（原则上，可以任意生成）基于小尺度流体力学模拟校准的大的 Ly-α 模拟，这些模拟可用于管道测试和协方差矩阵估计。作为一个初步的概念验证，我们生成了大尺度聚类模拟，其盒子长度为 V = 23(h−1Gpc)^3，单元格大小约为 ≈1.95 h−1Mpc，并基于

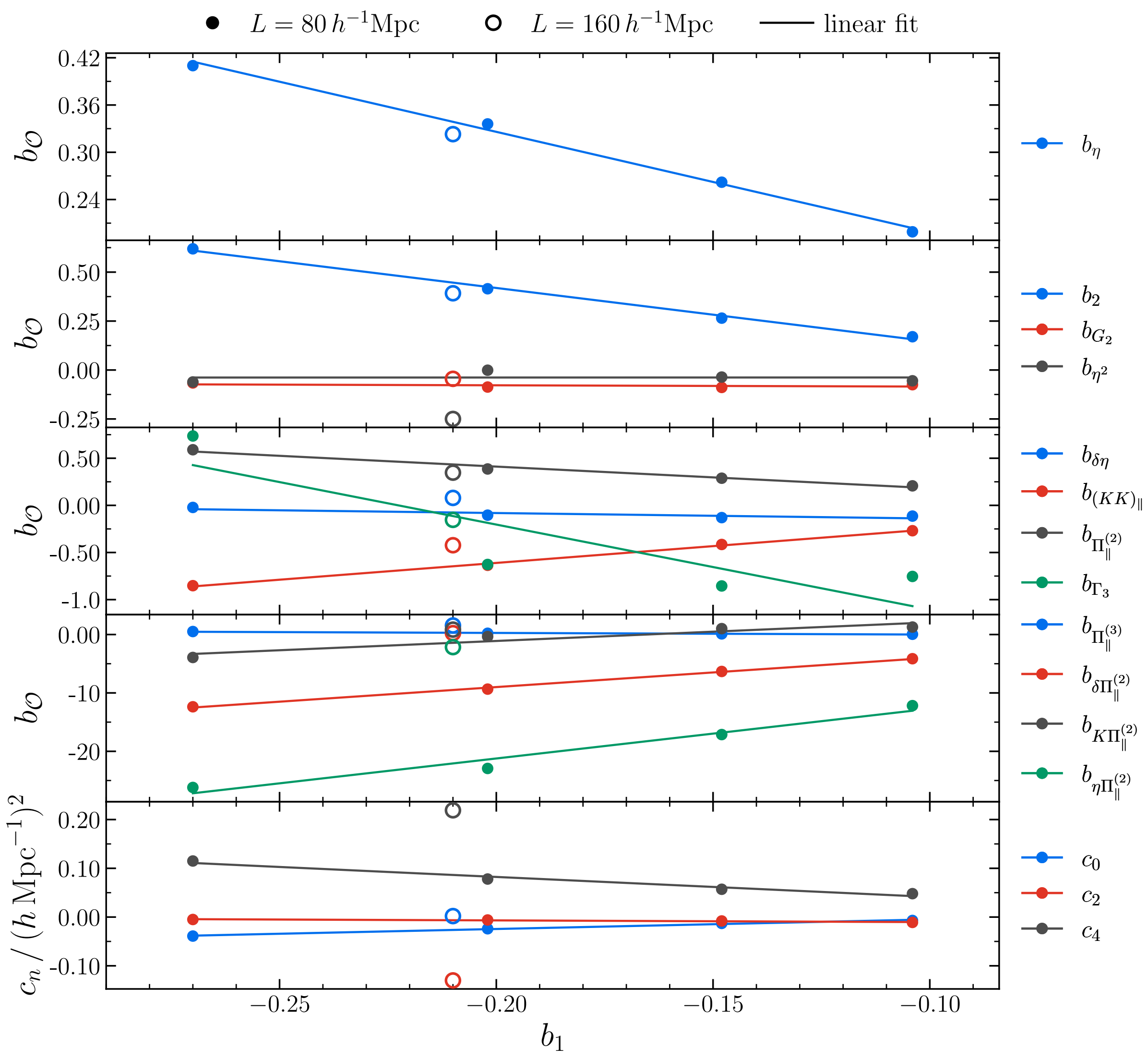

FIG. 20. **Bias parameter relation:** Comparison of best-fit bias parameters and counterterms obtained from field-level fits to the Sherwood $L = 160\,h^{-1}$ Mpc (empty circles) at $z = 2.8$ and $L = 80\,h^{-1}$ Mpc boxes at four redshifts (from left to right for $z = 3.2$, 2.8, 2.4, 2.0 as filled circles). Note that we do not show error bars as the field-level fits are best-fit values performed for a single realization. The solid lines are linear fits to the evolution of the bias parameters and counter terms of the smaller box. We remind the reader that our convention of b_η is related to the literature by a negative sign. The coefficients for the linear fits are tabulated in Tab. VI.

Sherwood simulations.[9]

In summary, we perform the following steps to create the three-dimensional Ly-α forest flux decrement or dark matter halo density fields.

1. We fit the transfer functions given in Eq. (33) for the Ly-α forest and Eq. (34) for dark matter halos at the field level on hydrodynamic simulations. These fits benefit from cosmic variance cancellation as we use the same set of initial conditions for the forward model as for the simulations.

2. Next, we compute (and fit) the perturbative β_1 transfer function predicted by the EFT one-loop model obtained from `ClASS-PT` and fit a polynomial to the remaining $\beta_{\mathcal{O}}$ transfer functions. Similarly, we fit a polynomial to the measured mean-square model error ($P_{\rm err}$).

---

[9] We find consistent results when changing the resolution by a factor of two.

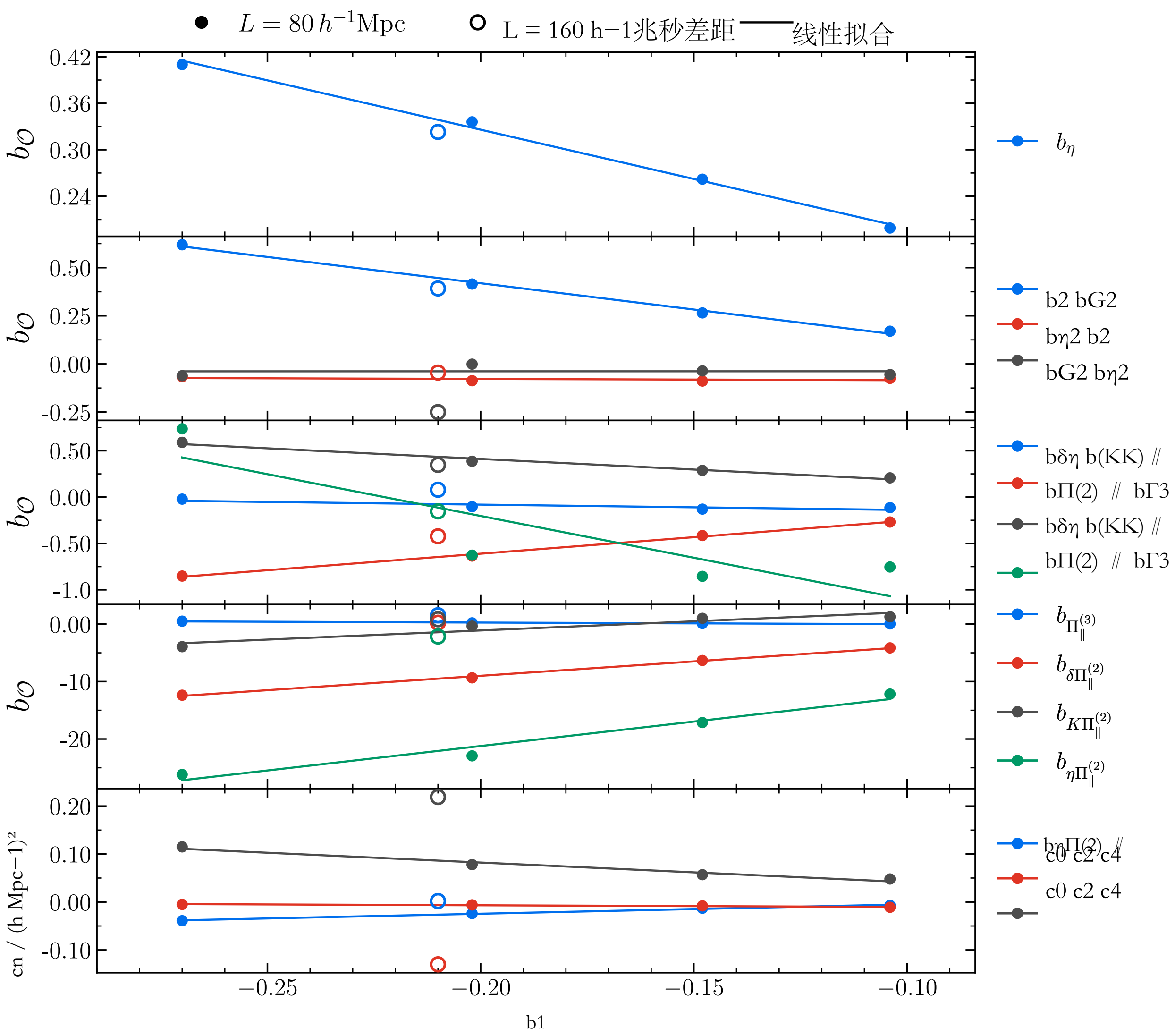

图 20. 偏差参数关系：比较从场级拟合中得到的最佳拟合偏差参数和反项，针对 Sherwood L = 160 h−1Mpc (空心圆) 在 z = 2.8 的情况，以及 L = 80 h−1Mpc 盒在四个红移下的情况（从左到右为 z = 3.2、2.8、2.4、2.0 的实心圆）。注意，我们没有显示误差条，因为场级拟合是针对单个实现得到的最佳拟合值。实线表示对较小盒子偏差参数和反项演化的线性拟合。我们提醒读者，我们的 bη 约定与文献相比带有负号。线性拟合的系数列于表 VI。

舍伍德模拟.9

总之，我们执行以下步骤来创建三维Ly-α森林通量减弱或暗区
物质晕密度场。

1. 我们在流体动力学模拟的场级上拟合方程 (33) 给出的 Ly-α 森林传递函数以及方程 (34) 给出的暗物质晕传递函数。这些拟合得益于宇宙方差的抵消，因为我们使用了相同的集合
用于正向模型的初始条件与用于模拟的初始条件相同。

2. 接下来，我们计算（并拟合）由 ClASS-PT 得到的 EFT 单回路模型所预测的微扰 β1 传递函数，并对剩余的 βO 传递函数拟合一个多项式。同样，我们对测量到的均方模型误差 (Perr) 拟合一个多项式。

---

9 当我们将分辨率改变一倍时，结果是一致的。

3. To generate large-scale clustering mocks we first generate a new set of initial conditions with a given box size (here: $L = 2\,h^{-1}$ Gpc) and grid resolution (here: $1.95\,h^{-1}$ Mpc). These are used to construct the shifted fields $\tilde{\delta}$ which are subsequently orthogonalized using the Gram-Schmidt procedure. Note that since the dependence on the specific ICs is captured by the shifted fields, we can apply the transfer functions and their smooth polynomial fits to realizations with different ICs.
4. We multiply each orthogonal shifted field $\tilde{\delta}^{\perp}$ by the corresponding best-fit transfer function obtained in step (1) which yields the part of the Ly-α forest (or halo) field that is correlated with the ICs.
5. In addition to the signal, we compute the stochastic part by generating a field with the same box size and resolution with a power spectrum matching the mean-squared model error, given by the polynomial fits from step (2).

The final mock is the sum of the signal and noise component. We generate these mocks for both the Ly-α forest, calibrated on Sherwood simulations, and high-redshift LAE and LBG samples, calibrated on Astrid simulations, respectively. We validate our methodology to construct large-scale clustering mocks by performing full-shape and BAO analyses at fixed cosmology for (i) the Ly-α forest alone, and (ii) a joint analysis of the Ly-α forest with LAEs/LBGs.

We stress that the mocks we generate contain the deterministic non-Gaussian information. We use a Gaussian approximation for the stochastic noise only, because it allows for a straightforward implementation of the scale- and angle-dependence of the error power spectrum. In principle, one may consider a more complicated model for the distribution of the noise field, which may include a combination of Poissonian and Gaussian components. However, using a scale-dependent Gaussian noise is sufficient for the modeling of the two-point statistics, which the main goal of our current study.

A. Validation procedure: Full-shape & BAO fits

The key quantity of interest of DESI Ly-α and galaxy clustering analyses is to measure the position of the BAO peak [50–52]. The BAO scaling parameters are obtained by setting constraints on distortions created from a difference between the fiducial and the true cosmology when converting observed to 3D coordinates, the so-called Alcock-Paczyński effect [174], parameterized by $\boldsymbol{\alpha} \equiv \{\alpha_\parallel, \alpha_\perp\}$. The radial and transverse parameters encode the Hubble parameter $H(z)$, the sound horizon at the redshift of decoupling $r_s(z_d)$ and the angular diameter distance $D_A(z)$ at the effective redshift z of the sample by re-mapping true to observed positions $k \to k'$, $\mu \to \mu'$ with

$$k' = \frac{k}{\alpha_\perp}\left[1 + \mu^2\left(\frac{1}{F^2} - 1\right)\right]^{1/2}, \qquad \mu' = \frac{\mu}{F}\left[1 + \mu^2\left(\frac{1}{F^2} - 1\right)\right]^{-1/2}, \tag{56}$$

where we define $F \equiv \alpha_\parallel/\alpha_\perp$. Now, the BAO scaling factors are defined as [175]

$$\alpha_\parallel \equiv \frac{H^{\rm fid}(z) r_s^{\rm fid}(z_d)}{H(z) r_s(z_d)}, \quad \alpha_\perp \equiv \frac{D_A(z) r_s^{\rm fid}(z_d)}{D_A^{\rm fid}(z) r_s(z_d)}, \tag{57}$$

with $r_s^{\rm fid}(z_d)$ being the fiducial value of the sound horizon scale at the drag epoch. We "observe" unprimed quantities which are evaluated at the redshift of the mock. The superscripts *fid* and *tem* refer to quantities in the fiducial and template (here: Sherwood) cosmology. The final expression for the fits is a function of the rescaled coordinates k' and μ' and an additional normalization (volume) factor $P(k,\mu) = P_m(k',\mu')/(\alpha_\perp^2 \alpha_\parallel)$.

For the model of the power spectrum $P(k,\mu)$ we employ the one-loop EFT model of the Ly-α forest which consists of four key components

$$P^{\rm th.}(k,\mu) = P^{\rm tree}(k,\mu) + P^{1-\rm loop}(k,\mu) + P^{\rm ct}(k,\mu) + P^{\rm st}(k,\mu), \tag{58}$$

where k is the Fourier wavenumber and μ the angle of $k = \{k_\parallel, \mathbf{k}_\perp\}$ to the line-of-sight, $\mu \equiv k_\parallel/k$. The infrared resummed linear theory power spectrum[10] is given by

$$P^{\rm tree}(k,\mu) = K_1^2(\mathbf{k}) P_{\rm lin}(k), \qquad K_1(\mathbf{k}) \equiv (b_1 - b_\eta f \mu^2) \tag{59}$$

---

[10] For the fits we use time-sliced perturbation theory to decompose the power spectrum into a smooth (nw) and oscillatory (w) component [145, 171]. Both components are treated independently throughout the modeling procedure and are combined only at the final stage.

---

3. 为了生成大规模聚类模拟，我们首先生成一组具有给定盒子尺寸（此处：L = 2 h−1Gpc）和网格分辨率（此处：1.95 h−1Mpc）的初始条件。这些条件用于构建移位场 ˜ δ，随后使用 Gram-Schmidt 程序对其进行正交化。请注意，由于对特定初始条件的依赖由移位场捕捉，我们可以应用传递函数及其平滑多项式。

适合具有不同IC的实现。

4. 我们将每个正交平移场 ˜ δ⊥ 乘以在步骤 (1) 中获得的对应最佳拟合传递函数
这产生了与初始条件（ICs）相关的莱曼-α森林（或晕）场的部分。

5. 除了信号之外，我们还通过生成一个具有相同盒子大小和分辨率的场来计算随机部分，其功率谱与均方模型误差相匹配，该误差由多项式拟合给出
步骤（2）。

最终的模拟是信号和噪声成分的总和。我们为 Ly-α 森林（基于 Sherwood 模拟校准）以及高红移 LAE 和 LBG 样本（分别基于 Astrid 模拟校准）生成这些模拟。我们通过在固定宇宙学条件下对 (i) 仅 Ly-α 森林和 (ii) Ly-α 森林与 LAEs/LBGs 联合分析进行全形状和 BAO 分析，来验证构建大尺度聚类模拟的方法。

我们强调，我们生成的模拟包含确定性的非高斯信息。我们仅对随机噪声使用高斯近似，因为它可以方便地实现误差功率谱的尺度和角度依赖性。原则上，可以考虑更复杂的噪声场分布模型，其中可能包括泊松和高斯成分的组合。然而，对于两点统计的建模而言，使用尺度依赖的高斯噪声已经足够，这也是我们当前研究的主要目标。

A. 验证程序：全形状和BAO拟合

DESI Ly-α 和星系聚类分析中最关键的关注量是测量 BAO 峰的位置 [50 – 52]。BAO 缩放参数是通过对在将观测到的坐标转换为三维坐标时，由参考宇宙学与真实宇宙学的差异产生的扭曲设置约束来获得的，这就是所谓的 Alcock-Paczyński 效应 [174]，其参数化为 α ≡ {α⫽, α⊥}。径向和横向参数通过将真实位置重新映射为观测位置 k → k′, μ → μ′ 来编码样本有效红移 z 下的哈勃参数 H(z)、解耦红移处的声波视界 rs(zd) 和角直径距离 DA(z)。

$$k' = \frac{k}{\alpha_\perp} \quad 1 + \mu^2 \quad \frac{1}{F^2} - 1 \quad ^{1/2}, \qquad \mu' = \frac{\mu}{F} \quad 1 + \mu^2 \quad \frac{1}{F^2} - 1 \quad ^{-1/2}, \tag{56}$$

我们在这里定义 F ≡ α⫽ / α⊥。现在，BAO 缩放因子定义为 [175]

$$\alpha_\parallel \equiv \frac{H^{\rm fid}(z) r_s^{\rm fid}(z_d)}{H(z) r_s(z_d)}, \quad \alpha_\perp \equiv \frac{D_A(z) r_s^{\rm fid}(z_d)}{D_A^{\rm fid}(z) r_s(z_d)}, \tag{57}$$

其中 rfid s(zd) 是阻力历元处声波视界尺度的基准值。我们"观测"的是在模拟红移处计算的非原始量。上标 fid 和 tem 分别表示基准宇宙学和模板（此处：Sherwood）宇宙学中的量。拟合的最终表达式是重缩放坐标 k′ 和 μ′ 的函数，并包含一个额外的归一化（体积）因子 P(k, μ) = Pm(k′, μ′)/(α2 ⊥α⫽)。对于功率谱 P(k, μ) 的模型，我们采用 Ly-α 林的一环 EFT 模型，该模型由四个关键组成部分构成

$$P^{\rm th.}(k,\mu) = P^{\rm tree}(k,\mu) + P^{1-\rm loop}(k,\mu) + P^{\rm ct}(k,\mu) + P^{\rm st}(k,\mu), \tag{58}$$

其中 k 是傅里叶波数，μ 是 k = {k⫽, k⊥} 与视线的夹角，μ ≡ k⫽/k。红外重求和的线性理论功率谱10 给出为

$$P^{\rm tree}(k,\mu) = K_1^2(\mathbf{k}) P_{\rm lin}(k), \qquad K_1(\mathbf{k}) \equiv (b_1 - b_\eta f \mu^2) \tag{59}$$

---

[10] 对于拟合，我们使用时间切片微扰理论将功率谱分解为平滑（nw）和振荡（w）分量 [145, 171]。在整个建模过程中，这两个分量都是独立处理的，并且仅在最后阶段合并。

阶段

where f is the (linear) growth rate. The first higher-order correction to the tree-level power spectrum is

$$P^{1-\mathrm{loop}}(k,\mu) = 2\int_{\mathbf{q}} K_2^2(\mathbf{q},\mathbf{k}-\mathbf{q})P_{\mathrm{lin}}(|\mathbf{k}-\mathbf{q}|)P_{\mathrm{lin}}(q) + 6K_1(\mathbf{k})P_{\mathrm{lin}}(k)\int_{\mathbf{q}} K_3(\mathbf{k},-\mathbf{q},\mathbf{q})P_{\mathrm{lin}}(q)\,. \tag{60}$$

with higher order redshift-space kernels, $K_{2,3}$ [see Eq. (3.19) in 136] and we use the notation $\int_{\mathbf{q}} \equiv \int \frac{d^3q}{(2\pi)^3}$ to denote the three-dimensional integral over $\mathbf{q}$. The counter terms

$$P^{\mathrm{ct}}(k,\mu) = -2(c_0 + c_2\mu^2 + c_4\mu^4)K_1(\mathbf{k})k^2P_{\mathrm{lin}}(k)\,, \tag{61}$$

scale as $k^2P_{\mathrm{lin}}(k)$ and the stochastic contributions

$$P^{\mathrm{st}}(k,\mu) = P_{\mathrm{shot}} + a_0\frac{k^2}{{k_{\mathrm{NL}}}^2} + a_2\frac{k^2\mu^2}{{k_{\mathrm{NL}}}^2}\,, \tag{62}$$

scale as a constant shot noise piece and a scale-and angle-dependent term capturing small-scale clustering. We refer the reader to Refs. [77, 78, 127] for a fuller presentation of the methodology.

To jointly fit the Ly-α forest and high-redshift galaxy samples (here: LBGs and LAEs) we use a one-loop EFT model, following the procedure described in Refs. [78, 127]. In linear theory, this corresponds to computing the geometric mean of both tree-level tracers, $P_{\times}^{\mathrm{tree}}(k,\mu) = (b_1 - b_\eta f\mu^2)(b_1^q + f\mu^2)P_{\mathrm{lin}}(k)$, where the subscript $\times$ denotes the cross-correlation, and superscript q represents the quasar (or halo) tracer, i.e. b_1^q is the linear bias parameter for quasars. In practice, we remove the line-of-sight dependent terms by remapping the bias parameters (see, e.g., [86, 135, 136]) which yields for the cross-correlation the following form

$$\begin{aligned} P_{\times}^{\mathrm{th.}}(k,\mu) =& K_1(\mathbf{k})K_1^q(\mathbf{k})P_{\mathrm{lin}}(k) + 2\int_{\mathbf{q}} K_2(\mathbf{q},\mathbf{k}-\mathbf{q})K_2^q(\mathbf{q},\mathbf{k}-\mathbf{q})P_{\mathrm{lin}}(|\mathbf{k}-\mathbf{q}|)P_{\mathrm{lin}}(q) \\ &+ 3P_{\mathrm{lin}}(k)\int_{\mathbf{q}}[K_1(\mathbf{k})K_3^q(\mathbf{k},-\mathbf{q},\mathbf{q}) + K_1^q(\mathbf{k})K_3(\mathbf{k},-\mathbf{q},\mathbf{q})]P_{\mathrm{lin}}(q) - (c_0 + c_2\mu^2 + c_4\mu^4)K_1^q(\mathbf{k})k^2P_{\mathrm{lin}}(k) \\ &- (c_0^q + c_2^q\mu^2 + c_4^q\mu^4)K_1(\mathbf{k})k^2P_{\mathrm{lin}}(k) - c_x^q(f\mu k)^4(K_1^q(\mathbf{k}))^2P_{\mathrm{lin}}(k)\,, \end{aligned} \tag{63}$$

where $K_1^q(\mathbf{k}) \equiv (b_1^q + f\mu^2)$ and $K_{2,3}^q$ are the standard redshift space kernels for galaxies (see, e.g., [176]), and c_x is the next-to-leading order k^4 redshift-space counter term [127].

We fit for the bias and BAO scaling parameters by using a Gaussian likelihood to fit the one-loop Ly-α forest power spectrum, denoted by P^{model}, to the measured 2D power spectra, P_i^{data}. Therefore, we sample a χ^2 function

$$\chi^2 = \sum_i \frac{\left[P_i^{\mathrm{data}} - P^{\mathrm{model}}(k_i,\mu_i)\right]^2}{2\left(P_i^{\mathrm{data}}\right)^2/N_i}\,, \tag{64}$$

where N_i are the Fourier modes per bin, k_i the Fourier wavenumbers with the cosine of the angle to the line-of-sight, $\mu = k_\parallel/k$. The fits are done using five μ bins and a wavenumber spacing of $\Delta k = 0.003\,h\,\mathrm{Mpc}^{-1}$.

B. Ly-α forest mocks at fixed redshift

First, we take the best-fit transfer functions from the Sherwood simulations obtained from $L = 160\,h^{-1}\,\mathrm{Mpc}$ and $L = 80\,h^{-1}\,\mathrm{Mpc}$ boxes. To generate a large Ly-α forest mock, we use the readily available ICs from the Abacus simulations in a box of size $L = 2\,h^{-1}\,\mathrm{Gpc}$ with $N_{\mathrm{grid}} = 1024$, which is comparable to, yet slightly higher resolution than, current state-of-the-art simulations for cosmological analyses of the Ly-α forest in DESI ($\sim 2.5\,h^{-1}\,\mathrm{Mpc}$) [129].[11] The resulting mock is at redshift $z = 2.8$ and shown in Fig. 21. This shows a qualitative agreement between the power spectrum of the field-level simulation and the input high-resolution Sherwood simulation. For illustration purposes, we show for this case three angular wedges for which we find good qualitative agreement down to scales of $k \approx 0.7\,h\,\mathrm{Mpc}^{-1}$ between the power spectrum measured on the small-scale high-resolution simulation compared to the one measured from the field-level simulation.

---

[11] Note that the DESI Ly-α forest mocks require a volume of $V = 10^3(h^{-1}\,\mathrm{Gpc})^3$. We leave this engineering challenge to future work.

其中 f 是（线性）增长率。树级功率谱的第一个高阶修正是

$$P^{1-\mathrm{loop}}(k,\mu) = 2\int_{\mathbf{q}} \text{K2 2(q, k −q)Plin(|k −q|)Plin(q) + 6K1(k)Plin(k)} \int_{\mathbf{q}} \text{K3(k, −q, q)Plin(q) .} \tag{60}$$

使用更高阶的红移空间核 K2,3 [参见 136 式 (3.19)]，我们使用的符号为 $_{\mathbf{q}} \equiv \int \frac{d^3q}{(2\pi)^3}$ to denote
关于 q 的三重积分。对偶项

$$P^{\mathrm{ct}}(k,\mu) = -2(c_0 + c_2\mu^2 + c_4\mu^4)K_1(\mathbf{k})k^2P_{\mathrm{lin}}(k)\,, \tag{61}$$

按 k²P_lin(k) 缩放以及随机贡献

$$\text{Pst(k, μ) = Pshot + a0}\quad \frac{k^2}{{k_{\mathrm{NL}}}^2} + a_2\frac{k^2\mu^2}{{k_{\mathrm{NL}}}^2}\,, \tag{62}$$

将其视为一个恒定的散粒噪声部分以及一个随尺度和角度变化的项，用于描述小尺度聚类。我们建议读者参阅文献 [77, 78, 127] 以获得方法论的更完整介绍。

为了联合拟合Ly-α森林和高红移星系样本（此处：LBGs和LAEs），我们使用一环EFT模型，遵循文献[78, 127]中描述的程序。在线性理论下，这对应于计算两个树水平示踪器的几何平均，P_tree × (k, μ) = (b1 − bηfμ²)(bq₁ + fμ²)P_lin(k)，其中下标×表示交叉相关，上标q表示类星体（或晕）的示踪器，即bq₁是类星体的线性偏置参数。在实际操作中，我们通过重新映射偏置参数移除沿视线方向的依赖项（见，例如，[86, 135, 136]），这对于交叉相关得出如下形式

$$\begin{aligned} &\text{Pth. ×(k, μ) =K1(k)Kq 1(k)Plin(k) + 2} \int_{\mathbf{q}} K_2(\mathbf{q},\mathbf{k}-\mathbf{q})K_2^q(\mathbf{q},\mathbf{k}-\mathbf{q})P_{\mathrm{lin}}(|\mathbf{k}-\mathbf{q}|)P_{\mathrm{lin}}(q) \\ &+ 3P_{\mathrm{lin}}(k)\ {}_{\mathbf{q}}\ \text{[K1(k)Kq 3(k, −q, q) + Kq 1(k)K3(k, −q, q)]Plin(q) −(c0 + c2μ2 + c4μ4)Kq 1(k)k2Plin(k)} \\ &\text{−(cq 0+ cq 2μ2 + cq 4μ4)K1(k)k2Plin(k) −cq x(fμk)4(Kq} \\ &\text{1(k))2Plin(k) ,} \end{aligned} \tag{63}$$

其中 Kq 1(k) ≡(bq 1+ fμ2)，而 Kq 2,3 是星系的标准红移空间核（参见，例如，[176]），cx 是下一阶 k4 红移空间对撞项 [127]。我们通过使用高斯似然函数将一次环 Ly-α 森林功率谱 Pmodel 拟合到测量的二维功率谱 Pdata i，从而拟合偏差和 BAO 缩放参数。因此，我们对 χ2 函数进行采样

$$\chi^2 = \sum_i \frac{\left[P_i^{\mathrm{d}} \cdots - P^{\cdots}(k_i,\mu_i)\right]^2}{2\left(P_i^{\mathrm{data}}\right)^2/N_i}\,, \tag{64}$$

其中 Ni 是每个箱子的傅里叶模，ki 是傅里叶波数，带有与视线夹角的余弦 μ = k ∥ /k。拟合使用五个 μ 箱和波数间隔 Δk = 0.003 h Mpc−1 进行。

B. 固定红移下的Ly-α森林模拟

首先，我们从 Sherwood 模拟中获得的 L = 160 h−1Mpc 和 L = 80 h−1Mpc 箱体尺寸中取出最适合的传递函数。为了生成一个大的 Ly-α 森林模拟，我们使用 Abacus 模拟中现成的 ICs，在箱体尺寸为 L = 2 h−1Gpc、Ngrid = 1024 的情况下进行，该分辨率与当前用于 DESI 中 Ly-α 森林宇宙学分析的最先进模拟（约 2.5 h−1Mpc）相当，但略高。生成的模拟对应红移 z = 2.8，如图 21 所示。图中显示了场级模拟的功率谱与输入的高分辨率 Sherwood 模拟之间的定性一致性。为了说明，我们在此案例中展示了三个角向楔区，在这些区域中，我们发现在小尺度高分辨率模拟测得的功率谱与场级模拟测得的功率谱在 k ≈ 0.7 h Mpc−1 的尺度上有良好的定性一致性。

请注意，DESI Ly-α 林模拟需要体积为 V = 103(h−1 Gpc)³。我们将这一工程挑战留给未来的工作。

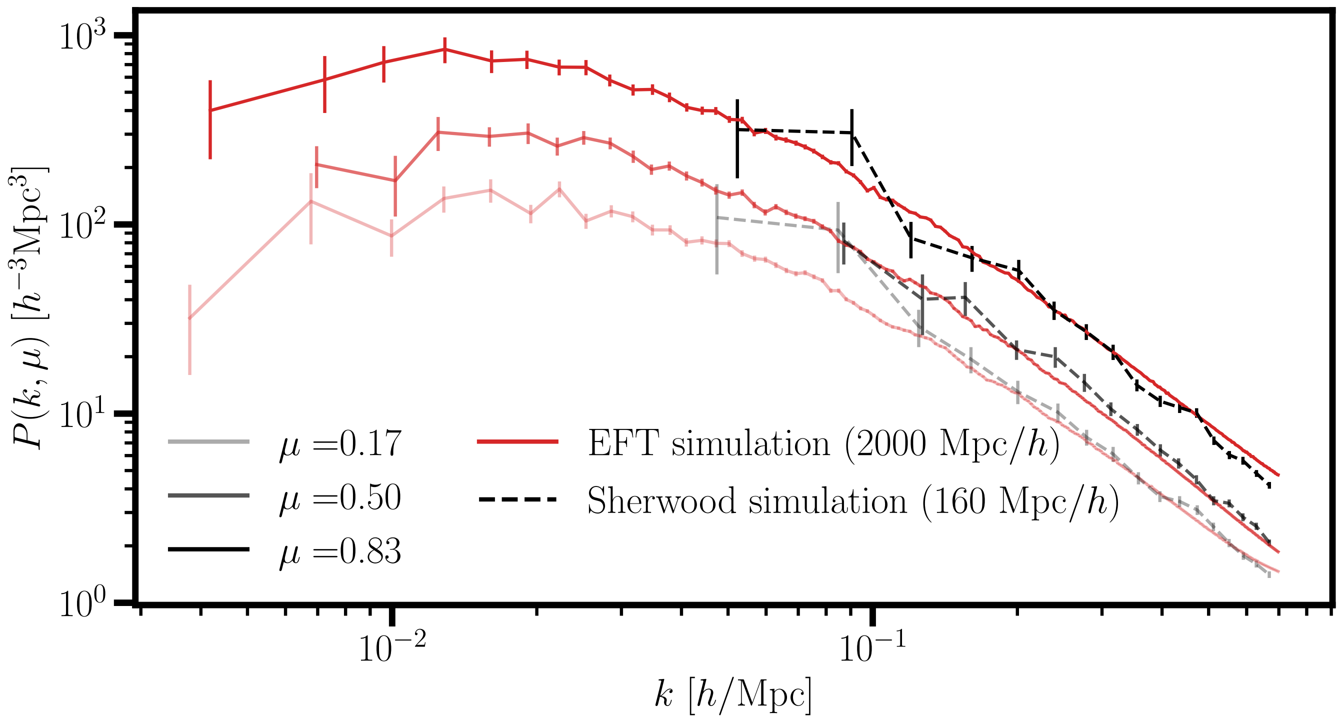

FIG. 21. **Large-scale Clustering Ly-α Mock:** Two-dimensional Power spectrum $P(k,\mu)$ in three angular bins of the measured power spectrum from the Sherwood hydrodynamic simulation (dashed black lines) compared to the measured power spectrum of the new Ly-α mock using the Sherwood transfer functions applied to the Abacus initial conditions. This demonstrates that the field-level methodology is able to generate large-scale clustering mocks with arbitrary box sizes (here: $L = 2\,h^{-1}\,\mathrm{Gpc}$) calibrated on small hydrodynamic simulations (Sherwood: $L = 160\,h^{-1}\,\mathrm{Mpc}$). The EFT simulation is generated at a resolution with $N_{\mathrm{cell}} = 1024$.

1. *Validation of Ly-α forest mocks at fixed redshift of $z = 2.8$*

We validate the field-level mock, shown in Fig. 21, by performing BAO fits to the measured power spectra, as detailed in Sec. A. In Fig. 22 we show the resulting best-fit power spectra using $k_{\max} = 0.5\,h\,\mathrm{Mpc}^{-1}$ (left panel) and the corresponding loop correction plot (right panel) for the large-scale clustering mock obtained from the transfer functions fitted to the $L = 160\,h^{-1}\,\mathrm{Mpc}$ in the top row and the $L = 80\,h^{-1}\,\mathrm{Mpc}$ in the bottom row. The model yields an excellent fit to the simulation at all scales (solid line) and linear bias (dashed lines) shows strong deviations on quasi-linear scales. The tree-level model deviation in the loop correction plot approximately matches our previous findings, i.e. beyond $k \approx 0.08\,h\,\mathrm{Mpc}^{-1}$ the corrections surpass the 5% threshold.

In Fig. 23, we present the corresponding drift plots to Fig. 22 obtained by fitting each box using two values of $k_{\max} = 0.3, 0.5\,h\,\mathrm{Mpc}^{-1}$ and obtain excellent agreement between the fits for the same realization. As expected, the one- and two-dimensional marginalized posteriors systematically shrink with increasing $k_{\max}$. The linear bias parameters b_1 and b_η values are fully consistent with our field-level fits discussed further in Sec. V and tabulated in Tab. V. We highlight two key results from these fits: First, we recover unbiased BAO scaling parameters at the 1σ level, demonstrating how these mocks can be used to validate cosmological fitting pipelines for DESI Ly-α forest analyses (see, e.g., [51, 61]). Second, the recovered bias parameters from the field-level mock are consistent with both, the field-level measurements (see Sec. VI) and the P3D measurements presented in Ref. [136]. The remaining differences with the latter are primarily driven by analysis choices: (i) we perform an anisotropic one-loop IR resummation instead of an isotropic tree-level treatment, and (ii) we fit for the BAO scaling parameters, which the $L = 160\,h^{-1}\,\mathrm{Mpc}$ box can only marginally resolve; and (iii) the fitted transfer functions $\beta(k,\mu)$ obtained from field-level analyses benefit from cosmic variance cancellation, and can therefore differ from P3D fits derived from small-volume simulations that are strongly affected by cosmic variance.

In the following we quantify the agreement between the recovered bias parameters from fits to the large-volume mock (Tab. VII) compared to the transfer function fits mapped onto the CLASS-PT basis (Tab. V). The well-constrained linear and velocity-bias parameters, b_1 and b_η, agree at the $\lesssim 1.7\sigma$ level, the only mildly elevated values arising for b_1

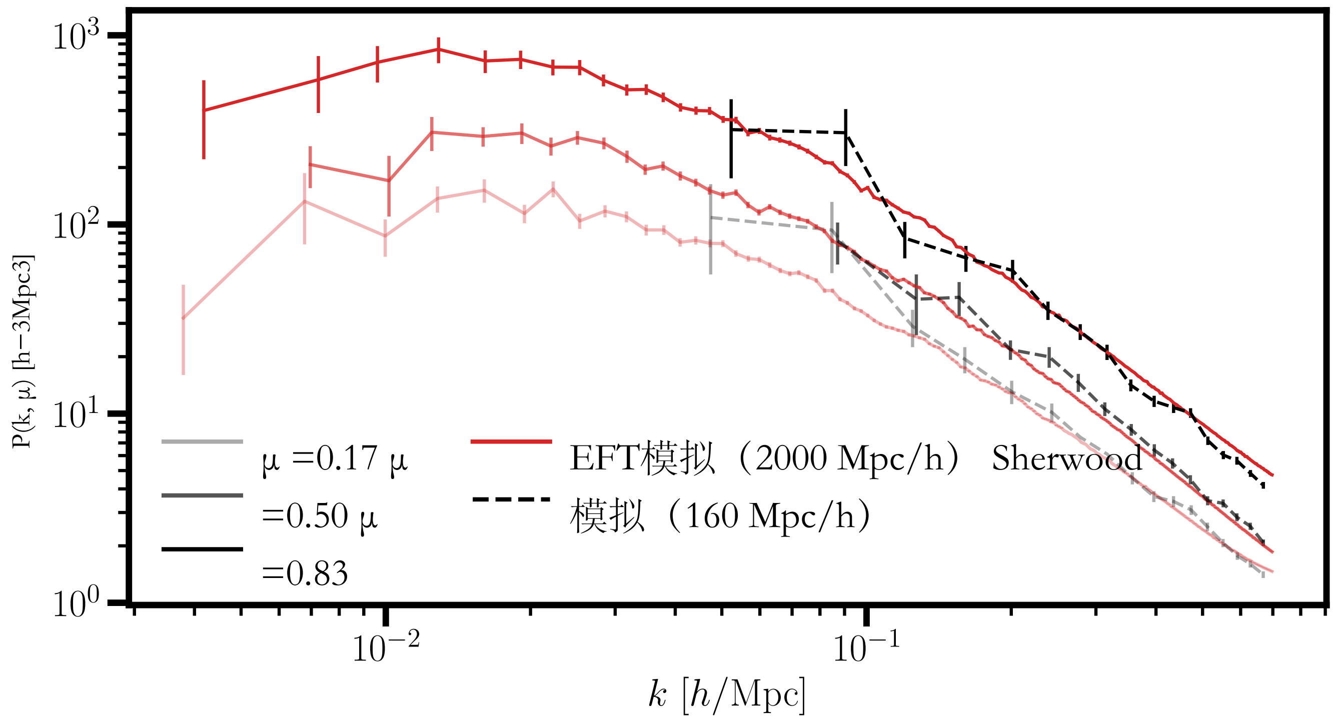

图21. 大规模Ly-α聚类模拟：二维功率谱P(k, μ)在三个角度区间的测量功率谱（来自Sherwood流体动力学模拟，黑色虚线）与使用Sherwood传输函数应用于Abacus初始条件的新Ly-α模拟测量功率谱的比较。这表明，场级方法能够生成任意盒子大小（此处：L = 2 h−1Gpc）的大规模聚类模拟，并以小型流体动力学模拟（Sherwood：L = 160 h−1Mpc）为标定。EFT模拟在分辨率为Ncell = 1024下生成。

1. 在固定红移 z = 2.8 下对 Ly-α 森林模拟的验证

我们通过对测量的功率谱进行BAO拟合来验证图21中所示的场级模拟，如附录A节所述。在图22中，我们展示了利用kmax = 0.5 h Mpc−1（左图）得到的最佳拟合功率谱，以及相应的环修正图（右图），这些是从顶部行L = 160 h −1Mpc和底部行L = 80 h−1Mpc的转移函数拟合得到的大尺度聚类模拟。模型在所有尺度上对模拟都给出了极好的拟合（实线），而线性偏置（虚线）在准线性尺度上显示出强偏离。环修正图中的树级模型偏离大致与我们之前的发现一致，即在k ≈ 0.08 h Mpc−1之后，修正超过了5%的阈值。在图23中，我们展示了对应图22的漂移图，这些图是通过使用两个kmax值（0.3, 0.5 h Mpc−1）拟合每个盒子获得的，并且在相同实现的拟合之间获得了极好的一致性。正如预期的那样，随着 kmax 的增加，一维和二维边缘后验分布系统性地收缩。线性偏差参数 b1 和 bη 的数值与我们在第五节中进一步讨论并在表 V 中列出的场级拟合结果完全一致。我们从这些拟合中强调两个关键结果：第一，我们在 1σ 范围内恢复了无偏的 BAO 尺度参数，展示了这些模拟如何用于验证 DESI Ly-α 林分析的宇宙学拟合流程（见例如，[51, 61]）。第二，从场级模拟中恢复的偏差参数与场级测量（见第六节）以及参考文献 [136] 中给出的 P3D 测量结果一致。与后者的其余差异主要由分析选择驱动：（i）我们执行各向异性的单环红外重求和，而不是各向同性的树级处理；（ii）我们拟合BAO缩放参数，而L = 160 h−1Mpc的盒子只能勉强分辨；（iii）从场级分析获得的拟合传递函数β(k, μ)受益于宇宙方差消除，因此可能与从强烈受到宇宙方差影响的小体积模拟中得出的P3D拟合不同。

在下文中，我们量化了从大体积模拟拟合中恢复的偏差参数（表 VII）与映射到 CLASS-PT 基础上的传递函数拟合（表 V）之间的一致性。约束良好的线性和速度偏差参数 b1 和 bη 在 ≲1.7σ 的水平上是一致的，唯一稍高的值出现在 b1 上。

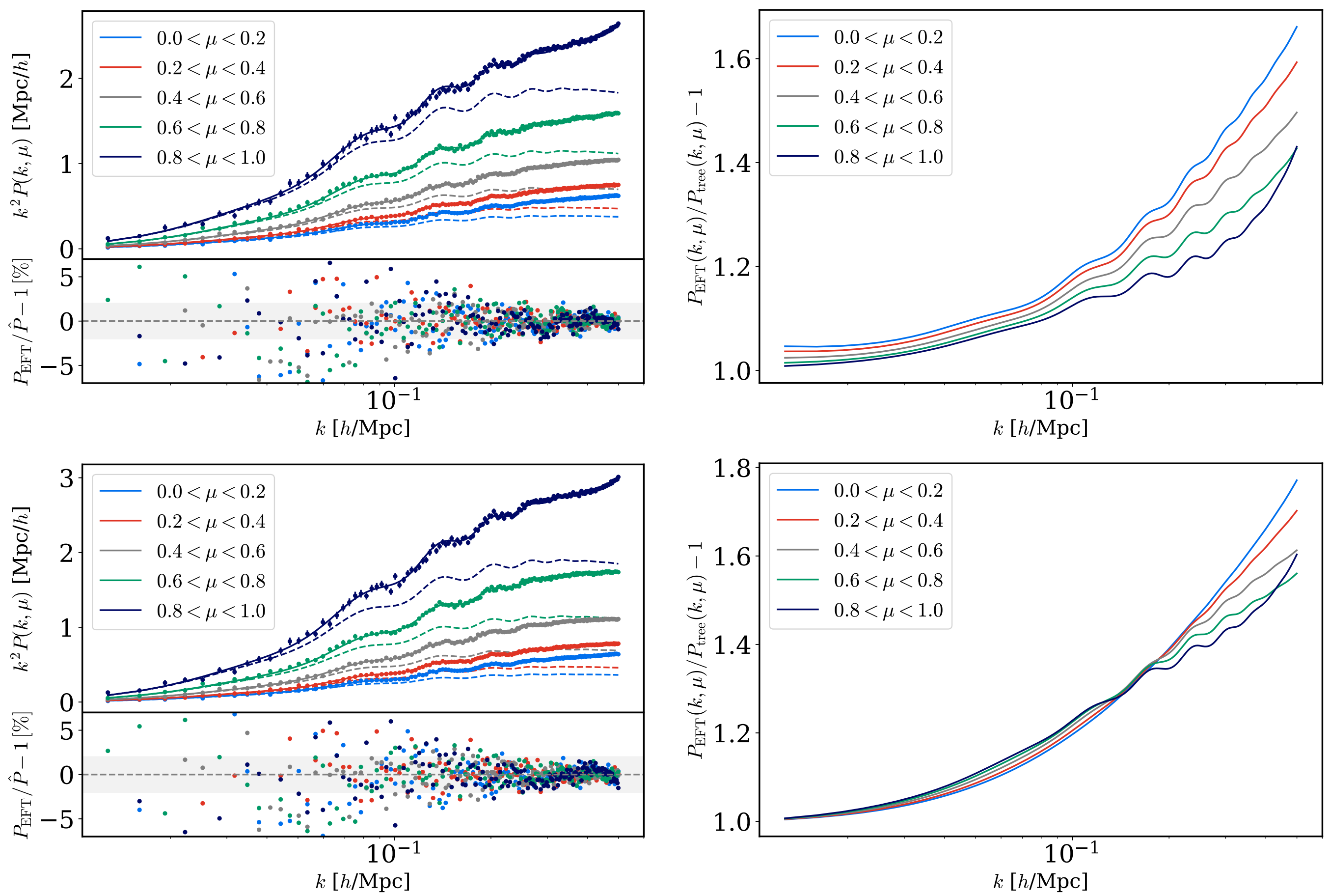

FIG. 22. **Validation of Ly-α Mocks:** EFT fits to the Ly-α auto-correlation (*left panel*) and corresponding loop-correction (*right panel*) shown in five angular bins, μ, with $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$ for a single realization. The top (bottom) panel uses the $L = 160\,(80)\ h^{-1}\,$Mpc box as input. The best-fit EFT model (solid line) is compared to $P^{\rm tree}$ (dashed line) and the data points show error bars which are obtained assuming a diagonal Gaussian covariance based on the number of expected Fourier modes per bin $P(k,\mu)\sqrt{2/N(k,\mu)}$. The bottom panel shows the ratio between model and data of which one is subtracted in addition to a gray band indicating the 2% error band to guide the eye.

at $L = 160\ h^{-1}\,$Mpc (1.4–1.7 σ). The quadratic bias b_2 is recovered within 1σ in every case. We thus find no significant tension between the two analyses. Note that we recover unbiased BAO parameters and that the model captures the shift *and* smearing of the BAO peak through the one-loop corrections and the IR resummation, respectively.

2. *Ly-α clustering mocks at $z_{\rm eff} = 2.33$*

Similarly, we perform a BAO fit of the Ly-α forest EFT mocks at $z = 2.33$, the effective redshift of the Ly-α forest sample of DESI DR1 [51]. The best-fit spectra are shown in Fig. 24 and the marginalized EFT nuisance and BAO parameters are tabulated in the last two columns of Tab. VII. We highlight two key findings for the $z = 2.33$ mock. First, we find excellent agreement between our linear bias parameters and the ones measured by DESI DR1 [51]. Second, We find unbiased constraints on the BAO scaling parameters:

$$\alpha_\parallel = 1.011 \pm 0.018, \qquad \alpha_\perp = 1.0021^{+0.0060}_{-0.0080}\,. \tag{65}$$

The resulting best-fit spectra are shown in Fig. 24. Importantly, the loop corrections are exceeding 20% at quasi linear scales of $k \approx 0.1\,h\,{\rm Mpc}^{-1}$ emphasizing the importance of higher order bias parameters.

The resulting 1D and 2D marginalized posteriors for the sampled parameters are shown in Fig. 25.

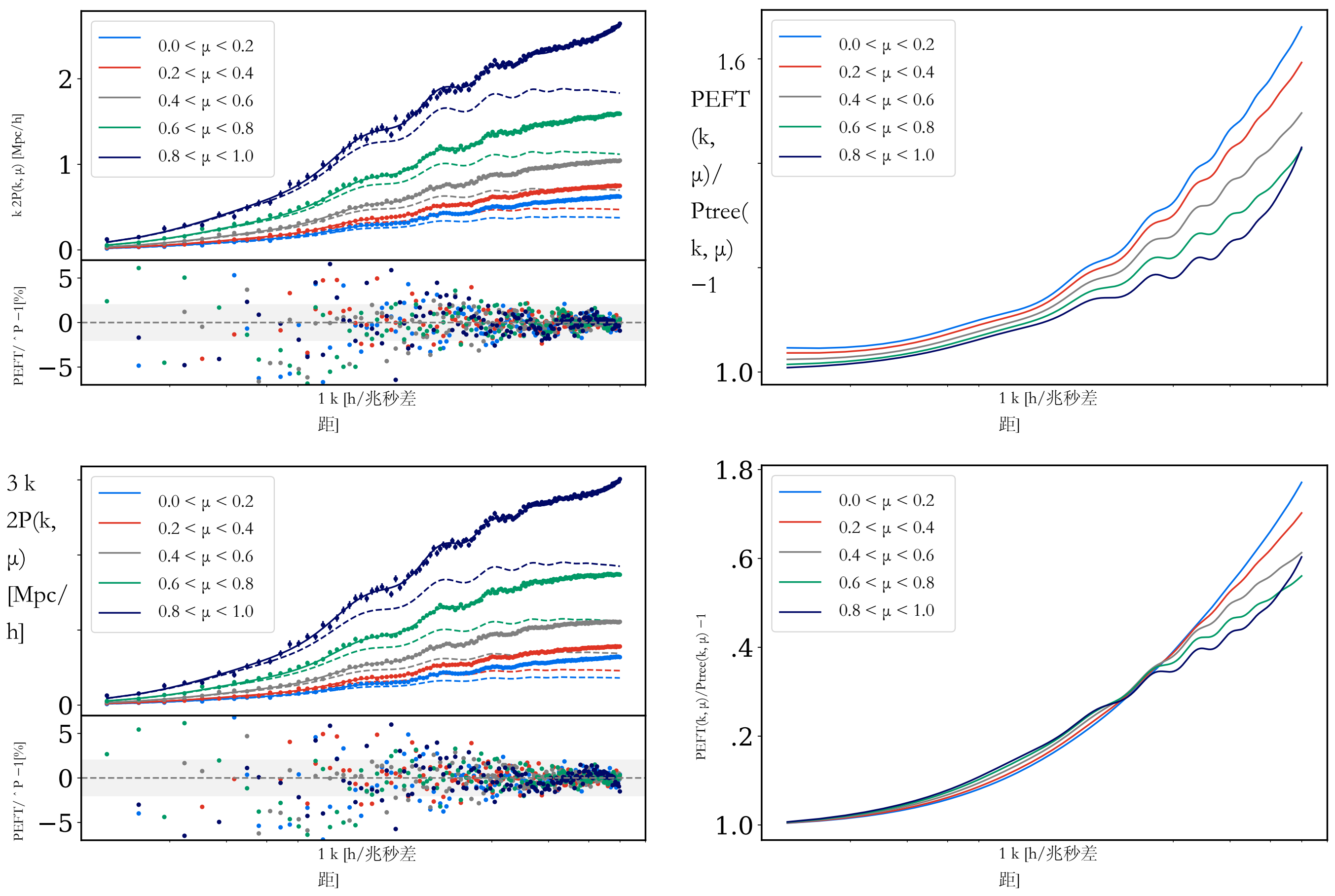

图 22. Ly-α 模拟验证：EFT 对 Ly-α 自相关（左图）及相应环修正（右图）的拟合，分为五个角度区间 μ，$k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$，针对单次实现。顶部（底部）图使用 $L = 160$（80）h^{-1}Mpc 的盒子作为输入。最佳拟合的 EFT 模型（实线）与 $P_{\rm tree}$（虚线）比较，数据点显示误差条，误差条假定基于每个区间的预期傅里叶模式数量的对角高斯协方差 $P(k, \mu)\,p^2/N(k, \mu)$ 得出。底部图显示模型与数据的比值，其中一个被减去，并附有灰色带表示 2% 的误差带以便于观察。

在 $L = 160\,h^{-1}$Mpc（1.4 – 1.7 σ）下。二次偏差 b_2 在每种情况下都在 1σ 范围内恢复。因此，我们没有发现两种分析之间存在显著的矛盾。请注意，我们恢复了无偏的 BAO 参数，并且模型分别通过单环修正和红外重求和捕捉到了 BAO 峰的位移和展宽。

2. 在有效红移 $z_{\rm eff} = 2.33$ 的 Ly-α 聚类模拟

同样地，我们对 $z = 2.33$ 的 Ly-α 森林 EFT 模拟进行了 BAO 拟合，这也是 DESI DR1 Ly-α 森林样本的有效红移 [51]。最佳拟合谱如图 24 所示，边缘化的 EFT 干扰项和 BAO 参数列在表 VII 的最后两列。我们强调对于 $z = 2.33$ 模拟的两个关键发现。首先，我们发现我们的线性偏置参数与 DESI DR1 测得的参数 [51] 非常一致。其次，我们发现对 BAO 缩放参数的约束是无偏的：

$$\alpha_\parallel = 1.011 \pm 0.018, \qquad \alpha_\perp = 1.0021^{+0.0060}_{-0.0080}\,. \tag{65}$$

得到的最佳拟合谱如图24所示。重要的是，在准线性尺度 $k \approx 0.1\,h\,{\rm Mpc}^{-1}$ 时，环路修正超过20%，强调了高阶偏差参数的重要性。得到的取样参数的一维和二维边际后验如图25所示。

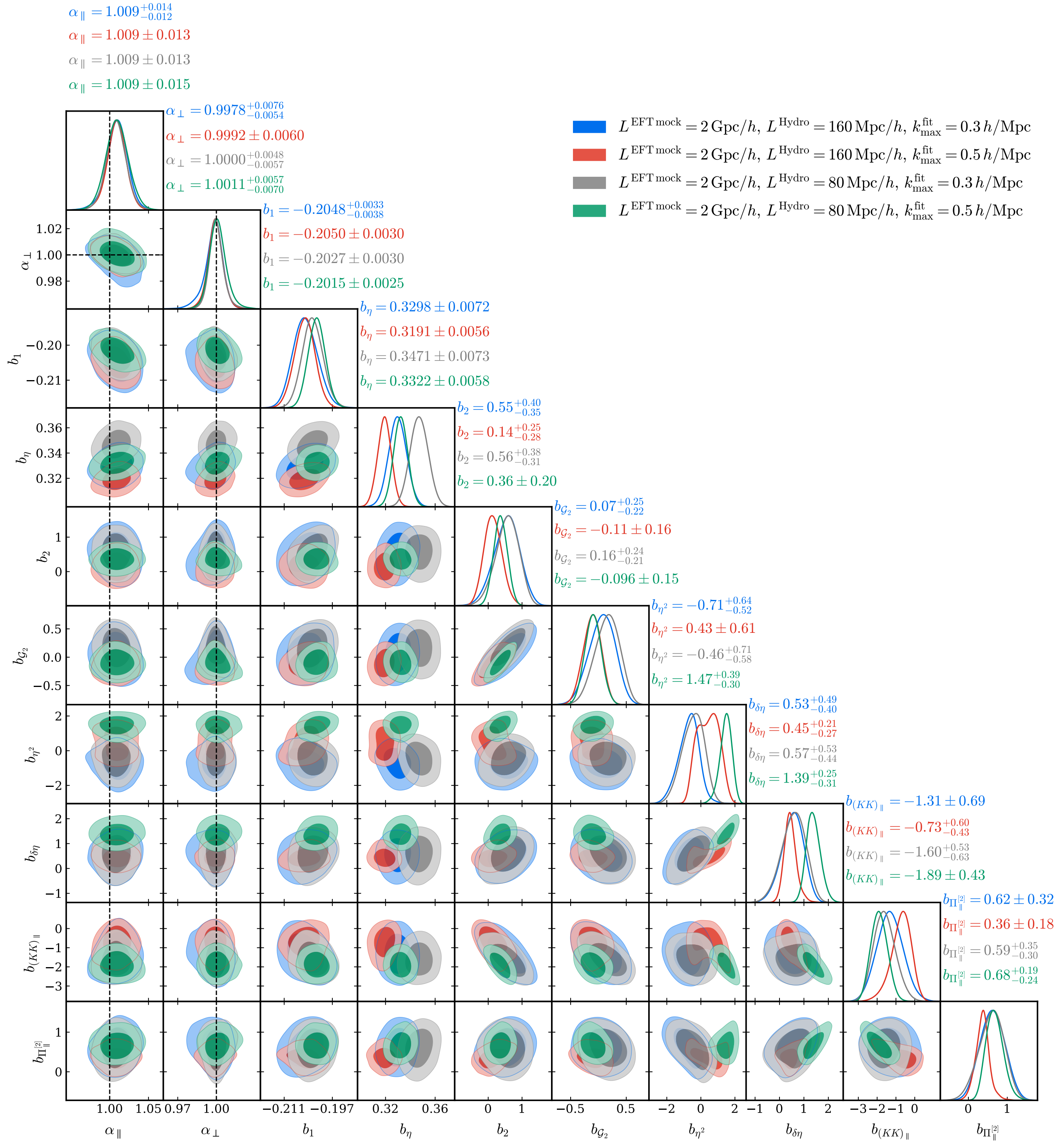

FIG. 23. **Drift plot of Ly-α Mock at** $z = 2.8$**:** Best-fit 1D and 2D marginalized posteriors for the two Sherwood simulations with different box sizes and for two different maximum wavenumbers $k_{\rm max} = 0.3,\,0.5\,h\,{\rm Mpc}^{-1}$ used in the fits centered at $z = 2.8$. We find consistent results within each set of simulations. Additionally, we recover unbiased BAO scaling parameters that are consistent with the fiducial values at the 1σ-level – the key parameter of interest for DESI analyses. Note that the box sizes denote the Sherwood simulations that have been used to perform the field-level fits of the transfer functions and that the BAO fits have been performed on power spectra measured from boxes of length $L = 2\,h^{-1}$ Gpc.

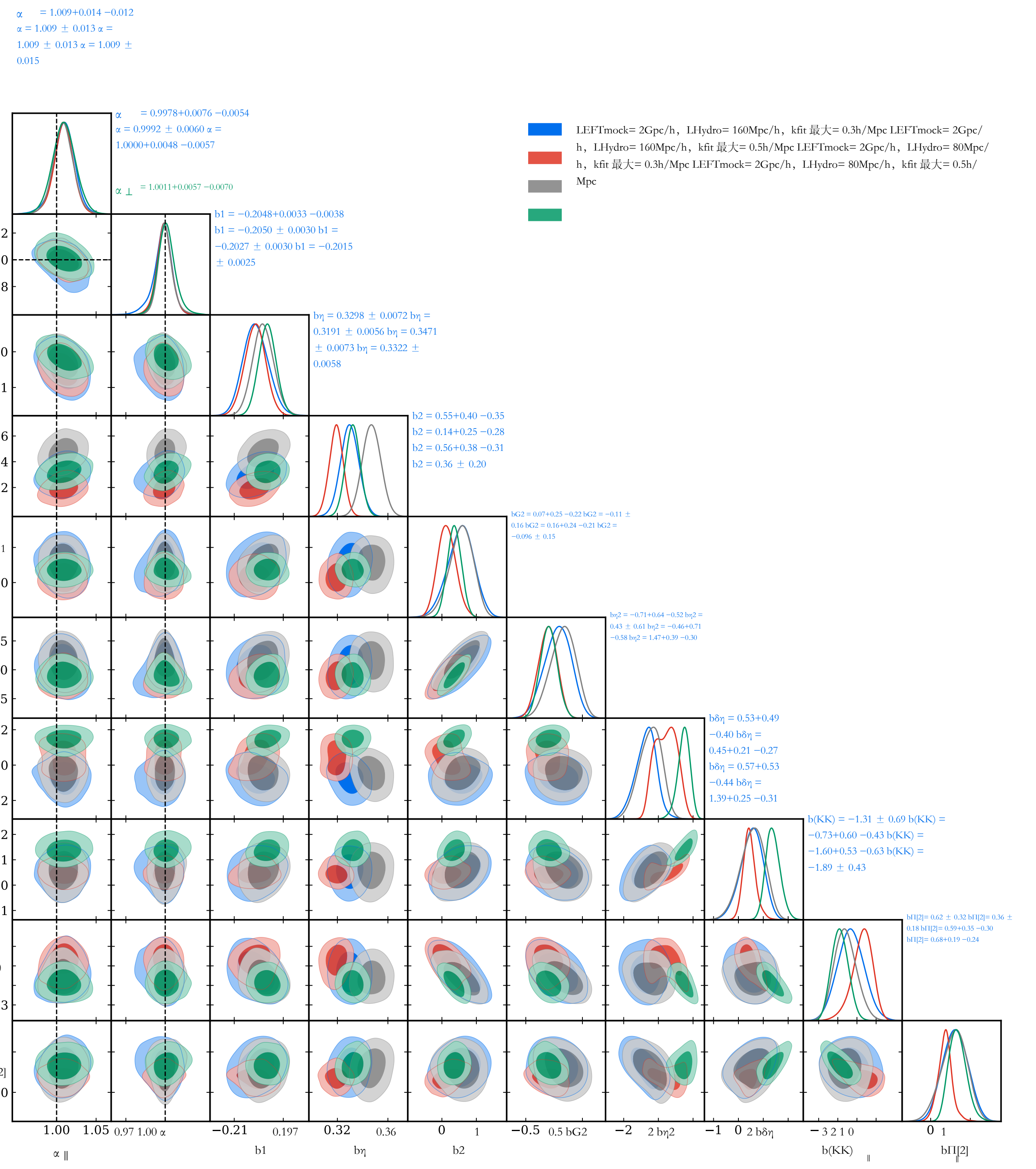

图23. 在 z = 2.8 时 Ly-α 模拟的漂移图：针对两个不同盒子尺寸的 Sherwood 模拟，以及用于拟合的两个不同最大波数 kmax = 0.3、0.5 h Mpc−1 的最佳拟合一维和二维边际后验分布，拟合中心为 z = 2.8。我们发现在每组模拟中结果一致。此外，我们恢复了无偏的 BAO 缩放参数，其与基准值在 1σ 水平上保持一致——这是 DESI 分析关注的关键参数。请注意，盒子尺寸表示用于执行传递函数场级拟合的 Sherwood 模拟，而 BAO 拟合是在长度为 L = 2 h−1Gpc 的盒子中测量的功率谱上进行的。

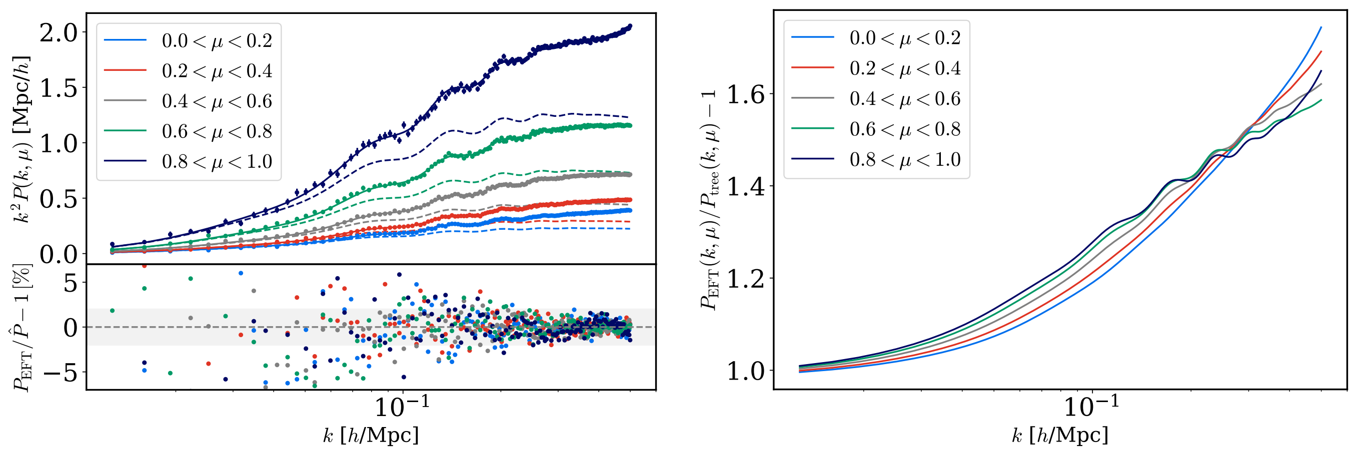

FIG. 24. **Validation of Ly-α Mocks at DESI redshift:** Same as Fig. 22 but at the effective redshift of DESI for the Ly-α forest using a $k_{\rm max} = 0.5\,h\,{\rm Mpc}^{-1}$ for the fits.

C. Cross-correlations between the Ly-α forest and high-redshift galaxies

A key source of cosmological information in Ly-α forest analyses stems from cross-correlating the Ly-α forest with quasar positions (see, e.g., [51, 52, 61, 126]). Near-future surveys such as DESI-II will measure large samples of high-redshift galaxies such as Lyman-break galaxies (LBGs) and Ly-α emitters (LAEs). Yet, their cosmological analyses in combination with the Ly-α forest is barely explored [177]. Here, we create large-scale clustering mocks of high-redshift galaxy samples, introduced in Sec. III. To this end, we use the transfer functions shown in Fig. 17 with a scale cut of $k_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$ which is driven by the scale and angular dependence of the error power spectrum [154]. We extend the previous section to create large-scale clustering mocks for LAEs and LBGs and perform full-shape and BAO fits for the cross-correlation of the Ly-α forest with mocks of these galaxy samples.

In Figs. 26-27 we show the resulting best-fit spectra of the joint power spectrum fits to the Ly-α forest and the high-redshift galaxy samples (LAEs and LBGs).[12] We show the best-fit spectra in the top row with the corresponding residuals and the loop corrections in the bottom row of each figure for the LBG S5 and the LAE S5 samples. The joint fits use $k^{\rm FF}_{\rm max} = 0.40\,h\,{\rm Mpc}^{-1}$, $k^{\rm Fg}_{\rm max} = 0.30\,h\,{\rm Mpc}^{-1}$, and $k^{\rm gg}_{\rm max} = 0.20\,h\,{\rm Mpc}^{-1}$. Following baseline expectation, the loop corrections vanish on large scales and exceed the 5% threshold beyond $k = 0.1\,h\,{\rm Mpc}^{-1}$ when including the point tracers. We find good agreement between the best-fit parameters obtained from the joint fits given in Tab. VIII to the ones summarized in table III in [154]. In addition to this, the BAO fits confirm, first, the usefulness of the presented mocks as we find unbiased constraints on the BAO scaling parameters from the joint fits of the Ly-α forest and high-redshift galaxies at the 1σ level which can be used for inference pipeline validation. Second, the EFT model is applicable for cross-correlations with high-z galaxies, even for highly-biased samples such as LBGs.

In this section we have demonstrated a new procedure to generate large-scale clustering mocks that can be used to validate cosmological inference pipelines. Note that the generated perturbative mocks are at fixed redshifts. To put these simulations on a light cone one would evolve all the fields and the gravitational potential with redshift and place an observer at, e.g., the center of the cartesian grid. Since the $160\,h^{-1}\,{\rm Mpc}$ Sherwood simulation only has a single redshift, we would again use the $L = 80\,h^{-1}\,{\rm Mpc}$ suite of simulations and fit to each snapshot the transfer functions at the field-level, see App. A. We would interpolate the resulting polynomial fits to the desired redshift allowing us to compute a realization of the Ly-α field at arbitrary redshift (within the redshift coverage of the snapshots) similar to what we have done in Sec. C. We leave the construction of light cone mocks to future work.

The galaxy bias parameters recovered from our joint fits agree well with the field-level measurements of [154] (their Table III), who fit the same LBG and LAE selections in the Astrid hydrodynamical simulations at $z = 3$. For the linear bias we find $b_1^{\rm q} \simeq 3.8$–4.1 (LBGs) and $b_1^{\rm q} \simeq 2.0$–2.1 (LAEs), reproducing the expected LBG/LAE hierarchy and lying within or just below their Astrid range ($b_1 \simeq 3.9$–4.2 and 1.9–2.7, respectively); our ODIN value $b_1^{\rm q} = 1.97$ is moreover consistent with the observed $b_1^{\rm q} \approx 2.0$. The tidal bias is similarly consistent: we obtain order-unity, negative values, $b^{\rm q}_{\mathcal{G}_2} \in [-1.44, -0.71]$, in agreement with [154] given the large recovered uncertainties on these parameters. The quadratic bias shows the largest offset: for the LBG samples we measure $b_2^{\rm q} \simeq 10$–12, about twice the reported $\simeq 6$.

---

[12] We leave the inclusion of the cross-covariance between the two tracers to future work.

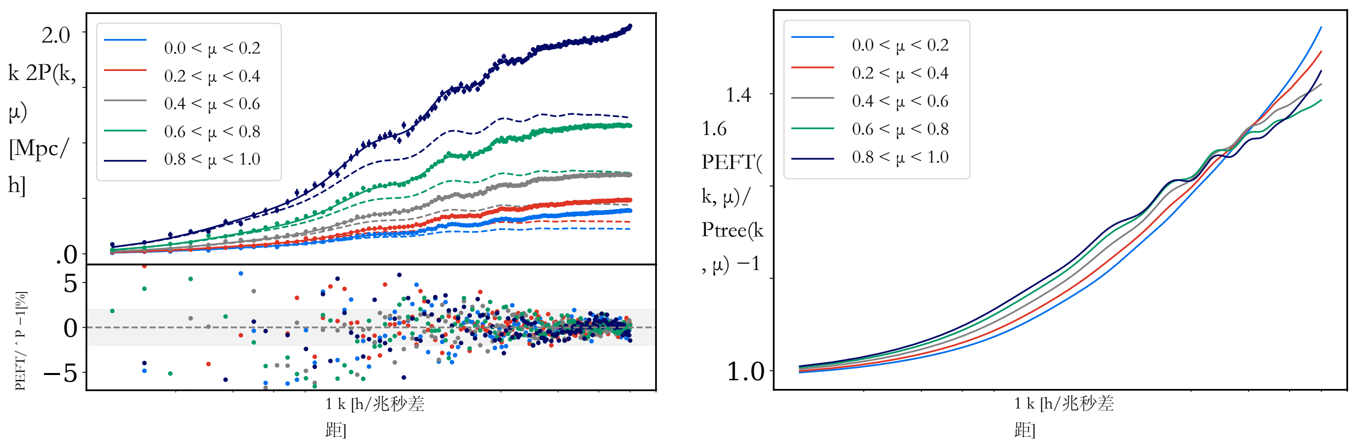

图24. 在DESI红移下对Ly-α模拟的验证：与图22相同，但在DESI对Ly-α森林的有效红移下使用kmax = 0.5 h Mpc−1进行拟合。

C. Ly-α 森林与高红移星系之间的互相关

在 Ly-α 森林分析中，一个关键的宇宙学信息来源来自将 Ly-α 森林与类星体位置进行交叉相关（参见，例如 [51, 52, 61, 126]）。近期的调查如 DESI-II 将测量大量高红移星系样本，如 Lyman-断点星系（LBGs）和 Ly-α 发射星系（LAEs）。然而，将这些星系与 Ly-α 森林结合进行的宇宙学分析几乎没有被探索 [177]。在这里，我们创建了高红移星系样本的大尺度聚类模拟，如第 III 节所介绍。为此，我们使用图 17 中显示的传递函数，并设定尺度截止 kmax = 0.3 h Mpc−1，这一选择受误差功率谱的尺度和角度依赖性驱动 [154]。我们扩展上一节的内容，为 LAEs 和 LBGs 创建大尺度聚类模拟，并对 Ly-α 森林与这些星系样本模拟的交叉相关进行全形状和 BAO 拟合。

在图26-27中，我们展示了对Ly-α森林和高红移星系样本（LAEs和LBGs）进行联合功率谱拟合后得到的最佳拟合谱。我们在每个图的上方展示了最佳拟合谱，在下方展示了相应的残差和环路修正，用于LBG S5和LAE S5样本。联合拟合使用kFF max = 0.40 h Mpc−1, kFg max = 0.30 h Mpc−1, 和 kgg max = 0.20 h Mpc−1。按照基线预期，环路修正将在大尺度上消失，并在包括点追踪器时超过k = 0.1 h Mpc−1之后的5%阈值。我们发现从联合拟合得到的最佳拟合参数（见表VIII）与[154]中表III总结的参数之间具有良好的一致性。除此之外，BAO 拟合结果首先确认了所展示的模拟的实用性，因为我们从 Ly-α 森林和高红移星系的联合拟合中，在 1σ 水平上得到了对 BAO 缩放参数的无偏约束，这可以用于推断流程的验证。其次，EFT 模型适用于与高红移星系的交叉相关，即使对于高度偏倚的样本如 LBGs 也是适用的。

在本节中，我们演示了一种生成大规模聚类模拟的新方法，可用于验证宇宙学推断流程。请注意，生成的微扰模拟是在固定红移下的。要将这些模拟放在光锥上，需要随着红移演化所有场和引力势，并在例如笛卡尔网格的中心放置一个观测者。由于 160 h−1Mpc 的 Sherwood 模拟只有单一红移，我们将再次使用 L = 80 h−1Mpc 的模拟套件，并对每个快照在场级别上拟合传递函数，见附录 A。我们会将得到的多项式拟合插值到所需红移，从而使我们能够计算在任意红移下（在快照的红移覆盖范围内）Ly-α 场的实现，类似于我们在 C 节所做的工作。光锥模拟的构建留作未来工作。

从我们的联合拟合中恢复的星系偏差参数与 [154]（其表 III）的场级测量结果非常一致，他们在 z = 3 的 Astrid 水力学模拟中拟合了相同的 LBG 和 LAE 选择。对于线性偏差，我们得到 bq1≃3.8 − 4.1（LBG）和 bq1≃2.0 − 2.1（LAE），重现了预期的 LBG/LAE 层次，并位于或略低于他们的 Astrid 范围（分别为 b1 ≃3.9 − 4.2 和 1.9 − 2.7）；我们的 ODIN 值 bq1= 1.97 也与观测到的 bq1≈2.0 一致。潮汐偏差同样一致：我们得到数量级为 1 的负值 bqG2∈[−1.44, −0.71]，考虑到这些参数复原时的不确定性较大，这与 [154] 相符。二次偏差显示出最大的偏移：对于 LBG 样本，我们测得 bq2≃10 − 12，大约是报告值≃6 的两倍。

---

[12] 我们将两个示踪器之间的交叉协方差的包含留待未来的工作。

Redshift	$z = 2.8$				$z = 2.33$	
Sherwood	$L = 160\ h^{-1}\,\mathrm{Mpc}$		$L = 80\ h^{-1}\,\mathrm{Mpc}$		$L = 80\ h^{-1}\,\mathrm{Mpc}$	
$k_{\max}^{\mathrm{fit}}$	0.3	0.5	0.3	0.5	0.3	0.5
Sampled parameters						
b_1	$-0.2048^{+0.0033}_{-0.0038}$	-0.2050 ± 0.0030	-0.2027 ± 0.0030	-0.2015 ± 0.0025	-0.1399 ± 0.0023	-0.1396 ± 0.0019
b_η	0.3298 ± 0.0072	0.3191 ± 0.0056	0.3471 ± 0.0073	0.3322 ± 0.0058	0.2589 ± 0.0059	0.2472 ± 0.0045
b_2	$0.55^{+0.40}_{-0.35}$	$0.14^{+0.25}_{-0.28}$	$0.56^{+0.38}_{-0.31}$	0.36 ± 0.20	$0.32^{+0.28}_{-0.24}$	0.28 ± 0.13
$b_{\mathcal{G}_2}$	$0.07^{+0.25}_{-0.22}$	-0.11 ± 0.16	$0.16^{+0.24}_{-0.21}$	-0.096 ± 0.15	$0.09^{+0.19}_{-0.15}$	-0.06 ± 0.10
b_{η^2}	$-0.71^{+0.64}_{-0.52}$	0.43 ± 0.61	$-0.46^{+0.71}_{-0.58}$	$1.47^{+0.39}_{-0.30}$	-0.33 ± 0.48	$1.01^{+0.24}_{-0.19}$
$b_{\delta\eta}$	$0.53^{+0.49}_{-0.40}$	$0.45^{+0.21}_{-0.27}$	$0.57^{+0.53}_{-0.44}$	$1.39^{+0.25}_{-0.31}$	$0.40^{+0.42}_{-0.38}$	$1.03^{+0.17}_{-0.21}$
$b_{(KK)_\parallel}$	-1.31 ± 0.69	$-0.73^{+0.60}_{-0.43}$	$-1.60^{+0.53}_{-0.63}$	-1.89 ± 0.43	$-1.10^{+0.39}_{-0.51}$	$-1.34^{+0.27}_{-0.34}$
$b_{\Pi^{[2]}_\parallel}$	0.62 ± 0.32	0.36 ± 0.18	$0.59^{+0.35}_{-0.30}$	$0.68^{+0.19}_{-0.24}$	$0.45^{+0.26}_{-0.20}$	$0.43^{+0.15}_{-0.19}$
$\alpha_\parallel$	$1.009^{+0.014}_{-0.012}$	1.009 ± 0.013	1.009 ± 0.013	1.009 ± 0.015	1.010 ± 0.014	1.011 ± 0.018
$\alpha_\perp$	$0.9978^{+0.0076}_{-0.0054}$	0.9992 ± 0.0060	$1.0000^{+0.0048}_{-0.0057}$	$1.0011^{+0.0057}_{-0.0070}$	$1.0001^{+0.0046}_{-0.0057}$	$1.0021^{+0.0060}_{-0.0080}$
Analytically marginalized						
$b_{\Pi^{[3]}_\parallel}$	1.3854 ± 0.3592	2.2579 ± 0.1161	2.3390 ± 0.3818	3.9789 ± 0.1261	1.6803 ± 0.2466	2.7891 ± 0.0766
$b_{\delta\Pi^{[2]}_\parallel}$	-0.2006 ± 0.8437	-0.1935 ± 0.4518	-0.7482 ± 0.8638	-1.5778 ± 0.4703	-1.1724 ± 0.7020	-0.7901 ± 0.3159
$b_{(K\Pi^{[2]})_\parallel}$	-1.2460 ± 0.8426	-3.3381 ± 0.3855	-2.9227 ± 0.8766	-5.5891 ± 0.3947	-2.4507 ± 0.5353	-3.4554 ± 0.2444
$b_{\eta\Pi^{[2]}_\parallel}$	-0.4614 ± 1.7829	-0.4338 ± 1.2428	-0.4876 ± 1.8040	-0.5198 ± 1.2934	-0.4975 ± 1.7575	0.7903 ± 0.9370
b_{Γ_3}	0.5259 ± 0.3569	0.2269 ± 0.2086	0.5012 ± 0.3592	0.6851 ± 0.2144	0.6367 ± 0.2737	0.4511 ± 0.1472
P_{shot}	-0.7920 ± 0.5576	-0.9231 ± 0.2003	-2.8145 ± 0.5607	-2.4783 ± 0.1957	-2.8691 ± 0.3647	-2.2519 ± 0.1173
a_0	0.7054 ± 5.0299	12.2297 ± 3.4545	0.9571 ± 4.9795	16.1193 ± 3.4851	1.2038 ± 4.9625	17.7688 ± 2.6652
a_2	0.2653 ± 5.0186	2.7929 ± 4.8289	0.0609 ± 4.9918	3.6760 ± 4.8137	-0.0266 ± 5.0211	1.9947 ± 4.7306
c_0	0.3195 ± 0.2075	0.1045 ± 0.0756	0.7356 ± 0.2084	0.6500 ± 0.0754	0.9325 ± 0.1623	0.6460 ± 0.0535
c_2	0.0144 ± 0.4208	-0.0773 ± 0.1190	-0.8157 ± 0.4414	-1.3573 ± 0.1221	-1.1771 ± 0.3688	-1.3204 ± 0.1009
c_4	-0.1286 ± 0.1724	1.0108 ± 0.0797	0.3731 ± 0.1792	2.3128 ± 0.0826	0.3669 ± 0.1878	1.9862 ± 0.0731
χ^2_{marg}	341.2	590.7	341.8	581.0	338.1	592.8

TABLE VII. **Ly-α forest auto-correlation fits:** Marginalized best-fit EFT parameters obtained from fits to the Ly-α forest auto-power spectrum of large-volume ($L = 2{,}000\ h^{-1}\,\mathrm{Mpc}$) mocks calibrated on Sherwood hydrodynamic simulations, for hydrodynamic box sizes $L = 160$ and $80\ h^{-1}\,\mathrm{Mpc}$ at redshifts $z = 2.8$ (*first four columns*) and interpolated to $z = 2.33$ (*last two columns*). The top section lists the sampled parameters; the bottom section lists the parameters that we analytically marginalize over and recover *a posteriori* from the chains. The counterterms c_0, c_2, c_4 are quoted in units of $[h^{-1}\mathrm{Mpc}]^2$ and $k_{\max}$ is in $h\,\mathrm{Mpc}^{-1}$. The $k_{\max} = 0.3\,(0.5)\ h\,\mathrm{Mpc}^{-1}$ fits use 460 (780) data points. The linear bias parameters are consistent at the 1–2σ level with fits from the three-dimensional power spectrum [136].

Note that the shotnoise value for the galaxies is very small as we already subtract it prior to performing the fits.

VIII. SUMMARY AND CONCLUSIONS

The Ly-α forest is a powerful tracer of the large-scale structure of our Universe. It probes a cosmological volume that remains inaccessible to galaxy surveys until Stage-V spectroscopy [131], thereby extending the reach of perturbative methods. Owing to its higher redshift range ($2 \leq z \leq 5$), the Ly-α forest also exhibits rotated degeneracy directions among cosmological parameters. A key advantage of the Ly-α forest is the well-understood physics relating the neutral hydrogen distribution to the underlying dark matter, which has enabled accurate hydrodynamical simulations down to sub-kpc scales [66, 68, 73, 178, 179]. These simulations are commonly used to calibrate linear-theory-based models augmented by phenomenological fitting functions for the power spectrum (or two-point correlation function) down to scales of a few $h\,\mathrm{Mpc}^{-1}$ [73–75, 148], resulting in remarkably robust cosmological constraints (see, e.g., [51, 56, 58]).

红移	z = 2.8				z = 2.33	
舍伍德	L = 160 h−1兆秒差距		L = 80 h−1兆秒差距		L = 80 h−1兆秒差距	
$k_{\max}^{\mathrm{fit}}$	0.3	0.5	0.3	0.5	0.3	0.5
采样参数 b1 −0.2048+0.0033						
−0.0038		−0.2050 ± 0.0030	−0.2027 ± 0.0030	−0.2015 ± 0.0025	−0.1399 ± 0.0023	−0.1396 ± 0.0019
η	0.3298 ± 0.0072	0.3191 ± 0.0056	0.3471 ± 0.0073	0.3322 ± 0.0058	0.2589 ± 0.0059	0.2472 ± 0.0045
b_2	0.55+0.40 −0.35	0.14+0.25 −0.28	0.56+0.38 −0.31		0.32+0.28 −0.24	
$b_{\mathcal{G}_2}$	0.07+0.25 −0.22	− ± 6	0.16+0.24 −0.21	−0.096 0.15	0.09+0.19 −0.15	−0.06 0.10
b_{η^2}	−0.71+0.64 −0.52	0.43 0.61	−0.46+0.71 −0.58	1.47+0.39 −0.30	− 8	1.01+0.24 −0.19
$b_{\delta\eta}$	0.53+0.49 −0.40	0.45+0.21 −0.27	0.57+0.53 −0.44	1.39+0.25 −0.31	0.40+0.42 −0.38	1.03+0.17 −0.21
$b_{(KK)_\parallel}$	− ± 9	−0.73+0.60 −0.43	−1.60+0.53 −0.63	− 3	−1.10+0.39 −0.51	−1.34+0.27 −0.34
$b_{\Pi^{[2]}_\parallel}$	0.62 ± 0.32	8	0.59+0.35 −0.30	0.68+0.19 −0.24	0.45+0.26 −0.20	0.43+0.15 −0.19
$\alpha_\parallel$	1.009+0.014	1.009 0.013	1.009 0.013	1.009 0.015	1.010 0.014	1.011 0.018
$\alpha_\perp$	−0.012 0.9978+0.0076	0.9992 ± 0.0060	1.0000+0.0048	1.0011+0.0057	1.0001+0.0046	1.0021+0.0060
	−0.0054		−0.0057	−0.0070	−0.0057	−0.0080
解析边际化 bΠ[3] // 1.3854						
± 0.3592		2.2579 ± 0.1161	2.3390 ± 0.3818	3.9789 ± 0.1261	1.6803 ± 0.2466	2.7891 ± 0.0766
$b_{\delta\Pi^{[2]}_\parallel}$	−0.2006 ± 0.8437	−0.1935 ± 0.4518	−0.7482 ± 0.8638	−1.5778 ± 0.4703	−1.1724 ± 0.7020	−0.7901 ± 0.3159
b(KΠ[2]) //	−1.2460 ± 0.8426	−3.3381 ± 0.3855	−2.9227 ± 0.8766	−5.5891 ± 0.3947	−2.4507 ± 0.5353	−3.4554 ± 0.2444
$b_{\eta\Pi^{[2]}_\parallel}$	−0.4614 ± 1.7829	−0.4338 ± 1.2428	−0.4876 ± 1.8040	−0.5198 ± 1.2934	−0.4975 ± 1.7575	0.7903 ± 0.9370
b_{Γ_3}	0.5259 ± 0.3569	0.2269 ± 0.2086	0.5012 ± 0.3592	0.6851 ± 0.2144	0.6367 ± 0.2737	0.4511 ± 0.1472
P_{shot}	−0.7920 ± 0.5576	−0.9231 ± 0.2003	−2.8145 ± 0.5607	−2.4783 ± 0.1957	−2.8691 ± 0.3647	−2.2519 ± 0.1173
a_0	0.7054 ± 5.0299	12.2297 ± 3.4545	0.9571 ± 4.9795	16.1193 ± 3.4851	1.2038 ± 4.9625	17.7688 ± 2.6652
a_2	0.2653 ± 5.0186	2.7929 ± 4.8289	0.0609 ± 4.9918	3.6760 ± 4.8137	−0.0266 ± 5.0211	1.9947 ± 4.7306
c_0	0.3195 ± 0.2075	0.1045 ± 0.0756	0.7356 ± 0.2084	0.6500 ± 0.0754	0.9325 ± 0.1623	0.6460 ± 0.0535
c_2	0.0144 ± 0.4208	−0.0773 ± 0.1190	−0.8157 ± 0.4414	−1.3573 ± 0.1221	−1.1771 ± 0.3688	−1.3204 ± 0.1009
c_4	−0.1286 ± 0.1724	1.0108 ± 0.0797	0.3731 ± 0.1792	2.3128 ± 0.0826	0.3669 ± 0.1878	1.9862 ± 0.0731
χ^2_{marg}						

表 VII. Ly-α 森林自相关拟合：通过对大体积（L = 2,000 h−1Mpc）模拟的 Ly-α 森林自功率谱拟合得到的边际最佳 EFT 参数，这些模拟经过 Sherwood 流体力学模拟校准，对于流体力学盒子大小 L = 160 和 80 h−1Mpc，在红移 z = 2.8（前四列），并插值至 z = 2.33（最后两列）。顶部部分列出了采样参数；底部部分列出了我们进行解析边际化，并从链中后验恢复的参数。反项 c0、c2、c4 的单位为 [h−1Mpc]²，kmax 的单位为 h Mpc−1。kmax = 0.3 (0.5) h Mpc−1 拟合使用了 460 (780) 个数据点。线性偏差参数与三维功率谱的拟合结果在 1 − 2σ 水平上是一致的 [136]。

请注意，星系的散粒噪声值非常小，因为我们在进行拟合之前已经将其减去。

八. 总结与结论

Ly-α森林是我们宇宙大尺度结构的有力探针。它探测的宇宙体积在V阶段光谱学之前对星系巡天仍然无法触及，从而扩展了摄动方法的应用范围。由于其较高的红移范围（2 ≤ z ≤ 5），Ly-α森林在宇宙学参数之间还表现出旋转的简并方向。Ly-α森林的一个关键优势是将中性氢分布与底层暗物质关联的物理机制被很好理解，这使得对亚千秒差距（sub-kpc）尺度的精确流体动力学模拟成为可能。这些模拟通常用于校准基于线性理论的模型，并结合用于功率谱（或两点相关函数）的现象学拟合函数，适用于降至几个 h Mpc−1 的尺度，从而获得非常稳健的宇宙学约束。

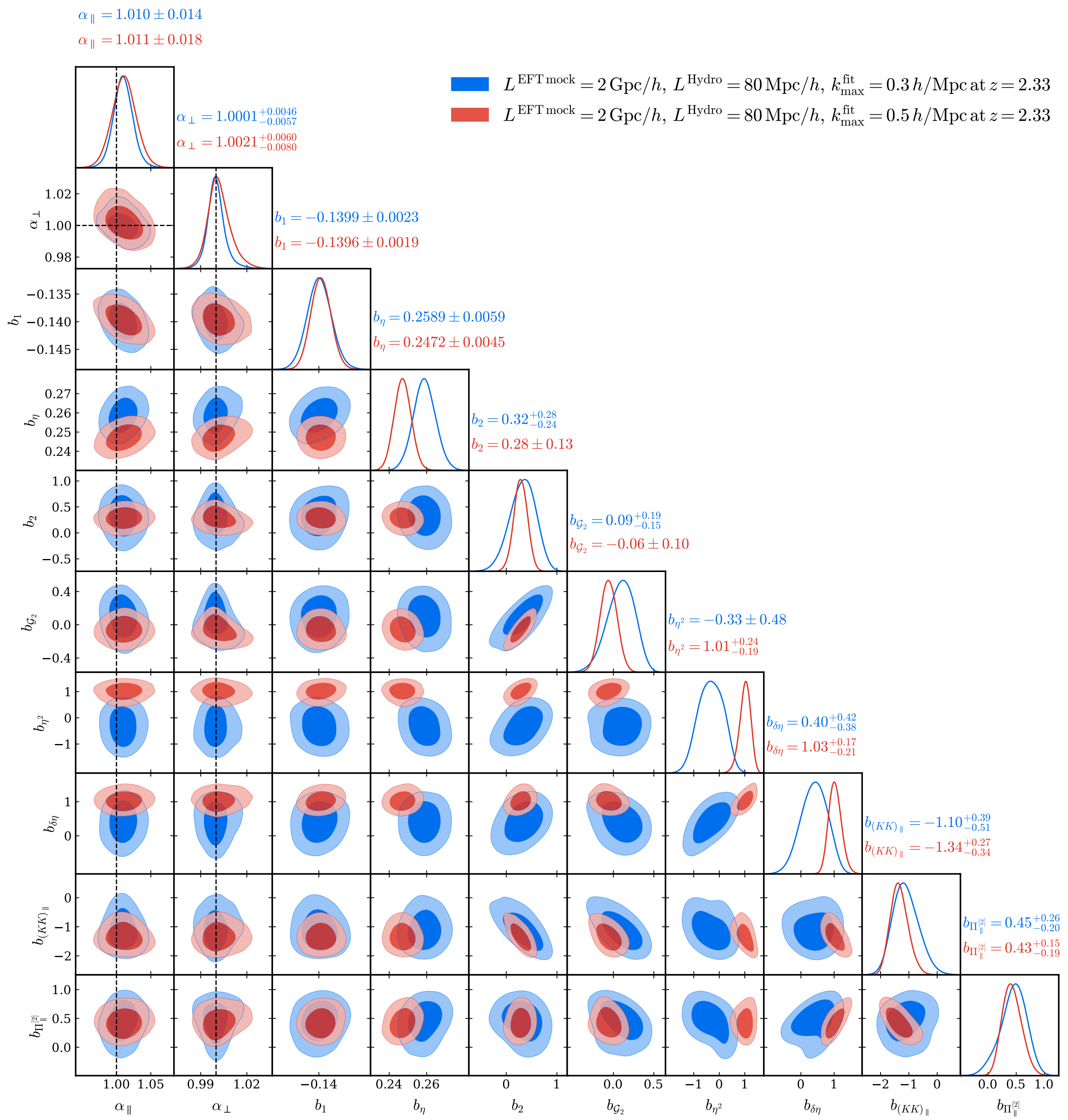

FIG. 25. **Drift plot of Ly-α Mock at $z = 2.33$:** Best-fit 1D and 2D marginalized posteriors for the $l = 80\,h^{-1}$ Mpc Sherwood simulation using two different maximum wavenumbers for the full-shape fits $k_{\rm max} = 0.3, 0.5\,h\,{\rm Mpc}^{-1}$. Crucially, we recover *unbiased BAO scaling parameters.*

However, fits to the BAO scaling parameters – the primary observables for constraining the cosmic expansion history with DESI – using quasi-linear theory models are biased at the ≈ 0.3% level [77, 78]. This is comparable to the forecasted cumulative precision of $\lesssim 0.2\%$ when combining all tracers and redshift bins [49] indicating that existing modeling frameworks are approaching their limits.

Recent advances in theoretical modeling using the effective field theory (EFT) of large-scale structure, extended to the Ly-α forest [86, 127, 136], enable a consistent description of the Ly-α forest from large to intermediate scales.

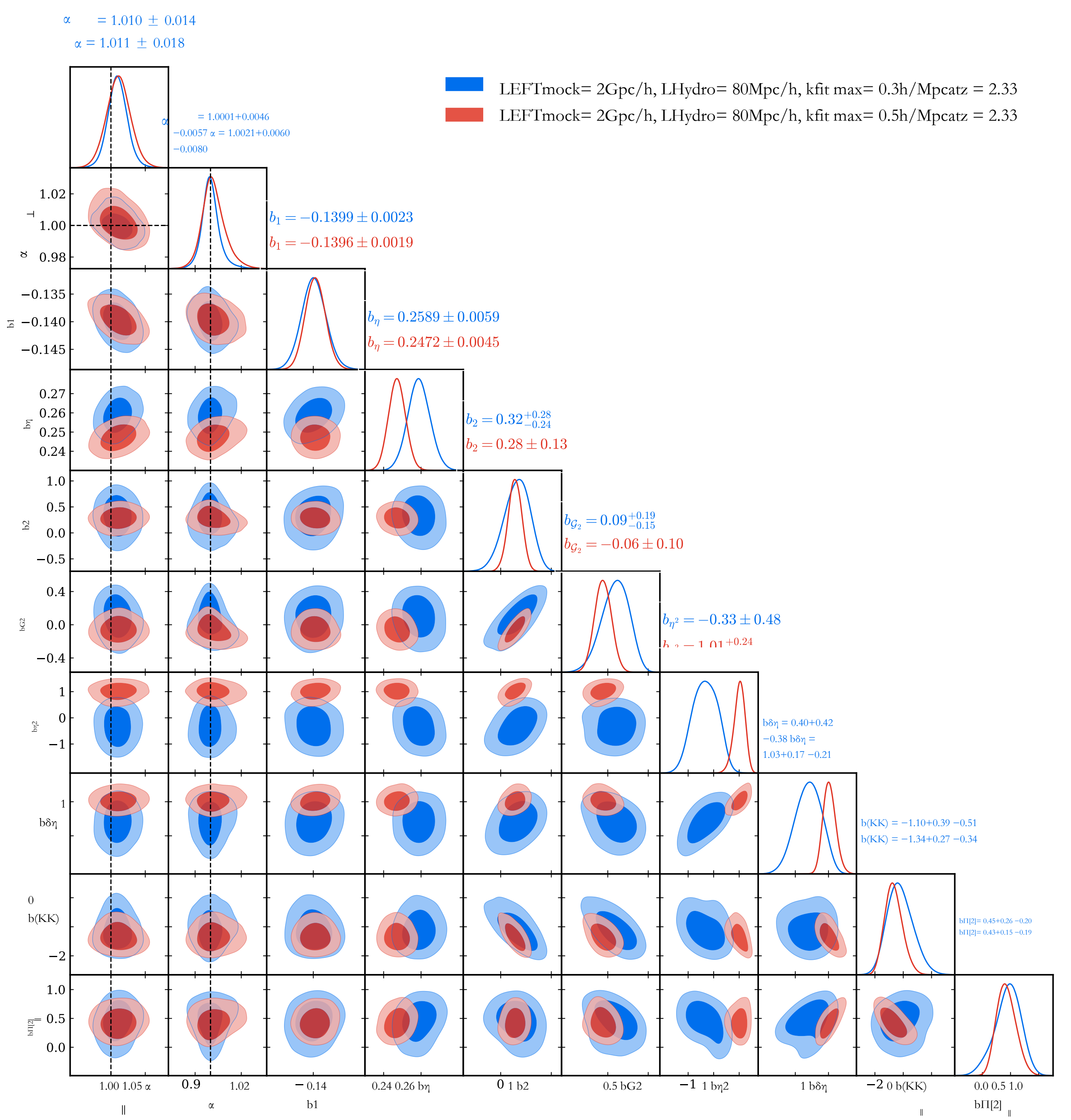

图25. 在 z = 2.33 时 Ly-α 模拟的漂移图：针对 l = 80 h−1Mpc Sherwood 模拟使用两种不同最大波数进行全形状拟合 kmax = 0.3、0.5 h Mpc−1 的最佳拟合一维和二维边际后验分布。关键是，我们恢复了无偏的 BAO 缩放参数。

然而，使用准线性理论模型对BAO标度参数的拟合——这是使用DESI约束宇宙膨胀历史的主要观测量——存在大约0.3%的偏差 [77, 78]。这与在结合所有示踪器和红移区间时预测的累计精度 < ~ 0.2% [49] 相当，表明现有的建模框架正接近其极限。最近在使用大尺度结构有效场理论（EFT）进行理论建模方面的进展，并扩展到Ly-α森林 [86, 127, 136]，使得能够从大尺度到中等尺度一致地描述Ly-α森林。

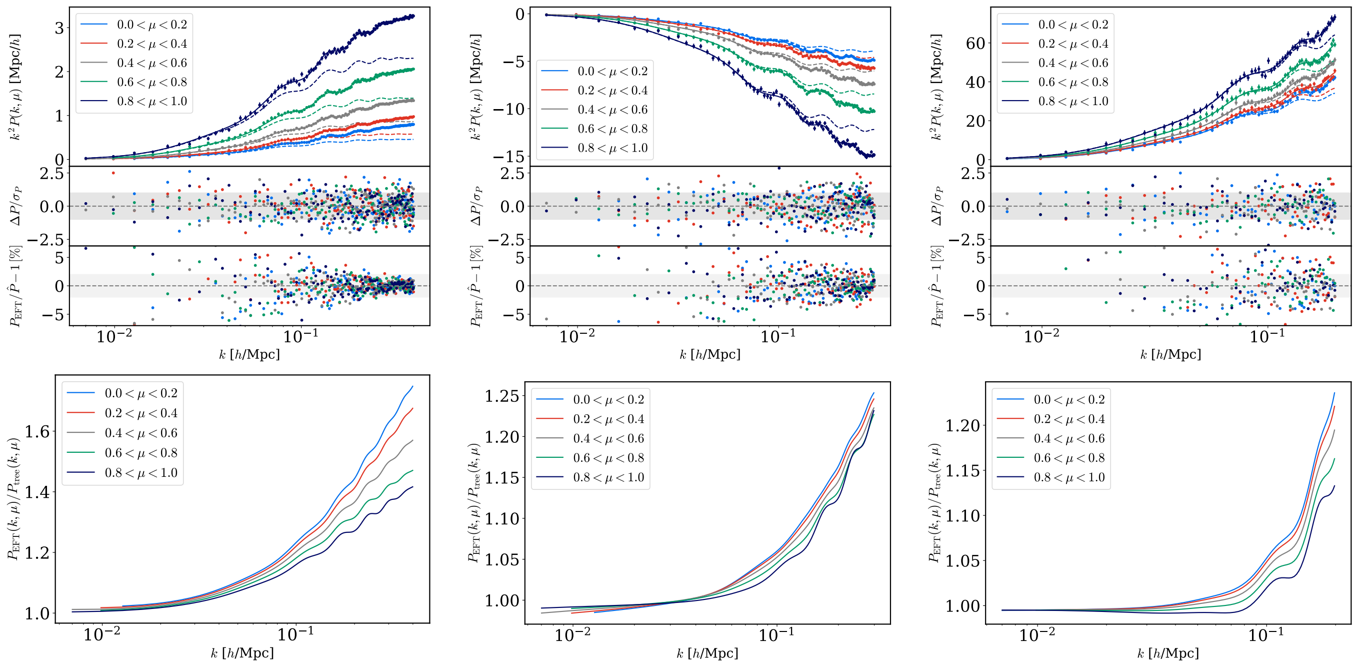

FIG. 26. **Validation of Ly-α × LAE S5 Mock:** Same as Fig. 22 but for the joint fit of the Ly-α auto-correlation (*left column*) with the cross-correlation of Ly-α and galaxies (*center column*) and the galaxy auto-correlation (*right column*) using LAEs of the mimicked “S5” sample. The top row shows the best-fit EFT model (*solid line*) which is compared to $P^{\rm tree}$ (*dashed line*) and the data points with corresponding error bars. The two bottom panels show, first, the normalized and fractional residuals. For both, a gray band indicating the 2% (or 1σ) band is shown to guide the eye. The bottom row shows the loop corrections which are the ratio between the best-fit EFT one-loop power spectrum and the tree-level power spectrum. The joint fits use the following maximum wavenumbers for the fits: $k^{\rm FF}_{\rm max} = 0.40\,h\,{\rm Mpc}^{-1}$, $k^{\rm Fg}_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$, and $k^{\rm gg}_{\rm max} = 0.2\,h\,{\rm Mpc}^{-1}$ for each realization. Note that the Ly-α forest transfer functions are computed using a Sherwood simulation with box size $L = 80\,h^{-1}\,{\rm Mpc}$ and the mock is centered at $z = 3$.

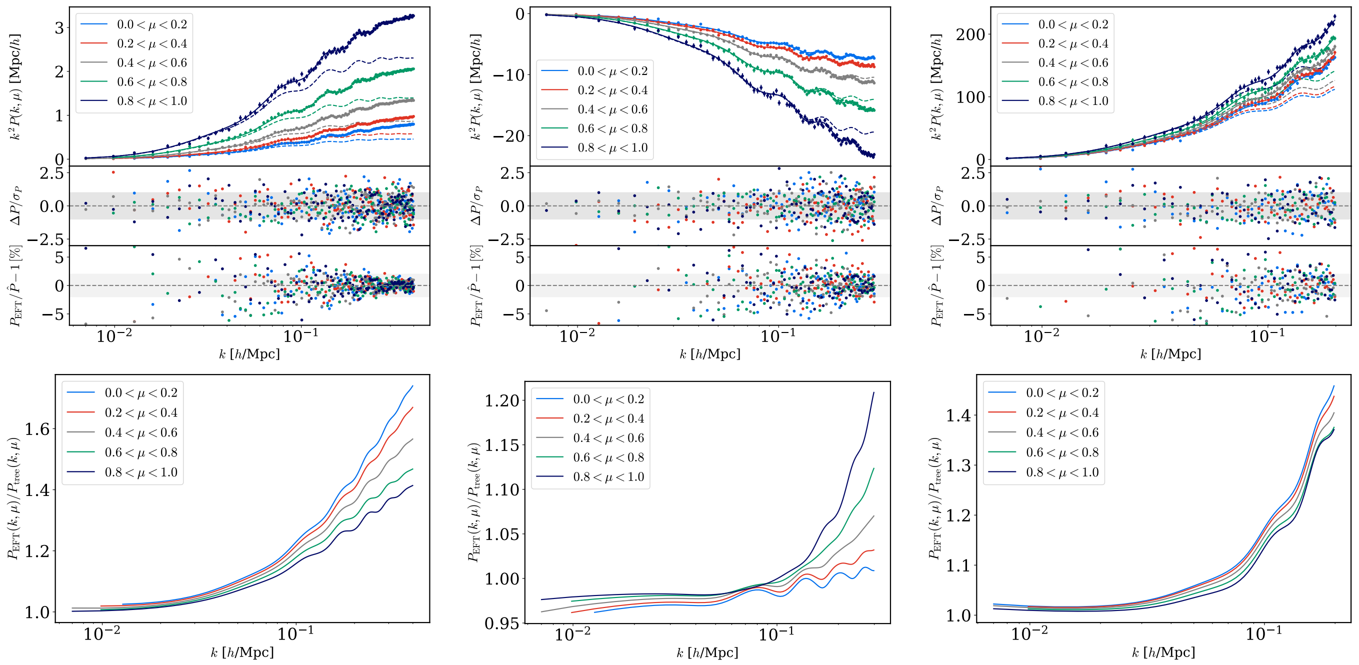

FIG. 27. **Validation of Ly-α × LBG S5 Mock:** Same as Fig. 26 but for S5 LBG sample.

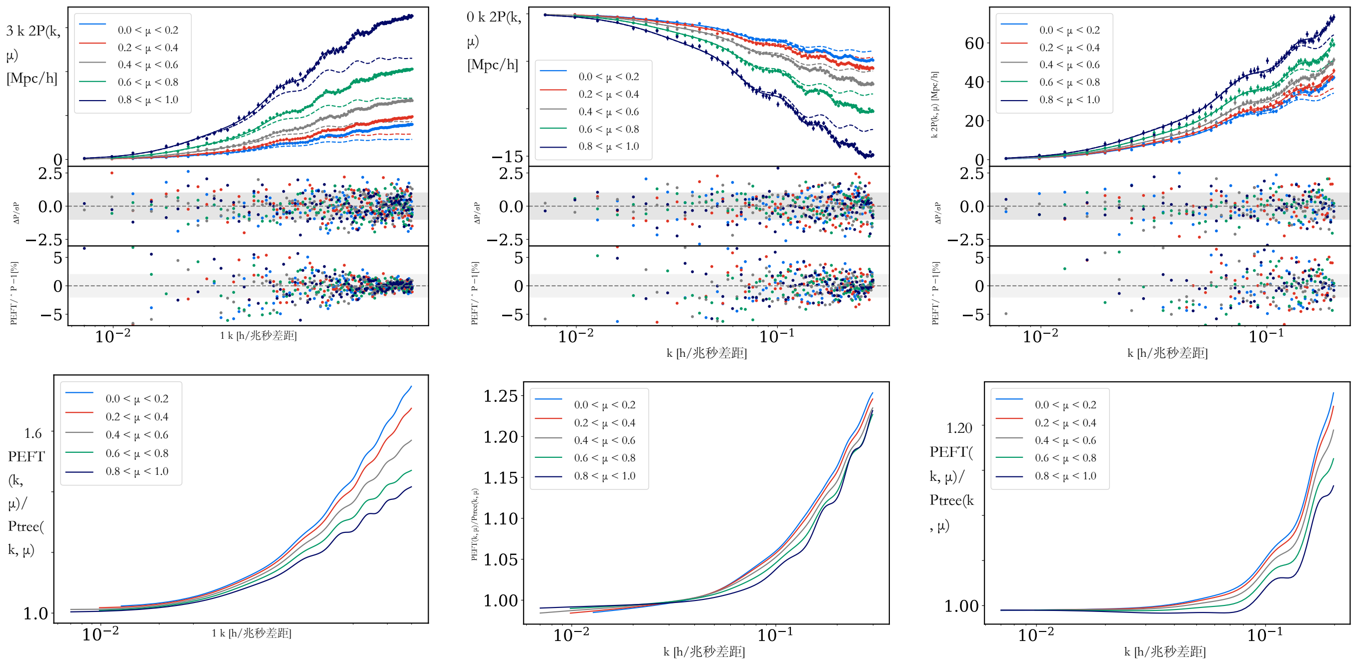

图26. Ly-α × LAE S5模拟验证：与图22相同，但针对Ly-α自相关（左列）与Ly-α和星系的交叉相关（中列）以及星系自相关（右列）的联合拟合，使用模拟的“S5”样本的LAEs。顶行显示最佳拟合的EFT模型（实线），并与Ptree（虚线）以及带有相应误差条的数据点进行比较。底部两个面板首先显示归一化和百分比残差。对于这两项，灰色带显示2%（或1σ）区间以便参考。最底行显示环修正，它们是最佳拟合EFT一环功率谱与树级功率谱的比值。联合拟合使用的最大波数为：kFF max= 0.40 h Mpc−1, kFg max= 0.3 h Mpc−1, 和 kgg max= 0.2 h Mpc−1，针对每个实现。请注意，Ly-α森林的传递函数是使用盒子大小为 L = 80 h−1Mpc 的 Sherwood 模拟计算的，并且模拟的中心红移为 z = 3。

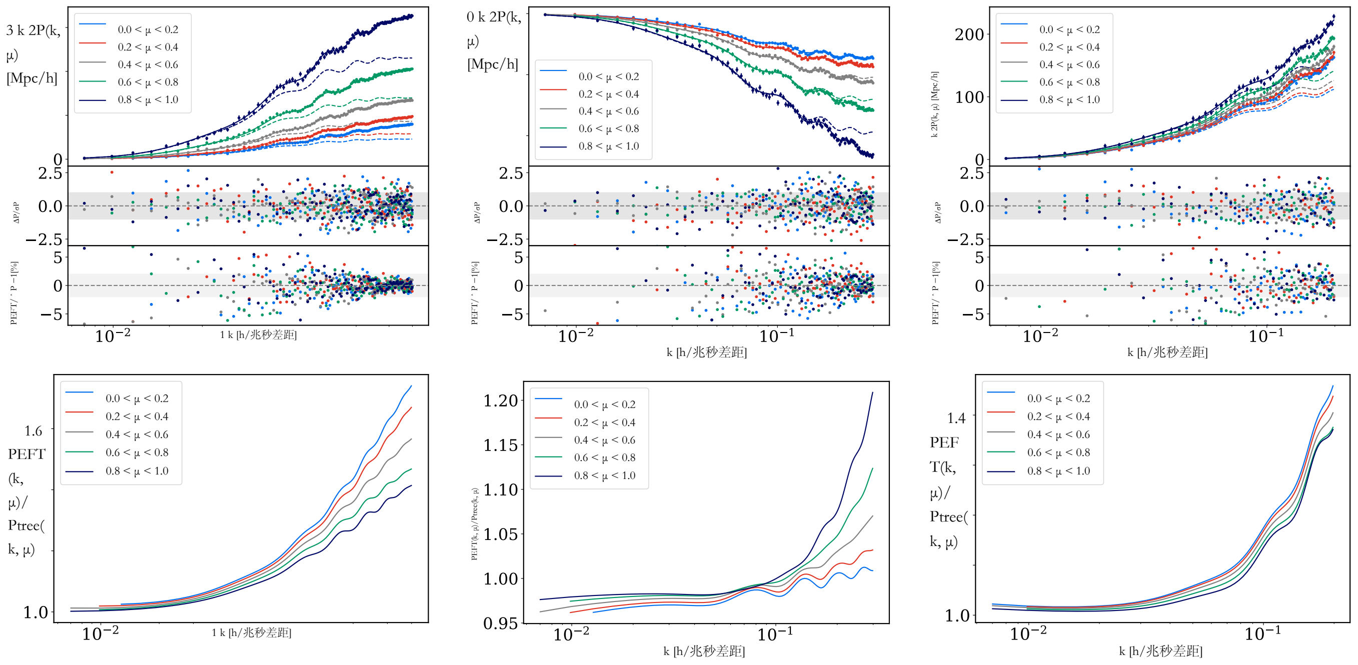

图27. Ly-α × LBG S5 模拟的验证：与图26相同，但针对S5 LBG样本。

Sample	**LBG**		**LAE**	
$b_{\mathcal{O}}$	CARS	S5	S5	ODIN
$b_1^{\rm q}$	4.073 ± 0.031	$3.798^{+0.027}_{-0.028}$	$2.094^{+0.022}_{-0.017}$	$1.969^{+0.023}_{-0.038}$
$b_2^{\rm q}$	$11.71^{+2.89}_{-2.83}$	$10.400^{+2.930}_{-2.938}$	$2.745^{+2.947}_{-1.615}$	$-5.436^{+0.266}_{-3.694}$
$b_{\mathcal{G}_2}^{\rm q}$	$-0.7115^{+0.8136}_{-0.8095}$	$-1.4360^{+0.7184}_{-0.7140}$	$-1.1070^{+0.5132}_{-0.5050}$	$-1.0540^{+0.6946}_{-0.7010}$
b_1	-0.236 ± 0.002	-0.2359 ± 0.0020	-0.2354 ± 0.0020	-0.2369 ± 0.0020
b_η	0.3810 ± 0.0052	0.3818 ± 0.0052	0.3820 ± 0.0053	$0.3882^{+0.0056}_{-0.0057}$
b_2	$0.5299^{+0.1680}_{-0.1355}$	$0.5349^{+0.1690}_{-0.1304}$	$0.3768^{+0.2397}_{-0.1390}$	$0.24820^{+0.21780}_{-0.18153}$
$b_{\mathcal{G}_2}$	$-0.17110^{+0.09106}_{-0.09010}$	$-0.2119^{+0.0905}_{-0.0903}$	$-0.122400^{+0.119932}_{-0.088200}$	$-0.08786^{+0.17645}_{-0.13754}$
b_{η^2}	$-0.8555^{+0.2447}_{-0.2035}$	$-0.8688^{+0.2247}_{-0.1972}$	$-0.8933^{+0.3710}_{-0.2787}$	$-0.3790^{+0.3311}_{-0.2789}$
$b_{\delta\eta}$	$0.8217^{+0.2643}_{-0.2159}$	$0.8855^{+0.2635}_{-0.2066}$	$0.7630^{+0.2920}_{-0.2158}$	$0.8819^{+0.2681}_{-0.2205}$
$b_{(KK)_\parallel}$	$-0.05146^{+0.15616}_{-0.15654}$	$-0.09012^{+0.15319}_{-0.15138}$	$0.02497^{+0.22013}_{-0.16717}$	$-0.3620^{+0.2957}_{-0.2878}$
$b_{\Pi^{[2]}_\parallel}$	$0.6957^{+0.1911}_{-0.2289}$	$0.7214^{+0.1875}_{-0.2204}$	$0.7328^{+0.2263}_{-0.2257}$	$0.5261^{+0.1542}_{-0.2165}$
$\alpha_\parallel$	$1.005^{+0.006}_{-0.005}$	$1.0040^{+0.0050}_{-0.0052}$	$1.007^{+0.005}_{-0.006}$	$1.0030^{+0.0050}_{-0.0057}$
$\alpha_\perp$	$1.0010^{+0.0040}_{-0.0027}$	$1.0010^{+0.0040}_{-0.0026}$	$0.9985^{+0.0035}_{-0.0032}$	1.000 ± 0.003
$b_{\Pi^{[3]}_\parallel}$	2.526 ± 0.22	2.438 ± 0.22	2.416 ± 0.218	1.917 ± 0.227
$b_{\delta\Pi^{[2]}_\parallel}$	-0.1866 ± 0.673	-0.3883 ± 0.625	-0.578 ± 0.673	-1.455 ± 0.662
$b_{(K\Pi^{[2]})_\parallel}$	-5.57 ± 0.515	-5.59 ± 0.498	-6.077 ± 0.52	-5.323 ± 0.511
$b_{\eta\Pi^{[2]}_\parallel}$	0.2743 ± 1.72	-0.06847 ± 1.65	-0.3481 ± 1.74	-0.2983 ± 1.68
b_{Γ_3}	1.67 ± 0.254	1.808 ± 0.256	1.786 ± 0.265	2.104 ± 0.267
$P_{\rm shot}$	-0.2746 ± 0.3	-0.2858 ± 0.292	-0.182 ± 0.334	-0.5699 ± 0.338
a_0	0.165 ± 0.73	0.1384 ± 0.729	0.1246 ± 0.754	0.136 ± 0.715
a_2	-0.007086 ± 0.74	-0.04853 ± 0.736	0.008977 ± 0.749	0.0487 ± 0.741
c_0	-0.04172 ± 0.0924	-0.03431 ± 0.0896	-0.04338 ± 0.105	0.08935 ± 0.107
c_2	0.1545 ± 0.156	0.153 ± 0.159	0.07084 ± 0.178	-0.5302 ± 0.184
c_4	-0.1949 ± 0.11	-0.2521 ± 0.106	-0.1554 ± 0.117	0.1098 ± 0.117
$b_4^\times$	-11.26 ± 3.81	-3.662 ± 3.68	4.088 ± 3.62	11.2 ± 3.65
$P_{\rm shot}^{\rm q}$	-1.011 ± 0.0809	-2.773 ± 0.126	-0.2253 ± 0.0332	-0.001836 ± 0.0273
$a_0^{\rm q}$	71.65 ± 19.2	87.29 ± 31.3	38.71 ± 7.31	6.731 ± 6.55
$a_2^{\rm q}$	-166.6 ± 67.1	-208.1 ± 105	-0.06663 ± 24.7	-7.069 ± 21
$c_0^{\rm q}$	-2.699 ± 1.7	-2.007 ± 1.53	3.008 ± 1.02	1.615 ± 0.954
$c_2^{\rm q}$	10.28 ± 3.37	12.15 ± 3.13	4.394 ± 2.04	-2.152 ± 1.89
$c_4^{\rm q}$	2.279 ± 2.73	-0.0832 ± 2.53	1.991 ± 1.64	-1.039 ± 1.54
$b_{\Gamma_3}^{\rm q}$	2.252 ± 0.0336	2.38 ± 0.033	2.038 ± 0.0354	1.28 ± 0.0354
$b_4^{\rm q}$	-78.02 ± 43.2	-38.43 ± 38.5	6.535 ± 35.6	63.54 ± 37.1
$\chi^2_{\rm marg}$	1204.7	1242.2	1179.3	1226.4

TABLE VIII. **Joint Ly-α and LAE/LBG fits:** Marginalized best-fit EFT parameters obtained from joint fits of the auto- and cross-correlation of the Ly-α forest and the LAE/LBG samples. The top section shows the sampled parameters and the middle (bottom) section the ones that we analytically marginalize over for the auto- (cross-) correlation. We use 1400 data points, sample over 13 parameters explicitly (*top block*) and analytically marginalize over 20 (*bottom two blocks*) yielding a marginalized χ^2 that is ≈ 0.85 for each fit.

In a companion *Letter* [87], we demonstrated that the same bias expansion can be used to successfully model the Ly-α forest flux decrement and halo densities at the field level. In this work, we present the theoretical framework underlying this approach and summarize our main conclusions as follows.

- **Field-level modeling of the Ly-α forest.** The EFT framework can accurately model the Ly-α forest flux

样本	LBG		莱	
b				
$b_1^{\rm q}$	4.073 ± 0.031	$3.798^{+0.027}_{-0.028}$	$2.094^{+0.022}_{-0.017}$	$1.969^{+0.023}_{-0.038}$
$b_2^{\rm q}$	$11.71^{+2.89}_{-2.83}$	$10.400^{+2.930}_{-2.938}$	$2.745^{+2.947}_{-1.615}$	$-5.436^{+0.266}_{-3.694}$
$b_{\mathcal{G}_2}^{\rm q}$	$-0.7115^{+0.8136}_{-0.8095}$	$-1.4360^{+0.7184}_{-0.7140}$	$-1.1070^{+0.5132}_{-0.5050}$	$-1.0540^{+0.6946}_{-0.7010}$
b_1	-0.236 ± 0.002	-0.2359 ± 0.0020	-0.2354 ± 0.0020	-0.2369 ± 0.0020
b_η	0.3810 ± 0.0052	0.3818 ± 0.0052	0.3820 ± 0.0053	$0.3882^{+0.0056}$
b_2	0.5299+0.1680	0.5349+0.1690	0.3768+0.2397	[illegible]
$b_{\mathcal{G}_2}$	[illegible]	[illegible]	[illegible]	[illegible]
b_{η^2}	[illegible]	[illegible]	[illegible]	[illegible]
$b_{\delta\eta}$	[illegible]	[illegible]	[illegible]	[illegible]
$b_{(KK)_\parallel}$	[illegible]	[illegible]	[illegible]	[illegible]
$b_{\Pi^{[2]}_\parallel}$	[illegible]	[illegible]	[illegible]	[illegible]
$\alpha_\parallel$	[illegible]	[illegible]	[illegible]	[illegible]
$\alpha_\perp$	[illegible]	[illegible]	[illegible]	−0.0057±
$b_{\Pi^{[3]}_\parallel}$	[illegible]	[illegible]	[illegible]	1.917 ± 0.227
$b_{\delta\Pi^{[2]}_\parallel}$	−0.1866 ± 0.673	−0.3883 ± 0.625	−0.578 ± 0.673	−1.455 ± 0.662
b(KΠ[2]) //	-5.57 ± 0.515	-5.59 ± 0.498	-6.077 ± 0.52	−5.323 ± 0.511
$b_{\eta\Pi^{[2]}_\parallel}$	0.2743 ± 1.72	−0.06847 ± 1.65	−0.3481 ± 1.74	−0.2983 ± 1.68
b_{Γ_3}	1.67 ± 0.254	1.808 0.256	1.786 0.265	2.104 0.267
$P_{\rm shot}$	-0.2746 ± 0.3	−0.2858 ± 0.292	−0.182 ± 0.334	−0.5699 ± 0.338
a_0	0.165 0.73	0.1384 ± 0.729	0.1246 ± 0.754	
a_2	−0.007086 ± 0.74	−0.04853 ± 0.736	0.008977 ± 0.749	0.0487 ± 0.741
c_0	−0.04172 ± 0.0924	−0.03431 ± 0.0896	−0.04338 ± 0.105	0.08935 ± 0.107
c_2	0.1545 ± 0.156		0.07084 ± 0.178	−0.5302 ± 0.184
c_4	−0.1949 ± 0.11	−0.2521 ± 0.106	−0.1554 ± 0.117	0.1098 ± 0.117
$b_4^\times$				
$P_{\rm shot}^{\rm q}$	−1.011 ± 0.0809	−2.773 ± 0.126	−0.2253 ± 0.0332	−0.001836 ± 0.0273
$a_0^{\rm q}$				
$a_2^{\rm q}$	-166.6 ± 67.1	-208.1 ± 105	−0.06663 ± 24.7	-7.069 ± 21
$c_0^{\rm q}$	-2.699 ± 1.7	-2.007 ± 1.53		1.615 ± 0.954
$c_2^{\rm q}$	10.28 ± 3.37	12.15 3.13	4.394 ± 2.04	-2.152 ± 1.89
$c_4^{\rm q}$	2.279 2.73	−0.0832 ± 2.53	1.991 1.64	-1.039 ± 1.54
$b_{\Gamma_3}^{\rm q}$	2.252 ± 0.0336		2.038 ± 0.0354	1.28 ± 0.0354
$b_4^{\rm q}$		-38.43 ± 38.5		63.54 ± 37.1
$\chi^2_{\rm marg}$	1204.7	1242.2	1179.3	1226.4

表 VIII. 联合 Ly-α 与 LAE/LBG 拟合：通过 Ly-α 林和 LAE/LBG 样本的自相关和交叉相关联合拟合获得的边际最佳拟合 EFT 参数。顶部部分显示采样的参数，中间（底部）部分显示我们在自相关（交叉相关）中进行解析边际化的参数。我们使用 1400 个数据点，对 13 个参数进行显式采样（顶部块），并对 20 个参数进行解析边际化（底部两个块），每次拟合得到的边际 χ² 约为 0.85。

在一封配套的信件[87]中，我们证明了相同的偏差展开可以成功地用于在场尺度上模拟莱曼-α森林通量减弱和晕密度。在这项工作中，我们提出了这一方法的理论框架，并总结了我们的主要结论如下。

- **Ly-α森林的场级建模。EFT框架可以准确地模拟Ly-α森林通量**

decrement at the field level. Using an EFT perturbative model, we find agreement at the $\leq 5\%$ level between the modeled (and fitted) and simulated power spectra down to $k_{\max} \leq 1.0\,h\,\mathrm{Mpc}^{-1}$ for the 3D and 1D power spectra, for a range of redshifts $z = 2.0 - 3.2$. The corresponding one-point probability distribution function agrees down to cell sizes of $\approx 2\,h^{-1}\,\mathrm{Mpc}$, capturing information beyond the two-point function. Importantly, this forward model reproduces *all* amplitudes and phases across *all* modes, representing a substantially more stringent test of accuracy than comparisons based solely on mode-averaged statistics. More quantitatively, the forward-modeled and simulated 1D and 3D fields are correlated at the $> 99\%$ level for $k \lesssim 0.3\,h\,\mathrm{Mpc}^{-1}$ and at the $> 95\%$ level down to $k \lesssim 1.0\,h\,\mathrm{Mpc}^{-1}$. A high precision description of the Ly-α data for $k \gtrsim 1\,h\,\mathrm{Mpc}^{-1}$ must include a consistent modeling of the Ly-α stochasticity (the analog of the one-halo term), which we have quantified in detail for the first time. Note that the redshifts studied in our current analysis are lower than those currently utilized in the actual EFT-based analyses of the Ly-α P1D data [15, 17, 20, 180]. A steep increase of the non-linear wavenumber (proxy to the EFT cutoff) at $z > 3$ suggests that the field-level EFT modeling should be applicable to an even wider range of scales at these redshifts [20, 85].

- **Cross-correlations with halos.** The same EFT framework – without the line-of-sight-dependent operators specific to the Ly-α forest – can accurately model the cross-correlation with massive dark matter halos, used here as proxies for high-redshift galaxies, up to the shot-noise limit. For both simulations, we find agreement at the $\leq 5\%$ level between the modeled and simulated power spectra down to $k_{\max} \leq 1.0\,h\,\mathrm{Mpc}^{-1}$ for the 3D cross power, using all available halo masses. Abacus uses a mass range of $10.8 \leq \log_{10}(M/(h^{-1}M_\odot)) \leq 14.2$ approximately yielding a linear bias matching DESI observations $b_q \approx 3.3$ and a number density of $\sim 1.75 \times 10^{-4}\ (h^{-1}\,\mathrm{Mpc})^{-3}$. Sherwood uses halo in a mass range of $10^9 \leq M_\odot \lesssim 10^{14}$.
- **Linear theory modeling at the field level.** We find that a linear theory forward model fails to generate the Ly-α forest flux decrement at the field level. While the model and simulation power spectra agree at the 5–10% level down to $k \lesssim 0.1 - 0.3\,h\,\mathrm{Mpc}^{-1}$, the corresponding error power spectrum only captures these large scales at the $3 - 5\%$ level in this regime. This implies that fits to the power spectrum (or two-point correlation function) using a linear theory model effectively fit noise. The discrepancy is even more pronounced for the one-point probability distribution functions, which differ drastically between the linear theory fields and the simulations, even for large cell sizes of $R = 10$ and $30\,h^{-1}\,\mathrm{Mpc}$. Indeed, one would obtain a formally perfect match to the power spectrum using the forward model $\delta_F^{\mathrm{model}} = (P_{\mathrm{F}}(k,\mu)/P_{\mathrm{lin}}(k))^{1/2}\delta_1(\boldsymbol{k})$ where $P_{\mathrm{F}}(k,\mu)$ is the non-linear Ly-α forest power spectrum. At the field level, the form of this construction is equivalent to that of the linear model employed here, yet it fails beyond scales of $k \gtrsim 0.1\,h\,\mathrm{Mpc}^{-1}$. These findings highlight the necessity of a higher-order bias expansion to capture non-linearities in the data already on quasi-linear scales.
- **Large-volume mock generation and DESI-II relevance.** We construct large-volume ($V = 2^3(\,h^{-1}\,\mathrm{Gpc})^3$) clustering mocks for the Ly-α forest and its cross-correlation with high-redshift galaxies, such as Ly-α emitters (LAEs), a key tracer for the near-future DESI-II survey. These mocks are calibrated on hydrodynamical simulations and exploit the computational efficiency of the perturbative forward model presented here, enabling the generation of large ensembles over cosmological volumes. This capability is particularly relevant given that current Ly-α forest analyses [51, 61] rely on approximate lognormal mocks [130] that fail to capture the physics shaping the BAO feature [77, 128]. To validate these semi-analytic mocks, we perform a BAO analysis at fixed cosmology and recover unbiased BAO scaling parameters, making them applicable for validating inference pipelines.

Our approach bridges the gap between theory and simulation by enabling efficient modeling of the flux decrement and the dark matter halo field directly at the field level, rather than restricting the analysis to summary statistics. This framework naturally opens several promising avenues for future work: To extend the perturbative reach of our approach, transfer functions can be computed fully within the EFT framework; beyond β_1, this will require two-loop calculations along the lines developed for halos in Ref. [181]. The present semi-analytic mocks are well suited for validating inference pipelines and constructing simulation-based covariance matrices for DESI analyses. Extending these mocks to the light cone is therefore a natural next step paired with a detailed investigation of their properties, as well as data-driven forecasts of effects such as radiative transfer measured via the cross-correlation between the Ly-α forest and high-redshift galaxies. Further directions include the development of simulation-based priors for EFT-based full-shape analyses along the lines of [125, 153, 154, 172, 182–185], and the application of our technique to the modeling of the Ly-α field at high redshifts, $3.2 \lesssim z \lesssim 5.4$. The latter is especially important since the Ly-α data at these redshifts is especially powerful at constraining new physics with EFT, making the development of robust simulation-based priors for such analysis a priority [15, 17, 20, 180]. Finally, it would be interesting to study cosmological parameter inference directly from Fourier modes [186–205]. While the latter remains debated for galaxy surveys, particularly regarding the information content beyond the power spectrum and bispectrum [205, 206],

---

在场级上递减。使用EFT微扰模型，我们发现模型（和拟合的）与模拟的功率谱在kmax ⩽ 1.0 h Mpc−1下，对于3D和1D功率谱，在红移范围z = 2.0 −3.2之间，达到了⩽5%的匹配水平。相应的一点概率分布函数在单元格大小≈2 h−1Mpc下也能达成一致，捕获了超越二点函数的信息。重要的是，这个前向模型能够再现所有模态的振幅和相位，比仅基于模态平均统计量的比较更严格地检验了精度。更具体地说，对于k < ~0.3 h Mpc−1，前向建模和模拟的1D及3D场的相关性超过99%；对于k < ~1.0 h Mpc−1，相关性仍超过95%。对于 k ≳1 h Mpc−1 的 Ly-α 数据，高精度的描述必须包括对 Ly-α 随机性的统一建模（类似于单晕项），而我们首次对其进行了详细量化。请注意，我们当前分析研究的红移低于目前实际基于 EFT 的 Ly-α P1D 数据分析中使用的红移 [15, 17, 20, 180]。在 z > 3 时，非线性波数（EFT 截断的代理）的急剧增加表明，在这些红移下，场级 EFT 建模应该适用于更广泛的尺度 [20, 85]。

- 与晕的互相关。相同的有效场论（EFT）框架——不含特定于Ly-α森林的视线方向依赖算符——可以准确地模拟与大质量暗物质晕的互相关，这里将其用作高红移星系的代理，直到达到拍噪限制。对于两种模拟，我们发现，在使用所有可用晕质量的情况下，建模与模拟的功率谱在三维互功率 kmax ⩽1.0 h Mpc−1 范围内一致性达到⩽5%。Abacus 使用的质量范围为 10.8 ⩽log10(M/(h−1M⊙)) ⩽14.2，大致对应线性偏差与 DESI 观测匹配 bq ≈3.3，数密度约为 ~1.75×10−4 (h−1Mpc)−3。Sherwood 使用的晕质量范围为 10^9 ⩽ M⊙ < ~10^14。在场水平上的线性理论建模。我们发现线性理论前向模型无法在场水平上生成 Ly-α 森林通量减弱。虽然模型和模拟的功率谱在 k < ~0.1–0.3 h Mpc−1 时在 5 – 10% 级别上相符，但相应的误差功率谱在该范围内仅以 3–5% 的水平捕捉到这些大尺度。这意味着使用线性理论模型拟合功率谱（或二点相关函数）实际上是在拟合噪声。对于单点概率分布函数，这种差异更加明显，即使在 R = 10 和 30 h−1Mpc 的大单元格情况下，线性理论场与模拟的分布也存在巨大差异。实际上，使用前向模型 δ_model_F = (P_F(k, μ)/P_lin(k))^(1/2) δ_1(k)（其中 P_F(k, μ) 是非线性 Ly-α 林功率谱）可以在形式上获得功率谱的完全匹配。在场层面上，这种构造的形式等价于这里使用的线性模型，但在 k > ~0.1 h Mpc−1 的尺度上，它会失效。这些发现凸显了为了捕捉已经在准线性尺度上的数据中的非线性，有必要进行高阶偏差展开。
- 大规模模拟生成与DESI-II相关性。我们为Ly-α森林及其与高红移星系（如Ly-α发射体（LAEs））的交叉相关构建了大体积（V = 23(h−1Gpc)³）聚类模拟，这些高红移星系是近期DESI-II调查的重要追踪对象。这些模拟是在水动力学模拟上校准的，并利用本文提出的微扰前向模型的计算效率，使得能够在宇宙学体积上生成大规模样本。这一能力尤其相关，因为当前的Ly-α森林分析[51, 61]依赖于近似对数正态模拟[130]，这些模拟未能捕捉形成BAO特征的物理过程[77, 128]。为了验证这些半解析模拟，我们在固定宇宙学参数下进行了BAO分析，并恢复了无偏的BAO缩放参数，从而使它们可用于验证推理流程。

我们的方法通过在场级别上直接对通量减量和暗物质晕场进行高效建模，从而弥合了理论与模拟之间的差距，而不仅仅将分析局限于总结统计量。该框架自然地为未来的工作开辟了几个有前景的方向：为了扩展我们方法的微扰适用范围，可以在EFT框架内完整计算传递函数；在β1之外，这将需要沿着文献[181]中为晕发展出的两环计算方法。当前的半解析模拟非常适合用于验证推断管道以及为DESI分析构建基于模拟的协方差矩阵。因此，将这些模拟扩展到光锥是一个自然的下一步，同时需要对其性质进行详细研究，以及通过高红移星系与Ly-α森林的交叉相关来进行如辐射转移等效应的数据驱动预测。进一步的方向包括沿着 [125, 153, 154, 172, 182 – 185] 的思路，为基于 EFT 的全形分析开发基于模拟的先验，并将我们的方法应用于高红移（3.2 ≲ z ≲ 5.4）Ly-α 场的建模。后者尤其重要，因为在这些红移下的 Ly-α 数据在用 EFT 约束新物理方面尤其有力，使得为此类分析开发稳健的基于模拟的先验成为优先事项 [15, 17, 20, 180]。最后，直接从傅里叶模态研究宇宙学参数推断也是有趣的 [186 – 205]。尽管后者在星系调查中仍存在争议，特别是在功率谱和双谱之外的信息内容方面 [205, 206]，

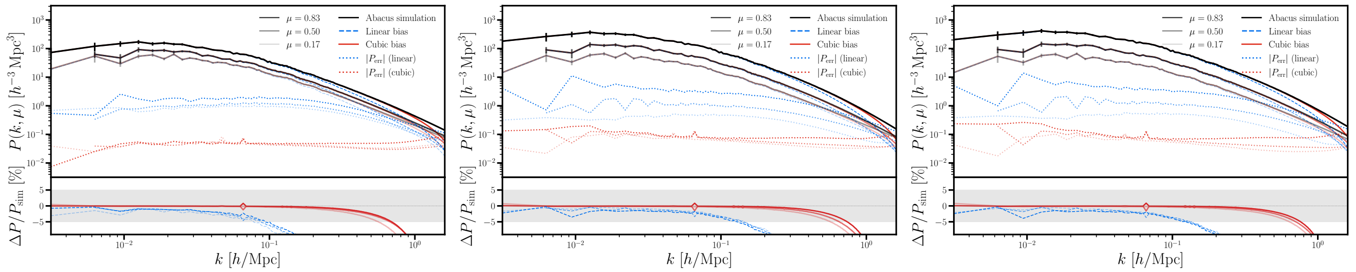

FIG. 28. **Model-dependence of Power Spectra:** Same as Fig. 2 for three different FGPA implementations. Models one, two and three are shown from left to right comparing the linear to the cubic model as well as their error power spectra. Our perturbative forward model is robust to the details of the Abacus simulations, finding a consistent agreement at the 5% level between the forward model and the input simulation. The scale and orientation dependence of the error power spectrum does, however, change with increasing b_η (which we use as proxy for a larger amount of RSD in the simulations). The bottom panel displays the percent difference between the simulation and model power spectra. A gray band highlights the ±5% region in the bottom panel.

it is well motivated and worth investigating for the Ly-α forest. We leave these exciting research directions for future investigation.

ACKNOWLEDGMENTS

We thank Martin White and Pat McDonald for fruitful discussions. We are also grateful to Jamie Bolton, Jonás Chaves-Montero, Jahmour Givans, Julien Guy, Naim Göksel Karaçaylı, and Andreu Font-Ribera for valuable discussions and assistance with the Sherwood files.

This work is supported by the National Science Foundation under Cooperative Agreement PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, `http://iaifi.org/`). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. JMS acknowledges that, in part, support for this work was provided by The Brinson Foundation through a Brinson Prize. KA acknowledges supports from Fostering Joint International Research (B) under Contract No. 21KK0050 and the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant No. JP24K17056. Support for this work was provided by NASA through the NASA Hubble Fellowship grant HST-HF2-51572.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555.

Appendix A: Transfer functions from Sherwood simulations

For the mocks correlating the Ly-α forest with high redshift galaxies in Sec. C, we use the $L = 80\,h^{-1}$ Mpc snapshots since these are available at a series of redshifts $z = 2.0, 2.4, 2.8, 3.2$ which we interpolate between redshifts.

Appendix B: Field-level fits on different Abacus FGPA mocks

To assess the robustness of the outlined analytical forward model, we apply it to three additional implementations of the FGPA procedure (see [102] for more details) representing different combinations of bias parameters. These have been fitted using linear theory in Ref. [102] and using the one-loop power spectrum in Ref. [78]. Whilst we have so far used model 'one' which is similar to model 'two', i.e. having a low value of b_η we also compare to realizations (model 'three' and 'four') with large values of b_η. Qualitatively the results look very similar and are shown for the Ly-α power spectrum in Fig. 28 and their cross-correlation coefficient in Fig. 33. The values obtained from the polynomial fits are tabulated in Tab. XI. In Fig. 32 we show the model dependence for the P1D performance. Note that we find the largest differences between the four models for the P1D performance. In particular, for models III and IV where the velocity bias value is larger (and closer to the data) the linear theory model performs better than for both models with low b_η.

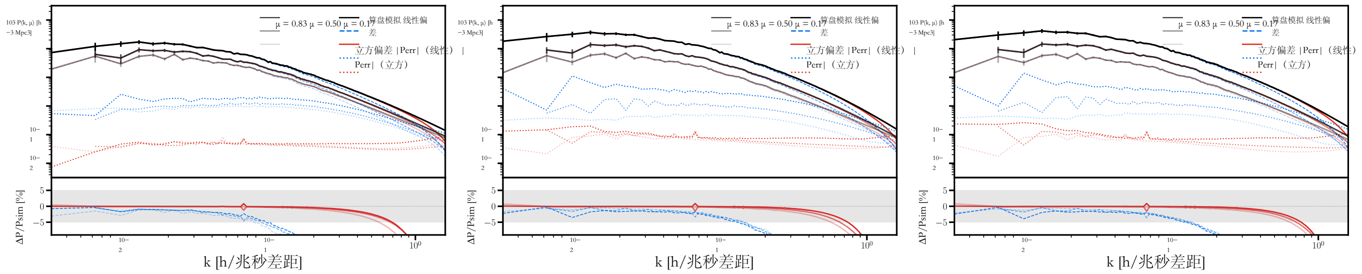

图28. 功率谱的模型依赖性：同图2，但用于三种不同的FGPA实现。三个模型从左到右分别展示了线性模型与三次模型及其误差功率谱的比较。我们的微扰前向模型对Abacus模拟的细节具有鲁棒性，在前向模型与输入模拟之间保持约5%的吻合度。然而，误差功率谱的尺度和方向依赖性会随着bη的增加而变化（我们将bη用作模拟中更大量RSD的代理）。底部面板显示了模拟与模型功率谱之间的百分比差异。灰色带在底部面板中突出显示了±5%的区域。

这是有充分理由和价值去研究Ly-α森林的。我们将这些令人兴奋的研究方向留待将来的研究。

致谢

我们感谢 Martin White 和 Pat McDonald 富有成效的讨论。我们也感谢 Jamie Bolton、Jonás Chaves-Montero、Jahmour Givans、Julien Guy、Naim Göksel Karaçaylı 和 Andreu Font-Ribera 提供宝贵的讨论和在 Sherwood 文件方面的帮助。

本工作得到国家科学基金会（National Science Foundation）在合作协议 PHY-2019786 之下的支持（NSF 人工智能与基础相互作用研究所，http://iaifi.org/）。本研究使用了美国能源部科学办公室用户设施——国家能源研究科学计算中心（NERSC）的资源，该中心根据合同号 DE-AC02-05CH11231 运营。JMS 承认，本工作部分支持来自布林森基金会（The Brinson Foundation）通过布林森奖提供。KA 承认支持来自合同号 21KK0050 的联合国际研究促进项目（B）以及日本学术振兴会（JSPS）科研资助金 KAKENHI Grant No. JP24K17056。本工作的支持还来自 NASA，通过空间望远镜科学研究所颁发的 NASA 哈勃奖学金 HST-HF2-51572.001，该研究所由美国参与天文研究大学协会运营，为 NASA 提供服务，合同号 NAS5-26555。

附录 A：来自 Sherwood 模拟的传递函数

对于在C节中将Ly-α森林与高红移星系相关联的模拟，我们使用了L = 80 h−1Mpc的快照，因为这些快照在一系列红移z = 2.0、2.4、2.8、3.2下可用，我们在红移之间进行插值。

附录 B：不同 Abacus FGPA 模拟器上的字段级拟合

为了评估所概述分析前向模型的稳健性，我们将其应用于FGPA程序的另外三种实现（详见[102]），这些实现代表了不同的偏差参数组合。这些模型在参考文献[102]中使用线性理论拟合，并在参考文献[78]中使用单圈功率谱进行拟合。虽然到目前为止我们使用的是类似于模型‘二’的模型‘一’，即具有较低的bη值，我们也将其与具有较大bη值的实现（模型‘三’和‘四’）进行比较。从定性上看，结果非常相似，并分别在图28中显示了Ly-α功率谱，在图33中显示了它们的互相关系数。通过多项式拟合得到的数值列在表XI中。在图32中，我们展示了P1D性能的模型依赖性。需要注意的是，对于P1D性能，我们发现四个模型之间的差异最大。特别是对于速度偏差值较大（且更接近数据）的模型III和IV，线性理论模型的表现优于速度偏差较低的两种模型。

TF	c_0	c_{01}	c_1	c_{12}	c_{14}	c_4	c_{22}	c_{44}
$\mathbf{z = 2.0}$								
β_1	−0.123	−0.265	0.077	0.733	−0.290	−0.005	−0.546	0.298
β_2	0.104	0.124	−0.028	−0.076	0.014	−0.001	0.026	−0.039
$\beta_{\mathcal{G}_2}$	−0.088	−0.360	0.026	0.531	−0.153	−0.023	−0.336	0.139
β_3	−0.018	−0.007	0.032	0.090	−0.108	−0.010	−0.046	0.106
$\beta_{KK_\parallel}$	−0.115	−0.127	−0.126	−0.206	0.362	0.064	0.123	−0.190
β_η	−0.172	−0.215	−0.046	1.351	−0.599	0.046	−0.816	0.233
β_{η^2}	0.043	0.038	−0.122	0.300	−0.001	0.046	−0.272	0.095
$\beta_{\delta\eta}$	0.051	−0.346	−0.077	1.878	−0.604	0.094	−1.957	1.188
$\mathbf{z = 2.4}$								
β_1	−0.171	−0.385	0.090	1.042	−0.350	0.009	−0.777	0.368
β_2	0.158	0.180	−0.040	−0.078	0.006	−0.001	−0.016	0.012
$\beta_{\mathcal{G}_2}$	−0.118	−0.500	0.017	0.899	−0.314	−0.020	−0.618	0.324
β_3	−0.036	0.006	0.053	0.043	−0.110	−0.012	−0.009	0.133
$\beta_{KK_\parallel}$	−0.157	−0.215	−0.259	−0.340	0.642	0.121	0.110	−0.216
β_η	−0.331	−0.254	−0.083	2.338	−1.127	0.069	−1.384	0.420
β_{η^2}	0.103	0.024	−0.235	0.696	0.059	0.086	−0.710	0.329
$\beta_{\delta\eta}$	0.013	−0.492	−0.033	2.802	−0.873	0.091	−2.998	1.818
$\mathbf{z = 2.8}$								
β_1	−0.231	−0.536	0.100	1.483	−0.422	0.033	−1.113	0.451
β_2	0.247	0.250	−0.061	−0.080	−0.013	0.001	−0.069	0.074
β_{G2}	−0.168	−0.591	0.024	1.149	−0.494	−0.030	−0.822	0.556
β_3	−0.075	0.007	0.090	−0.005	−0.050	−0.017	0.051	0.098
$\beta_{KK\|}$	−0.156	−0.380	−0.553	−0.046	0.803	0.234	−0.373	0.090
β_η	−0.607	−0.239	−0.194	4.289	−2.365	0.097	−2.494	0.937
β_{η^2}	0.230	−0.077	−0.406	1.504	0.220	0.144	−1.533	0.704
$\beta_{\delta\eta}$	−0.105	−0.589	0.083	3.623	−1.154	0.062	−4.100	2.611
$\mathbf{z = 3.2}$								
β_1	−0.310	−0.714	0.113	2.108	−0.521	0.069	−1.602	0.566
β_2	0.393	0.322	−0.103	−0.038	−0.063	0.012	−0.168	0.145
$\beta_{\mathcal{G}_2}$	−0.240	−0.587	0.046	1.117	−0.622	−0.058	−0.785	0.726
β_3	−0.169	0.035	0.192	−0.252	0.147	−0.051	0.312	−0.086
$\beta_{KK_\parallel}$	−0.050	−0.703	−1.146	1.645	0.393	0.476	−2.141	1.082
β_η	−0.968	−0.132	−0.627	7.547	−4.960	0.190	−3.648	1.495
β_{η^2}	0.461	−0.335	−0.639	2.958	0.543	0.210	−2.974	1.356
$\beta_{\delta\eta}$	−0.332	−0.650	0.260	4.658	−1.677	0.053	−5.621	3.806

TABLE IX. **Best-Fit Sherwood Transfer Function:** Sherwood snapshots at $z = 2.0$, 2.4, 2.8 and $z = 3.2$ for the 80 h^{-1} Mpc box.

z	a_0	a_2	a_3	a_4	a_{22}	a_{44}
2.0	0.018	0.590	−1.908	1.482	0.229	0.013
2.4	0.028	1.181	−3.816	3.037	0.295	0.058
2.8	0.089	0.422	−1.158	0.694	0.272	−0.265
3.2	0.030	5.860	−17.706	14.405	0.835	−1.620

TABLE X. **Error Power Spectrum Sherwood:** Fitted parameters for $P_{\rm err}$ for the different snapshots in the $L = 80\,h^{-1}$ Mpc Sherwood simulation.

TF	c_0	c_{01}	c_1	c_{12}	c_{14}	c_4	c_{22}	c_{44}
$\mathbf{z = 2.0}$								
β_1	−0.123	−0.265	0.077	0.733	−0.290	−0.005	−0.546	0.298
β_2	0.104	0.124	−0.028	−0.076				−0.039
$\beta_{\mathcal{G}_2}$	−0.088	−		0.531	−0.153	−0.023	−0.336	0.139
β_3	0.018	0.007	0.032	0.090	−0.108	−0.010	−0.046	0.106
$\beta_{KK_\parallel}$	−0.115	−0.127	−0.126	−0.206				−0.190
β_η	−0.172	−0.215	−0.046		−0.599	0.046	−0.816	0.233
β_{η^2}				0.300	−0.001	0.046	−0.272	0.095
$\beta_{\delta\eta}$	0.051	−0.346	−0.077	1.878	−0.604	0.094	−1.957	1.188
$\mathbf{z = 2.4}$								
β_1	−0.171	−0.385	0.090	1.042	−0.350	0.009	0.777	0.368
β_2	0.158	0.180	−0.040	−0.078	0.006	−0.001	−0.016	0.012
$\beta_{\mathcal{G}_2}$	−0.118	−		0.899	−0.314	−0.020	−0.618	0.324
β_3	0.036	0.006	0.053	0.043	−0.110	−0.012	−0.009	0.133
$\beta_{KK_\parallel}$	−0.157	−0.215	−0.259	−0.340				−0.216
β_η	−0.331	−0.254	−0.083		−1.127	0.069	−1.384	0.420
β_{η^2}				0.696	0.059	0.086	−0.710	0.329
$\beta_{\delta\eta}$	0.013	−0.492	−0.033	2.802	−0.873	0.091	−2.998	1.818
$\mathbf{z = 2.8}$								
β_1	−0.231	−0.536	0.100	1.483	0.422	0.033	−1.113	0.451
β_2	0.247	0.250	−0.061	−0.080	−0.013	0.001	0.069	0.074
β_{G2}	−0.168	−		1.149	−0.494	−0.030	−0.822	0.556
β_3	0.075	0.007	0.090	−0.005	−0.050	−0.017		0.098
$\beta_{KK\|}$	−0.156	−0.380	−0.553	−0.046			−0.373	0.090
β_η	−0.607	−0.239	−0.194		−2.365	0.097	−2.494	0.937
β_{η^2}	0.230	−0.077	−0.406	1.504	0.220	0.144	−1.533	0.704
$\beta_{\delta\eta}$	−			3.623	−1.154	0.062	−4.100	2.611
$\mathbf{z = 3.2}$								
β_1	−0.310	−0.714	0.113	2.108	0.521	0.069	−1.602	0.566
β_2	0.393	0.322	−0.103	−0.038	−0.063	0.012	0.168	0.145
$\beta_{\mathcal{G}_2}$	−0.240	−		1.117	−0.622	−0.058	−0.785	0.726
β_3	0.169	0.035	0.192	−				−0.086
$\beta_{KK_\parallel}$	−0.050	−0.703	−1.146	1.645	0.393	0.476	−2.141	1.082
β_η	−0.968	−0.132	−0.627	7.547	−4.960	0.190	−3.648	1.495
β_{η^2}	0.461	−0.335	−0.639	2.958	0.543	0.210	−2.974	1.356
$\beta_{\delta\eta}$	−			4.658	−1.677	0.053	−5.621	3.806

表 IX. 最佳拟合 Sherwood 传输函数：80 h−1Mpc 盒子在 z = 2.0、2.4、2.8 和 z = 3.2 的 Sherwood 快照。

z	a_0	a_2	a_3	a_4	a_{22}	a_{44}
2.0	0.018	0.590	−1.908	1.482	0.229	0.013
2.4	0.028	1.181	−3.816	3.037	0.295	0.058
2.8	0.089	0.422	−1.158	0.694	0.272	−0.265
3.2	0.030	5.860	−17.706	14.405	0.835	−1.620

表 X. 错误功率谱 Sherwood：L = 80 h−1Mpc Sherwood 模拟中不同快照的 Perr 拟合参数。

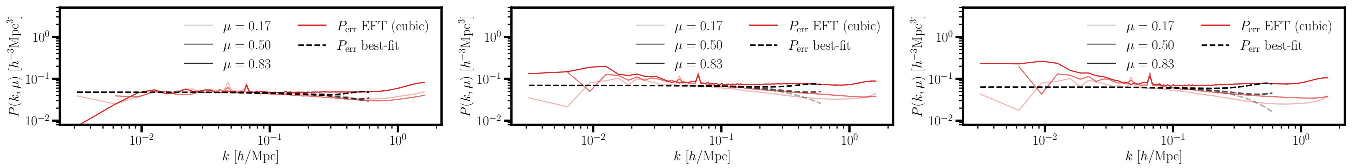

FIG. 29. **Model-dependence of best-fit** $P_{\rm err}$**:** Same as Fig. 28 showing the best-fit model for the error power spectrum for models two, three and four. The fits use $k_{\rm max} = 0.6\,h\,{\rm Mpc}^{-1}$ and show a mild onset of scale and orientation dependence of the error power spectrum at around $0.2 - 0.3\,h\,{\rm Mpc}^{-1}$.

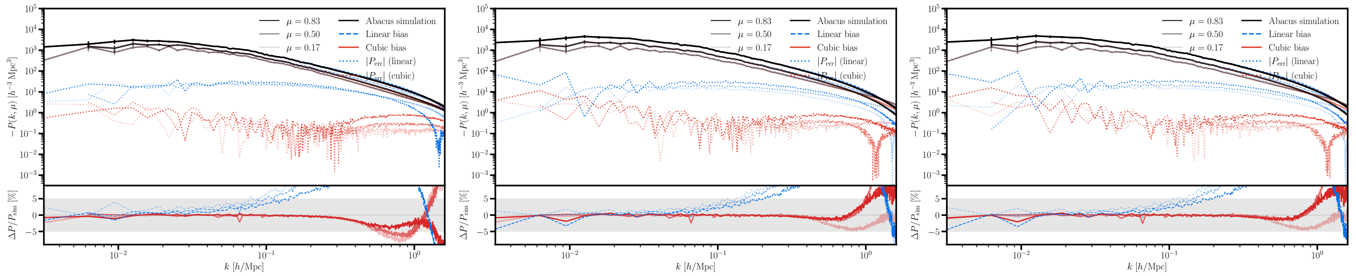

FIG. 30. **Model-dependence of Power Spectra:** Same as Fig. 5 for three different FGPA implementations, denoted by models II-IV.

TF	c_0	c_{01}	c_1	c_{12}	c_{14}	c_4	c_{22}	c_{44}
Model II								
β_1	−0.134	−0.142	0.073	0.253	−0.052	−0.008	−0.156	0.055
β_2	0.097	0.060	−0.017	0.074	−0.080	−0.018	−0.058	−0.045
$\beta_{\mathcal{G}_2}$	−0.075	−0.151	0.018	0.041	0.077	0.017	0.002	0.097
β_3	−0.006	−0.011	−0.016	−0.108	0.087	0.016	0.092	0.008
$\beta_{KK_\parallel}$	−0.018	−0.152	−0.181	0.563	−0.659	−0.056	−0.897	0.841
β_η	−0.199	0.066	0.121	−0.241	0.101	−0.011	−0.043	0.178
β_{η^2}	0.064	−0.034	−0.098	0.071	0.097	0.014	0.019	0.036
$\beta_{\delta\eta}$	−0.057	0.005	−0.016	−0.244	−0.008	−0.007	0.478	−0.296
Model III								
β_1	−0.139	−0.295	0.050	0.407	−0.057	−0.002	−0.099	−0.022
β_2	0.060	0.105	−0.022	0.244	−0.160	−0.008	−0.129	−0.082
$\beta_{\mathcal{G}_2}$	−0.072	−0.300	0.041	−0.110	0.205	0.009	0.149	0.100
β_3	−0.002	−0.012	−0.012	−0.020	0.018	0.011	0.074	−0.003
$\beta_{KK_\parallel}$	−0.168	−0.515	0.106	0.461	−0.137	−0.093	−1.051	0.777
β_η	−0.252	−0.046	0.265	−0.442	0.245	−0.091	0.255	−0.030
β_{η^2}	−0.032	0.010	−0.163	0.415	−0.076	0.052	−0.008	−0.073
$\beta_{\delta\eta}$	0.045	0.054	0.027	−0.211	−0.290	−0.032	0.223	0.030
Model IV								
β_1	−0.135	−0.329	0.051	0.436	−0.052	−0.002	−0.120	−0.017
β_2	0.064	0.119	−0.025	0.257	−0.164	−0.008	−0.142	−0.088
$\beta_{\mathcal{G}_2}$	−0.072	−0.331	0.040	−0.100	0.221	0.010	0.147	0.111
β_3	0.000	−0.012	−0.015	−0.032	0.022	0.011	0.083	0.002
$\beta_{KK_\parallel}$	−0.152	−0.594	0.078	0.727	−0.223	−0.078	−1.337	0.893
β_η	−0.286	−0.044	0.323	−0.752	0.396	−0.116	0.509	−0.143
β_{η^2}	−0.011	−0.001	−0.164	0.476	−0.113	0.049	−0.048	−0.045
$\beta_{\delta\eta}$	0.043	0.068	0.039	−0.273	−0.364	−0.037	0.275	0.037

TABLE XI. **Coefficients of Best-Fit Abacus Transfer Functions Models:** Same as Tab. IV for models II-IV.

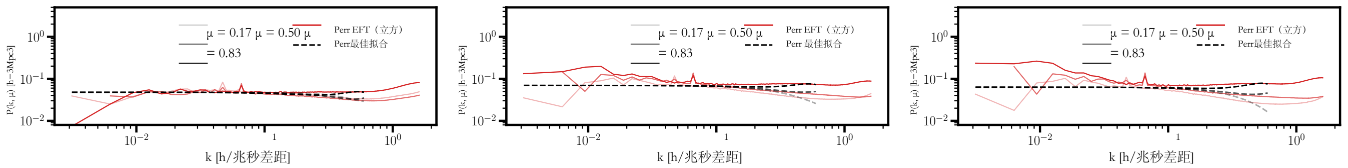

图29. 最佳拟合Perr的模型依赖性：与图28相同，显示模型二、三和四的误差功率谱的最佳拟合模型。拟合使用 kmax = 0.6 h Mpc −1，并显示在约 0.2–0.3 h Mpc−1 处误差功率谱尺度和方向依赖性的轻微出现。

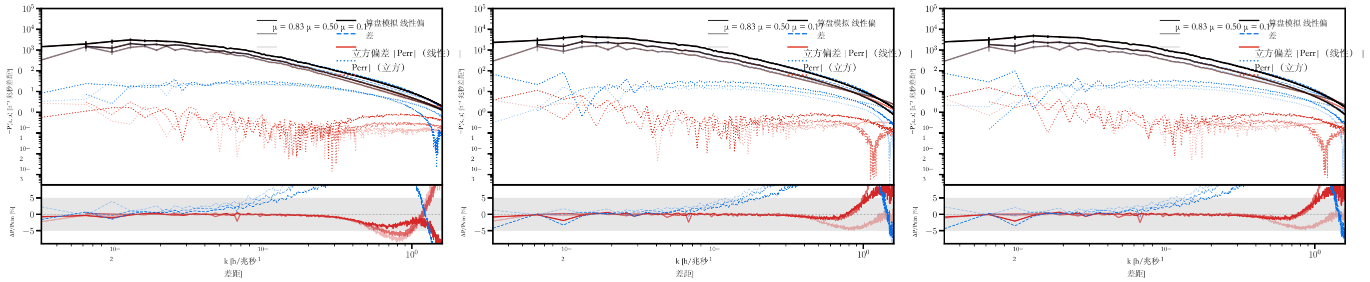

图30. 功率谱的模型依赖性：与图5相同，对于三种不同的FGPA实现，分别用模型II-IV表示。

TF	c_0	c_{01}	c_1	c_{12}	c_{14}	c_4	c_{22}	c_{44}
型号 II								
β1	−0.134	−0.142	0.073	0.253	−0.052	−0.008	−0.156	0.055
β_2	0.097	0.060	−0.017	0.074	−0.080	−0.018	−0.058	−0.045
$\beta_{\mathcal{G}_2}$	0.075	0.151	0.018					
β_3	−0.006	−0.011	−0.016	−0.108	0.087	0.016	0.092	0.008
$\beta_{KK_\parallel}$	−0.018	−0.152	−0.181	0.563	−0.659	−0.056	−0.897	0.841
β_η				−	0.101	−0.011	−0.043	0.178
β_{η^2}	0.064	−0.034	−0.098	0.071				0.036
$\beta_{\delta\eta}$	−	0.005	−0.016	−0.244	−0.008	−0.007	0.478	−0.296
模型 III								
β1	−0.139	−0.295	0.050	0.407	−0.057	−0.002	−0.099	−0.022
β_2	0.060	0.105	−0.022	0.244	−0.160	−0.008	−0.129	−0.082
$\beta_{\mathcal{G}_2}$	0.072	0.300	0.041					
β_3	−0.002	−0.012	−0.012	−0.020	0.018	0.011	0.074	−0.003
$\beta_{KK_\parallel}$				0.461	−0.137	−0.093	−1.051	0.777
β_η	−0.252	−0.046	0.265	−				0.030
β_{η^2}	−0.032	0.010	−0.163	0.415	0.076	0.052	−0.008	−0.073
$\beta_{\delta\eta}$	0.045	0.054	0.027	−0.211	−0.290	−0.032		
型号 IV								
β1	−0.135	−0.329	0.051	0.436	−0.052	−0.002	−0.120	−0.017
β_2	0.064	0.119	−0.025	0.257	−0.164	−0.008	−0.142	−0.088
$\beta_{\mathcal{G}_2}$	−0.072	0.331	0.040					
β_3	0.000	−0.012	−0.015	−0.032	0.022	0.011	0.083	0.002
$\beta_{KK_\parallel}$	−			0.727	−0.223	−0.078	−1.337	0.893
β_η	0.286	0.044	0.323	−				0.143
β_{η^2}	−0.011	−0.001	−0.164	0.476	0.113	0.049	−0.048	−0.045
$\beta_{\delta\eta}$			0.039	−0.273	−0.364	−0.037		

表 XI. 最佳拟合算筹传递函数模型系数：与表 IV 相同，适用于模型 II-IV。

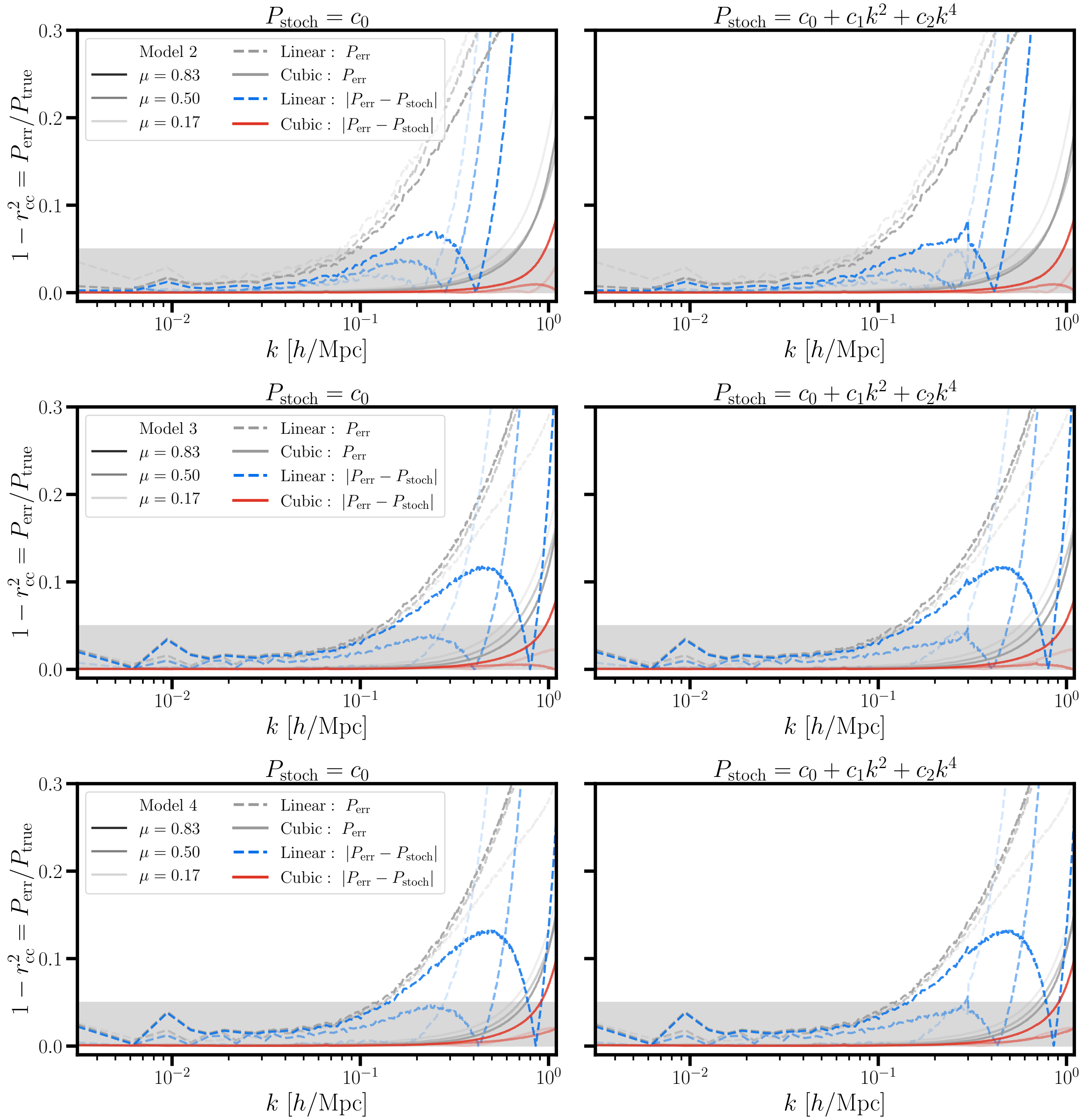

FIG. 31. **Model-dependence of $1 - r_{\rm cc}^2$:** Same as Fig. 10 showing the cross-correlation coefficient for models II-IV.

---

[1] S. S. Vogt, S. L. Allen, B. C. Bigelow, L. Bresee, B. Brown, T. Cantrall et al., *HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope*, in *Instrumentation in Astronomy VIII*, D. L. Crawford and E. R. Craine, eds., vol. 2198 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 362, June, 1994, DOI.

[2] J. M. O'Meara, N. Lehner, J. C. Howk and J. X. Prochaska, *The Third Data Release of the KODIAQ Survey*, AJ **161** (2021) 45 [2010.09061].

[3] H. Dekker, S. D'Odorico, A. Kaufer, B. Delabre and H. Kotzlowski, *Design, construction, and performance of UVES, the echelle spectrograph for the UT2 Kueyen Telescope at the ESO Paranal Observatory*, in *Optical and IR Telescope*

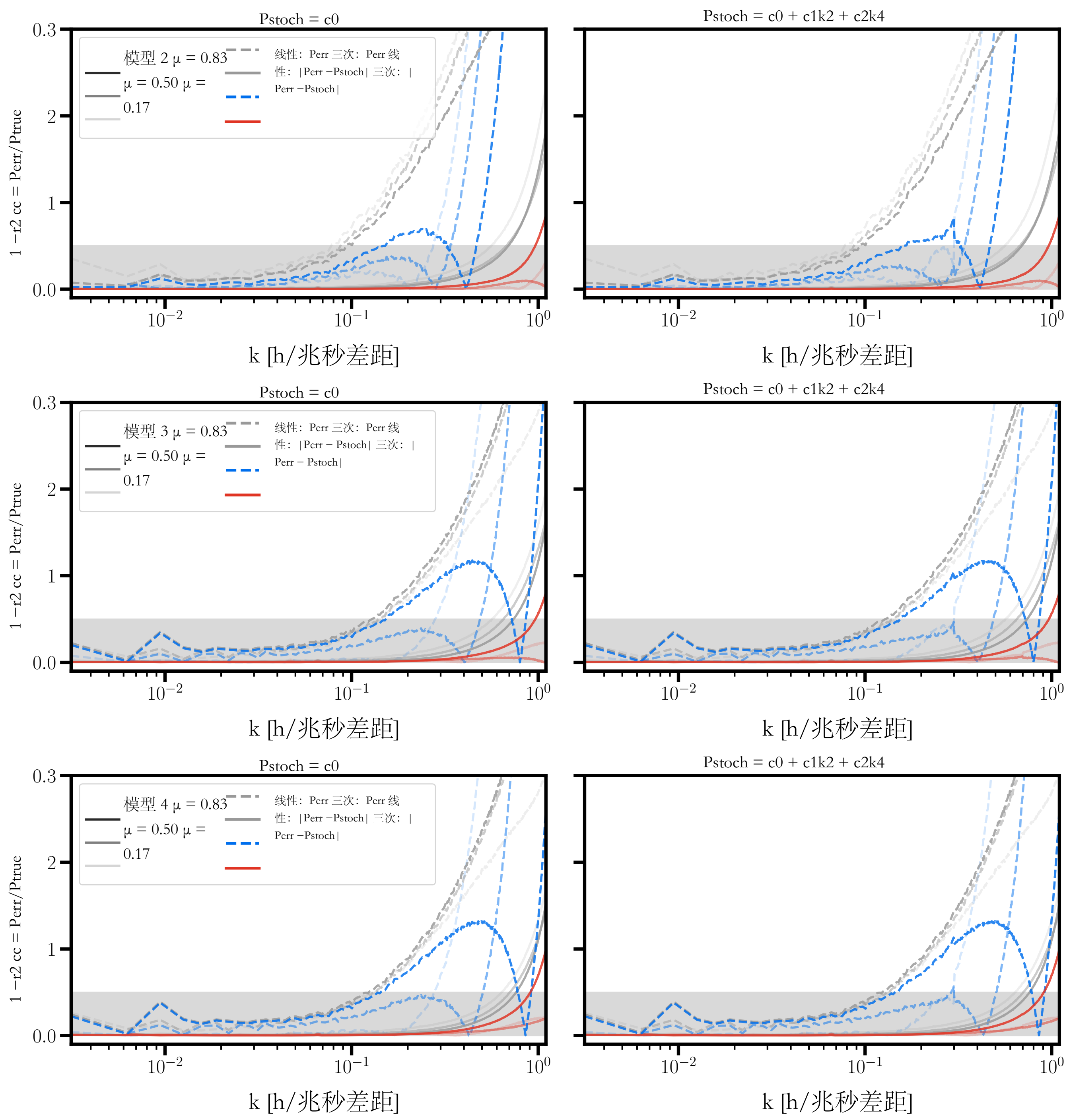

图 31. 1 − r^2_cc 的模型依赖性：与图 10 相同，显示模型 II-IV 的交叉相关系数。

---

[1] S. S. Vogt, S. L. Allen, B. C. Bigelow, L. Bresee, B. Brown, T. Cantrall 等, HIRES：凯克10米望远镜上的高分辨率艾歇尔光谱仪, 收录于《天文学仪器 VIII》中, D. L. Crawford 和 E. R. Craine 编, 社会光学与光电子仪器工程师协会（SPIE）会议系列第2198卷, 第362页, 1994年6月, DOI. [2] J. M. O'Meara, N. Lehner, J. C. Howk 和 J. X. Prochaska, KODIAQ 调查的第三次数据发布, AJ 161

(2021) 45 [2010.09061].

[3] H. Dekker, S. D'Odorico, A. Kaufer, B. Delabre 和 H. Kotzlowski，UVES 的设计、建造及性能，UVES 是 ESO Paranal 天文台 UT2 Kueyen 望远镜的惠普光栅光谱仪，载于《光学与红外望远镜》。

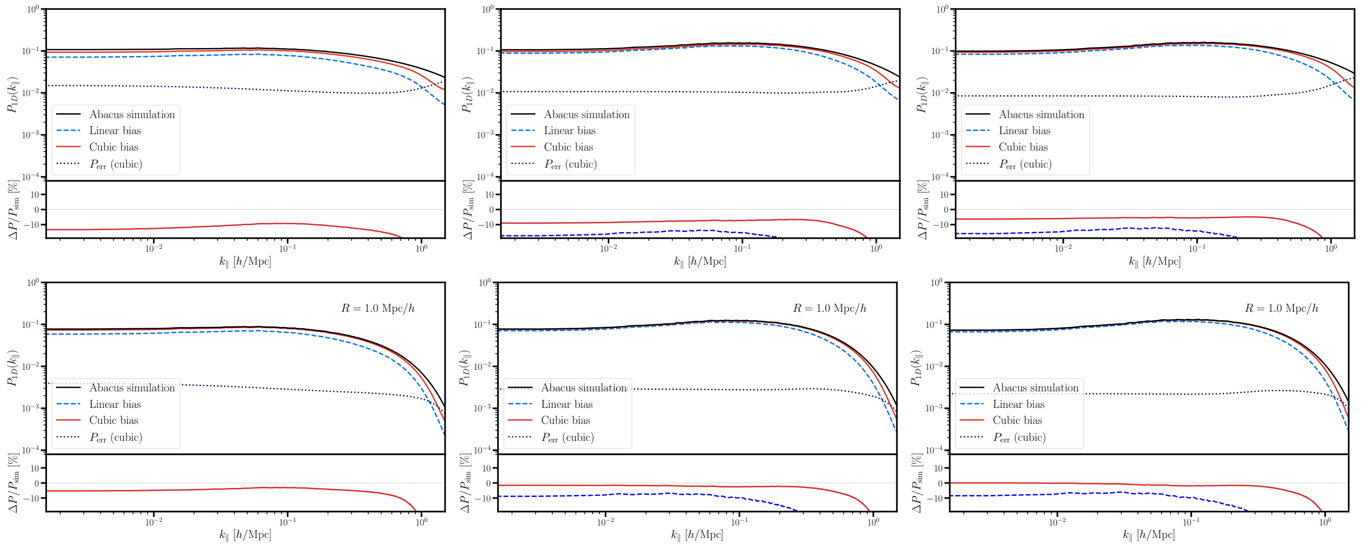

FIG. 32. Model-dependence of P1D: Same as Fig. 28 showing the one-dimensional power spectrum (P1D) for models II, III and IV (from left to right panel) comparing the smoothing scale $R = 1\,h^{-1}$ Mpc with the case of no smoothing.

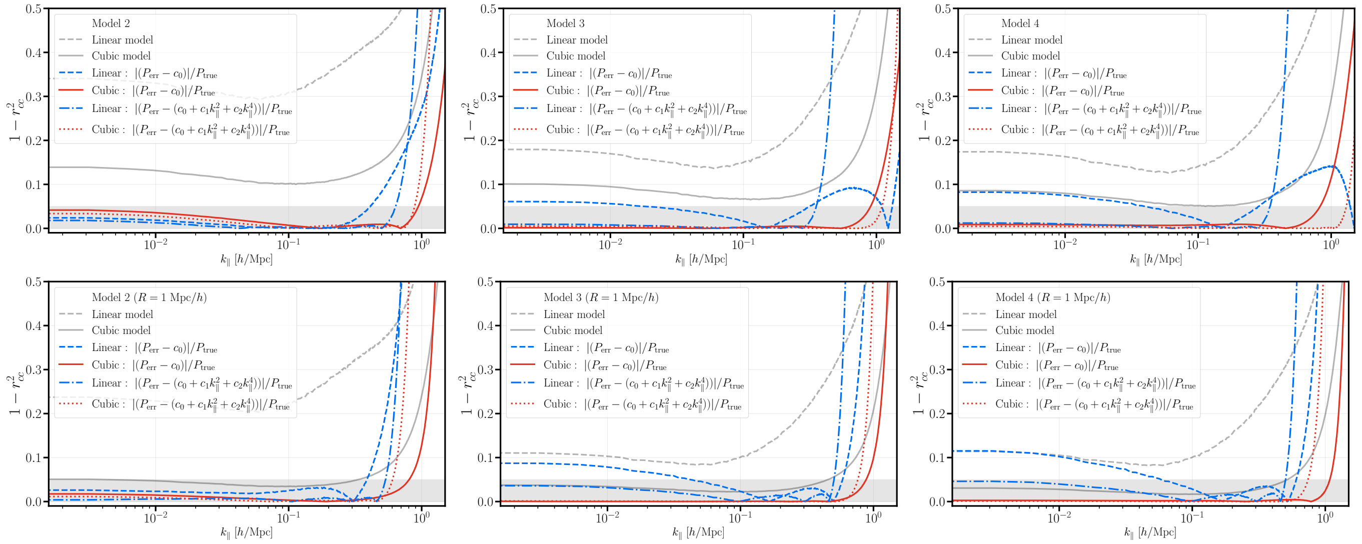

FIG. 33. Model-dependence of Cross-Correlation Coefficient: Same as Fig. 28 showing the cross-correlation coefficients for models two, three and four.

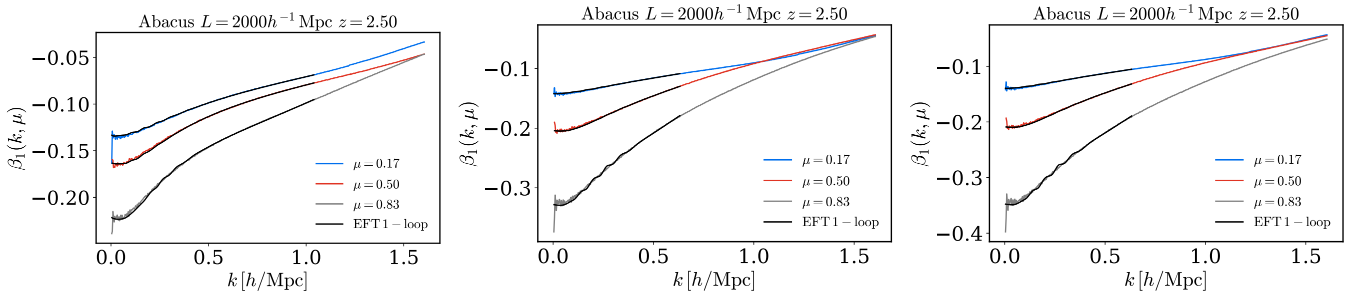

FIG. 34. Best-fit EFT TF Abacus: Same as Fig. 19 for models II-IV for Abacus from left to right. Note that models III and IV use a lower $k_{\rm max} = 0.75\,h\,{\rm Mpc}^{-1}$ for the fits. The perturbative transfer functions fit the data to high accuracy.

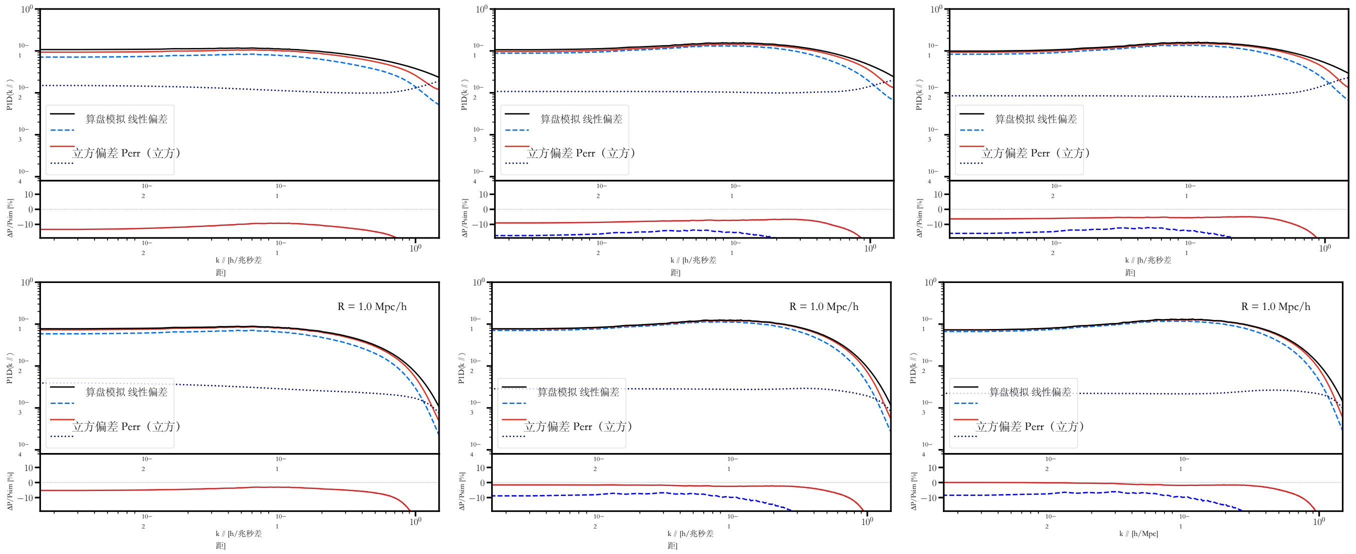

图32. P1D的模型依赖性：与图28相同，显示模型II、III和IV（从左到右面板）的单维功率谱（P1D），比较平滑尺度R = 1 h^{-1}Mpc与未平滑情况。

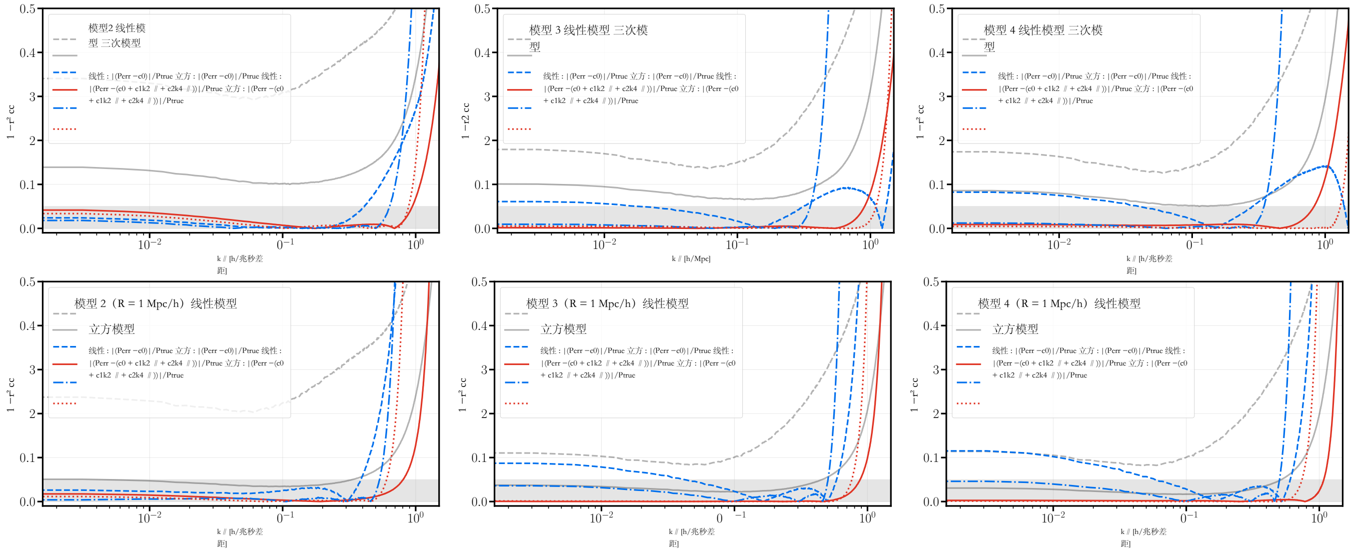

图33. 互相关系数的模型依赖性：与图28相同，显示模型二、三和四的互相关系数。

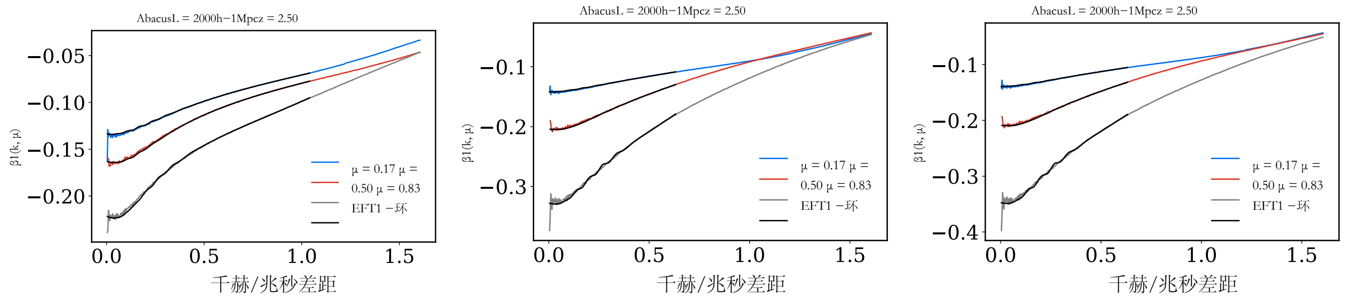

图34. 最佳拟合EFT TF Abacus：与图19相同，显示Abacus模型II-IV，从左到右。注意模型III和IV在拟合中使用了较低的kmax = 0.75 h Mpc −1。微扰转移函数对数据的拟合具有很高的精度。

Instrumentation and Detectors, M. Iye and A. F. Moorwood, eds., vol. 4008 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 534–545, Aug., 2000, DOI.

[4] M. T. Murphy, G. G. Kacprzak, G. A. D. Savorgnan and R. F. Carswell, *The UVES Spectral Quasar Absorption Database (SQUAD) data release 1: the first 10 million seconds*, MNRAS **482** (2019) 3458 [1810.06136].

[5] U. Seljak, A. Makarov, P. McDonald, S. F. Anderson, N. A. Bahcall, J. Brinkmann et al., *Cosmological parameter analysis including SDSS Lyα forest and galaxy bias: Constraints on the primordial spectrum of fluctuations, neutrino mass, and dark energy*, Phys. Rev. D **71** (2005) 103515 [astro-ph/0407372].

[6] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese and A. Riotto, *Constraining warm dark matter candidates including sterile neutrinos and light gravitinos with WMAP and the Lyman-α forest*, Phys. Rev. D **71** (2005) 063534 [astro-ph/0501562].

[7] P. McDonald, U. Seljak, S. Burles, D. J. Schlegel, D. H. Weinberg, R. Cen et al., *The Lyα Forest Power Spectrum from the Sloan Digital Sky Survey*, ApJS **163** (2006) 80 [astro-ph/0405013].

[8] N. Palanque-Delabrouille, C. Yèche, A. Borde, J.-M. Le Goff, G. Rossi, M. Viel et al., *The one-dimensional Lyα forest power spectrum from BOSS*, A&A **559** (2013) A85 [1306.5896].

[9] S. Chabanier, N. Palanque-Delabrouille, C. Yèche, J.-M. Le Goff, E. Armengaud, J. Bautista et al., *The one-dimensional power spectrum from the SDSS DR14 Lyα forests*, J. Cosmology Astropart. Phys. **2019** (2019) 017 [1812.03554].

[10] C. Pedersen, A. Font-Ribera, K. K. Rogers, P. McDonald, H. V. Peiris, A. Pontzen et al., *An emulator for the Lyman-α forest in beyond-ΛCDM cosmologies*, *JCAP* **05** (2021) 033 [2011.15127].

[11] C. Ravoux, M. L. Abdul Karim, E. Armengaud, M. Walther, N. G. Karaçaylı, P. Martini et al., *The Dark Energy Spectroscopic Instrument: one-dimensional power spectrum from first Ly α forest samples with Fast Fourier Transform*, MNRAS **526** (2023) 5118 [2306.06311].

[12] N. G. Karaçaylı, P. Martini, J. Guy, C. Ravoux, M. L. A. Karim, E. Armengaud et al., *Optimal 1D Lyα Forest Power Spectrum Estimation - III. DESI early data*, MNRAS (2024) [2306.06316].

[13] M. Viel, M. G. Haehnelt and V. Springel, *The effect of neutrinos on the matter distribution as probed by the intergalactic medium*, J. Cosmology Astropart. Phys. **2010** (2010) 015 [1003.2422].

[14] N. Palanque-Delabrouille, C. Yèche, N. Schöneberg, J. Lesgourgues, M. Walther, S. Chabanier et al., *Hints, neutrino bounds, and WDM constraints from SDSS DR14 Lyman-α and Planck full-survey data*, J. Cosmology Astropart. Phys. **2020** (2020) 038 [1911.09073].

[15] M. M. Ivanov, M. W. Toomey and N. G. Karaçaylı, *Fundamental Physics with the Lyman-Alpha Forest: Constraints on the Growth of Structure and Neutrino Masses from SDSS with Effective Field Theory*, *Phys. Rev. Lett.* **134** (2025) 091001 [2405.13208].

[16] A. He, R. An, M. M. Ivanov and V. Gluscevic, *Self-Interacting Neutrinos in Light of Large-Scale Structure Data*, 2309.03956.

[17] A. He, M. M. Ivanov, S. Bird, R. An and V. Gluscevic, *A Fresh Look at Neutrino Self-Interactions With the Lyman-α Forest: Constraints from EFT and PRIYA*, 2503.15592.

[18] N. Afshordi, P. McDonald and D. N. Spergel, *Primordial Black Holes as Dark Matter: The Power Spectrum and Evaporation of Early Structures*, ApJ **594** (2003) L71 [astro-ph/0302035].

[19] R. Murgia, G. Scelfo, M. Viel and A. Raccanelli, *Lyman-α Forest Constraints on Primordial Black Holes as Dark Matter*, *Phys. Rev. Lett.* **123** (2019) 071102 [1903.10509].

[20] M. M. Ivanov and S. Trifinopoulos, *Effective Field Theory Constraints on Primordial Black Holes from the High-Redshift Lyman-α Forest*, 2508.04767.

[21] M. Viel, G. D. Becker, J. S. Bolton and M. G. Haehnelt, *Warm dark matter as a solution to the small scale crisis: New constraints from high redshift Lyman-α forest data*, Phys. Rev. D **88** (2013) 043502 [1306.2314].

[22] J. Baur, N. Palanque-Delabrouille, C. Yèche, C. Magneville and M. Viel, *Lyman-alpha forests cool warm dark matter*, J. Cosmology Astropart. Phys. **2016** (2016) 012 [1512.01981].

[23] V. Iršič, M. Viel, M. G. Haehnelt, J. S. Bolton, S. Cristiani, G. D. Becker et al., *New Constraints on the free-streaming of warm dark matter from intermediate and small scale Lyman-α forest data*, *ArXiv e-prints* (2017) [1702.01764].

[24] T. Kobayashi, R. Murgia, A. De Simone, V. Iršič and M. Viel, *Lyman-α constraints on ultralight scalar dark matter: Implications for the early and late universe*, Phys. Rev. D **96** (2017) 123514 [1708.00015].

[25] E. Armengaud, N. Palanque-Delabrouille, C. Yèche, D. J. E. Marsh and J. Baur, *Constraining the mass of light bosonic dark matter using SDSS Lyman-α forest*, MNRAS **471** (2017) 4606 [1703.09126].

[26] R. Murgia, V. Iršič and M. Viel, *Novel constraints on noncold, nonthermal dark matter from Lyman-α forest data*, Phys. Rev. D **98** (2018) 083540 [1806.08371].

[27] A. Garzilli, A. Magalich, T. Theuns, C. S. Frenk, C. Weniger, O. Ruchayskiy et al., *The Lyman-α forest as a diagnostic of the nature of the dark matter*, MNRAS **489** (2019) 3456 [1809.06585].

[28] V. Iršič, H. Xiao and M. McQuinn, *Early structure formation constraints on the ultralight axion in the postinflation scenario*, Phys. Rev. D **101** (2020) 123518 [1911.11150].

[29] K. K. Rogers, C. Dvorkin and H. V. Peiris, *Limits on the Light Dark Matter-Proton Cross Section from Cosmic Large-Scale Structure*, Phys. Rev. Lett. **128** (2022) 171301 [2111.10386].

[30] B. Villasenor, B. Robertson, P. Madau and E. Schneider, *New constraints on warm dark matter from the Lyman-α forest power spectrum*, Phys. Rev. D **108** (2023) 023502 [2209.14220].

[31] V. Iršič, M. Viel, M. G. Haehnelt, J. S. Bolton, M. Molaro, E. Puchwein et al., *Unveiling Dark Matter free-streaming at the smallest scales with high redshift Lyman-alpha forest*, *arXiv e-prints* (2023) arXiv:2309.04533 [2309.04533].

仪器与探测器，M. Iye 和 A. F. Moorwood 编辑，第 4008 卷，光学工程学会（SPIE）会议系列，第 534 – 545 页，2000 年 8 月，DOI。

[4] M. T. Murphy, G. G. Kacprzak, G. A. D. Savorgnan 和 R. F. Carswell, 《UVES 光谱类星体吸收数据库（SQUAD）数据发布 1：前 1000 万秒》, MNRAS 482 (2019) 3458 [1810.06136].

[5] U. Seljak, A. Makarov, P. McDonald, S. F. Anderson, N. A. Bahcall, J. Brinkmann 等人，宇宙学参数分析，包括 SDSS Lyα 森林和星系偏差：对初始涨落光谱、中微子的约束 **质量和暗能量，物理评论 D 71 (2005) 103515 [astro-ph/0407372]。**

[6] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese 和 A. Riotto, 利用 WMAP 和莱曼-α 森林约束包括无味中微子和轻引力子在内的温暗物质候选者, 物理评论 D 71 (2005) 063534 [astro-ph/0501562].

[7] P. McDonald, U. Seljak, S. Burles, D. J. Schlegel, D. H. Weinberg, R. Cen 等，《来自 Lyα 森林的功率谱》 **斯隆数字天空调查，ApJS 163 (2006) 80 [astro-ph/0405013]。**

[8] N. Palanque-Delabrouille, C. Y` eche, A. Borde, J.-M. Le Goff, G. Rossi, M. Viel 等，《一维 Lyα 森林》 **来自 BOSS 的功率谱，A&A 559 (2013) A85 [1306.5896]。**

[9] S. Chabanier, N. Palanque-Delabrouille, C. Y` eche, J.-M. Le Goff, E. Armengaud, J. Bautista 等，《来自 SDSS DR14 Lyα 森林的一维功率谱》，《宇宙学与天体粒子物理学杂志》2019 (2019) 017 [1812.03554]。

[10] C. Pedersen, A. Font-Ribera, K. K. Rogers, P. McDonald, H. V. Peiris, A. Pontzen 等，Lyman-α 的仿真器 **超越ΛCDM宇宙学中的森林，JCAP 05 (2021) 033 [2011.15127]。**

[11] C. Ravoux, M. L. Abdul Karim, E. Armengaud, M. Walther, N. G. Kara¸ caylı, P. Martini 等, 暗能量光谱仪：利用快速傅里叶变换从首批Ly α 森林样本提取的一维功率谱 **MNRAS 526 (2023) 5118 [2306.06311]。**

[12] N. G. Kara¸ caylı, P. Martini, J. Guy, C. Ravoux, M. L. A. Karim, E. Armengaud 等，最佳一维 Lyα 森林功率频谱估计 - III. DESI 早期数据, MNRAS (2024) [2306.06316].

[13] M. Viel, M. G. Haehnelt 和 V. Springel，中微子对通过……探测的物质分布的影响 **星际介质，J. 宇宙学与天体物理学 2010 (2010) 015 [1003.2422]。**

[14] N. Palanque-Delabrouille, C. Y` eche, N. Schöneberg, J. Lesgourgues, M. Walther, S. Chabanier 等人, 来自 SDSS DR14 Lyman-α 和 Planck 全测量数据的提示、中微子界限及 WDM 限制，《宇宙学与天体粒子物理学杂志》 **2020 (2020) 038 [1911.09073].**

[15] M. M. Ivanov, M. W. Toomey 和 N. G. Kara¸ caylı, 利用莱曼-α森林的基础物理学：通过有效场论从SDSS对结构生长和中微子质量的约束, 物理评论快报 134 (2025) 091001 [2405.13208].

[16] A. He, R. An, M. M. Ivanov 和 V. Gluscevic，《结合大尺度结构数据的自相互作用中微子》，2309.03956.

[17] A. He, M. M. Ivanov, S. Bird, R. An 和 V. Gluscevic, 用 Lyman-α 对中微子自相互作用的新研究森林：来自EFT和PRIYA的约束，2503.15592。

[18] N. Afshordi, P. McDonald 和 D. N. Spergel，原初黑洞作为暗物质：功率谱和 **早期结构的蒸发，天体物理学期刊 594 (2003) L71 [astro-ph/0302035]。**

[19] R. Murgia, G. Scelfo, M. Viel 和 A. Raccanelli，赖曼-α森林对原初黑洞作为暗物质的约束 **Matter, Phys. Rev. Lett. 123 (2019) 071102 [1903.10509].**

[20] M. M. 伊万诺夫 和 S. 特里菲诺普洛斯，《来自高红移的原初黑洞有效场论约束》莱曼-α 森林，2508.04767。

[21] M. Viel, G. D. Becker, J. S. Bolton 和 M. G. Haehnelt，温暗物质作为小尺度危机的解决方案：新 **来自高红移莱曼-α森林数据的约束，物理评论 D 88 (2013) 043502 [1306.2314]。**

[22] J. Baur, N. Palanque-Delabrouille, C. Y` eche, C. Magneville 和 M. Viel, 莱曼α森林冷暖暗物质, J. **宇宙学与天体物理学 2016 (2016) 012 [1512.01981]。**

[23] V. Ir ˇ si ˇ c, M. Viel, M. G. Haehnelt, J. S. Bolton, S. Cristiani, G. D. Becker 等, 来自中等和小尺度莱曼-α森林数据对温暗物质自由流动的新约束, ArXiv 电子预印本 (2017) [1702.01764]。 [24] T. Kobayashi, R. Murgia, A. De Simone, V. Ir ˇ si ˇ c 和 M. Viel, 莱曼-α 对超轻标量暗物质的约束：

早期和晚期宇宙的影响，物理评论 D 96 (2017) 123514 [1708.00015]。

[25] E. Armengaud, N. Palanque-Delabrouille, C. Y` eche, D. J. E. Marsh 和 J. Baur，限制轻玻色子的质量 **使用 SDSS Lyman-α 森林研究暗物质，MNRAS 471 (2017) 4606 [1703.09126]。**

[26] R. Murgia, V. Ir ˇ si ˇ c 和 M. Viel，来自莱曼-α森林数据的非冷非热暗物质的新约束, **物理评论 D 98 (2018) 083540 [1806.08371]。**

[27] A. Garzilli, A. Magalich, T. Theuns, C. S. Frenk, C. Weniger, O. Ruchayskiy 等，《Lyman-α 森林作为诊断工具》 **关于暗物质的性质，MNRAS 489 (2019) 3456 [1809.06585]。**

[28] V. Ir ˇ si ˇ c, H. Xiao 和 M. McQuinn，关于超轻轴子在后暴胀时期的早期结构形成约束 **情景，Phys. Rev. D 101 (2020) 123518 [1911.11150]。**

[29] K. K. Rogers, C. Dvorkin 和 H. V. Peiris，从宇宙中对轻暗物质-质子截面的限制 **大尺度结构，《物理评论快报》128 (2022) 171301 [2111.10386]。**

[30] B. Villasenor, B. Robertson, P. Madau 和 E. Schneider，来自莱曼-α的关于温暗物质的新约束 **森林功率谱，物理评论 D 108 (2023) 023502 [2209.14220]。**

[31] V. Ir ˇ si ˇ c, M. Viel, M. G. Haehnelt, J. S. Bolton, M. Molaro, E. Puchwein 等，高红移莱曼-α森林揭示最小尺度的暗物质自由流动，arXiv 电子预印本 (2023) arXiv:2309.04533 [2309.04533]。

[32] M. Zaldarriaga, *Searching for Fluctuations in the Intergalactic Medium Temperature Using the Lyα Forest*, ApJ **564** (2002) 153 [astro-ph/0102205].

[33] A. A. Meiksin, *The physics of the intergalactic medium*, *Reviews of Modern Physics* **81** (2009) 1405 [0711.3358].

[34] M. McQuinn, *The Evolution of the Intergalactic Medium*, ARA&A **54** (2016) 313 [1512.00086].

[35] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese and A. Riotto, *Can Sterile Neutrinos Be Ruled Out as Warm Dark Matter Candidates?*, Phys. Rev. Lett. **97** (2006) 071301 [astro-ph/0605706].

[36] M. Walther, J. Oñorbe, J. F. Hennawi and Z. Lukić, *New Constraints on IGM Thermal Evolution from the Lyα Forest Power Spectrum*, ApJ **872** (2019) 13 [1808.04367].

[37] J. S. Bolton, M. Viel, T. S. Kim, M. G. Haehnelt and R. F. Carswell, *Possible evidence for an inverted temperature-density relation in the intergalactic medium from the flux distribution of the Lyα forest*, MNRAS **386** (2008) 1131 [0711.2064].

[38] A. Garzilli, J. S. Bolton, T. S. Kim, S. Leach and M. Viel, *The intergalactic medium thermal history at redshift $z = 1.7$-3.2 from the Lyα forest: a comparison of measurements using wavelets and the flux distribution*, MNRAS **424** (2012) 1723 [1202.3577].

[39] P. Gaikwad, R. Srianand, V. Khaire and T. R. Choudhury, *Effect of non-equilibrium ionization on derived physical conditions of the high-z intergalactic medium*, MNRAS **490** (2019) 1588 [1812.01016].

[40] E. Boera, G. D. Becker, J. S. Bolton and F. Nasir, *Revealing Reionization with the Thermal History of the Intergalactic Medium: New Constraints from the Lyα Flux Power Spectrum*, ApJ **872** (2019) 101 [1809.06980].

[41] P. Gaikwad, R. Srianand, M. G. Haehnelt and T. R. Choudhury, *A consistent and robust measurement of the thermal state of the IGM at $2 \leq z \leq 4$ from a large sample of Ly α forest spectra: evidence for late and rapid He II reionization*, MNRAS **506** (2021) 4389 [2009.00016].

[42] B. Wilson, V. Iršič and M. McQuinn, *A measurement of the Ly β forest power spectrum and its cross with the Ly α forest in X-Shooter XQ-100*, MNRAS **509** (2022) 2423 [2106.04837].

[43] B. Villasenor, B. Robertson, P. Madau and E. Schneider, *Inferring the Thermal History of the Intergalactic Medium from the Properties of the Hydrogen and Helium Lyα Forest*, ApJ **933** (2022) 59 [2111.00019].

[44] S. Goldstein, J. C. Hill, V. Iršič and B. D. Sherwin, *Canonical Hubble-Tension-Resolving Early Dark Energy Cosmologies are Inconsistent with the Lyman-α Forest*, 2303.00746.

[45] L. Fuß and M. Garny, *Decaying Dark Matter and Lyman-α forest constraints*, 2210.06117.

[46] M. Garny, T. Konstandin, L. Sagunski and S. Tulin, *Lyman-α forest constraints on interacting dark sectors*, *JCAP* **09** (2018) 011 [1805.12203].

[47] U. Seljak, A. Slosar and P. McDonald, *Cosmological parameters from combining the Lyman-alpha forest with CMB, galaxy clustering and SN constraints*, *JCAP* **10** (2006) 014 [astro-ph/0604335].

[48] K. S. Dawson, J.-P. Kneib, W. J. Percival, S. Alam, F. D. Albareti, S. F. Anderson et al., *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data*, AJ **151** (2016) 44 [1508.04473].

[49] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen et al., *The DESI Experiment Part I: Science,Targeting, and Survey Design*, *arXiv e-prints* (2016) arXiv:1611.00036 [1611.00036].

[50] DESI Collaboration, A. G. Adame, J. Aguilar, S. Ahlen, S. Alam, D. M. Alexander et al., *DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations*, *arXiv e-prints* (2024) arXiv:2404.03002 [2404.03002].

[51] DESI Collaboration, A. G. Adame, J. Aguilar, S. Ahlen, S. Alam, D. M. Alexander et al., *DESI 2024 IV: Baryon Acoustic Oscillations from the Lyman Alpha Forest*, *arXiv e-prints* (2024) arXiv:2404.03001 [2404.03001].

[52] DESI collaboration, *DESI DR2 results. I. Baryon acoustic oscillations from the Lyman alpha forest*, Phys. Rev. D **112** (2025) 083514 [2503.14739].

[53] A. Cuceu, A. Font-Ribera, B. Joachimi and S. Nadathur, *Cosmology beyond BAO from the 3D distribution of the Lyman-α forest*, MNRAS **506** (2021) 5439 [2103.14075].

[54] F. Gerardi, A. Cuceu, A. Font-Ribera, B. Joachimi and P. Lemos, *Direct cosmological inference from three-dimensional correlations of the Lyman α forest*, *Mon. Not. Roy. Astron. Soc.* **518** (2022) 2567 [2209.11263].

[55] P. McDonald and D. J. Eisenstein, *Dark energy and curvature from a future baryonic acoustic oscillation survey using the Lyman-α forest*, Phys. Rev. D **76** (2007) 063009 [astro-ph/0607122].

[56] A. Slosar, V. Iršič, D. Kirkby, S. Bailey, N. G. Busca, T. Delubac et al., *Measurement of baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9*, J. Cosmology Astropart. Phys. **2013** (2013) 026 [1301.3459].

[57] N. G. Busca, T. Delubac, J. Rich, S. Bailey, A. Font-Ribera, D. Kirkby et al., *Baryon acoustic oscillations in the Lyα forest of BOSS quasars*, A&A **552** (2013) A96 [1211.2616].

[58] H. du Mas des Bourboux, J. Rich, A. Font-Ribera, V. de Sainte Agathe, J. Farr, T. Etourneau et al., *The Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations with Lyα Forests*, ApJ **901** (2020) 153 [2007.08995].

[59] A. Cuceu, A. Font-Ribera, S. Nadathur, B. Joachimi and P. Martini, *Constraints on the Cosmic Expansion Rate at Redshift 2.3 from the Lyman-α Forest*, Phys. Rev. Lett. **130** (2023) 191003 [2209.13942].

[60] C. Gordon, A. Cuceu, J. Chaves-Montero, A. Font-Ribera, A. Xochitl González-Morales, J. Aguilar et al., *3D Correlations in the Lyman-α Forest from Early DESI Data*, *arXiv e-prints* (2023) arXiv:2308.10950 [2308.10950].

[61] A. Cuceu et al., *DESI DR1 Lyα forest: 3D full-shape analysis and cosmological constraints*, 2509.15308.

[62] R. Cen, J. Miralda-Escudé, J. P. Ostriker and M. Rauch, *Gravitational Collapse of Small-Scale Structure as the Origin of the Lyman-Alpha Forest*, ApJ **437** (1994) L9 [astro-ph/9409017].

[32] M. Zaldarriaga，使用莱曼α森林搜索星际介质温度的波动，天体物理杂志 564
(2002) 153 [astro-ph/0102205].

[33] A. A. Meiksin，《银河际介质的物理学》，《现代物理评论》81 (2009) 1405 [0711.3358]。[34] M. McQuinn，《银河际介质的演化》，《天文与天体物理年评》54 (2016) 313 [1512.00086]。

[35] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese 和 A. Riotto，绝育中微子能否被排除作为温暖
暗物质候选者？, 物理评论快报 97 (2006) 071301 [astro-ph/0605706].

[36] M. Walther, J. O˜ norbe, J. F. Hennawi 和 Z. Luki´ c，通过 Lyα 森林对IGM热演化的新约束
功率谱，天体物理学杂志 872 (2019) 13 [1808.04367]。

[37] J. S. Bolton, M. Viel, T. S. Kim, M. G. Haehnelt 和 R. F. Carswell，通过 Lyα 森林的通量分布，可能发现了星际介质中温度-密度关系倒置的证据，MNRAS 386
(2008) 1131 [0711.2064].

[38] A. Garzilli, J. S. Bolton, T. S. Kim, S. Leach 和 M. Viel, 从 Lyα 森林研究红移 z = 1.7-3.2 的星际介质热历史：使用小波和通量分布测量的比较，MNRAS 424
(2012) 1723 [1202.3577].

[39] P. Gaikwad, R. Srianand, V. Khaire 和 T. R. Choudhury，非平衡电离对推导物理量的影响
高红移星际介质的条件，MNRAS 490 (2019) 1588 [1812.01016]。

[40] E. Boera, G. D. Becker, J. S. Bolton 和 F. Nasir，通过星际介质的热历史揭示再电离
媒介：来自 Lyα 通量功率谱的新约束，ApJ 872 (2019) 101 [1809.06980]。

[41] P. Gaikwad、R. Srianand、M. G. Haehnelt 和 T. R. Choudhury，通过大量 Ly α 森林光谱样本对 2 ⩽ z ⩽ 4 范围内 IGM 热状态进行一致且稳健的测量：晚期且快速 He II 再电离的证据，MNRAS 506 (2021) 4389 [2009.00016]。 [42] B. Wilson、V. Ir ˇ si ˇ c 和 M. McQuinn，Ly β 森林功率谱及其与 Ly α 的交叉测量
X-Shooter XQ-100 中的森林，MNRAS 509 (2022) 2423 [2106.04837]。

[43] B. Villasenor, B. Robertson, P. Madau 和 E. Schneider，推断星际介质的热历史
来自氢和氦 Lyα 森林的性质，ApJ 933 (2022) 59 [2111.00019]。[44] S. Goldstein, J. C. Hill, V. Ir ˇ si ˇ c 和 B. D. Sherwin, 标准哈勃张力解决早期暗能量
宇宙学与莱曼-α森林不一致，2303.00746。

[45] L. Fuß 和 M. Garny，衰变暗物质与莱曼-α森林约束，2210.06117。

[46] M. Garny, T. Konstandin, L. Sagunski 和 S. Tulin，关于相互作用暗部门的莱曼-α森林约束，JCAP 09
(2018) 011 [1805.12203].

[47] U. Seljak, A. Slosar 和 P. McDonald，通过结合莱曼α森林与宇宙微波背景获得的宇宙学参数，
星系聚类和超新星约束，JCAP 10 (2006) 014 [astro-ph/0604335]。

[48] K. S. Dawson, J.-P. Kneib, W. J. Percival, S. Alam, F. D. Albareti, S. F. Anderson 等， 《SDSS-IV 扩展重子振荡光谱调查：概览与早期数据》 ，天文杂志 151 (2016) 44 [1508.04473]。

[49] DESI 合作组, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen 等人, DESI 实验第 I 部分:
科学、目标和调查设计，arXiv 电子预印本 (2016) arXiv:1611.00036 [1611.00036]。

[50] DESI 合作组， A. G. Adame, J. Aguilar, S. Ahlen, S. Alam, D. M. Alexander 等，DESI 2024 VI：来自重子声学振荡测量的宇宙学约束，arXiv 电子预印本 (2024) arXiv:2404.03002
[2404.03002].

[51] DESI 合作组, A. G. Adame, J. Aguilar, S. Ahlen, S. Alam, D. M. Alexander 等, DESI 2024 IV: 来自 Lyman Alpha 森林的重子声波振荡，arXiv 电子预印本 (2024) arXiv:2404.03001 [2404.03001]。

[52] DESI 合作组，DESI DR2 结果。I. 来自莱曼α森林的重子声学振荡，物理评论 D 112
(2025) 083514 [2503.14739].

[53] A. Cuceu, A. Font-Ribera, B. Joachimi 和 S. Nadathur，通过三维分布的宇宙学超越 BAO
莱曼-α森林，MNRAS 506 (2021) 5439 [2103.14075]。

[54] F. Gerardi, A. Cuceu, A. Font-Ribera, B. Joachimi 和 P. Lemos，来自三维的直接宇宙学推断
莱曼α森林的相关性，《皇家天文学会月刊》518 (2022) 2567 [2209.11263]。

[55] P. McDonald 和 D. J. Eisenstein，通过未来的重子声学振荡调查研究暗能量和曲率
莱曼-α森林, 物理评论 D 76 (2007) 063009 [astro-ph/0607122]。

[56] A. Slosar, V. Ir ˇ si ˇ c, D. Kirkby, S. Bailey, N. G. Busca, T. Delubac 等人, 在 BOSS 数据发布 9 中测量 Lyman-α 森林波动的重子声学振荡, 宇宙学与宇宙粒子物理学杂志 2013 (2013) 026 [1301.3459]。 [57] N. G. Busca, T. Delubac, J. Rich, S. Bailey, A. Font-Ribera, D. Kirkby 等人, Lyα 中的重子声学振荡
BOSS类星体的森林，A&A 552 (2013) A96 [1211.2616]。

[58] H. du Mas des Bourboux, J. Rich, A. Font-Ribera, V. de Sainte Agathe, J. Farr, T. Etourneau 等， 《完成的 SDSS-IV 扩展重子振荡光谱调查：使用 Lyα 森林的重子声学振荡》 ，天体物理学杂志 901
(2020) 153 [2007.08995].

[59] A. Cuceu, A. Font-Ribera, S. Nadathur, B. Joachimi 和 P. Martini，关于宇宙膨胀速率的约束于
从莱曼-α森林测得的红移 2.3，《物理评论快报》130 (2023) 191003 [2209.13942]。

[60] C. Gordon, A. Cuceu, J. Chaves-Montero, A. Font-Ribera, A. Xochitl González-Morales, J. Aguilar 等，早期 DESI 数据的莱曼-α 森林三维相关性，arXiv 电子预印本 (2023) arXiv:2308.10950 [2308.10950]。 [61] A. Cuceu 等，DESI DR1 莱曼-α 森林：三维全形状分析及宇宙学约束，2509.15308。

[62] R. Cen, J. Miralda-Escudé, J. P. Ostriker 和 M. Rauch，小尺度结构的引力坍缩作为起源
莱曼-阿尔法森林, ApJ 437 (1994) L9 [astro-ph/9409017].

[63] J. Miralda-Escudé, R. Cen, J. P. Ostriker and M. Rauch, *The Ly alpha Forest from Gravitational Collapse in the Cold Dark Matter + Lambda Model*, ApJ **471** (1996) 582 [astro-ph/9511013].

[64] A. Meiksin, G. Bryan and M. Machacek, *Hydrodynamical simulations of the Lyα forest: data comparisons*, MNRAS **327** (2001) 296 [astro-ph/0102367].

[65] J. Sexton, Z. Lukic, A. Almgren, C. Daley, B. Friesen, A. Myers et al., *Nyx: A massively parallel amr code for computational cosmology*, *Journal of Open Source Software* **6** (2021) 3068.

[66] J. S. Bolton, E. Puchwein, D. Sijacki, M. G. Haehnelt, T.-S. Kim, A. Meiksin et al., *The Sherwood simulation suite: overview and data comparisons with the Lyman α forest at redshifts 2 ¡ z ¡ 5*, MNRAS **464** (2017) 897 [1605.03462].

[67] R. Weinberger, V. Springel, L. Hernquist, A. Pillepich, F. Marinacci, R. Pakmor et al., *Simulating galaxy formation with black hole driven thermal and kinetic feedback*, MNRAS **465** (2017) 3291 [1607.03486].

[68] S. Bird, M. Fernandez, M.-F. Ho, M. Qezlou, R. Monadi, Y. Ni et al., *PRIYA: a new suite of Lyman-α forest simulations for cosmology*, *JCAP* **10** (2023) 037 [2306.05471].

[69] A. Pillepich et al., *Simulating Galaxy Formation with the IllustrisTNG Model*, *Mon. Not. Roy. Astron. Soc.* **473** (2018) 4077 [1703.02970].

[70] F. Villaescusa-Navarro, D. Anglés-Alcázar, S. Genel, D. N. Spergel, R. S. Somerville, R. Dave et al., *The CAMELS Project: Cosmology and Astrophysics with Machine-learning Simulations*, ApJ **915** (2021) 71 [2010.00619].

[71] M. Walther, E. Armengaud, C. Ravoux, N. Palanque-Delabrouille, C. Yèche and Z. Lukić, *Simulating intergalactic gas for DESI-like small scale Lymanα forest observations*, J. Cosmology Astropart. Phys. **2021** (2021) 059 [2012.04008].

[72] S. Chabanier, C. Ravoux, L. Latrille, J. Sexton, E. Armengaud, J. Bautista et al., *The ACCEL2 project: simulating Lyman-α forest in large-volume hydrodynamical simulations*, *Mon. Not. Roy. Astron. Soc.* **534** (2024) 2674 [2407.04473].

[73] P. McDonald, *Toward a measurement of the cosmological geometry at Z 2: predicting lyman-alpha forest correlation in three dimensions, and the potential of future data sets*, *Astrophys. J.* **585** (2003) 34 [astro-ph/0108064].

[74] A. Arinyo-i Prats, J. Miralda-Escudé, M. Viel and R. Cen, *The Non-Linear Power Spectrum of the Lyman Alpha Forest*, *JCAP* **12** (2015) 017 [1506.04519].

[75] J. J. Givans, A. Font-Ribera, A. Slosar, L. Seeyave, C. Pedersen, K. K. Rogers et al., *Non-linearities in the Lyman-α forest and in its cross-correlation with dark matter halos*, *JCAP* **09** (2022) 070 [2205.00962].

[76] H. du Mas des Bourboux, J. Rich, A. Font-Ribera, V. de Sainte Agathe, J. Farr, T. Etourneau et al., *The Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations with Lyα Forests*, ApJ **901** (2020) 153 [2007.08995].

[77] R. de Belsunce, S.-F. Chen, M. M. Ivanov, C. Ravoux, S. Chabanier, J. Sexton et al., *The ACCEL2 Project: Precision Measurements of EFT Parameters and BAO Peak Shifts for the Lyman-α Forest*, *arXiv e-prints* (2024) arXiv:2412.06892 [2412.06892].

[78] B. Hadzhiyska, R. de Belsunce, A. Cuceu, J. Guy, M. M. Ivanov, H. Coquinot et al., *Measuring and unbiasing the BAO shift in the Lyman-Alpha forest with AbacusSummit*, 2503.13442.

[79] M. M. Pieri, S. Bonoli, J. Chaves-Montero, I. Pâris, M. Fumagalli, J. S. Bolton et al., *WEAVE-QSO: A Massive Intergalactic Medium Survey for the William Herschel Telescope*, in *SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, C. Reylé, J. Richard, L. Cambrésy, M. Deleuil, E. Pécontal, L. Tresse et al., eds., pp. 259–266, Dec., 2016, 1611.09388, DOI.

[80] J. Greene, R. Bezanson, M. Ouchi, J. Silverman and the PFS Galaxy Evolution Working Group, *The Prime Focus Spectrograph Galaxy Evolution Survey*, *arXiv e-prints* (2022) arXiv:2206.14908 [2206.14908].

[81] R. S. de Jong, O. Agertz, A. A. Berbel, J. Aird, D. A. Alexander, A. Amarsi et al., *4MOST: Project overview and information for the First Call for Proposals*, *The Messenger* **175** (2019) 3 [1903.02464].

[82] J. S. Bolton, E. Puchwein, D. Sijacki, M. G. Haehnelt, T.-S. Kim, A. Meiksin et al., *The Sherwood simulation suite: overview and data comparisons with the Lyman α forest at redshifts $2 \leq z \leq 5$*, *Mon. Not. Roy. Astron. Soc.* **464** (2017) 897 [1605.03462].

[83] S. Bird, Y. Ni, T. Di Matteo, R. Croft, Y. Feng and N. Chen, *The ASTRID simulation: galaxy formation and reionization*, MNRAS **512** (2022) 3703 [2111.01160].

[84] Y. Ni, T. Di Matteo, S. Bird, R. Croft, Y. Feng, N. Chen et al., *The ASTRID simulation: the evolution of supermassive black holes*, MNRAS **513** (2022) 670 [2110.14154].

[85] M. M. Ivanov, *Lyman alpha forest power spectrum in effective field theory*, *Phys. Rev. D* **109** (2024) 023507 [2309.10133].

[86] R. de Belsunce, J. M. Sullivan and P. McDonald, *Compressed 3D Lyman-α forest bispectrum*, *Phys. Rev. D* **113** (2026) 063550 [2510.23597].

[87] R. de Belsunce, M. M. Ivanov, J. M. Sullivan, K. Akitsu and S.-F. Chen, *Modeling the Cosmological Lyman-α Forest at the Field Level*, 2507.00284.

[88] P. McDonald and A. Roy, *Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS*, *JCAP* **0908** (2009) 020 [0902.0991].

[89] D. Baumann, A. Nicolis, L. Senatore and M. Zaldarriaga, *Cosmological Non-Linearities as an Effective Fluid*, *JCAP* **1207** (2012) 051 [1004.2488].

[90] J. J. M. Carrasco, S. Foreman, D. Green and L. Senatore, *The Effective Field Theory of Large Scale Structures at Two Loops*, *JCAP* **07** (2014) 057 [1310.0464].

[91] M. M. Ivanov, *Effective Field Theory for Large Scale Structure*, 2212.08488.

[92] P. McDonald, J. Miralda-Escude, M. Rauch, W. L. W. Sargent, T. A. Barlow, R. Cen et al., *The Observed probability distribution function, power spectrum, and correlation function of the transmitted flux in the Lyman-alpha forest*,

---

[63] J. Miralda-Escudé, R. Cen, J. P. Ostriker 和 M. Rauch，由引力塌缩产生的冷宇宙中的莱曼α森林
暗物质 + Λ 模型, ApJ 471 (1996) 582 [astro-ph/9511013].

[64] A. Meiksin, G. Bryan 和 M. Machacek，Lyα森林的流体动力学模拟：数据比较，MNRAS 327
(2001) 296 [astro-ph/0102367].

[65] J. Sexton, Z. Lukic, A. Almgren, C. Daley, B. Friesen, A. Myers 等，Nyx：一个用于大规模并行 AMR 的代码
计算宇宙学，《开源软件杂志》6（2021）3068。

[66] J. S. Bolton, E. Puchwein, D. Sijacki, M. G. Haehnelt, T.-S. Kim, A. Meiksin 等，《Sherwood 模拟套件：概述及与红移 $2 < z < 5$ 莱曼 α 森林的数据比较》，MNRAS 464 (2017) 897 [1605.03462]。 [67] R. Weinberger, V. Springel, L. Hernquist, A. Pillepich, F. Marinacci, R. Pakmor 等，《模拟星系形成与》

黑洞驱动的热反馈和动能反馈，MNRAS 465 (2017) 3291 [1607.03486]。

[68] S. Bird, M. Fernandez, M.-F. Ho, M. Qezlou, R. Monadi, Y. Ni 等, PRIYA：一套新的 Lyman-α 森林工具
用于宇宙学的模拟，JCAP 10 (2023) 037 [2306.05471]。

[69] A. Pillepich 等人，《使用 IllustrisTNG 模型模拟星系形成》，《皇家天文学会月刊》473 (2018)
4077 [1703.02970].

[70] F. Villaescusa-Navarro, D. Anglés-Alcázar, S. Genel, D. N. Spergel, R. S. Somerville, R. Dave 等，《CAMELS 项目：利用机器学习模拟的宇宙学与天体物理学》，《天体物理学杂志》915 (2021) 71 [2010.00619]。

[71] M. Walther, E. Armengaud, C. Ravoux, N. Palanque-Delabrouille, C. Y` eche 和 Z. Luki´ c, 为类似 DESI 的小尺度 Lymanα 森林观测模拟星际气体, J. Cosmology Astropart. Phys. 2021 (2021) 059 [2012.04008]。 [72] S. Chabanier, C. Ravoux, L. Latrille, J. Sexton, E. Armengaud, J. Bautista 等, ACCEL2 项目: 模拟

大体积流体动力学模拟中的Lyman-α森林，《皇家天文学会月刊》534 (2024) 2674 [2407.04473]。[73] P. McDonald，迈向测量Z≈2的宇宙几何：预测Lyman-α森林相关性
三维，以及未来数据集的潜力，天体物理学杂志. J. 585 (2003) 34 [astro-ph/0108064]。[74] A. Arinyo-i Prats, J. Miralda-Escudé, M. Viel 和 R. Cen，莱曼α森林的非线性功率谱
JCAP 12 (2015) 017 [1506.04519]。

[75] J. J. Givans, A. Font-Ribera, A. Slosar, L. Seeyave, C. Pedersen, K. K. Rogers 等，Lyman-α中的非线性
森林及其与暗物质晕的交叉相关，JCAP 09 (2022) 070 [2205.00962]。

[76] H. du Mas des Bourboux, J. Rich, A. Font-Ribera, V. de Sainte Agathe, J. Farr, T. Etourneau 等人，《完成的 SDSS-IV 扩展重子声波振荡光谱调查：利用 Lyα 森林的重子声波振荡》，天体物理学杂志 901
(2020) 153 [2007.08995].

[77] R. de Belsunce, S.-F. Chen, M. M. Ivanov, C. Ravoux, S. Chabanier, J. Sexton 等，《ACCEL2 项目：精密》
Lyman-α森林的EFT参数测量与BAO峰位移，arXiv电子印本 (2024) arXiv:2412.06892 [2412.06892]。

[78] B. Hadzhiyska, R. de Belsunce, A. Cuceu, J. Guy, M. M. Ivanov, H. Coquinot 等，测量与去偏的BAO
使用 AbacusSummit 的 Lyman-Alpha 森林中的位移，2503.13442。

[79] M. M. Pieri, S. Bonoli, J. Chaves-Montero, I. Pˆ aris, M. Fumagalli, J. S. Bolton 等，WEAVE-QSO：威廉 · 赫歇尔望远镜的大规模星际介质调查，载于 SF2A-2016：法国天文学与天体物理学会年会论文集，C. Reyl´ e, J. Richard, L. Cambr´ esy, M. Deleuil, E. P´ econtal,

L. Tresse 等, 编, 第 259 – 266 页, 2016 年 12 月, 1611.09388, DOI.

[80] J. Greene, R. Bezanson, M. Ouchi, J. Silverman 及 PFS 银河系演化工作组，《主焦点》
光谱仪银河演化调查，arXiv 电子预印本 (2022) arXiv:2206.14908 [2206.14908]。

[81] R. S. de Jong, O. Agertz, A. A. Berbel, J. Aird, D. A. Alexander, A. Amarsi 等，4MOST：项目概述及
关于第一次征集提案的信息，《信使》175 (2019) 3 [1903.02464]。

[82] J. S. Bolton, E. Puchwein, D. Sijacki, M. G. Haehnelt, T.-S. Kim, A. Meiksin 等, Sherwood 模拟套件：概述及与红移 $2 \leqslant z \leqslant 5$ 的莱曼 α 森林数据比较, 月刊皇家天文学会, 464 (2017) 897 [1605.03462]。 [83] S. Bird, Y. Ni, T. Di Matteo, R. Croft, Y. Feng 和 N. Chen, ASTRID 模拟：星系形成与

再电离，MNRAS 512 (2022) 3703 [2111.01160]。

[84] Y. Ni, T. Di Matteo, S. Bird, R. Croft, Y. Feng, N. Chen 等, ASTRID 模拟：超大质量的演化
黑洞，MNRAS 513 (2022) 670 [2110.14154]。

[85] M. M. Ivanov，拉曼阿尔法森林有效场理论中的功率谱，物理评论 D 109 (2024) 023507
[2309.10133].

[86] R. de Belsunce, J. M. Sullivan 和 P. McDonald, 压缩的三维 Lyman-α 森林双谱, 物理评论 D 113 (2026)
063550 [2510.23597].

[87] R. de Belsunce, M. M. Ivanov, J. M. Sullivan, K. Akitsu 和 S.-F. Chen, 宇宙学莱曼-α森林建模于
字段级别，2507.00284。

[88] P. McDonald 和 A. Roy，暗物质示踪器的聚类：为即将到来的精确大尺度结构时代推广偏差，
JCAP 0908 (2009) 020 [0902.0991]。

[89] D. Baumann, A. Nicolis, L. Senatore 和 M. Zaldarriaga, 宇宙学非线性作为有效流体, JCAP
1207 (2012) 051 [1004.2488].

[90] J. J. M. Carrasco, S. Foreman, D. Green 和 L. Senatore，大尺度结构的有效场论在二阶
环路，JCAP 07 (2014) 057 [1310.0464]。

[91] M. M. Ivanov，大尺度结构的有效场论，2212.08488。

[92] P. McDonald, J. Miralda-Escude, M. Rauch, W. L. W. Sargent, T. A. Barlow, R. Cen 等, 莱曼-α森林中透射通量的观测概率分布函数、功率谱和相关函数,

Astrophys. J. **543** (2000) 1 [astro-ph/9911196].

[93] J. J. Givans and C. M. Hirata, *Redshift-space streaming velocity effects on the Lyman-α forest baryon acoustic oscillation scale*, *Phys. Rev. D* **102** (2020) 023515 [2002.12296].

[94] V. Desjacques, D. Jeong and F. Schmidt, *The Galaxy Power Spectrum and Bispectrum in Redshift Space*, *JCAP* **1812** (2018) 035 [1806.04015].

[95] S.-F. Chen, Z. Vlah and M. White, *The Lyα forest flux correlation function: a perturbation theory perspective*, *JCAP* **05** (2021) 053 [2103.13498].

[96] M. M. Ivanov, M. Simonović and M. Zaldarriaga, *Cosmological Parameters from the BOSS Galaxy Power Spectrum*, *JCAP* **05** (2020) 042 [1909.05277].

[97] G. D'Amico, J. Gleyzes, N. Kokron, D. Markovic, L. Senatore, P. Zhang et al., *The Cosmological Analysis of the SDSS/BOSS data from the Effective Field Theory of Large-Scale Structure*, 1909.05271.

[98] S.-F. Chen, Z. Vlah and M. White, *A new analysis of galaxy 2-point functions in the BOSS survey, including full-shape information and post-reconstruction BAO*, *JCAP* **02** (2022) 008 [2110.05530].

[99] O. H. E. Philcox and M. M. Ivanov, *BOSS DR12 full-shape cosmology: ΛCDM constraints from the large-scale galaxy power spectrum and bispectrum monopole*, *Phys. Rev. D* **105** (2022) 043517 [2112.04515].

[100] S.-F. Chen, M. White, J. DeRose and N. Kokron, *Cosmological analysis of three-dimensional BOSS galaxy clustering and Planck CMB lensing cross correlations via Lagrangian perturbation theory*, *JCAP* **07** (2022) 041 [2204.10392].

[101] S.-F. Chen, M. M. Ivanov, O. H. E. Philcox and L. Wenzl, *Suppression without Thawing: Constraining Structure Formation and Dark Energy with Galaxy Clustering*, 2406.13388.

[102] B. Hadzhiyska, A. Font-Ribera, A. Cuceu, S. Chabanier, J. Aguilar, D. Brooks et al., *Planting a Lyman alpha forest on ABACUSSUMMIT*, MNRAS **524** (2023) 1008 [2305.08899].

[103] R. A. C. Croft, *Characterization of Lyman Alpha Spectra and Predictions of Structure Formation Models: A Flux Statistics Approach*, in *Eighteenth Texas Symposium on Relativistic Astrophysics*, A. V. Olinto, J. A. Frieman and D. N. Schramm, eds., p. 664, 1998, astro-ph/9701166.

[104] M. Qezlou, A. B. Newman, G. C. Rudie and S. Bird, *Characterizing Protoclusters and Protogroups at z 2.5 Using Lyα Tomography*, ApJ **930** (2022) 109 [2112.03930].

[105] S. Peirani, D. H. Weinberg, S. Colombi, J. Blaizot, Y. Dubois and C. Pichon, *LyMAS: Predicting Large-scale Lyα Forest Statistics from the Dark Matter Density Field*, ApJ **784** (2014) 11 [1306.1533].

[106] S. Peirani, S. Prunet, S. Colombi, C. Pichon, D. H. Weinberg, C. Laigle et al., *LyMAS reloaded: improving the predictions of the large-scale Lyman-α forest statistics from dark matter density and velocity fields*, MNRAS **514** (2022) 3222 [2204.06365].

[107] D. Sorini, J. Oñorbe, Z. Lukić and J. F. Hennawi, *Modeling the Lyα Forest in Collisionless Simulations*, ApJ **827** (2016) 97 [1602.08099].

[108] F. Sinigaglia, F.-S. Kitaura, A. Balaguera-Antolínez, I. Shimizu, K. Nagamine, M. Sánchez-Benavente et al., *Mapping the Three-dimensional Lyα Forest Large-scale Structure in Real and Redshift Space*, ApJ **927** (2022) 230.

[109] F. Sinigaglia, F. S. Kitaura, K. Nagamine, Y. Oku and A. Balaguera-Antolínez, *Field-level Lyman-α forest modeling in redshift space via augmented nonlocal Fluctuating Gunn-Peterson Approximation*, A&A **682** (2024) A21 [2305.10428].

[110] B. Horowitz, M. Dornfest, Z. Lukić and P. Harrington, *hyphy: Deep Generative Conditional Posterior Mapping of Hydrodynamical Physics*, *Astrophys. J.* **941** (2022) 42 [2106.12675].

[111] P. Harrington, M. Mustafa, M. Dornfest, B. Horowitz and Z. Lukić, *Fast, High-fidelity Lyα Forests with Convolutional Neural Networks*, *Astrophys. J.* **929** (2022) 160 [2106.12662].

[112] C. Jacobus, S. Chabanier, P. Harrington, J. Emberson, Z. Lukić and S. Habib, *A Gigaparsec-Scale Hydrodynamic Volume Reconstructed with Deep Learning*, 2411.16920.

[113] B. Horowitz, C. Cuesta-Lazaro and O. Yehia, *BaryonBridge: Stochastic Interpolant Model for Fast Hydrodynamical Simulations*, 2510.19224.

[114] F. Hafezianzadeh, X. Zhang, Y. Ni, R. A. C. Croft, T. DiMatteo, M. Qezlou et al., *An AI super-resolution field emulator for cosmological hydrodynamics: the Lyman-α forest*, 2507.16189.

[115] S. Bird, K. K. Rogers, H. V. Peiris, L. Verde, A. Font-Ribera and A. Pontzen, *An Emulator for the Lyman-alpha Forest*, *JCAP* **02** (2019) 050 [1812.04654].

[116] C. Pedersen, A. Font-Ribera, K. K. Rogers, P. McDonald, H. V. Peiris, A. Pontzen et al., *An emulator for the Lyman-α forest in beyond-ΛCDM cosmologies*, *JCAP* **05** (2021) 033 [2011.15127].

[117] M.-F. Ho, S. Bird, M. A. Fernandez and C. R. Shelton, *MF-Box: multifidelity and multiscale emulation for the matter power spectrum*, *Mon. Not. Roy. Astron. Soc.* **526** (2023) 2903 [2306.03144].

[118] J. Chaves-Montero et al., *ForestFlow: predicting the Lyman-α forest clustering from linear to nonlinear scales*, *Astron. Astrophys.* **694** (2025) A187 [2409.05682].

[119] A. Pontzen, A. Slosar, N. Roth and H. V. Peiris, *Inverted initial conditions: exploring the growth of cosmic structure and voids*, *Phys. Rev. D* **93** (2016) 103519 [1511.04090].

[120] U. Seljak, *Bias, redshift space distortions and primordial nongaussianity of nonlinear transformations: application to Lyman alpha forest*, *JCAP* **03** (2012) 004 [1201.0594].

[121] A. M. Cieplak and A. Slosar, *Towards physics responsible for large-scale Lyman-α forest bias parameters*, *JCAP* **03** (2016) 016 [1509.07875].

[122] M. Schmittfull, M. Simonović, V. Assassi and M. Zaldarriaga, *Modeling Biased Tracers at the Field Level*, *Phys. Rev. D* **100** (2019) 043514 [1811.10640].

天体物理学杂志. J. 543 (2000) 1 [astro-ph/9911196].

[93] J. J. Givans 和 C. M. Hirata，红移空间流动速度对莱曼-α森林重子声学的影响
振荡尺度, 物理评论 D 102 (2020) 023515 [2002.12296].

[94] V. Desjacques, D. Jeong 和 F. Schmidt,《红移空间中的星系功率谱与双谱》, JCAP 1812
(2018) 035 [1806.04015].

[95] 陈思凡, Z. Vlah 和 M. White, Lyα 林通量相关函数：微扰理论视角, JCAP 05
(2021) 053 [2103.13498].

[96] M. M. Ivanov、M. Simonović 和 M. Zaldarriaga，《来自 BOSS 星系功率谱的宇宙学参数》
JCAP 05 (2020) 042 [1909.05277]。

[97] G. D'Amico, J. Gleyzes, N. Kokron, D. Markovic, L. Senatore, P. Zhang 等，《宇宙学分析》
来自大尺度结构有效场论的 SDSS/BOSS 数据, 1909.05271.

[98] S.-F. Chen, Z. Vlah 和 M. White，在 BOSS 调查中对银河系二点函数进行的新分析，包括全形状
信息与重建后的BAO，JCAP 02 (2022) 008 [2110.05530]。

[99] O. H. E. Philcox 和 M. M. Ivanov，BOSS DR12 全形状宇宙学：来自大尺度星系的 ΛCDM 约束
功率谱和双谱单极子, 物理评论 D 105 (2022) 043517 [2112.04515]。

[100] S.-F. Chen, M. White, J. DeRose 和 N. Kokron，通过拉格朗日微扰理论对三维 BOSS 星系聚类和 Planck CMB 引力透镜交叉相关进行
宇宙学分析，JCAP 07 (2022) 041 [2204.10392]。 [101] S.-F. Chen, M. M. Ivanov, O. H. E. Philcox 和 L. Wenzl，无解冻下的抑制：约束结构

星系聚类下的形成与暗能量，2406.13388。

[102] B. Hadzhiyska, A. Font-Ribera, A. Cuceu, S. Chabanier, J. Aguilar, D. Brooks 等人，在……上种植莱曼α森林
ABACUSSUMMIT,《皇家天文学会月刊》524 (2023) 1008 [2305.08899]。

[103] R. A. C. Croft，《赖曼α谱的表征及结构形成模型的预测：一种通量统计方法》，收录于《第十八届德克萨斯相对论天体物理学研讨
会》，A. V. Olinto、J. A. Frieman 和 D. N. 编辑
Schramm，编辑，第664页，1998年，astro-ph/9701166。

[104] M. Qezlou, A. B. Newman, G. C. Rudie 和 S. Bird，使用 Lyα 描述 z 2.5 的原星团和原群体
断层扫描，天体物理学杂志 930 (2022) 109
[2112.03930]。

[105] S. Peirani, D. H. Weinberg, S. Colombi, J. Blaizot, Y. Dubois 和 C. Pichon, LyMAS：预测大尺度 Lyα 森林
来自暗物质密度场的统计数据，ApJ 784 (2014) 11 [1306.1533]。

[106] S. Peirani, S. Prunet, S. Colombi, C. Pichon, D. H. Weinberg, C. Laigle 等，LyMAS 重新加载：通过暗物质密度和速度场改进大尺度
Lyman-α 森林统计的预测，MNRAS 514 (2022)

3222 [2204.06365].

[107] D. Sorini, J. O˜ norbe, Z. Luki´ c 和 J. F. Hennawi，在无碰撞模拟中对 Lyα 森林进行建模，ApJ 827

(2016) 97 [1602.08099].

[108] F. Sinigaglia, F.-S. Kitaura, A. Balaguera-Antolínez, I. Shimizu, K. Nagamine, M. Sánchez-Benavente 等, 绘图
三维 Lyman-α 森林在真实空间和红移空间的大尺度结构，ApJ 927 (2022) 230。[109] F. Sinigaglia, F. S. Kitaura, K. Nagamine, Y. Oku 和 A.
Balaguera-Antolínez，通过增强的非局域波动 Gunn-Peterson 近似在红移空间进行场级 Lyman-α 森林建模，A&A 682 (2024) A21
[2305.10428]。[110] B. Horowitz, M. Dornfest, Z. Lukić 和 P. Harrington，hyphy：深度生成条件后验映射
流体动力学物理学，天体物理学杂志 941 (2022) 42 [2106.12675]。

[111] P. Harrington, M. Mustafa, M. Dornfest, B. Horowitz 和 Z. Lukić, 使用卷积实现快速、高保真 Lyα 森林
神经网络, 天体物理学杂志, 929 (2022) 160 [2106.12662].

[112] C. Jacobus, S. Chabanier, P. Harrington, J. Emberson, Z. Lukić 和 S. Habib，一项千兆秒差距尺度的流体力学研究
使用深度学习重建的体积，2411.16920。

[113] B. Horowitz, C. Cuesta-Lazaro 和 O. Yehia, BaryonBridge: 用于快速流体力学的随机插值模型
模拟，2510.19224。

[114] F. Hafezianzadeh, X. Zhang, Y. Ni, R. A. C. Croft, T. DiMatteo, M. Qezlou 等人，《一种 AI 超分辨率场模拟器》
用于宇宙学流体动力学：莱曼-α森林, 2507.16189.

[115] S. Bird, K. K. Rogers, H. V. Peiris, L. Verde, A. Font-Ribera 和 A. Pontzen，《Lyman-alpha 森林的仿真器》，
JCAP 02 (2019) 050 [1812.04654]。

[116] C. Pedersen, A. Font-Ribera, K. K. Rogers, P. McDonald, H. V. Peiris, A. Pontzen 等，Lyman-α 的仿真器
超越ΛCDM宇宙学中的森林，JCAP 05 (2021) 033 [2011.15127]。

[117] M.-F. Ho, S. Bird, M. A. Fernandez 和 C. R. Shelton，MF-Box：用于物质的多保真度和多尺度模拟
功率谱，Mon. Not. Roy. Astron. Soc. 526 (2023) 2903 [2306.03144]。

[118] J. Chaves-Montero 等，ForestFlow：从线性到非线性尺度预测莱曼-α森林聚类，天文学。
天体物理学. 694 (2025) A187 [2409.05682].

[119] A. Pontzen, A. Slosar, N. Roth 和 H. V. Peiris，倒置初始条件：探索宇宙结构的形成
以及空洞，Phys. Rev. D 93 (2016) 103519 [1511.04090]。

[120] U. Seljak，偏差、红移空间畸变和非线性变换的原始非高斯性：应用于
Lyman α 森林, JCAP 03 (2012) 004 [1201.0594].

[121] A. M. Cieplak 和 A. Slosar, 迈向负责大尺度莱曼-α森林偏置参数的物理学, JCAP 03
(2016) 016 [1509.07875].

[122] M. Schmittfull, M. Simonović, V. Assassi 和 M. Zaldarriaga,《在场水平上建模有偏示踪器》，物理评论 D
100 (2019) 043514 [1811.10640].

[123] M. Schmittfull, M. Simonović, M. M. Ivanov, O. H. E. Philcox and M. Zaldarriaga, *Modeling Galaxies in Redshift Space at the Field Level*, *JCAP* **05** (2021) 059 [2012.03334].

[124] A. Obuljen, M. Simonović, A. Schneider and R. Feldmann, *Modeling HI at the field level*, *Phys. Rev. D* **108** (2023) 083528 [2207.12398].

[125] M. M. Ivanov, A. Obuljen, C. Cuesta-Lazaro and M. W. Toomey, *Full-shape analysis with simulation-based priors: cosmological parameters and the structure growth anomaly*, 2409.10609.

[126] A. Font-Ribera et al., *The large-scale Quasar-Lyman \alpha\ Forest Cross-Correlation from BOSS*, *JCAP* **05** (2013) 018 [1303.1937].

[127] A. Chudaykin and M. M. Ivanov, *Lyman Alpha Forest - Halo Cross-Correlations in Effective Field Theory*, *arXiv e-prints* (2025) arXiv:2501.04770 [2501.04770].

[128] S.-F. Chen et al., *Baryon Acoustic Oscillation Theory and Modelling Systematics for the DESI 2024 results*, 2402.14070.

[129] J. Farr, A. Font-Ribera, H. du Mas des Bourboux, A. Muñoz-Gutiérrez, F. J. Sánchez, A. Pontzen et al., *LyaCoLoRe: synthetic datasets for current and future Lyman-α forest BAO surveys*, J. Cosmology Astropart. Phys. **2020** (2020) 068 [1912.02763].

[130] C. Ramírez-Pérez, J. Sanchez, D. Alonso and A. Font-Ribera, *CoLoRe: fast cosmological realisations over large volumes with multiple tracers*, *JCAP* **05** (2022) 002 [2111.05069].

[131] D. J. Schlegel, J. A. Kollmeier, G. Aldering, S. Bailey, C. Baltay, C. Bebek et al., *The MegaMapper: A Stage-5 Spectroscopic Instrument Concept for the Study of Inflation and Dark Energy*, *arXiv e-prints* (2022) arXiv:2209.04322 [2209.04322].

[132] N. Kaiser, *Clustering in real space and in redshift space*, MNRAS **227** (1987) 1.

[133] M. Mirbabayi, F. Schmidt and M. Zaldarriaga, *Biased Tracers and Time Evolution*, *JCAP* **1507** (2015) 030 [1412.5169].

[134] E. Pajer and M. Zaldarriaga, *On the Renormalization of the Effective Field Theory of Large Scale Structures*, *JCAP* **08** (2013) 037 [1301.7182].

[135] V. Desjacques, D. Jeong and F. Schmidt, *Large-Scale Galaxy Bias*, *Phys. Rept.* **733** (2018) 1 [1611.09787].

[136] M. M. Ivanov, *Lyman alpha forest power spectrum in effective field theory*, Phys. Rev. D **109** (2024) 023507 [2309.10133].

[137] Y. B. Zel'Dovich, *Reprint of 1970A&A.....5...84Z. Gravitational instability: an approximate theory for large density perturbations.*, *Astronomy and Astrophysics* **500** (1970) 13.

[138] R. Scoccimarro and J. Frieman, *Loop corrections in nonlinear cosmological perturbation theory*, *Astrophys. J. Suppl.* **105** (1996) 37 [astro-ph/9509047].

[139] R. Scoccimarro and J. Frieman, *Loop corrections in nonlinear cosmological perturbation theory 2. Two point statistics and selfsimilarity*, *Astrophys. J.* **473** (1996) 620 [astro-ph/9602070].

[140] R. Scoccimarro, *Cosmological perturbations: Entering the nonlinear regime*, *Astrophys. J.* **487** (1997) 1 [astro-ph/9612207].

[141] M. Crocce and R. Scoccimarro, *Nonlinear Evolution of Baryon Acoustic Oscillations*, *Phys. Rev.* **D77** (2008) 023533 [0704.2783].

[142] T. Baldauf, M. Mirbabayi, M. Simonović and M. Zaldarriaga, *Equivalence Principle and the Baryon Acoustic Peak*, *Phys. Rev.* **D92** (2015) 043514 [1504.04366].

[143] L. Senatore and M. Zaldarriaga, *The IR-resummed Effective Field Theory of Large Scale Structures*, *JCAP* **1502** (2015) 013 [1404.5954].

[144] D. Blas, M. Garny, M. M. Ivanov and S. Sibiryakov, *Time-Sliced Perturbation Theory for Large Scale Structure I: General Formalism*, *JCAP* **1607** (2016) 052 [1512.05807].

[145] D. Blas, M. Garny, M. M. Ivanov and S. Sibiryakov, *Time-Sliced Perturbation Theory II: Baryon Acoustic Oscillations and Infrared Resummation*, *JCAP* **1607** (2016) 028 [1605.02149].

[146] M. M. Ivanov and S. Sibiryakov, *Infrared Resummation for Biased Tracers in Redshift Space*, *JCAP* **1807** (2018) 053 [1804.05080].

[147] A. Vasudevan, M. M. Ivanov, S. Sibiryakov and J. Lesgourgues, *Time-sliced perturbation theory with primordial non-Gaussianity and effects of large bulk flows on inflationary oscillating features*, *JCAP* **09** (2019) 037 [1906.08697].

[148] M. McQuinn and M. White, *On estimating ly α forest correlations between multiple sightlines*, *Monthly Notices of the Royal Astronomical Society* **415** (2011) 2257.

[149] B. Hadzhiyska et al., *Mitigating the noise of DESI mocks using analytic control variates*, *Open J. Astrophys.* **6** (2023) 2308.12343 [2308.12343].

[150] N. A. Maksimova, L. H. Garrison, D. J. Eisenstein, B. Hadzhiyska, S. Bose and T. P. Satterthwaite, *ABACUSSUMMIT: a massive set of high-accuracy, high-resolution N-body simulations*, MNRAS **508** (2021) 4017 [2110.11398].

[151] V. Springel, *The cosmological simulation code GADGET-2*, MNRAS **364** (2005) 1105 [astro-ph/0505010].

[152] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown et al., *Planck 2013 results. XVI. Cosmological parameters*, A&A **571** (2014) A16 [1303.5076].

[153] M. M. Ivanov et al., *The Millennium and Astrid galaxies in effective field theory: comparison with galaxy-halo connection models at the field level*, 2412.01888.

[154] J. M. Sullivan, C. Cuesta-Lazaro, M. M. Ivanov, Y. Ni, S. Bose, B. Hadzhiyska et al., *High-redshift Millennium and Astrid galaxies in effective field theory at the field level*, 2505.03626.

[155] M. White et al., *The clustering of Lyman Alpha Emitting galaxies at $z = 2-3$*, *JCAP* **08** (2024) 020 [2406.01803].

[156] H. Ebina and M. White, *Cosmology before noon with multiple galaxy populations*, *JCAP* **06** (2024) 052 [2401.13166].

[123] M. Schmittfull, M. Simonovi´ c, M. M. Ivanov, O. H. E. Philcox 和 M. Zaldarriaga，红移空间中的星系建模
在场级别，JCAP 05 (2021) 059 [2012.03334]。

[124] A. Obuljen, M. Simonović, A. Schneider 和 R. Feldmann,《在场层面上的HI建模》，物理评论 D 108 (2023)
083528 [2207.12398].

[125] M. M. Ivanov, A. Obuljen, C. Cuesta-Lazaro 和 M. W. Toomey，基于仿真的先验的全形态分析：
宇宙学参数与结构增长异常，2409.10609。

[126] A. Font-Ribera 等人，《BOSS 大尺度类星体-赖曼 α 森交叉相关》，JCAP 05 (2013) 018 [1303.1937]。 [127] A. Chudaykin 和 M. M. Ivanov，《有效场论中的赖曼 α 森-晕交叉相关》，arXiv
电子预印本 (2025) arXiv:2501.04770 [2501.04770]。

[128] 陈思发 等,《DESI 2024 结果的重子声学振荡理论与建模系统误差》, 2402.14070. [129] J. Farr, A. Font-Ribera, H. du Mas des Bourboux, A. Muñoz-Gutiérrez, F. J. Sánchez, A. Pontzen 等,《LyaCoLoRe：用于当前和未来莱曼-α森林 BAO 调查的合成数据集》, 宇宙学与天体粒子物理杂志, 2020 (2020) 068
[1912.02763].

[130] C. Ramírez-Pérez, J. Sanchez, D. Alonso 和 A. Font-Ribera, CoLoRe：大体积的快速宇宙学实现
使用多重示踪剂，JCAP 05 (2022) 002 [2111.05069]。

[131] D. J. Schlegel, J. A. Kollmeier, G. Aldering, S. Bailey, C. Baltay, C. Bebek 等，《MegaMapper：用于研究暴涨和暗能量的第五阶段光谱仪概念》，arXiv 电子预印本 (2022) arXiv:2209.04322
[2209.04322].

[132] N. Kaiser, 实空间和红移空间中的聚类,《皇家天文学会月报》227 (1987) 1.

[133] M. Mirbabayi, F. Schmidt 和 M. Zaldarriaga, 偏置迹象与时间演化, JCAP 1507 (2015) 030 [1412.5169]. [134] E. Pajer 和 M. Zaldarriaga, 关于大尺度结构有效场论的重整化, JCAP 08
(2013) 037 [1301.7182].

[135] V. Desjacques, D. Jeong 和 F. Schmidt，大尺度星系偏差，物理学报告 733 (2018) 1 [1611.09787]。[136] M. M. Ivanov，拉曼α森林功率谱在有效场论中的研究，物理评论 D 109 (2024) 023507
[2309.10133].

[137] Y. B. Zel’Dovich, 1970A&A.....5...84Z 转载。引力不稳定性：大密度的近似理论
扰动。，天文学与天体物理学 500 (1970) 13.

[138] R. Scoccimarro 和 J. Frieman，非线性宇宙学扰动理论中的环修正，《天体物理学杂志增刊》105
(1996) 37 [astro-ph/9509047].

[139] R. Scoccimarro 和 J. Frieman，非线性宇宙学摄动理论中的环路修正 2. 二点统计
以及自相似性，《天体物理学杂志》473 (1996) 620 [astro-ph/9602070]。

[140] R. Scoccimarro，宇宙学微扰：进入非线性阶段，《天体物理学杂志》487 (1997) 1
[astro-ph/9612207].

[141] M. Crocce 和 R. Scoccimarro，重子声学振荡的非线性演化，《物理评论 D》77 (2008) 023533
[0704.2783].

[142] T. Baldauf, M. Mirbabayi, M. Simonović 和 M. Zaldarriaga，《等效原理与重子声学峰》，
物理评论 D92 (2015) 043514 [1504.04366]。

[143] L. Senatore 和 M. Zaldarriaga，《大尺度结构的红外重整有效场论》，JCAP 1502 (2015)
013 [1404.5954].

[144] D. Blas, M. Garny, M. M. Ivanov 和 S. Sibiryakov，大尺度结构的时间切片微扰理论 I：
一般形式主义，JCAP 1607 (2016) 052 [1512.05807]。

[145] D. Blas, M. Garny, M. M. Ivanov 和 S. Sibiryakov，时间切片微扰理论 II：重子声学振荡
以及红外重求和，JCAP 1607 (2016) 028 [1605.02149]。

[146] M. M. 伊万诺夫 和 S. 西比里亚科夫,《红移空间中偏倚示踪体的红外重和》，JCAP 1807 (2018) 053
[1804.05080].

[147] A. Vasudevan, M. M. Ivanov, S. Sibiryakov 和 J. Lesgourgues, 带有原始非高斯性的时间切片微扰理论以及大体流对膨胀振荡特征的影响, JCAP 09 (2019) 037 [1906.08697]。 [148] M. McQuinn 和 M. White, 关于估计多个视线之间的 Ly α 森林相关性, Monthly Notices of the
皇家天文学会 415 (2011) 2257。

[149] B. Hadzhiyska 等人，使用解析控制变量减轻 DESI 模拟的噪声，《开放天体物理学杂志》6 (2023)
2308.12343 [2308.12343]。

[150] N. A. Maksimova, L. H. Garrison, D. J. Eisenstein, B. Hadzhiyska, S. Bose 和 T. P. Satterthwaite, ABACUSSUMMIT:
一套大规模的高精度、高分辨率 N 体模拟，MNRAS 508 (2021) 4017 [2110.11398]。[151] V. Springel, 宇宙学模拟代码 GADGET-2, MNRAS 364 (2005) 1105 [astro-ph/0505010]。[152] Planck 合作组, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown 等, Planck 2013
结果。XVI. 宇宙学参数,《天文学与天体物理学》571 (2014) A16 [1303.5076]。

[153] M. M. Ivanov 等,《有效场论中的千禧和阿斯特里德星系：与星系晕的比较》
在字段级别的连接模型，2412.01888。

[154] J. M. Sullivan, C. Cuesta-Lazaro, M. M. Ivanov, Y. Ni, S. Bose, B. Hadzhiyska 等，高红移千禧年与
Astrid 在场水平上的有效场论中的星系，2505.03626。

[155] M. White 等，红移 $z = 2-3$ 的莱曼α发射星系的聚类，《JCAP》08 (2024) 020 [2406.01803]。[156] H. Ebina 和 M. White，多种星系群体下的中午前宇宙学，《JCAP》06 (2024) 052 [2401.13166]。

[157] H. Hildebrandt, J. Pielorz, T. Erben, L. van Waerbeke, P. Simon and P. Capak, *CARS: the CFHTLS-Archive-Research Survey. II. Weighing dark matter halos of Lyman-break galaxies at* $z = 3\text{-}5$, A&A **498** (2009) 725 [0903.3951].

[158] A. Arinyo-i Prats, J. Miralda-Escudé, M. Viel and R. Cen, *The Non-Linear Power Spectrum of the Lyman Alpha Forest*, *JCAP* **12** (2015) 017 [1506.04519].

[159] R. de Belsunce, B. Hadzhiyska and M. M. Ivanov, *Bridging Simulations and EFT: A Hybrid Model of the Lyman-Alpha Forest Field*, 2512.13681.

[160] S.-F. Chen, Z. Vlah and M. White, *Consistent Modeling of Velocity Statistics and Redshift-Space Distortions in One-Loop Perturbation Theory*, *JCAP* **07** (2020) 062 [2005.00523].

[161] D. Blas, M. Garny and T. Konstandin, *On the non-linear scale of cosmological perturbation theory*, *JCAP* **09** (2013) 024 [1304.1546].

[162] Z. Vlah, U. Seljak, M. Y. Chu and Y. Feng, *Perturbation theory, effective field theory, and oscillations in the power spectrum*, *JCAP* **1603** (2016) 057 [1509.02120].

[163] R. A. Porto, L. Senatore and M. Zaldarriaga, *The Lagrangian-space Effective Field Theory of large scale structures*, J. Cosmology Astropart. Phys. **2014** (2014) 022 [1311.2168].

[164] Z. Vlah, M. White and A. Aviles, *A Lagrangian effective field theory*, J. Cosmology Astropart. Phys. **2015** (2015) 014 [1506.05264].

[165] S.-F. Chen, Z. Vlah and M. White, *Modeling features in the redshift-space halo power spectrum with perturbation theory*, J. Cosmology Astropart. Phys. **2020** (2020) 035 [2007.00704].

[166] S.-F. Chen, Z. Vlah, E. Castorina and M. White, *Redshift-Space Distortions in Lagrangian Perturbation Theory*, *JCAP* **03** (2021) 100 [2012.04636].

[167] S.-F. Chen, Z. Vlah and M. White, *The bispectrum in Lagrangian perturbation theory*, J. Cosmology Astropart. Phys. **2024** (2024) 012 [2406.00103].

[168] T. Baldauf, E. Schaan and M. Zaldarriaga, *On the reach of perturbative descriptions for dark matter displacement fields*, *JCAP* **03** (2016) 017 [1505.07098].

[169] T. Baldauf, E. Schaan and M. Zaldarriaga, *On the reach of perturbative methods for dark matter density fields*, *JCAP* **03** (2016) 007 [1507.02255].

[170] V. Iršič and M. McQuinn, *Absorber Model: the Halo-like model for the Lyman-α forest*, *JCAP* **04** (2018) 026 [1801.02671].

[171] A. Chudaykin, M. M. Ivanov, O. H. E. Philcox and M. Simonović, *Nonlinear perturbation theory extension of the Boltzmann code CLASS*, *Phys. Rev. D* **102** (2020) 063533 [2004.10607].

[172] M. M. Ivanov, C. Cuesta-Lazaro, S. Mishra-Sharma, A. Obuljen and M. W. Toomey, *Full-shape analysis with simulation-based priors: constraints on single field inflation from BOSS*, 2402.13310.

[173] F. Sinigaglia, F.-S. Kitaura, K. Nagamine, Y. Oku and A. Balaguera-Antolínez, *Field-level Lyman-α forest modeling in redshift space via augmented nonlocal Fluctuating Gunn-Peterson Approximation*, *Astron. Astrophys.* **682** (2024) A21 [2305.10428].

[174] C. Alcock and B. Paczynski, *An evolution free test for non-zero cosmological constant*, *Nature* **281** (1979) 358.

[175] N. Padmanabhan and M. White, *Calibrating the baryon oscillation ruler for matter and halos*, Phys. Rev. D **80** (2009) 063508 [0906.1198].

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

[177] H. K. Herrera-Alcantar et al., *The Lyman-α Forest from LBGs: First 3D Correlation Measurement with DESI and Prospects for Cosmology*, 2507.21852.

[178] A. S. Almgren, J. B. Bell, M. J. Lijewski, Z. Lukić and E. Van Andel, *Nyx: A Massively Parallel AMR Code for Computational Cosmology*, ApJ **765** (2013) 39 [1301.4498].

[179] Z. Lukić, C. W. Stark, P. Nugent, M. White, A. A. Meiksin and A. Almgren, *The Lyman α forest in optically thin hydrodynamical simulations*, MNRAS **446** (2015) 3697 [1406.6361].

[180] P. Parashari, V. Gluscevic, Y. Zhang, S. Bird, M. M. Ivanov and A. He, *Lyα forest bounds on sterile neutrino production via neutrino self-interactions*, 2602.17821.

[181] M. M. Abidi and T. Baldauf, *Cubic Halo Bias in Eulerian and Lagrangian Space*, *JCAP* **1807** (2018) 029 [1802.07622].

[182] G. Cabass, O. H. E. Philcox, M. M. Ivanov, K. Akitsu, S.-F. Chen, M. Simonović et al., *BOSS Constraints on Massive Particles during Inflation: The Cosmological Collider in Action*, 2404.01894.

[183] K. Akitsu, Y. Li and T. Okumura, *Quadratic shape biases in three-dimensional halo intrinsic alignments*, *JCAP* **08** (2023) 068 [2306.00969].

[184] K. Akitsu, *Mapping the galaxy-halo connection to the galaxy bias: implication to the HOD-informed prior*, 2410.08998.

[185] A. Chudaykin, M. M. Ivanov and O. H. E. Philcox, *Reanalyzing DESI DR1: 5. Cosmological Constraints with Simulation-Based Priors*, 2602.18554.

[186] J. Jasche and G. Lavaux, *Physical Bayesian modelling of the non-linear matter distribution: New insights into the nearby universe*, A&A **625** (2019) A64 [1806.11117].

[187] G. Lavaux, J. Jasche and F. Leclercq, *Systematic-free inference of the cosmic matter density field from SDSS3-BOSS data*, 1909.06396.

[188] N.-M. Nguyen, F. Schmidt, G. Lavaux and J. Jasche, *Impacts of the physical data model on the forward inference of initial conditions from biased tracers*, *JCAP* **03** (2021) 058 [2011.06587].

[189] M. Millea, E. Anderes and B. D. Wandelt, *Sampling-based inference of the primordial CMB and gravitational lensing*, *Phys. Rev. D* **102** (2020) 123542 [2002.00965].

---

[157] H. Hildebrandt, J. Pielorz, T. Erben, L. van Waerbeke, P. Simon 和 P. Capak, CARS: CFHTLS-档案研究调查. II. 测量红移 z = 3-5 的 Lyman-断裂星系的暗物质晕质量, A&A 498 (2009) 725 [0903.3951].

[158] A. Arinyo-i Prats, J. Miralda-Escudé, M. Viel 和 R. Cen，莱曼-α森林的非线性功率谱，**JCAP 12 (2015) 017 [1506.04519]。**

[159] R. de Belsunce, B. Hadzhiyska 和 M. M. Ivanov，连接模拟与有效场论：Lyman-Alpha 的混合模型森林田野，2512.13681。

[160] S.-F. 陈, Z. Vlah 和 M. White, 速度统计与红移空间畸变的一致建模在**单圈微扰理论，JCAP 07 (2020) 062 [2005.00523]。**

[161] D. Blas, M. Garny 和 T. Konstandin, 关于宇宙学微扰理论的非线性尺度, JCAP 09 (2013) 024 [1304.1546].

[162] Z. Vlah, U. Seljak, M. Y. Chu 和 Y. Feng，微扰理论、有效场论以及功率中的振荡**谱，JCAP 1603 (2016) 057 [1509.02120]。**

[163] R. A. Porto, L. Senatore 和 M. Zaldarriaga，大尺度结构的拉格朗日空间有效场论，J. **宇宙学与天体物理学 2014 (2014) 022 [1311.2168]。**

[164] Z. Vlah, M. White 和 A. Aviles,《拉格朗日有效场论》,《宇宙学与天体粒子物理学杂志》2015 (2015) 014 [1506.05264].

[165] S.-F. Chen, Z. Vlah 和 M. White，使用摄动理论对红移空间晕功率谱中的特征建模，**J. 宇宙学与天体物理学 2020 (2020) 035 [2007.00704]。**

[166] S.-F. Chen, Z. Vlah, E. Castorina 和 M. White，《拉格朗日微扰理论中的红移空间畸变》，JCAP **03 (2021) 100 [2012.04636].**

[167] 陈士丰, Z. Vlah 和 M. White, 拉格朗日微扰理论中的双谱, 宇宙学与天体物理期刊 **2024 (2024) 012 [2406.00103].**

[168] T. Baldauf, E. Schaan 和 M. Zaldarriaga, 关于扰动描述在暗物质位移场中的适用范围，**JCAP 03 (2016) 017 [1505.07098]。**

[169] T. Baldauf, E. Schaan 和 M. Zaldarriaga, 关于扰动方法在暗物质密度场中的适用范围, JCAP **03 (2016) 007 [1507.02255].**

[170] V. Irˇ siˇ c 和 M. McQuinn，吸收器模型：类晕模型用于赖曼-α森林，JCAP 04 (2018) 026 [1801.02671].

[171] A. Chudaykin, M. M. Ivanov, O. H. E. Philcox 和 M. Simonovi´ c，非线性微扰理论扩展**玻尔兹曼代码 CLASS, 物理评论 D 102 (2020) 063533 [2004.10607]。**

[172] M. M. Ivanov, C. Cuesta-Lazaro, S. Mishra-Sharma, A. Obuljen 和 M. W. Toomey，全形态分析与基于模拟的先验：来自 BOSS 的单场膨胀约束，2402.13310。

[173] F. Sinigaglia, F.-S. Kitaura, K. Nagamine, Y. Oku 和 A. Balaguera-Antolínez，通过增强非局域波动甘-彼得森近似在红移空间中进行场级莱曼-α森林建模，Astron. Astrophys. 682 (2024) A21 [2305.10428].

[174] C. Alcock 和 B. Paczynski,《非零宇宙常数的无演化检验》,《自然》281 (1979) 358。[175] N. Padmanabhan 和 M. White,《校准物质和晕的重子振荡标尺》,《物理评论D》80 (2009)

063508 [0906.1198].

[176] F. Bernardeau, S. Colombi, E. Gaztanaga 和 R. Scoccimarro, 宇宙的大尺度结构与宇宙学**微扰理论, 物理报告 367 (2002) 1 [astro-ph/0112551].**

[177] H. K. Herrera-Alcantar 等人，《来自 LBG 的 Lyman-α 森林：使用 DESI 的首次三维相关性测量》宇宙学的前景，2507.21852。

[178] A. S. Almgren, J. B. Bell, M. J. Lijewski, Z. Lukić 和 E. Van Andel，Nyx：一个大规模并行自适应网格细化（AMR）代码用于**计算宇宙学,《天体物理学杂志》765 (2013) 39 [1301.4498]。**

[179] Z. Lukić, C. W. Stark, P. Nugent, M. White, A. A. Meiksin 和 A. Almgren, 光学稀薄中的莱曼 α 森林**流体动力学模拟，MNRAS 446 (2015) 3697 [1406.6361]。**

[180] P. Parashari, V. Gluscevic, Y. Zhang, S. Bird, M. M. Ivanov 和 A. He, Lyα 林对惰性中微子的限制通过中微子自相互作用的产生，2602.17821。

[181] M. M. Abidi 和 T. Baldauf，《欧拉与拉格朗日空间中的立方晕偏差》，JCAP 1807 (2018) 029 [1802.07622]。 [182] G. Cabass, O. H. E. Philcox, M. M. Ivanov, K. Akitsu, S.-F. Chen, M. Simonović 等，《BOSS 对大质量的限制》膨胀时期的粒子：宇宙对撞机的实际作用, 2404.01894.

[183] K. Akitsu, Y. Li 和 T. Okumura, 三维晕固有排列中的二次形状偏差, JCAP 08

(2023) 068 [2306.00969].

[184] K. Akitsu，将星系-晕关联映射到星系偏差：对HOD知情先验的影响，2410.08998。[185] A. Chudaykin, M. M. Ivanov和O. H. E. Philcox，重新分析DESI DR1：5. 宇宙学约束基于模拟的先验，2602.18554。

[186] J. Jasche 和 G. Lavaux，非线性物质分布的物理贝叶斯建模：新的见解**近邻宇宙，A&A 625 (2019) A64 [1806.11117]。**

[187] G. Lavaux, J. Jasche 和 F. Leclercq, 从 SDSS3-BOSS 系统无关地推断宇宙物质密度场数据，1909.06396。

[188] N.-M. Nguyen, F. Schmidt, G. Lavaux 和 J. Jasche, 物理数据模型对前向推断的影响**来自有偏示踪剂的初始条件，JCAP 03 (2021) 058 [2011.06587]。**

[189] M. Millea, E. Anderes 和 B. D. Wandelt，基于采样的原初 CMB 和引力透镜推断，**物理评论 D 102 (2020) 123542 [2002.00965]。**

[190] N. Porqueres, O. Hahn, J. Jasche and G. Lavaux, *A hierarchical field-level inference approach to reconstruction from sparse Lyman-α forest data*, *Astron. Astrophys.* **642** (2020) A139 [2005.12928].

[191] E. Tsaprazi, N.-M. Nguyen, J. Jasche, F. Schmidt and G. Lavaux, *Field-level inference of galaxy intrinsic alignment from the SDSS-III BOSS survey*, *JCAP* **08** (2022) 003 [2112.04484].

[192] A. Andrews, J. Jasche, G. Lavaux and F. Schmidt, *Bayesian field-level inference of primordial non-Gaussianity using next-generation galaxy surveys*, *Mon. Not. Roy. Astron. Soc.* **520** (2023) 5746 [2203.08838].

[193] I. Babić, F. Schmidt and B. Tucci, *BAO scale inference from biased tracers using the EFT likelihood*, *JCAP* **08** (2022) 007 [2203.06177].

[194] A. E. Bayer, C. Modi and S. Ferraro, *Joint velocity and density reconstruction of the Universe with nonlinear differentiable forward modeling*, *JCAP* **06** (2023) 046 [2210.15649].

[195] S. S. Boruah, E. Rozo and P. Fiedorowicz, *Map-based cosmology inference with lognormal cosmic shear maps*, *Mon. Not. Roy. Astron. Soc.* **516** (2022) 4111 [2204.13216].

[196] A. J. Zhou, X. Li, S. Dodelson and R. Mandelbaum, *Accurate field-level weak lensing inference for precision cosmology*, *Phys. Rev. D* **110** (2024) 023539 [2312.08934].

[197] J. Stadler, F. Schmidt and M. Reinecke, *Cosmology inference at the field level from biased tracers in redshift-space*, *JCAP* **10** (2023) 069 [2303.09876].

[198] N. Porqueres, A. Heavens, D. Mortlock, G. Lavaux and T. L. Makinen, *Field-level inference of cosmic shear with intrinsic alignments and baryons*, 2304.04785.

[199] BEYOND-2PT collaboration, *A Parameter-Masked Mock Data Challenge for Beyond-Two-Point Galaxy Clustering Statistics*, *Astrophys. J.* **990** (2025) 99 [2405.02252].

[200] J. Stadler, F. Schmidt, M. Reinecke and M. Esposito, *Fast, accurate and perturbative forward modeling of galaxy clustering. Part II. Redshift space*, *JCAP* **11** (2025) 055 [2411.04513].

[201] J. Stadler, F. Schmidt and M. Reinecke, *Fast, accurate and perturbative forward modeling of galaxy clustering. Part I. Galaxies in the restframe*, *JCAP* **04** (2025) 089 [2409.10937].

[202] I. Babić, F. Schmidt and B. Tucci, *Forward vs Backward: Improving BAO Constraints with Field-Level Inference*, 2505.13588.

[203] M. Peron, T. Nishimichi, M. Pietroni and A. Taruya, *Renormalized perturbation theory at field-level: the LSS bootstrap in GridSPT*, *JCAP* **10** (2025) 098 [2506.07105].

[204] N.-M. Nguyen, F. Schmidt, B. Tucci, M. Reinecke and A. Kostić, *How Much Information Can Be Extracted from Galaxy Clustering at the Field Level?*, *Phys. Rev. Lett.* **133** (2024) 221006 [2403.03220].

[205] K. Akitsu, M. Simonović, S.-F. Chen, G. Cabass and M. Zaldarriaga, *Cosmology inference with perturbative forward modeling at the field level: a comparison with joint power spectrum and bispectrum analyses*, 2509.09673.

[206] F. Spezzati, M. Marinucci and M. Simonović, *Equivalence of the field-level inference and conventional analyses on large scales*, 2507.05378.

---

[190] N. Porqueres, O. Hahn, J. Jasche 和 G. Lavaux，一种用于重建的分层场级推断方法稀疏的莱曼-α森林数据，天文学与天体物理学，642 (2020) A139 [2005.12928]。

[191] E. Tsaprazi, N.-M. Nguyen, J. Jasche, F. Schmidt 和 G. Lavaux，星系固有排列的场级推断来自 SDSS-III BOSS 调查, JCAP 08 (2022) 003 [2112.04484]。

[192] A. Andrews, J. Jasche, G. Lavaux 和 F. Schmidt，使用贝叶斯场级推断原初非高斯性下一代银河系调查，蒙特利尔皇家天文学会月刊，520 (2023) 5746 [2203.08838]。

[193] I. Babi´ c, F. Schmidt 和 B. Tucci，使用 EFT 似然从有偏示踪器推断 BAO 尺度，JCAP 08 (2022) 007 [2203.06177].

[194] A. E. Bayer, C. Modi 和 S. Ferraro, 使用非线性方法对宇宙的速度和密度进行联合重建可微正向建模，JCAP 06 (2023) 046 [2210.15649]。

[195] S. S. Boruah, E. Rozo 和 P. Fiedorowicz，基于图的宇宙学推断与对数正态宇宙剪切图，Mon. Not. 皇家天文学会学报 516 (2022) 4111 [2204.13216]。

[196] 周安捷, 李晓, S. Dodelson 和 R. Mandelbaum, 精准宇宙学的精确场级弱透镜推断, 物理评论 D 110 (2024) 023539 [2312.08934]。

[197] J. Stadler, F. Schmidt 和 M. Reinecke, 从红移空间中偏置示踪器的场级推断宇宙学，JCAP 10 (2023) 069 [2303.09876]。

[198] N. Porqueres, A. Heavens, D. Mortlock, G. Lavaux 和 T. L. Makinen，基于场级的宇宙剪切推断内在排列与重子，2304.04785。

[199] 超越二维协作：一种用于超越二维星系聚类的参数屏蔽模拟数据挑战统计学，天体物理学杂志 990 (2025) 99 [2405.02252]。

[200] J. Stadler, F. Schmidt, M. Reinecke 和 M. Esposito，快速、精确且微扰的星系正向建模聚类。第二部分。红移空间，JCAP 11 (2025) 055 [2411.04513]。

[201] J. Stadler, F. Schmidt 和 M. Reinecke，快速、精确且微扰的星系聚类前向建模。第一部分。静止系中的星系，JCAP 04 (2025) 089 [2409.10937]。

[202] I. Babi´ c, F. Schmidt 和 B. Tucci, 《前向与后向：通过场级推断改善BAO约束，》 2505.13588.

[203] M. Peron, T. Nishimichi, M. Pietroni 和 A. Taruya，场级的重整化微扰理论：LSS 自举发表于 GridSPT, JCAP 10 (2025) 098 [2506.07105].

[204] N.-M. Nguyen, F. Schmidt, B. Tucci, M. Reinecke 和 A. Kostić，可以从中提取多少信息场级别的银河系聚类？, 物理评论快报 133 (2024) 221006 [2403.03220]。

[205] K. Akitsu, M. Simonovi´ c, S.-F. Chen, G. Cabass 和 M. Zaldarriaga, 基于扰动前向建模的场级宇宙学推断：与联合功率谱和双谱分析的比较, 2509.09673.

[206] F. Spezzati, M. Marinucci 和 M. Simonović，大规模研究中场级推理与传统分析的等效性秤，2507.05378。