Draft version March 30, 2015
Preprint typeset using LaTeX style emulateapj v. 12/16/11

FIBER COLLISION CORRECTION METHOD FOR BOSS POWER SPECTRUM

ChangHoon Hahn[1], Michael R. Blanton[1], Roman Scoccimarro[1]

Draft version March 30, 2015

ABSTRACT

Fiber-fed multi-object spectroscopic surveys, with their ability to collect an unprecedented number of redshifts, dominate cosmological studies. However, the physical size of the fibers limit these surveys from successfully collecting redshifts from galaxies within the angular fiber-collision scale, which ultimately lead to significant systematic effects on galaxy clustering measurements. Using simulated mock catalogues, which have been extensively used in interpreting large-scale structure clustering results (such as the correlation function and power spectrum) of the Sloan Digital Sky Survey, we investigate the effects of fiber-collisions on the galaxy power spectrum in particular. We compare the methods currently used to correct for fiber-collisions in the literature and assess their success in reconstructing the true power spectrum for mock catalogues.

We present our fiber-collision correction method, which statistically reconstructs the clustering of fiber-collided pairs, using the distribution of the line-of-sight displacements between resolved fiber-collided pairs, and properly accounts for fiber-collisions in the shot-noise correction term of the power spectrum estimator. Using our method, we successfully reproduce the true power spectrum for the mock catalogues with residuals of $\sim 1\%$ at $k \sim 0.3$ h/Mpc and $< 10\%$ at $k \sim 0.8$ h/Mpc. Our comparisons to other correction methods demonstrate that our method most accurately reconstructs the true power spectrum from fiber-collided data especially at small scales ($k > 0.1$ h/Mpc), with the added advantage that the method can be validated and calibrated in actual observations. With statistical precision no longer obstructing galaxy clustering measurements, our correction method will allows us constrain systematic effects of fiber-collisions and extend power spectrum measurements to smaller scales.

Subject headings: cosmology: observations – cosmology: large-scale structure of universe – galaxies: distances and redshift – galaxies: halos – galaxies: statistics

1. INTRODUCTION

Through the millions of redshifts obtained through redshift surveys, cosmological measurements such as galaxy clustering statistics are no longer dominated by uncertainties from statistical precision, but from systematic effects of the measurements. These redshift surveys, such as the 2dF Galaxy Redshift Survey (2dFGRS; Colless 1999) and the Sloan Digital Sky Survey-III Baryon Oscillation Spectroscopic Survey (SDSS-III BOSS; Anderson et al. 2012; Dawson et al. 2013) and future redshift surveys such as the Extended Baryon Oscillation Spectroscopic Survey (eBOSS; **WHO DO CITE?**) use fiber fed spectrographs.

For each galaxy, a fiber is used to obtain a spectroscopic redshift. However, due to the physical size of the fibers even with efficient targeting strategies (e.g. Blanton et al. 2003 for SDSS), if galaxies are located within the fiber-collision angular scale from one another on the sky, separate fibers cannot be placed adjacently to separately extract the spectroscopic information (Yoon et al. 2008 **CITE MORE**). In these situations, the redshifts of fiber-collided galaxies are not individually resolved and only a single redshift is measured. With redshifts of galaxies in close angular proximities missing from the sample, any clustering statistic probing these scales will be systematically affected.

As our cosmological surveys extend further to higher redshifts, the systematic effect becomes worse as the fiber-collision angular scale corresponds to a larger comoving scale, thereby affecting our measurements on larger scales. SDSS-III BOSS, in particular, has an angular fiber-collision scale of 62". This corresponds to ~ 0.4 Mpc/h at the center of the survey redshift redshift range, and fiber-collided galaxies account for $\sim 5\%$ of the galaxy sample (Anderson et al. 2012). Although this accounts for a relatively small fraction of redshifts, its effect on clustering measurements such as the power spectrum is significant and must be understood as these measurements probe to smaller scales and as higher order statistics such as the three-point correlation function or the bispectrum are used in cosmological studies. Perhaps with the use of robotic fiber positioner technology in future surveys such as the Dark Energy Survey Instrument (DESI; Schlegel et al. 2011; Morales et al. 2012; Makarem et al. 2014) will render fiber-collisions obsolete. However, until then accounting for the systematic effects remains an unavoidable challenge for clustering statistics.

To correct for fiber collisions, one common approach used in clustering statistics is the nearest neighbor method (Zehavi et al. 2002; Berlind et al. 2006; Anderson et al. 2012). For fiber-collided galaxies without resolved redshifts, the method assigns the statistical weight of the fiber-collided galaxy to its nearest angular neighbor. This provides a reasonable correction for the fiber-collision effects at scales much larger than the fiber-collision scales; however the correction falls short near the fiber-collision scale. In fact, as Zehavi et al. (2005) find, fiber-collisions affect the two-point correla-

[1] Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003; chh327@nyu.edu

tion function (2PCF) measurements even on scales significantly larger than the fiber-collision scale (> 1 Mpc/h). For power spectrum measurements in BOSS, the nearest neighbor method has often been supplemented with adjustments in the constant shot-noise term of the power spectrum estimator to correct for fiber-collisions (Beutler et al. 2014; Gil-Marín et al. 2014). While this shot-noise term adjustment has been calibrated within the mock catalogues used to interpret the BOSS clustering results, validating these methods for observations remains a challenge.

Recently Guo et al. (2012), focusing on SDSS-III BOSS like samples, proposed a fiber-collision correction method for the 2PCF that is able to reasonably correct for fiber-collisions above and below the collision scale. Guo et al. (2012) estimates the total contribution of fiber-collided galaxies to the two point correlation function by examining the pair statistics in overlapping tiling regions of the survey, where a smaller fraction of galaxies suffer from fiber-collisions. Unfortunately, this method cannot be applied to the Fourier clustering measurements due to the complex geometry of the overlapping regions. As a result, we propose a fiber-collision correction method that, first, improves upon the nearest neighbor method by using the distribution of the line-of-sight displacement between resolved fiber-collided galaxies to statistically reconstruct the clustering of fiber-collided galaxies; and proper accounts for fiber-collisions in the shot-noise correction term of the power spectrum estimator.

In Section 2, we present a brief description of the simulated mock catalogues with realistic applied fiber-collisions and power spectrum estimator used throughout the paper. Then during our discussion of the two components of our correction method in Section 3, we quantify the effects of fiber-collisions of power spectrum measurements and test and compare common correction methods. Afterwards, we present our correction method in Section 4.1, test it on the mock catalogues, and compare it to other correction methods. Finally in Section 5 we summarize and discuss our results.

2. FIBER-COLLIDED MOCK CATALOGUES

For various purposes such as calculating covariance matrices or estimating cosmic variance, simulated mock catalogues play a crucial role in interpreting clustering measurements of observed galaxies (Scoccimarro & Sheth 2002; Anderson et al. 2012; Manera et al. 2013). They also provide a means of understanding systematic effects such as fiber-collisions (Guo et al. 2012; Manera et al. 2013 **LIST MORE**) since systematic effects can be simulated on the mock catalogues. This allows us to test how these effects influence clustering measurements and devise correction methods that attempt to account for these effects.

A direct way of understanding the effects of fiber-collisions on clustering statistics in observations is to first apply fiber-collisions to mock catalogues then compare the clustering statistics obtained from mock catalogues with and without the fiber collisions. Correction methods for fiber-collisions can then be applied to the fiber-collided mocks and the merit of the correction method can be quantified by how successfully they reproduce the clustering statistics of the original mock catalogues without fiber-collisions. The optimal correction method can then be applied to the observed data with some assurance that it accounts for fiber-collisions and improves the clustering measurements.

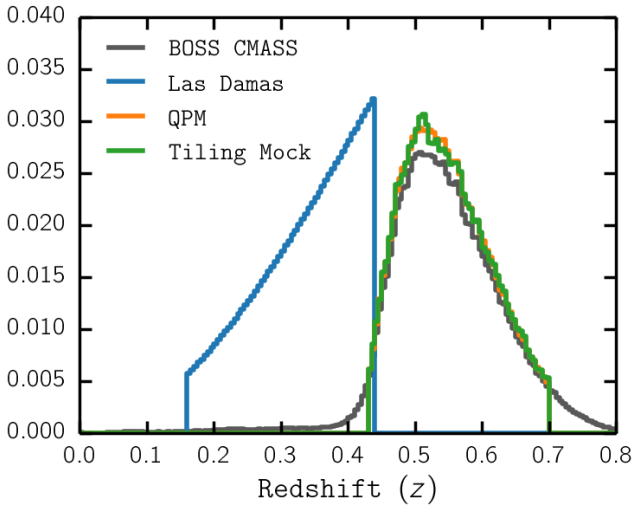

FIG. 1.— Normalized galaxy redshift distribution for Las Damas (blue), QPM (orange), and Tiling Mock (green) mock catalogues. The normalized redshift distribution of BOSS DR12 CMASS is also plotted (black). Bin size of $\Delta z = 0.025$ was used to compute the distributions. Las Damas mock catalogue has a constant number density of galaxies throughout its redshift range ($0.16 < z < 0.44$). QPM and Tiling Mock catalogues, on the other hand, trace the BOSS CMASS redshift distribution ($0.43 < z < 0.7$).

When applying the fiber-collisions to the mock catalogues, it is essential to apply them in the same manner they affect the observational data. For BOSS, galaxies within 62" are fiber-collided (Anderson et al. 2012). In reality, this simple criteria is further complicated by the tiling scheme of observing plates that create overlapping regions, which have a higher success rate in resolving galaxy spectra within the fiber-collision angular scale (Guo et al. 2012). Furthermore, fiber-collisions are only one of the systematic effects that influence BOSS data. Systematic effects include the unique geometry of the BOSS survey, the variable completeness in different sectors of the survey, and redshift failures (Anderson et al. 2012).

Effects of fiber-collisions must be understood and interpreted in conjunction with these systematic effects. Therefore, in this paper, we use LasDamas (McBride et al. 2009, 2011), Quick Particle Mesh (White et al. 2014), and Tiling mock catalogues, which have already been extensively used in interpreting clustering results for SDSS and BOSS and are generated through different prescriptions. Therefore they provide complementary tests fiber-collisions and our correction method tailored for BOSS.

Each of these mock catalogues use some sort of N-body dark matter simulation along with a HOD prescription to populate the dark matter halos with "galaxies".

The LasDamas Galaxy mock catalogues is constructed from the Oriana N-body Large Suite of Dark Matter Simulations (LasDamas) with 1280^3 particles with a box with side length, $L_{\rm Box} = 2400$ Mpc/h. The N-body LasDamas simulations uses a spatially flat ΛCDM cosmology with $\Omega_{\rm m} = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_{\rm b} = 0.04$, $\sigma_8 = 0.8$, $n_{\rm s} = 1$ and $h = 0.7$. The dark matter halos are then identified

using a friends-of-friends (FOF) algorithm (Davis et al. 1985) with a linking length $b = 0.2$ times the mean interparticle separation. The over-densities of dark matter halos are then populated with galaxies using the halo occupation distribution (HOD) framework to construct the galaxy mock catalogues (McBride et al. 2009, 2011). The HOD prescription is specified so that the galaxy mock catalogues reproduce the galaxy number density and projected correlation function of the observed SDSS Luminous Red Galaxy sample with $M_g < -21.8$. The LasDamas galaxy mock catalogues have 40 realizations that spans the redshift range, $0.16 < z < 0.44$ and are restricted to the survey geometry of SDSS Data Release 7. They also include redshift space distortions from velocities, but do not model fiber-collisions.

To model the fiber-collisions of the observed galaxies in BOSS, we impose fiber-collisions at the angular scale of $62''$. Once we identify galaxies within $62''$ of each other, we arbitrary select one of the galaxies and assign the statistical weights of the other galaxies within $62''$ of it. Roughly $\sim 2.5\%$ of the galaxies in the LasDamas mock catalogues are affected by fiber-collisions.

Next, the QPM mock galaxy catalogues uses a "quick particle mesh" method, which uses a low resolution particle-mesh N-body solver to evolve particles within a period simulation volume. The particles are assigned halo masses in order to match the halo mass function and large-scale bias of halos of high resolution simulations. Afterward the Tinker et al. (2012) HOD parameterization is used to populate the halos. The mock galaxy sample is then trimmed to the BOSS CMASS survey footprint, downsamped based on angular sky completeness (sector completeness) and radial selection. Furthermore, QPM mocks model the fiber-collisions of the BOSS CMASS sample ($62''$ angular fiber-collision scale). The QPM galaxy mock catalogue uses the following ΛCDM cosmology: $\Omega_{\rm m} = 0.29$, $\Omega_\Lambda = 0.71$, $\sigma_8 = 0.8$, $n_{\rm s} = 0.97$ and $h = 0.7$. We use 1000 realizations of the QPM galaxy mock catalogue. For a detailed description of the QPM galaxy mock catalogues we refer readers to White et al. (2014).

Finally the Tiling Mock catalogue is generated using **INSERT TILING MOCK DESCRIPTION HERE**

In Figure 2, we plot the redshift distribution for LasDamas, QPM, and Tiling Mock catalogues along with the redshift distribution of BOSS DR12 galaxies. The LasDamas mock catalogue, which is modeled after the SDSS LRG sample has a constant $\bar{n}(z)$ over the redshift $0.16 < z < 0.44$, while the QPM and Tiling mock redshift distributions closely trace the observed BOSS redshift distribution (roughly $0.43 < z < 0.7$).

2.1. *Power Spectrum Estimator*

In this paper, we focus on the effects of fiber-collisions on the galaxy power spectrum out of the various clustering measurements. Throughout the paper, when we measure the power spectrum, we use the standard Feldman et al. (1994) (hereafter FKP) estimator and the following formalisms from the FKP paper:

$$P(\mathbf{k}) =< |F(\mathbf{k})|^2 > - \frac{(1+\alpha)\int d^3r\, \bar{n}(\mathbf{r}) w^2_{\rm FKP}(\mathbf{r})}{\int d^3r\, \bar{n}^2(\mathbf{r}) w^2_{\rm FKP}(\mathbf{r})}, \quad (1)$$

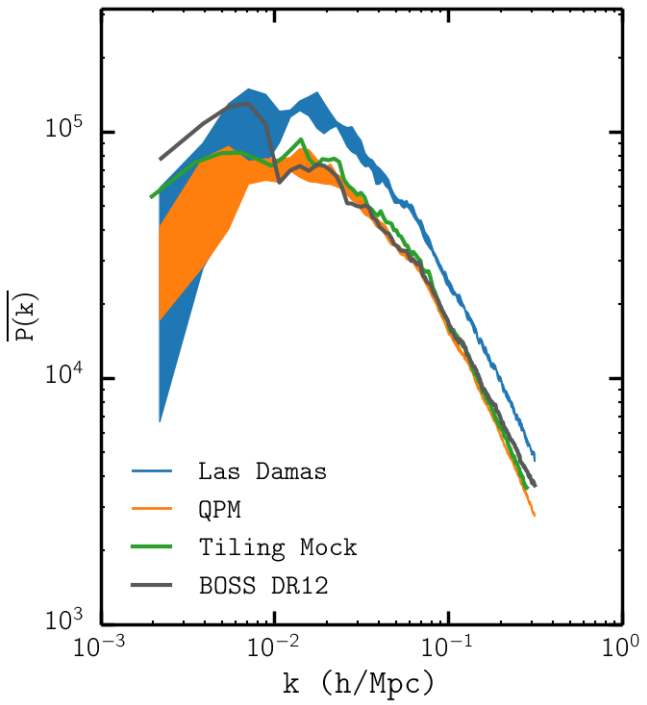

FIG. 2.— Average power spectrum $\overline{P(k)}$ for BOSS DR12 (black) and the mock catalogues Las Damas (blue), QPM (orange), and Tiling Mock (green), listed in Section 2 using the FKP $P(k)$ estimator. The width of the $P(k)$ for Las Damas and QPM represent the sample variance ($\Delta P(k)$) calculated from the multiple realizations of the mock catalogues. We emphasize that fiber-collisions were not applied to the mock catalogues for the $P(k)$ measurements. Meanwhile, galaxy weights (Equation 4) are applied to the BOSS DR12 $P(k)$ measurement. We remind readers that the LasDamas mock catalogue has an overall greater $\overline{P(k)}$ due to the fact that the catalogue spans the SDSS LRG sample redshift range instead of the BOSS redshift range of the other mock catalogues.

where

$$F(\mathbf{r}) = \frac{w_{\rm FKP}(\mathbf{r})[n_g(\mathbf{r}) - \alpha n_r(\mathbf{r})]}{[\int d^3r\, \bar{n}^2(\mathbf{r}) w^2_{\rm FKP}(\mathbf{r})]^{\frac{1}{2}}} \quad (2)$$

and the second term represents the constant shot-noise contribution to the power due to the discrete density field of our galaxies. n_g is the galaxy density, n_r is the density of the synthetic random catalogue, α is the ratio of the number of galaxies over the number of synthetic random galaxies, $\bar{n}$ is the mean density of the galaxies, and $w_{\rm FKP}$ is the minimum variance weight derived in Feldman et al. (1994):

$$w_{\rm FKP}(\mathbf{r}) = \frac{1}{1 + \bar{n}(\mathbf{r}) P_0} \quad (3)$$

where P_0 is the power spectrum amplitude at which the error is minimized. We use $P_0 = 20000\ {\rm Mpc}^3/h^3$ for our analysis, which corresponds to $k \sim 0.1\ h/{\rm Mpc}$.

For LasDamas, QPM, and Tiling mock catalogues we use the FKP estimator described above to measure the power spectrum in Figure 2. The fiber-collisions described in Section 2 are not applied to the mock catalogues used to calculated the $P(k)$ in Figure 2. Without fiber-collisions, we are able to use the exact formulation of Equation 1 in the FKP $P(k)$ estimator.

For the BOSS Data Release 12 $P(k)$ (black) in Figure 2, systematic weights are assigned to the galaxies in or-

der to account for sector completeness, redshift failures, and fiber-collisions. Each galaxy has a statistical weight determined by,

$$w_{\rm tot} = w_{\rm sys}(w_{\rm rf} + w_{\rm fc} - 1), \quad (4)$$

(Anderson et al. 2012; Beutler et al. 2014). As a result, we modify Equation 1 to obtain the power spectrum. Instead of the galaxy density term $n_g(\mathbf{r})$ we use a weighted galaxy density term $n'_g(\mathbf{r}) = n_g(\mathbf{r})w(\mathbf{r})$, which includes the systematic weights for each galaxy. Also, instead of $\alpha = N_{\rm gal}/N_{\rm ran}$ we use $\alpha' = N'_{\rm gal}/N_{\rm ran}$ where $N'_{\rm gal} = \sum_{\rm gal} w_{\rm tot}$.

For LasDamas and QPM, which have multiple realizations, we compute the sample variance of the power spectrum

$$\Delta P(k) = \sqrt{\frac{1}{N_{\rm realization}} \sum_{i}^{N_{\rm realization}} (P_i(k) - \overline{P(k)})^2}. \quad (5)$$

$N_{\rm realization}$ is the total number of realizations (40 for LasDamas and 1000 for QPM) and $P_i(k)$ is the power spectrum for each realization. $\Delta P(k)$ is represented by the width of the $P(k)$ in Figure 2.

3. FIBER-COLLISION CORRECTION METHOD

As mentioned above, a common approach in accounting for fiber-collisions in clustering measurements has been to use the nearest angular neighbor method (Zehavi et al. 2002; Berlind et al. 2006; Anderson et al. 2012). For galaxies without resolved spectroscopic redshifts due to fiber-collisions, the entire statistical weight of the galaxy is assigned to its nearest angular neighbor. This method assumes that all galaxies with angular fiber-collision scale (< 62" for BOSS) are correlated. In other words, the method assumes that galaxies within the fiber-collision angular scale reside in the same halo. Consequently, galaxies within the angular fiber-collision scale purely by coincidence (hereafter referred to as "chance alignments") are incorrectly assumed to be gravitationally correlated, so when their statistical weight is added to its nearest angular neighbor, they are in fact displaced significantly from its true position. For BOSS, the nearest angular neighbor method can displace galaxies by up to 500 Mpc. Even for fiber-collided galaxies that reside in the same group or cluster, upweighting the nearest neighbor ignores halo-scale line-of-sight displacements and places the fiber-collided galaxies on top of each other.

Although the nearest-neighbor method provides a reasonable estimate at large scales, as we demonstrate in Section 3.1, at small scales its effects on the power spectrum are significant and dominate sample variance. Afterwards, in Section 3.1, we present an improved method to account for correlated fiber-collided galaxies using the line-of-sight displacement between close galaxy pairs. Then in Section 3.2, we present adjustments to the FKP power spectrum estimator that properly accounts for fiber-collisions through the shot-noise correction term in the estimator.

3.1. *Line-of-Sight Displacement of Fiber-collided Pairs*

It is impossible to determine in observed galaxy data whether fiber-collided galaxies without resolved spectroscopic redshift are either correlated or chance alignments. Fortunately, using the close galaxy pairs with resolved redshifts (mainly in the overlapping regions mentioned in Section 2), it is possible to model the fraction of galaxy pairs that are corrected and uncorrelated.

Meanwhile, for simulated mock catalogues, described in Section 2, modeling the fiber-collided galaxy pairs is a much more direct task since fiber-collisions are post-processed after the simulated galaxy positions are generated. Hence all redshift values are resolved regardless of angular proximity. With both redshifts of the fiber-collided pair galaxies, the fiber-collisions we impose, or are already imposed, on the mock catalogues in Section 2 emulate the observed data by assigning nearest neighbor galaxy weights. So for a close galaxy pair, one galaxy ("the nearest neighbor") will have the statistical weight of both galaxies, while the other has a statistical weight of 0. Furthermore, having the redshift values of close galaxy pairs in mock catalogues, allows us to model the correlated and uncorrelated fiber-collided pairs.

In order to determine whether or not fiber-collided pairs are correlated, we examine the comoving line-of-sight displacement ($d_{\rm LOS}$) of fiber-collided galaxy pairs. We calculate the $d_{\rm LOS}$ by taking the difference of the line-of-sight comoving distance from the resolved redshifts:

$$d_{\rm LOS} = D_{\rm C}(z_1) - D_{\rm C}(z_2). \quad (6)$$

$D_{\rm C}(z)$ is the line-of-sight comoving distance at redshift z (Hogg 1999). z_1 and z_2 represent the resolved redshifts of the galaxies in the fiber-collided pair.

After calculating the $d_{\rm LOS}$ for every fiber-collided pairs, in Figure 3 we present the distribution of $d_{\rm LOS}$ for LasDamas (blue), QPM (orange), Tiling Mock (green), and BOSS DR12 (black). All of the $d_{\rm LOS}$ distributions contain two components: a peak contained within the range -20 Mpc $< d_{\rm LOS} < 20$ Mpc and a flat component (hereafter "tail" component) outside the peak that extends to $d_{\rm LOS} \sim \pm 500$ Mpc (entire range of the distribution not displayed in Figure 3). For the BOSS data, since we cannot compute the $d_{\rm LOS}$ for fiber-collided pairs without resolved spectra, the $d_{\rm LOS}$ distribution only reflects the $d_{\rm LOS}$ from fiber-colldied pairs *with* resolved spectroscopic redshifts, mostly from tile overlapping regions of the BOSS survey.

Galaxies within the same halo, due to their gravitational interactions at halo-scales, are more likely to be in close angular proximity with each other. These galaxies in over-dense regions cause the peak in the $d_{\rm LOS}$ distribution. While the "tail" component is composed of chance alignment galaxy pairs that are arbitrarily in close angular proximity even when they are not correlated.

Focusing on the peak of the distribution, we note that it closely traces an Gaussian function. Therefore, we fit a Gaussian of the form,

$$p(d_{\rm LOS}) = A\, e^{-d^2_{\rm LOS}/2\sigma^2} \quad (7)$$

for a mathematical prescription of the $d_{\rm LOS}$ distribution peak as a function of the displacement for each of the data catalogues in Figure 3. The best-fit σ and A parameters are obtained by fitting Equation 7 to the $d_{\rm LOS}$ distribution peaks using MPFIT (Markwardt 2009). We list the values of σ obtained for the mock catalogues in Table 1. Both Table 1 and Figure 3 illustrate that the

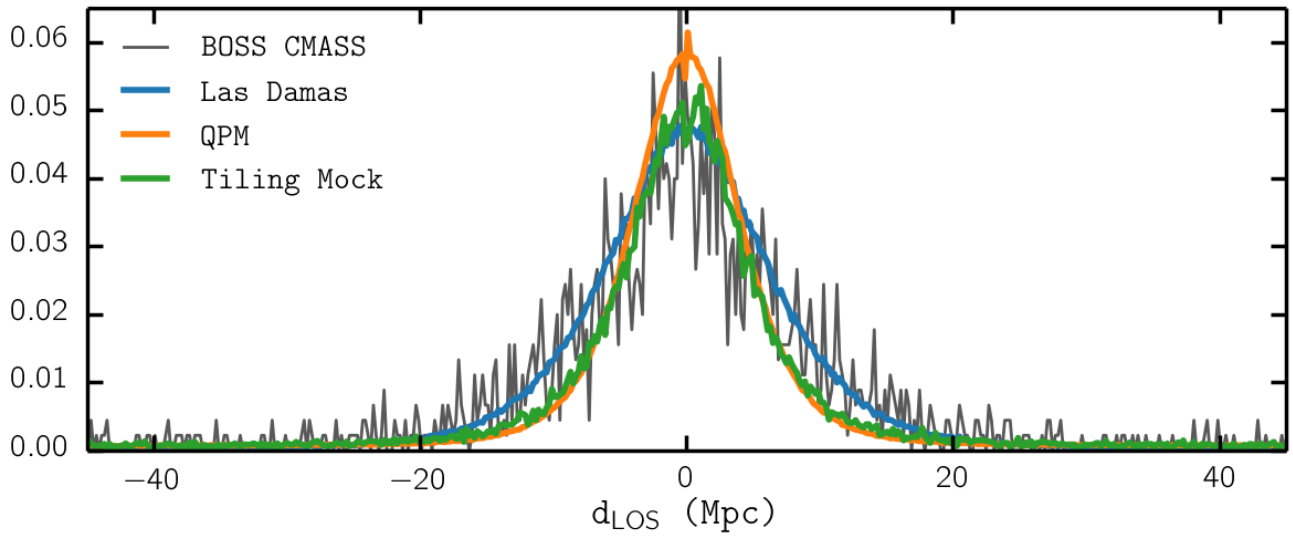

FIG. 3.— Normalized distribution of $d_{\rm LOS}$ for Las Damas (blue), QPM (orange), and Tiling Mock (green) mock catalogues. The normalized $d_{\rm LOS}$ distribution of BOSS DR12 is also plotted (black). Bin size of $\Delta d = 0.2$ was used to compute these $d_{\rm LOS}$ distributions. Although we focus on the peak of the distribution (discussed in 3.1), the $d_{\rm LOS}$ distribution extends to $\sim \pm 500$ Mpc (not shown above).

TABLE 1
$d_{\rm LOS}$ DISTRIBUTION BEST-FIT PARAMETERS

Catalogue	σ (Mpc)	$f_{\rm peak}$
Las Damas	6.85	0.71
QPM	4.98	0.65
Tiling Mock	5.38	0.57
BOSS DR12	7.6	0.65

Notes: Best-fit parameter σ and peak fraction $f_{\rm peak}$ for the $d_{\rm LOS}$ distributions in Figure 3.

$d_{\rm LOS}$ distributions for the mock catalogues closely trace the distribution of the BOSS DR12, which further goes to support the use of mock catalogues in our investigation of fiber-collisions.

Using the best-fit to the peak of the $d_{\rm LOS}$ distribution, we estimate the fraction of fiber-collided pairs that are within the peak:

$$f_{\rm peak} = \frac{\sum\limits_{\rm peak} p(d_{\rm LOS})}{N_{\rm pairs}}, \tag{8}$$

where $N_{\rm pairs}$ is the total number of fiber-collided pairs. As mentioned above, this portion of fiber-collided pairs are galaxy pairs that are correlated. We list the $f_{\rm peak}$ values for each of the data catalogues in Table 1. The $f_{\rm peak}$ values for the mock catalogue are also in good agreement with the BOSS DR12 $f_{\rm peak}$.

For the most commonly used nearest angular neighbor method (hereafter NN) to be completely correct, the $d_{\rm LOS}$ distribution in Figure 3 would have to be a delta function, which is clearly not the case. By simply incorporating the peak of the $d_{\rm LOS}$ distribution, we are able to significantly improve clustering statistics on small scales. Rather than placing the fiber-collided galaxy on top of its nearest angular neighbor as NN does, placing the fiber-collided galaxy at a displacement sampled from the $d_{\rm LOS}$ distribution, away from its nearest neighbor better accounts for the small scale in clustering measurements. On mock catalogues, we test this by first applying the NN to the mock catalogues. We then sample a displacement d_i from $p(d_{\rm LOS})$ (Equation 7) for each fiber-collided pair. Afterwards we "place" the fiber-collided galaxy a distance d_i from the nearest neighbor galaxy and reduce the statistical weight of the the nearest-neighbor by 1. We refer to this method as "$d_{\rm LOS}$-peak".

The $d_{\rm LOS}$-peak correction is the first part of our correction method we present in this paper. Nonetheless, we test it by applying it to the mock catalogues and comparing the resulting power spectrums. For each of the mock catalogues (Las Damas, QPM, and Tiling Mock) we use the FKP power spectrum estimator from before to compute the power spectrums of the original mock data, the mock data with the NN correction, and the mock data with the $d_{\rm LOS}$-peak correction. We note that both the NN and the $d_{\rm LOS}$-peak correction methods alter the redshift distribution of our samples by a negligible amount $< 1\%$. Therefore, the synthetic random catalogues and the $\bar{n}(z)$ values do not need to be adjusted in the FKP $P(k)$ calculations.

Since we are interested in how well fiber-collision correction methods can reproduce the power spectrum without fiber collisions (i.e. "true" power spectrum, unsullied by systematic effects), in Figure 4 we plot the ratio of the average power spectrum $\overline{P(k)}$ (averaged over the realizations) computed from corrected mock catalogues over the $\overline{P(k)}$ from the original mock catalogues: $\overline{P_{\rm corr}(k)}/\overline{P_{\rm true}(k)}$. The ratio for the NN method is plotted in blue while the ratio for the $d_{\rm LOS}$-peak correction is plotted in orange. We also plot $\Delta P_{\rm true}(k)/\overline{P_{\rm true}(k)}$ (black; dashed) in order to compare the ratios to sample variance of the mock catalogues. At large scales ($k < 0.1$ h/Mpc), for all mock catalogues, both fiber-collision methods are able to reasonably reconstruct the

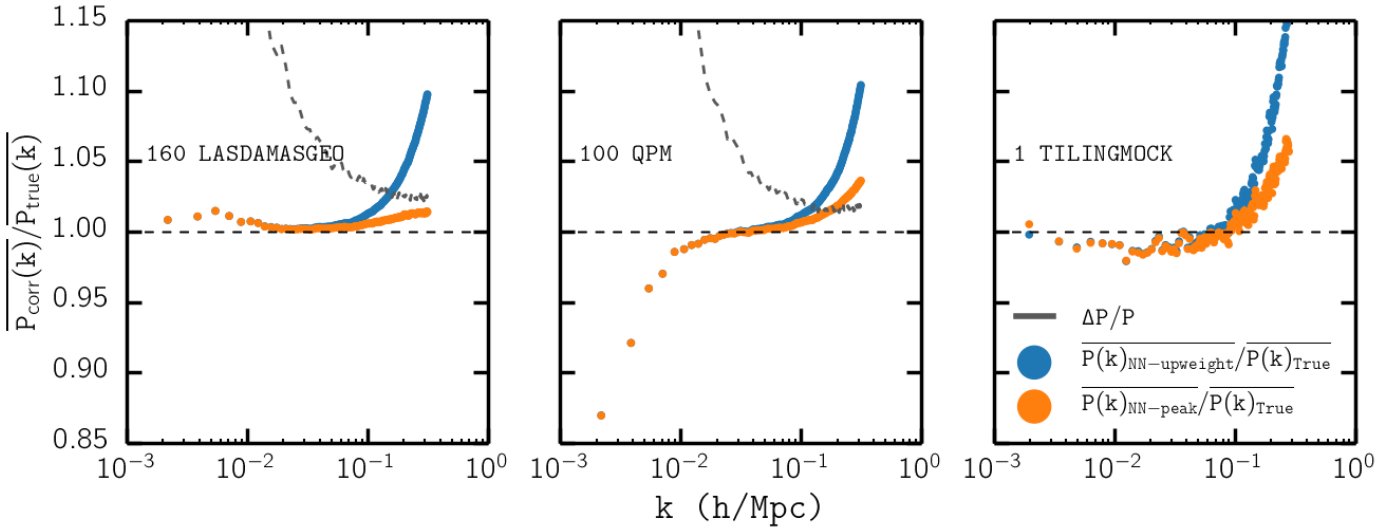

FIG. 4.— The ratio of $\overline{P(k)}$ using the nearest-neighbor (NN) correction (blue) and $d_{\rm LOS}$-peak correction (orange) over the true $\overline{P(k)}$ for LasDamas (left panel), QPM (middle panel), and Tiling Mock (right panel) catalogues. The ratio of the sample variance to the true $\overline{P(k)}$, $\Delta P_{\rm true}(k)/\overline{P_{\rm true}(k)}$ (black; dashed), is plotted for comparison. The NN correction and $d_{\rm LOS}$-peak correction methods are described in Section 3.1. Incorporating the $d_{\rm LOS}$ distribution significantly improves the correction of fiber-collisions; however, both methods dominate the sample variance at $k > 0.3$ h/Mpc.

true power spectrum. At small scales on the other than, the NN corrected $\overline{P(k)}$ is greater than $\overline{P_{\rm true}(k)}$ by $> 10\%$ at $k \sim 0.3$ h/Mpc. In comparison, the $d_{\rm LOS}$-peak corrected $\overline{P(k)}$ is $\sim 5\%$ greater than $P_{\rm true}(k)$ at $k \sim 0.3$ h/Mpc - a significant improvement. **include more description of even smaller scales when the figures are ready**.

Comparing the NN method and the $d_{\rm LOS}$-peak correction to the sample variance ratio $(\Delta P_{\rm true}(k)/\overline{P_{\rm true}(k)})$, reveal that neither of these methods sufficiently correct for fiber-collision at $k > 0.1$ h/Mpc. The $P(k)$s obtained from these methods are clearly dominated by the systematic effects of fiber-collisions on small scales. Although the $d_{\rm LOS}$-peak method we propose provides an improvement from the NN method through our statistical treatment of correlated pairs within the peak of the $d_{\rm LOS}$ distribution, there is still room for improvement since it does not properly account for chance alignments (uncorrelated fiber-collision pairs). Unfortunately, an extension of the $d_{\rm LOS}$-peak method to the entire $d_{\rm LOS}$ distribution range is not feasible is not feasible because such a correction results in an overall reduction of the $P(k)$. This is caused by the statistical sampling of the $d_{\rm LOS}$, which displaces fiber-collided pairs that are actually correlated pairs as chance alignments and the other way around. Therefore, in the following section we provide more robust correction method to account for the chance alignment fiber-collided pairs.

3.2. *Shot Noise Correction*

Both the NN method and the $d_{\rm LOS}$-peak correction do not sufficiently account for the systematic effects of fiber-collisions in the power spectrum. In this section we introduce the other part of our fiber-collision correction method that involve the shot-noise correction term of the FKP power spectrum estimator. By adjusting the shot noise term of the FKP estimator, we present a more robust correction method that accounts for the chance aligned fiber-collided pairs.

In Section 2, we described the FKP power spectrum estimator used to compute the power spectrums of Figure 2. In their estimator, to account for the discreteness of the density field, they subtract the shot-noise contribution to the power spectrum (Equation 1):

$$P_{\rm shot} = \frac{(1+\alpha)\int d^3r\ \bar{n}(\mathbf{r})w^2(\mathbf{r})}{\int d^3r\ \bar{n}^2(\mathbf{r})w^2(\mathbf{r})}. \quad (9)$$

In practice, the integrals in Equation 10 can be written in terms of discrete sums over the random catalogue. The spacial integral $\int d^3r\ \bar{n}(\mathbf{r})...$ can be computed as the discrete sum $\alpha\sum_{\rm ran}....$ With this conversion,

$$P^{\rm FKP}_{\rm shot} = \frac{(1+\alpha)\alpha \sum\limits_{\rm random} w^2_{\rm FKP}(\mathbf{r})}{\alpha \sum\limits_{\rm random} w^2_{\rm FKP}(\mathbf{r})}. \quad (10)$$

In the FKP derivation, they do not take systematic weights of galaxies into account.

The spatial integral $\int d^3r\ \bar{n}(\mathbf{r})...$ can also be computed using the mock catalogues rather than the synthetic random catalogues, as $\sum_{\rm gal} w_g...$ where w_g represents the statistical weight assigned to the galaxies (Cole et al. 2005; Yamamoto et al. 2006; Beutler et al. 2014; Gil-Marín et al. 2014). Then the shot-noise component can be expressed as:

$$P_{\rm shot} = \frac{\sum\limits_{\rm galaxy} w^2_{\rm FKP}w^2_g(\mathbf{r}) - \alpha^2 \sum\limits_{\rm random} w^2_{\rm FKP}(\mathbf{r})}{\alpha \sum\limits_{\rm random} w^2_{\rm FKP}(\mathbf{r})}. \quad (11)$$

In this form, systematic weights for galaxies can be accounted for in the w_g and α terms.

Following the latter formulation, most recently, both Beutler et al. (2014) and Gil-Marín et al. (2014) formu-

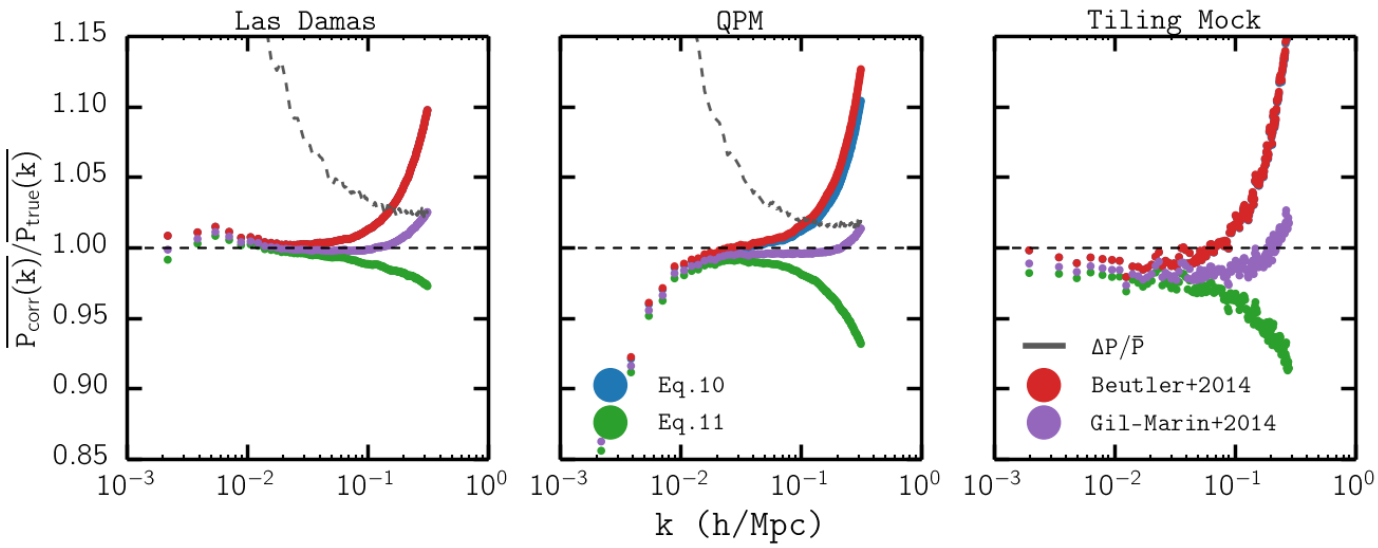

FIG. 5.— The ratio of $\overline{P_{\rm SN}(k)}$ computed using the shot-noise correction term given by Eq. 10 (blue), Eq. 15 (green), $P_{\rm shot}^{\rm Beutler+}$ (red), and $P_{\rm shot}^{\rm Gil-Marin+}$ (purple) over $\overline{P_{\rm true}(k)}$ for LasDamas (left panel), QPM (middle panel), and Tiling Mock (right panel) catalogues with nearest-neighbor fiber-collision weights. Also plotted for comparison, is $\Delta P_{\rm true}/P_{\rm true}(k)$. $\overline{P_{\rm SN}(k)}$ computed using Eq. 10 shot-noise correction is identical to the NN method $\overline{P_{\rm corr}(k)}$ presented in Figure 4 and significantly overestimates $\overline{P_{\rm true}(k)}$ at $k > 0.1$ h/Mpc. $\overline{P_{\rm SN}(k)}$ using $P_{\rm shot}^{\rm Beutler+}$ also overestimates $\overline{P_{\rm true}(k)}$ at $k > 0.1$ h/Mpc. On the other hand $\overline{P_{\rm SN}(k)}$ computed using Eq. 15 significantly underestimates the power spectrum at $k > 0.1$ h/Mpc. At these small scales, only $\overline{P_{\rm SN}(k)}$ computed using $P_{\rm shot}^{\rm Gil-Marin+}$ provides a reasonable correction for fiber-collisions ($\overline{P^{\rm Gil-Marin-SN}(k)}/\overline{P_{\rm true}(k)} - 1 \sim 2.5\%$), reducing its effect below sample variance.

late $P_{\rm shot}$ to account for systematic bias in their analysis of BOSS data. However, systematic weights are treated differently in each of their derivation. Keeping in mind the weights listed in Eq. 4 for the BOSS data, Beutler et al. (2014) derives

$$P_{\rm shot}^{\rm Beutler+} = \frac{\sum_{\rm galaxy} w_{\rm FKP}^2 w_{\rm tot}(\mathbf{r}) w_{\rm sys}(\mathbf{r}) - \alpha'^2 \sum_{\rm random} w_{\rm FKP}^2(\mathbf{r})}{\alpha' \sum_{\rm random} w_{\rm FKP}^2(\mathbf{r})}. \quad (12)$$

$\alpha' = \sum_{\rm gal} w_{\rm tot}/N_{\rm ran}$.

On the other hand, Gil-Marín et al. (2014), in a formulation designed to account for fiber-collisions derives $P_{\rm shot}$ using two separate components: one for correlated pairs and the other chance alignments. The shot-noise contribution to the power from correlated pairs (referred to as "true pairs" in Gil-Marín et al. 2014) is equivalent to Eq. 12 above. For chance alignments (referred to as "false pairs" in Gil-Marín et al. 2014), the shot-noise contribution is derived as,

$$P_{\rm shot}^{\rm False} = \frac{\sum_{\rm galaxy} w_{\rm FKP}^2 w_{\rm tot}^2(\mathbf{r}) - \alpha'^2 \sum_{\rm random} w_{\rm FKP}^2(\mathbf{r})}{\alpha' \sum_{\rm random} w_{\rm FKP}^2(\mathbf{r})}. \quad (13)$$

Then the total $P_{\rm shot}$ is calculated as a combination of $P_{\rm shot}^{\rm True}$ and $P_{\rm shot}^{\rm False}$:

$$P_{\rm shot}^{\rm Gil-Marin+} = (1 - x_{\rm PS}) P_{\rm shot}^{True} + x_{\rm PS} P_{\rm shot}^{False} \quad (14)$$

For the BOSS $P(k)$ Gil-Marín et al. (2014) quotes $x_{\rm PS} = 0.58$.

In our $P(k)$ calculations we use the FKP estimator with the shot-noise correction term computed as Eq. 11 with $w_{\rm g} = w_{\rm tot}$, which gives us,

$$P_{\rm shot}^{\rm Hahn+} = \frac{\sum_{\rm galaxy} w_{\rm FKP}^2 w_{\rm tot}^2(\mathbf{r}) - \alpha'^2 \sum_{\rm random} w_{\rm FKP}^2(\mathbf{r})}{\alpha' \sum_{\rm random} w_{\rm FKP}^2(\mathbf{r})}. \quad (15)$$

The difference between Eq. 15 and Eq. 12 is in the first term of the numerator, where Eq. 12 accounts for fiber-collisions in the $w_{\rm tot}$ term while Eq. 15 accounts for fiber-collisions in the $w_{\rm tot}^2$ term. Eq. 15 is equivalent to the shot-noise contribution from "false pairs" in Gil-Marín et al. (2014).

In order to compare the different $P_{\rm shot}$ derivations, we apply each of them ($P_{\rm shot}^{\rm FKP}$, $P_{\rm shot}^{\rm Hahn+}$, $P_{\rm shot}^{\rm Beutler+}$ and $P_{\rm shot}^{\rm Gil-Marin+}$) to the fiber-collided mock catalogues with nearest-neighbor weights. $P_{\rm shot}^{\rm FKP}$, $P_{\rm shot}^{\rm Hahn+}$, $P_{\rm shot}^{\rm Beutler+}$ and $P_{\rm shot}^{\rm Gil-Marin+}$ are calculated using Eq. 10, 15, 12 and 14, respectively. Since we are particularly interested in fiber-collisions, our comparison ignores the other contributions to systematic effects listed in Eq. 4 ($w_{\rm sys}$ and $w_{\rm rf}$). So in Equations 9 - 15, we use $w_{\rm sys} = 1$ and $w_{\rm tot} = w_{\rm fc}$. We compute the $P_{\rm SN}(k)$s using the various $P(k)$ estimators with different $P_{\rm shot}$ equations for each of the mock catalogues. Afterwards we compute the $\overline{P_{\rm SN}(k)}$s for the mock catalogues by averaging over the realizations, and then take the ratio over $\overline{P_{\rm true}(k)}$ calculated using the standard FKP $P(k)$ estimator.

In Figure 5, we present $\overline{P_{\rm SN}(k)}/\overline{P_{\rm true}(k)}$ computed using $P_{\rm shot}^{\rm FKP}$, $P_{\rm shot}^{\rm Hahn+}$, $P_{\rm shot}^{\rm Beutler+}$ and $P_{\rm shot}^{\rm Gil-Marin+}$ shot noise correction term equations. We note that $\overline{P_{\rm SN}(k)}$ calculated using $P_{\rm shot}^{\rm FKP}$ is the same quantity as the NN corrected $\overline{P(k)}$ in Figure 4. As before, it over estimates

the $\overline{P_{\rm true}(k)}$ by $> 10\%$ for all mocks at $k \sim 0.3$ h/Mpc. $\overline{P(k)}$ measured using $P_{\rm shot}^{\rm Beutler+}$ (Eq. 12) does not significantly improve this overestimation in any of the mock catalogues.

On the other hand, the $\overline{P(k)}$ using Eq. 15 significantly underestimates the $\overline{P_{\rm true}(k)}$. At $k \sim 0.3$ h/Mpc, $\overline{P^{\rm Hahn-SN}(k)}/\overline{P_{\rm true}(k)} - 1 =\sim -3\%, \sim -8\%$, and ~ -9 for LasDamas, QPM, and Tiling Mock respectively. Finally, $\overline{P(k)}$ measured using $P_{\rm shot}^{\rm Gil-Marin}$ reproduces $\overline{P_{\rm true}(k)}$ most accurately in the comparison of Figure 5. At $k \sim 0.3$ h/Mpc, $\overline{P^{\rm Gil-Marin-SN}(k)}/\overline{P_{\rm true}(k)} - 1 \sim 2.5\%$. As we plotted in Figure 4, we plot the sample variance $(\Delta P/\bar{P})$ for each of the mock catalogues in Figure 5. Only $\overline{P(k)}$ computed using $P_{\rm shot}^{\rm Gil-Marin}$ is able to reduce the effects of fiber-collisions below the sample variance.

The shot-noise correction term we present in Eq. 15 should only correct for the shot-noise contribution to the power spectrum from chance alignment fiber-collided pairs - not from correlated pairs. Because we apply the shot-noise correction to all fiber-collided pairs in Figure 5, the shot-noise correction term over-subtracting to underestimate the true power spectrum is expected. In the following section we present our fiber-collision correction method, that combines the $d_{\rm LOS}$-peak correction of Section 3.1 and Eq. 15 from this section.

4. RESULTS

4.1. *Fiber-Collision Correction Method*

In Section 3.1 we derive a method to better statistically reconstruct the clustering of correlated fiber-collided pairs using using the peak of the $d_{\rm LOS}$ distribution. Then in Section 3.2 we presented our formulation of the shot-noise correction term, which accounts for the added discreteness from fiber-collsions. We now combine these two methods in order to present our fiber-collision correction method.

We begin with once again with the fiber-collided mock catalogues with the nearest-neighbor weights. These mock catalogues accurately imitate the effects of fiber-collision on the actual observed BOSS data. Then as described in Section 3.1, we calculate the $d_{\rm LOS}$ distribution of all fiber-collided pairs in each mock catalogue. From the $d_{\rm LOS}$ distribution, we compute the fraction of fiber-collided pairs within the peak of the distribution ($f_{\rm peak}$; Section 3.1). We then select $f_{\rm peak}$ of the fiber-collided pairs and designate them as correlated pairs that lie in the peak of the $d_{\rm LOS}$ distribution. We refer to these fiber-collided pairs as "peak-assigned" pairs.

Each of these "peak-assigned" pairs, due to the nearest-neighbor weights, consist of a galaxy with $w_{\rm fc} > 1$ (the "nearest-neighbor" galaxy) and another galaxy with $w_{\rm fc} = 0$ (the "collided" galaxy). We disregard the collided galaxy. For each of the nearest-neighbor galaxies, we sample a value $d_{\rm peak-assigned}$ from the peak of the $d_{\rm LOS}$ distribution and place a galaxy with $w_{\rm fc} = 1$ at a comoving distance $d_{\rm peak-assigned}$ away from the nearest-neighbor galaxy. At the same time, we reduce $w_{\rm fc}$ of the nearest-neighbor galaxy (the one that was originally up-weighted) by 1. This process is repeated until the nearest-neighbor galaxy has $w_{\rm fc} = 1$.

For the remaining fiber-collided pairs that are not "peak-assigned" ($1 - f_{\rm peak}$ of the fiber-collided pairs), we keep the nearest-neighbor fiber collision weights. The resulting mock catalogue has $f_{\rm peak}$ fewer galaxies with $w_{\rm cp} > 1$ compared to the fiber-collided mock catalogues with nearest-neighbor weights that we started from; however, the total statistical weight of the mock catalogue is conserved.

Next, we use the FKP estimator with the Eq. 15 shot-noise correction term described in Section 3.2 to calculate the $P(k)$ for our newly corrected mock catalogues. In Figure 6 we present the ratio of the $\overline{P(k)}$ computed using our fiber-collision correction method over $P_{\rm true}(k)$. To assess the merit of the fiber-collision correction scheme, we once again over-plotted the sample variance of the mock catalogues. For Las Damas, at $k \sim 0.3$ h/Mpc, $P(k)/P_{\rm true}(k) - 1 \sim 0.5\%$ and throughout the k range, $0.3\% < P(k)/P_{\rm true}(k) - 1 < 1.3\%$. For QPM, at $k \sim 0.3$ h/Mpc, $P(k)/P_{\rm true}(k) - 1 \sim 0.04\%$ and throughout the k range, $0.3\% < P(k)/P_{\rm true}(k) - 1 < 1.3\%$. Moreover for the mock catalogues where we do have sample variance measurements (LasDamas and QPM), the $P(k)$ ratio is significantly below the sample variance throughout the entire k range probed.

4.2. *Goodness-of-fit*

The correction method we present in this paper corrects for the effect of fiber-collisions to the extent where $P_{\rm corr}(k)$ measurements are no longer dominated by systematic effects at the scales probed in our calculations ($k < 0.3$ h/Mpc). However, to further quantify how accurately our correction method is able to reproduce the true power spectrum, we quantify the goodness-of-fit for all the corrected $P(k)$s to $P_{\rm true}(k)$ by calculating the following χ^2:

$$\chi^2(k_{\rm max}) = \frac{1}{N_{k_{\rm max}}} \sum_{k<k_{\rm max}} \frac{(\overline{P_{\rm corr}(k)} - \overline{P_{\rm true}(k))^2}}{\Delta P_{\rm true}(k)^2}. \quad (16)$$

$N_{k_{\rm max}}$ is the number of k bins in the $P(k)$ calculation with $k < k_{\rm max}$. $\overline{P_{\rm corr}(k)}$ is the average fiber-collision corrected powerspectrum of all realizations for each catalogue. $\Delta P_{\rm true}(k)$ is the standard deviation of $P_{\rm true}(k)$ computed from all realizations of each mock catalogue. We calculate χ^2 as a function of $k_{\rm max}$ in order to determine the accumulated χ^2 to a specific scale ($k_{\rm max}$). In doing so, we determine the scale to which a power spectrum analysis can be extended to without fiber-collisions dominating the measurement.

For the Tiling Mock catalogue, since there is only one realization, the variance ($\Delta P_{\rm true}(k)$) cannot be computed independently. However, for an estimate of the χ^2, we estimate $\Delta P_{\rm true}^{\rm Tiling\ Mock}(k)$ from $\Delta P_{\rm true}(k)$ of the QPM mock catalogue scaled by the ratio of the power spectrums:

$$\Delta P_{\rm true}^{\rm Tiling}(k) = \Delta P_{\rm true}^{\rm QPM}(k) \times \frac{P_{\rm true}^{\rm Tiling}(k)}{P_{\rm true}^{\rm QPM}(k)}. \quad (17)$$

In Figure 7, we present the $\chi^2(k_{\rm max})$ for $P(k)$ using the following fiber collision correction methods: the NN correction, Beutler et al. (2014), Gil-Marín et al. (2014), and our method from Section 4.1. The comparison is made using the LasDamas (left panel), QPM (center panel) and Tiling (right panel) mock catalogues.

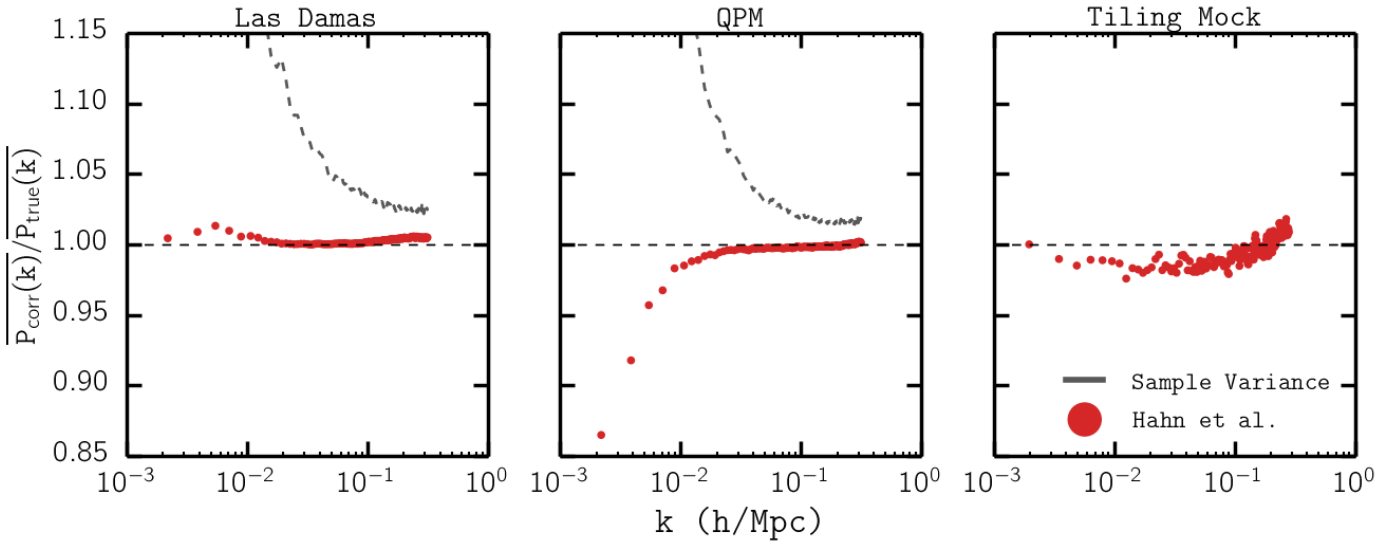

FIG. 6.— The ratio of $\overline{P(k)}$ computed using our fiber-collision correction method described in Section 4.1 over $\overline{P_{\rm true}(k)}$ for LasDamas (left panel), QPM (middle panel), and Tiling Mock (right panel) catalogues with nearest-neighbor fiber-collision weights (red). Also plotted for comparison, is $\Delta P_{\rm true}/P_{\rm true}(k)$. At $k \sim 0.3\ h$/Mpc, the $\overline{P(k)}$ ratio of our correction method $\sim 0.5\%$, which is significantly better than the correction methods of Anderson et al. (2012), Beutler et al. (2014) and Gil-Marín et al. (2014). Moreover, the $\overline{P(k)}$ ratio for our correction method is significantly lower than the $\Delta P_{\rm true}/P_{\rm true}(k)$ throughout the probed k values.

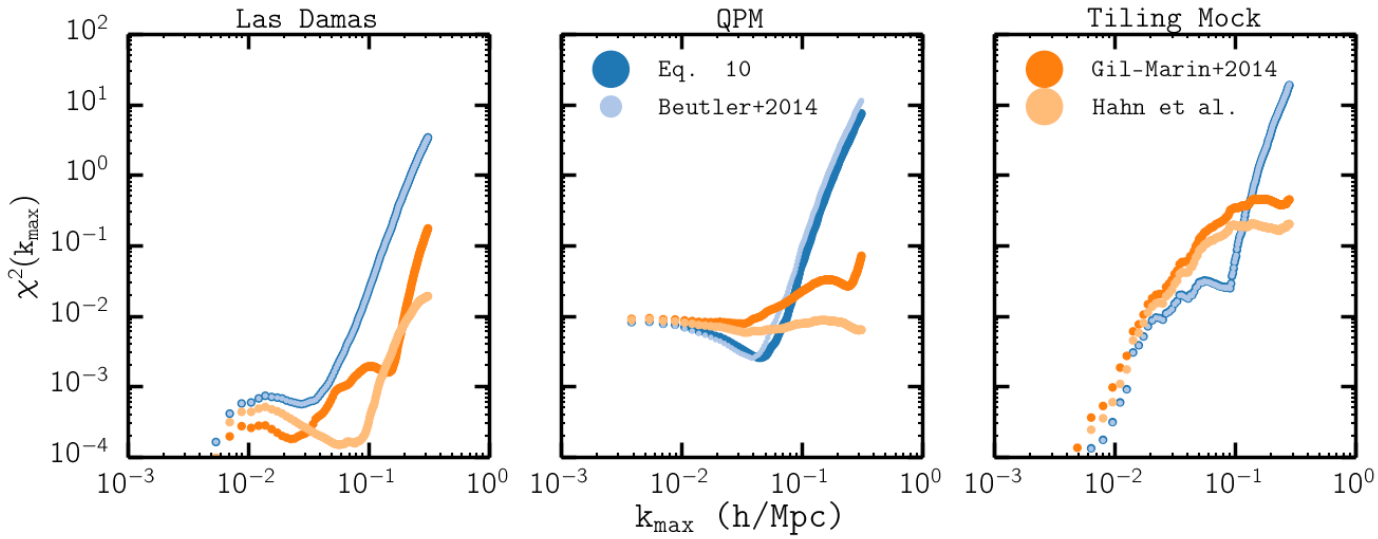

FIG. 7.— $\chi^2(k_{\rm max})$ for $P_{\rm NN-upweight}$ (blue), $P_{\rm Beutler}$ (light blue), $P_{\rm Gil-Marin}$ (orange), and $P_{\rm Hahn}$ (yellow) using the Las Damas (left panel), QPM (centerl panel), and Tiling (right panel) mock catalogues. Equation 16 was used to compute the $\chi^2(k_{\rm max})$.

First in the LasDamas panel, at scales larger than $k < 5 \times 10^{-2}\ h$/Mpc, $\chi^2 < 10^{-3}$ for all correction methods, which suggests that all correction methods sufficiently correct for fiber-collisions and reasonably reconstructs $P_{\rm true}(k)$ at these scale. In approaching smaller scales the χ^2 for all correction methods increase. However, throughout the entire $k_{\rm max}$ range, our correction method provides a notably lower χ^2 in comparison to the other correction methods. This is especially clear at $k \sim 0.3\ h$/Mpc, where $\chi^2_{\rm NN} \approx 0.33$, $\chi^2_{\rm Beutler} \approx 0.33$, and $\chi^2_{\rm Gil-Marin} \approx 0.17$ while $\chi^2 \approx 0.019$ for our method.

Next in the QPM panel (center), at scales larger than ($k < 2 \times 10^{-2}\ h$/Mpc) all corrected $P(k)$ once again provide a reasonable reconstruction of $P_{\rm true}(k)$ with $\chi^2 \approx 10^{-2}$. Then for $2 \times 10^{-2} < k_{\rm max} < 5 \times 10^{-2}\ h$/Mpc, the NN and Beutler et al. (2014) correction methods have the lower χ^2s than Gil-Marín et al. (2014) and our correction methods. However, as k increases, the χ^2s for NN and Beutler et al. (2014) quickly surpasses the χ^2s of Gil-Marín et al. (2014) and our method. Meanwhile, throughout the entire $k_{\rm max}$ range probed, our correction method has a significantly lower χ^2 than the correction method of Gil-Marín et al. (2014). The difference in χ^2 is again most remarkable at the smallest scales: $\chi^2(k_{\rm max} \sim 0.3) \approx 7.3,\ 11,\ 0.068$, and 0.0063 for NN, Beutler et al. (2014), Gil-Marín et al. (2014) and our

correction method respectively.

Finally in the right panel (Tiling Mock), we find that $\chi^2(k_{\rm max})$ increases in a near monotonic fashion as a function of $k_{\rm max}$. For $k_{\rm max} < 4 \times 10^{-2}$ h/Mpc, there are no significant differences between the χ^2s of any of the correction methods. Then at intermediate scales of $4 \times 10^{-2} < k_{\rm max} < 1.5 \times 10^{-2}$ h/Mpc, χ^2s of NN and Beutler et al. (2014) are notably lower than the χ^2s of Gil-Marín et al. (2014) and our method. However like the other mock catalogues at smaller scales, the χ^2s of NN and Beutler et al. (2014) quickly increase beyond the other χ^2s, ultimately both reach $\chi^2(k_{\rm max} \sim 0.3) \approx 13.5$. In the meantime, $\chi^2(k_{\rm max} \sim 0.3) \approx 0.32$ for Gil-Marín et al. (2014) and $\chi^2(k_{\rm max} \sim 0.3) \approx 0.15$ for our correction method. Throughout the entire $k_{\rm max}$ our method has a lower χ^2 than Gil-Marín et al. (2014). With only one realization of the Tiling Mocks and scaled ΔP, a detailed comparison of the χ^2 values is impossible. Yet the $\chi^2(k_{\rm max})$ results for our correction provide realistic assurance that the correction method can be applied to the actually BOSS data.

Overall, regardless of catalogue, for scales $k_{\rm max} < 3 \times 10^{-2}$ h/Mpc, the χ^2 for all fiber-collision corrected $P(k)$ are negligible. Then, with the exception of the Las Damas catalogue, at intermediate scales of $3 \times 10^{-2} < k_{\rm max} < 0.15$ h/Mpc the NN and Beutler et al. (2014) methods show the lowest χ^2s. However at these intermediate scales, all of the correction methods reasonably reconstruction of the true power spectrum with $\chi^2 < 1$. At the smallest scales, NN and Beutler et al. (2014) corrections methods no longer provide a sufficient fit to $\overline{P_{\rm true}(k)}$. A detailed comparison at the small scales ($k > 1.5 \times 10^{-1}$ h/Mpc) demonstrates that the fiber-collision correction we present in this paper significantly better reproduces the true power spectrum than any of the other fiber-collision correction methods. Furthermore, Figure 7 illustrates that throughout the entire $k_{\rm max}$, our correction method is able to reasonably reproduce $P_{\rm true}(k)$ at all scales.

5. SUMMARY AND DISCUSSION

Using simulated mock catalogues for SDSS data with realistically imposed fiber-collisions, we quantify the systematic effects of fiber-collisions on galaxy clustering measurements - in particular the power spectrum. Although fiber-collisions have little significant effect on the power spectrum at large scales, its effect quickly overtakes the sample variance at scales smaller than $k \sim 0.1$ h/Mpc. At the smallest scales that we explore in this paper, $k \sim 0.3$, fiber-collisions have over a 10% effect on the power spectrum. Consequently, at these small scales measurements of the power spectrum is dominated by the uncertainties from systematic effects of fiber-collisions.

Fortunately, through the fiber-collided mock catalogues we are able to model the distribution of the line-of-sight displacement between fiber-collided pairs. Using this model, we are able to statistically reconstruct the clustering fiber-collided galaxies that reside in the same halo. Combining this correction with a modification of the shot-noise term in the power spectrum estimator to properly account for the discreteness added by the nearest-neighbor weights of chance aligned fiber-collided pairs, we devise a new fiber-collision correction method that is able to reconstruct the true power spectrum from fiber-collided data even to $k \sim 0.3$ h/Mpc. Throughout the entire k range explored in our analysis, our correction method reconstructs $P_{\rm true}(k)$ safely within the sample variance.

Furthermore, we compare our method to the most common nearest-neighbor correction method and recent correction methods presented in Beutler et al. (2014) and Gil-Marín et al. (2014) to demonstrate that our method most successfully reproduces the true power spectrum at small scales and reliably at all scales for all the Las Damas, QPM, and Tiling mock catalogues. As an added advantage, our correction method can also be validated by actual data rather than only through the mock catalogues thereby making it's application to actual data further reliable.

The fiber-collision correction method we present will enable us to extend our galaxy clustering analysis to smaller scales for SDSS-III BOSS and future surveys such as eBOSS or any other large fiber-fed surveys the suffer from systematic effects of fiber-collisions. Our fiber-collision correction method can also be extend to higher order clustering statistics such as the quadrupole of the power spectrum and the bispectrum, which we will explore in future work.

We thank something something

REFERENCES

Anderson, L., Aubourg, E., Bailey, S., et al. 2012, MNRAS, 427, 3435

Berlind, A. A., Frieman, J., Weinberg, D. H., et al. 2006, ApJS, 167, 1

Beutler, F., Saito, S., Seo, H.-J., et al. 2014, MNRAS, 443, 1065

Blanton, M. R., Lin, H., Lupton, R. H., et al. 2003, AJ, 125, 2276

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505

Colless, M. 1999, Royal Society of London Philosophical Transactions Series A, 357, 105

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10

Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ, 426, 23

Gil-Marín, H., Noreña, J., Verde, L., et al. 2014, ArXiv e-prints, arXiv:1407.5668

Guo, H., Zehavi, I., & Zheng, Z. 2012, ApJ, 756, 127

Hogg, D. W. 1999, ArXiv Astrophysics e-prints, astro-ph/9905116

Makarem, L., Kneib, J.-P., Gillet, D., et al. 2014, A&A, 566, A84

Manera, M., Scoccimarro, R., Percival, W. J., et al. 2013, MNRAS, 428, 1036

Markwardt, C. B. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 411, Astronomical Data Analysis Software and Systems XVIII, ed. D. A. Bohlender, D. Durand, & P. Dowler, 251

McBride, C., Berlind, A., Scoccimarro, R., et al. 2009, in Bulletin of the American Astronomical Society, Vol. 41, American Astronomical Society Meeting Abstracts #213, 425.06

McBride, C., Berlind, A. A., Scoccimarro, R., et al. 2011, in Bulletin of the American Astronomical Society, Vol. 43, American Astronomical Society Meeting Abstracts #217, 249.07

Morales, I., Montero-Dorta, A. D., Azzaro, M., et al. 2012, MNRAS, 419, 1187

Schlegel, D., Abdalla, F., Abraham, T., et al. 2011, ArXiv e-prints, arXiv:1106.1706

Scoccimarro, R., & Sheth, R. K. 2002, MNRAS, 329, 629
Tinker, J. L., Sheldon, E. S., Wechsler, R. H., et al. 2012, ApJ, 745, 16
White, M., Tinker, J. L., & McBride, C. K. 2014, MNRAS, 437, 2594
Yamamoto, K., Nakamichi, M., Kamino, A., Bassett, B. A., & Nishioka, H. 2006, PASJ, 58, 93
Yoon, J. H., Schawinski, K., Sheen, Y.-K., Ree, C. H., & Yi, S. K. 2008, ApJS, 176, 414
Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, ApJ, 571, 172
Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, ApJ, 630, 1