The Astrophysical Journal, 580:42–53, 2002 November 20

LARGE-SCALE CORRELATION OF MASS AND GALAXIES WITH THE Lyα FOREST TRANSMITTED FLUX

Patrick McDonald,[1,2] Jordi Miralda-Escudé,[2] and Renyue Cen[3]

Received 2001 December 20; accepted 2002 July 19

ABSTRACT

We present predictions of the correlation between the Lyα forest absorption in quasar spectra and the mass within $\sim 5\,h^{-1}$ Mpc (comoving) of the line of sight, using fully hydrodynamic and hydro-PM numerical simulations of the cold dark matter model supported by present observations. The observed correlation based on galaxies and the Lyα forest can be directly compared to our theoretical results, assuming that galaxies are linearly biased on large scales. Specifically, we predict the average value of the mass fluctuation, $\langle \delta_m \rangle$, conditioned to a fixed value of the Lyα forest transmitted flux, δ_F, after they have been smoothed over a 10 h^{-1} Mpc cube and line-of-sight interval, respectively. We find that the relation between δ_F/σ_F and $\langle \delta_m \rangle/\sigma_m$ (where σ_m and σ_F are the rms dispersions) has a predicted slope of ~ 0.6 at $\delta_F = 0$ that is largely insensitive to the cosmological model and other Lyα forest parameters. Comparison of our predictions to observations should provide a fundamental test of our ideas on the nature of the Lyα forest and the distribution of galaxies and can yield a measurement of the bias factor of any type of galaxy that is observed in the vicinity of Lyα forest lines of sight.

Subject headings: cosmology: theory — intergalactic medium — large-scale structure of universe — quasars: absorption lines

On-line material: color figures

1. INTRODUCTION

The prevalent theory to explain the Lyα forest is that the absorption lines arise from density structures in a photoionized intergalactic medium that originate in the gravitational evolution of primordial fluctuations. Both semianalytic models and numerical simulations (Bi 1993; Cen et al. 1994; Zhang, Anninos, & Norman 1995; Zhang et al. 1998; Hernquist et al. 1996; Miralda-Escudé et al. 1996; Bi & Davidsen 1997; Schaye 2001) have shown that the predicted Lyα spectra appear remarkably similar to the observations. The good agreement of the predicted and observed flux distribution and power spectrum of the Lyα forest (Rauch et al. 1997; Croft et al. 1999, 2002; McDonald et al. 2000) and the large transverse size of the absorption systems (Bechtold et al. 1994; Dinshaw et al. 1994, 1997; Petitjean et al. 1998; Monier, Turnshek, & Hazard 1999; Dolan et al. 2000; López, Hagen, & Reimers 2000) are the basic tests that have so far been done and have supported the theory. This Lyα forest theory is also especially attractive because it is a consequence of the general cold dark matter (CDM) model with parameters that are well constrained from several other observations (e.g., Primack 2000). However, the theory also needs the additional hypothesis that the star formation and active nuclei within galaxies, and the possible winds that they may create, have a negligible effect on the intergalactic medium at low densities and the observed Lyα spectra, except for the photoionization due to the cosmic ionizing background.

Another important test of our ideas on the Lyα forest can be performed by observing the correlation of the Lyα transmitted flux with the galaxies in the vicinity of the line of sight. This type of observation can be particularly useful to test for any possible effects from galactic winds or other perturbing influences on the Lyα forest associated with galaxies. Some observations along these lines have been done for small scales and high column density absorbers, where the environment in the close vicinity of galaxies can be probed. These have shown a strong correlation between galaxies and gas halos, as expected if galaxies are located in halos that accrete intergalactic gas (Bergeron & Boissé 1991; Steidel, Dickinson, & Persson 1994; Lanzetta et al. 1995; Chen et al. 2001). Grogin & Geller (1998) and Penton, Stocke, & Shull (2002) have probed this correlation at low redshift and low column densities and have found that the weak absorption lines are often found in low-density regions of the galaxy distribution. Recently, Adelberger et al. (2002) have carried out the first analysis of this correlation on large scales (several comoving megaparsecs) and at high redshift (with galaxies detected with the Lyman break technique), using the transmitted flux as the quantity to correlate with the mean number of galaxies in a specified region.

Inspired by the observational results of Adelberger et al. (2002), this paper presents detailed theoretical predictions of the correlation of the Lyα forest with mass, assuming the standard Lyα forest theory where galactic winds do not alter the properties of the Lyα spectra. Specifically, predictions are made for the mean value of the mass given an observed value of the transmitted flux and the mean value of the transmitted flux for a fixed value of the mass, after both quantities have been smoothed over a certain region. The main difference between the functions that we analyze and the function shown by Adelberger et al. (2002) is that the observed objects are of course galaxies, and our

[1] Department of Physics, Princeton University, Princeton, NJ 08544; pmcdonal@feynman.princeton.edu.

[2] Department of Astronomy, Ohio State University, Columbus, OH 43210; jordi@astronomy.ohio-state.edu.

[3] Princeton University Observatory, Princeton University, Princeton, NJ 08544; cen@astro.princeton.edu.

simulations predict only the distribution of the mass. However, the predictions of the simulations can still be compared to the galaxy observations to determine the relation between galaxies and mass, which is especially straightforward if linear bias is a sufficient description of the distribution of galaxies relative to the mass on the large scales being probed. We examine the dependence of these statistical functions on the various parameters affecting the Lyα forest in § 3.

2. METHOD

We use several hydrodynamic particle-mesh (HPM; see Gnedin & Hui 1998) simulations to model the Lyα forest, after testing that the HPM approximation is sufficient for our purposes by comparing to a fully hydrodynamic simulation. The standard HPM simulations that will be used most often in this paper have box size 40 h^{-1} Mpc, with 512^3 particles. The cosmological model is CDM with a cosmological constant, in a flat universe with present matter density $\Omega_m = 0.4$, power spectrum index $n = 0.95$, and amplitude given by $\sigma_8 = 0.75$. All the results in this paper will be shown at $z = 3$, the typical redshift at which the observations have so far been done. The model that we use is consistent with the observed Lyα forest power spectrum: at the typical Lyα forest scale of $k = 0.008$ (km s^{-1})$^{-1}$, it has $\Delta^2_\rho(k) = 0.29$ at $z = 3$ [where $\Delta^2_\rho(k)$ is the contribution per unit $\ln k$ to the variance of the linear theory mass-density fluctuations], while the observed value is $\sim$0.26 (McDonald et al. 2000; Croft et al. 2002). The redshift interval corresponding to 40 h^{-1} Mpc in this model near $z = 3$ is $\Delta z = 0.068$. The density-temperature relation assumed for the gas is $T(\Delta) = T_{1.4}(\Delta/1.4)^{\gamma-1}$, where Δ is the gas density divided by the mean gas density. We use the parameters $T_{1.4} = 17,000$ K and $\gamma - 1 = 0.3$ in our standard simulation (the reason that we specify the value of the temperature $T_{1.4}$ at $\Delta = 1.4$ is because the error in the temperature was determined to be smallest at this density in McDonald et al. 2001). Several other similar simulations are used where we vary each one of the relevant parameters (see § 3), and we vary also the box size and resolution to verify the numerical convergence of the results. Unless otherwise indicated, all the results that we show for a given set of parameters are averages of four simulations with different random initial perturbations.

2.1. *Definitions*

We define the quantity

$$\delta_F \equiv 1 - \frac{F}{\bar{F}} , \tag{1}$$

where F is the fraction of transmitted flux in the Lyα forest after the spectrum has been smoothed along the line of sight with a filter, and $\bar{F}$ is the mean transmitted flux. We use a Gaussian smoothing filter in this paper, which is $W(kR) = \exp[-(kR)^2/2]$ in Fourier space. We also define

$$\delta_m \equiv \frac{\rho_m}{\bar{\rho}_m} - 1 \tag{2}$$

as the redshift-space mass-density perturbation, again smoothed over a specified region. We choose to smooth the mass fluctuation over cubes of 10 h^{-1} Mpc (comoving) in redshift space, centered on the Lyα forest line of sight, which is approximately the smoothing that is used in Adelberger et al. (2002). This is convenient to do in observations of galaxies, since galaxies are usually searched for in a square field of view and they can simply be divided into redshift bins. Note that the size of our cubes in velocity units, 1280 km s^{-1}, is more fundamental than the size in h^{-1} Mpc, because the temperature and observed flux power spectrum are both measured in km s^{-1}.

From each simulation we extract $(N/4)^2$ spectra per face of the cube, using a uniform grid of starting points for lines of sight, where N is the number of cells along an edge. Every N pixel spectrum is then smoothed using the Gaussian filter, and the entire N^3 cell redshift-space mass-density field is smoothed using the cubic filter. We compute our statistics using the resulting $3N(N/4)^2$ δ_m-δ_F pairs.

In general, the information that can be recovered from observations is the full joint probability distribution $P(\delta_F, \delta_g)$, where δ_g is the fluctuation in the number of galaxies. We show an example of the full distribution in Figure 1, for our standard model, box size, and resolution. A detailed comparison of the shape of this distribution with observations requires first testing the dependence of the theoretical prediction on the model, the box size, and the resolution, and the use of the HPM approximation. In the rest of this paper we focus on testing this dependence only for the more restricted function $\langle \delta_m | \delta_F \rangle$, which is the mean value of the mass fluctuation subject to the condition of a fixed value in the transmitted flux fluctuation, and for the analogous quantity $\langle \delta_F | \delta_m \rangle$. The observed galaxies may not follow exactly the same fluctuations as the mass, even when smoothed over a scale of 10 h^{-1} Mpc. However, if galaxies are linearly biased at this large scale, then $\delta_g = b\delta_m$, where δ_g is the fluctuation in the galaxies and b is the bias factor. If we now define the quantities $\tilde{\delta}_m = \delta_m/\sigma_m$ and $\tilde{\delta}_g = \delta_g/\sigma_g$, where σ_m and σ_g are the rms fluctuations of δ_m and δ_g, then $\tilde{\delta}_m = \tilde{\delta}_g$, and the correlation with δ_F is precisely predicted. For convenience, we also define the quantity $\tilde{\delta}_F = \delta_F/\sigma_F$, and we will present results for $\langle \tilde{\delta}_m | \tilde{\delta}_F \rangle$ and $\langle \tilde{\delta}_F | \tilde{\delta}_m \rangle$.

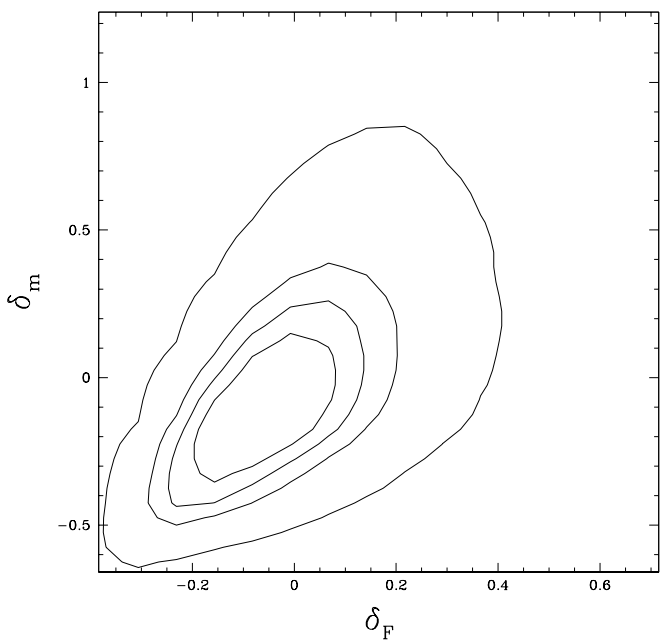

FIG. 1.—Contour plot of $P(\delta_F, \delta_g)$. The equiprobability lines shown enclose 33%, 50%, 67%, and 95% of the total.

2.2. *Smoothing*

To choose a value of the smoothing radius R for the smoothing filter of the Lyα forest transmitted flux, we compute first the correlation coefficient $\langle\delta_F\delta_m\rangle/\sigma_F\sigma_m$ in our standard simulation and show it as a function of R in Figure 2. The correlation is maximum at $R = 3.5\ h^{-1}$ Mpc. The dependence of the correlation on R is not surprising: if R is too small, the value of δ_F is altered by small-scale fluctuations that are not related to the value of δ_m and therefore act as noise, and if R is too big, then the fluctuations affecting δ_m are erased by smoothing in the value of δ_F.

A Gaussian filter with $R = 3.5\ h^{-1}$ Mpc will be used for the smoothing of the Lyα spectrum throughout the rest of this paper. However, before proceeding, we examine how much our results differ if we use instead a top-hat smoothing for the Lyα spectrum or if we vary the size of the cube over which the mass is smoothed.

Throughout the paper, our results will be presented as a set of four figures showing the four functions $\langle\delta_m|\delta_F\rangle$, $\langle\delta_F|\delta_m\rangle$, $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle$, and $\langle\tilde{\delta}_F|\tilde{\delta}_m\rangle$. In all these figures, the function $\langle\delta_m|\delta_F\rangle$ has been obtained by creating spectra along each row of cells in the simulation, smoothing the spectra along the line of sight, selecting all the pixels in the spectra where the value of δ_F is inside a given bin, and computing the average δ_m for these pixels, where δ_m is computed in redshift space. We use bins of width $\Delta\delta_F = 0.05$. Similarly, $\langle\delta_F|\delta_m\rangle$ is calculated by averaging the values of δ_F for pixels in bins with a given δ_m (again using $\Delta\delta_m = 0.05$ for the bin width).

As this work was being completed, we learned that Adelberger et al. (2002) use a top-hat filter. In Figures 3*a*–3*d*, we compare the results for our standard Gaussian filter (*thick, solid lines*) to the results obtained if we simply average the Lyα forest flux across the 10 h^{-1} Mpc extent of the cube with which the mass is smoothed, i.e., use a top-hat filter with $R = 5\ h^{-1}$ Mpc (*dotted lines*). The difference is not large, although noticeable, for the δ_m-δ_F relations and is practically zero for the $\tilde{\delta}_m$-$\tilde{\delta}_F$ relations.

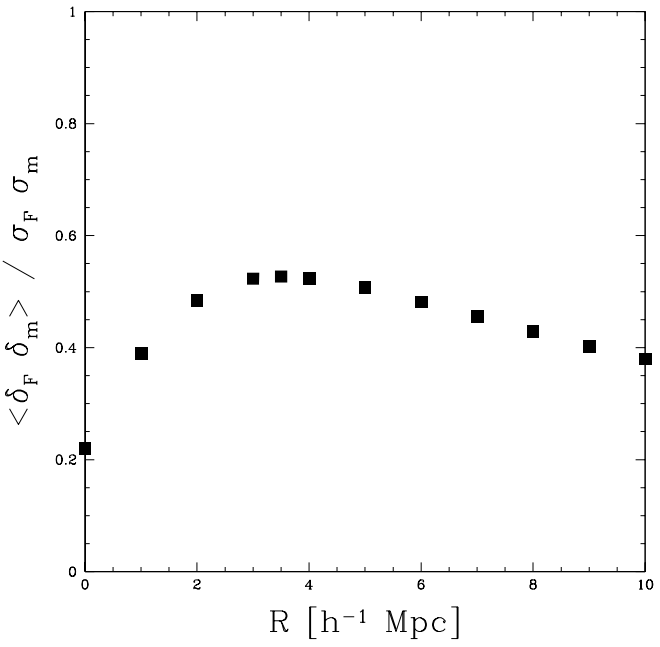

FIG. 2.—Correlation between mass and transmitted flux as a function of the smoothing applied on the spectrum, with the Gaussian filter $W(kR) = \exp[-(kR)/2]$. The squares show the correlation $\langle\delta_F\delta_m\rangle/\sigma_F\sigma_m$. Note that this figure is intended only to motivate our choice of R. The points should not be compared to observations or other simulations because we have not tested their sensitivity to the numerical effects and model parameters discussed later in the paper.

We also show in Figure 3 the effect of changing the size of the cube used to smooth the mass-density field from 10 to 12 h^{-1} Mpc (*thick, long-dashed lines*) or 8 h^{-1} Mpc (*thick, short-dashed lines*). In both of these cases we also change the smoothing length for the spectrum by the same factor. As the size of the smoothing cube is reduced, the mass fluctuations increase, and they are more strongly correlated with the Lyα forest, as expected. We see that for $\langle\delta_m|\delta_F\rangle$ it is important to match the dimensions of the cube in the simulations and observations to compare the two. Note, however, that the transverse angular size of the cube depends on the cosmological geometry. The other plots show less sensitivity to the cube size.

To promote comparison between our results and observations or other simulations, in Table 1 we give $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle$ and $\langle\tilde{\delta}_F|\tilde{\delta}_m\rangle$ for the three cube sizes, along with the σ_F and σ_m needed to convert from $\tilde{\delta}$ to δ. The rows in the table labeled 3s, 3ld, and 3sd correspond to cube sizes 10, 12, and 8 h^{-1} Mpc, respectively. For this table we have compressed the information in our curves by listing only the best-fit parameters, c_i, for the formulae $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle = c_0 + c_1\tilde{\delta}_F + c_2\tilde{\delta}_F^2$ and $\langle\tilde{\delta}_F|\tilde{\delta}_m\rangle = c_0 + c_1\tilde{\delta}_m + c_2\tilde{\delta}_m^2$. Examples of these fits are shown in Figures 3*c*–3*d*, where the thin lines show the fitted curves. The parameters were determined using the intervals $-2 < \tilde{\delta}_F < 2$ and $-1.2 < \tilde{\delta}_m < 2$ for $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle$ and $\langle\tilde{\delta}_F|\tilde{\delta}_m\rangle$, respectively. We do not attempt to improve the fit at the extremes in density (by using a more complicated formula) because, as we show later in this paper, the simulation predictions are uncertain at these densities (mostly because of finite box-size effects).

2.3. *Test of the HPM Approximation*

We resort to approximate HPM simulations (Gnedin & Hui 1998) because it is impractical at the present time to perform fully hydrodynamic simulations of the required size for all of the parameter variations that we would like to explore. However, we first test the accuracy of the HPM approximation by comparing to a state-of-the-art hydrodynamic simulation of a similar cosmological model. This simulation is Eulerian, with box size 25 h^{-1} Mpc divided into 768^3 cells for baryons, with 384^3 dark matter particles (see Cen et al. 2002 for a more complete description). We compare to a 25 h^{-1} Mpc, 512^3 particle HPM simulation with identical initial Fourier modes up to the Nyquist frequency of the mesh (as we show in the next subsection, the resolution of these simulations is sufficient for convergence of our chosen statistics, so we do not need to worry about exactly how to equate resolution between the two).

The comparison for our statistics is shown in Figure 4, where the solid lines are the HPM results and the dotted lines are the fully hydrodynamic results. The agreement is excellent except in the highest δ_F regions in the $\langle\delta_m|\delta_F\rangle$ and $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle$ comparisons (Figs. 4*a* and 4*c*). For reference, in Figures 4*a*–4*b* we plot the probability distribution functions (PDFs) of δ_F and δ_m, respectively (defined as the relative volume-weighted probability of finding δ within each of our usual bins). We see that the regions where the HPM

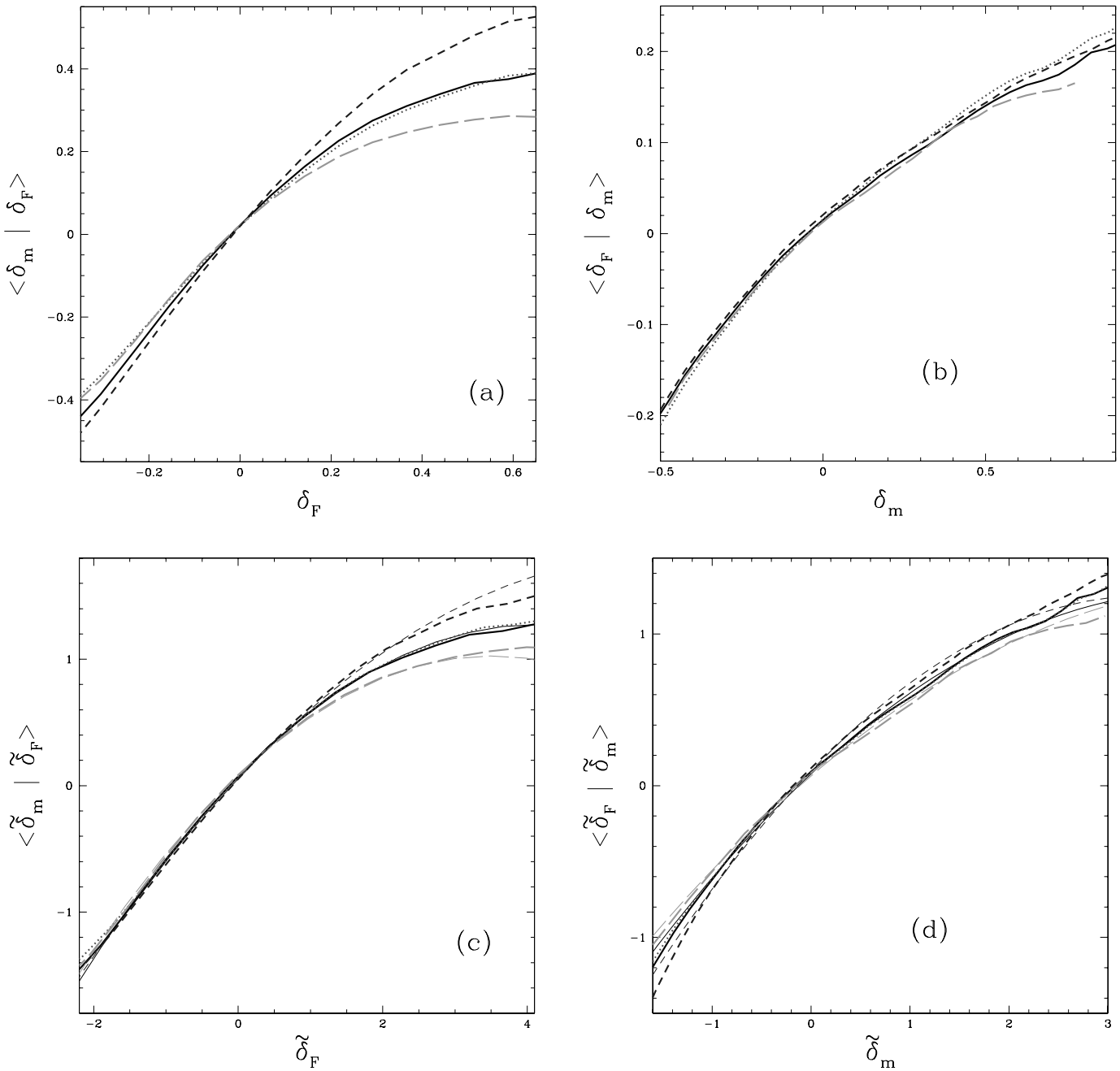

FIG. 3.—Comparison between our standard 3.5 h^{-1} Mpc Gaussian smoothing of the spectrum with 10 h^{-1} Mpc cube smoothing of the mass (*solid lines*) and 5 h^{-1} Mpc top-hat smoothing of the spectrum (*dotted lines*), 12 h^{-1} Mpc cube smoothing of the mass with 4.2 h^{-1} Mpc Gaussian smoothing of the spectrum (*long-dashed lines*), or 8 h^{-1} Mpc cube smoothing of the mass with 2.7 h^{-1} Mpc Gaussian smoothing of the spectrum (*short-dashed lines*). (*a*) Mean mass as a function of transmitted flux. (*b*) Mean flux as a function of mass. (*c–d*) Same as (*a–b*), except that δ_F and δ_m are divided by their rms fluctuations. The mass is always computed in redshift space. In (*c–d*), the thick lines show the simulation results, while the thin lines of similar type show the quadratic fits. [*See the electronic edition of the Journal for a color version of this figure.*]

approximation breaks down are extremely rare (at low δ_m in Fig. 2*b* the lines terminate at the lowest density found in the simulations). The values of σ_F are 0.138 and 0.134 for the hydrodynamic and HPM simulations, respectively, while the values of σ_m are 0.238 and 0.231.

2.4. *Convergence of the Results with Resolution and Box Size*

We start by testing the sensitivity of our numerical results to the resolution. For this purpose, we use boxes with size 20 h^{-1} Mpc in addition to our standard value of 40 h^{-1} Mpc.

The results for the four functions, $\langle \delta_m | \delta_F \rangle$, $\langle \delta_F | \delta_m \rangle$, $\langle \tilde{\delta}_m | \tilde{\delta}_F \rangle$, and $\langle \tilde{\delta}_F | \tilde{\delta}_m \rangle$, are shown in Figures 5*a–5d* for our standard cosmological model, with box size 20 h^{-1} Mpc and resolution of 512^3 and 256^3 particles for the long-dashed and short-dashed lines, respectively. The solid and dotted lines show results with box size 40 h^{-1} Mpc, again with 512^3 and 256^3 particles, respectively. The agreement between the 20 h^{-1} Mpc simulations differing only in the resolution is excellent, verifying that 40 h^{-1} Mpc simulations with 512^3

TABLE 1
SELECTED RESULTS

FIGURE AND LINE	σ_F	σ_m	$\langle\tilde{\delta}_m\|\tilde{\delta}_F\rangle$ c_0	c_1	c_2	$\langle\tilde{\delta}_F\|\tilde{\delta}_m\rangle$ c_0	c_1	c_2
3s[a]	0.16	0.31	0.073	0.58	−0.070	0.077	0.61	−0.077
3d[b]	0.17	0.31	0.065	0.58	−0.063	0.074	0.61	−0.075
3ld[c]	0.15	0.26	0.079	0.54	−0.077	0.064	0.56	−0.062
3sd[d]	0.19	0.37	0.056	0.60	−0.051	0.097	0.68	−0.10
7ld[e]	0.11	0.31	0.088	0.64	−0.084	0.050	0.63	−0.052
7sd[f]	0.19	0.31	0.064	0.55	−0.062	0.088	0.60	−0.091
9d[g]	0.18	0.40	0.057	0.58	−0.053	0.088	0.65	−0.094
9ld[h]	0.17	0.33	0.071	0.60	−0.067	0.086	0.64	−0.089

NOTES.—First column gives the figure and line to which the row of results applies; s = solid line, d = dotted line, ld = long-dashed line, and sd = short-dashed line. Formulae accurate to ∼0.1 in the intervals $-2 \lesssim \tilde{\delta}_F \lesssim 2$ and $-1.2 \lesssim \tilde{\delta}_m \lesssim 2$; σ_m and σ_F are intended only for use as conversion factors from $\tilde{\delta}$ to δ (alone they may be sensitive to resolution and box size in a way that we have not tested).

[a] Standard case.
[b] 5 h^{-1} Mpc top-hat smoothing of the spectra.
[c] 12 h^{-1} Mpc cube, 4.2 h^{-1} Mpc Gaussian.
[d] 8 h^{-1} Mpc cube, 2.7 h^{-1} Mpc Gaussian.
[e] $\bar{F} = 0.8$.
[f] $\bar{F} = 0.6$.
[g] $\sigma_8 = 1$.
[h] $n = 0.85$.

particles should be well resolved. The agreement between 40 h^{-1} Mpc simulations with different resolution is still quite good except at high δ_F, indicating that 80 h^{-1} Mpc, 512^3 particle simulations could still be useful.

Figure 5 shows a disturbingly large difference between the results for 20 and 40 h^{-1} Mpc boxes; however, assessing the effect of the box size is more difficult than for the resolution, because we cannot use the same set of initial mode amplitudes at each wavenumber, so the statistical fluctuations in the simulations introduce significant random differences in the result. Figures 6*a*–6*d* compare results for 40 h^{-1} Mpc boxes (*solid lines*), 20 h^{-1} Mpc boxes (*dotted lines*), and 80 h^{-1} Mpc boxes (*dashed lines*), all with 512^3 particles. The thick lines show the average of four simulations, and the thin lines show the results from each separate simulation to demonstrate the statistical error in the curves.

Figure 6*a* shows good agreement between the 40 and 80 h^{-1} Mpc boxes for $\langle\delta_m|\delta_F\rangle$, except at large δ_F, where the trend toward increasing mean mass with increasing box size from 20 to 40 to 80 h^{-1} Mpc is probably not a result of statistical fluctuations. The values of σ_m are 0.25, 0.31, and 0.34 for the 20, 40, and 80 h^{-1} Mpc boxes, respectively, so some nonconvergence of our statistics is not too surprising. In fact, the true difference between the 40 and 80 h^{-1} Mpc boxes is even larger than what is shown because the decreased resolution in the 80 h^{-1} Mpc box has the effect of suppressing $\langle\delta_m|\delta_F\rangle$ at large δ_F (see Fig. 5*a*). It is clear that much of the flattening of $\langle\delta_m|\delta_F\rangle$ that we see at large δ_F is an effect of the finite box size. Fortunately, the other statistics shown in Figures 6*b*–6*d* generally exhibit better convergence than $\langle\delta_m|\delta_F\rangle$.

The results in Figures 6*a*–6*d* should also serve as a cautionary reminder that statistical fluctuations from one simulation to another can be large and must be taken into account when comparing to observations or to other simulations, especially single, relatively small simulations. Surprisingly, the scatter between the separate simulations in the high-δ_F corner of the plot is not reduced by increasing the box size from 40 to 80 h^{-1} Mpc. This may be an indication that the statistical fluctuations in the result are still dominated by the longest wavelength modes in the box.

Based on the difference between the fully hydrodynamic and the HPM simulations and the imperfect convergence with box size, we estimate that the formulae that we give in Table 1 describing $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle$ and $\langle\tilde{\delta}_F|\tilde{\delta}_m\rangle$ are accurate to about 0.1 in the intervals $-2 \lesssim \tilde{\delta}_F \lesssim 2$ and $-1.2 \lesssim \tilde{\delta}_m \lesssim 2$, respectively (see Figs. 4*c*–4*d* and 6*c*–6*d*).

For reference, for our standard model, in a 40 h^{-1} Mpc box with our standard smoothing, the rms fluctuations in δ_F and δ_m are $\sigma_F = 0.16$ and $\sigma_m = 0.31$.

3. PARAMETER DEPENDENCE OF THE Lyα FOREST–MASS CORRELATION

We now examine the dependence of the four functions $\langle\delta_m|\delta_F\rangle$, $\langle\delta_F|\delta_m\rangle$, $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle$, and $\langle\tilde{\delta}_F|\tilde{\delta}_m\rangle$ on the most important parameters determining the properties of the Lyα forest. These are the mean transmitted flux, the mean temperature-density relation of the intergalactic gas (parameterized as $T = T_0\Delta^{\gamma-1}$), and the amplitude and power-law slope of the power spectrum. These parameters fully determine all the physical effects that are incorporated in an HPM simulation and the calculation of the Lyα spectra: the underlying dark matter fluctuations (assumed to be Gaussian initially), the smoothing of the gas distribution on the Jeans scale, and the neutral fraction of the gas and thermal broadening.

3.1. *Mean Transmitted Flux*

The mean transmitted flux of the Lyα forest is an increasing function of redshift. The value that we use in our standard model, $\bar{F} = 0.67$, is appropriate for $z = 3$. The variation of the mean transmitted flux is the most important factor accounting for the changes in the Lyα forest observed properties with redshift.

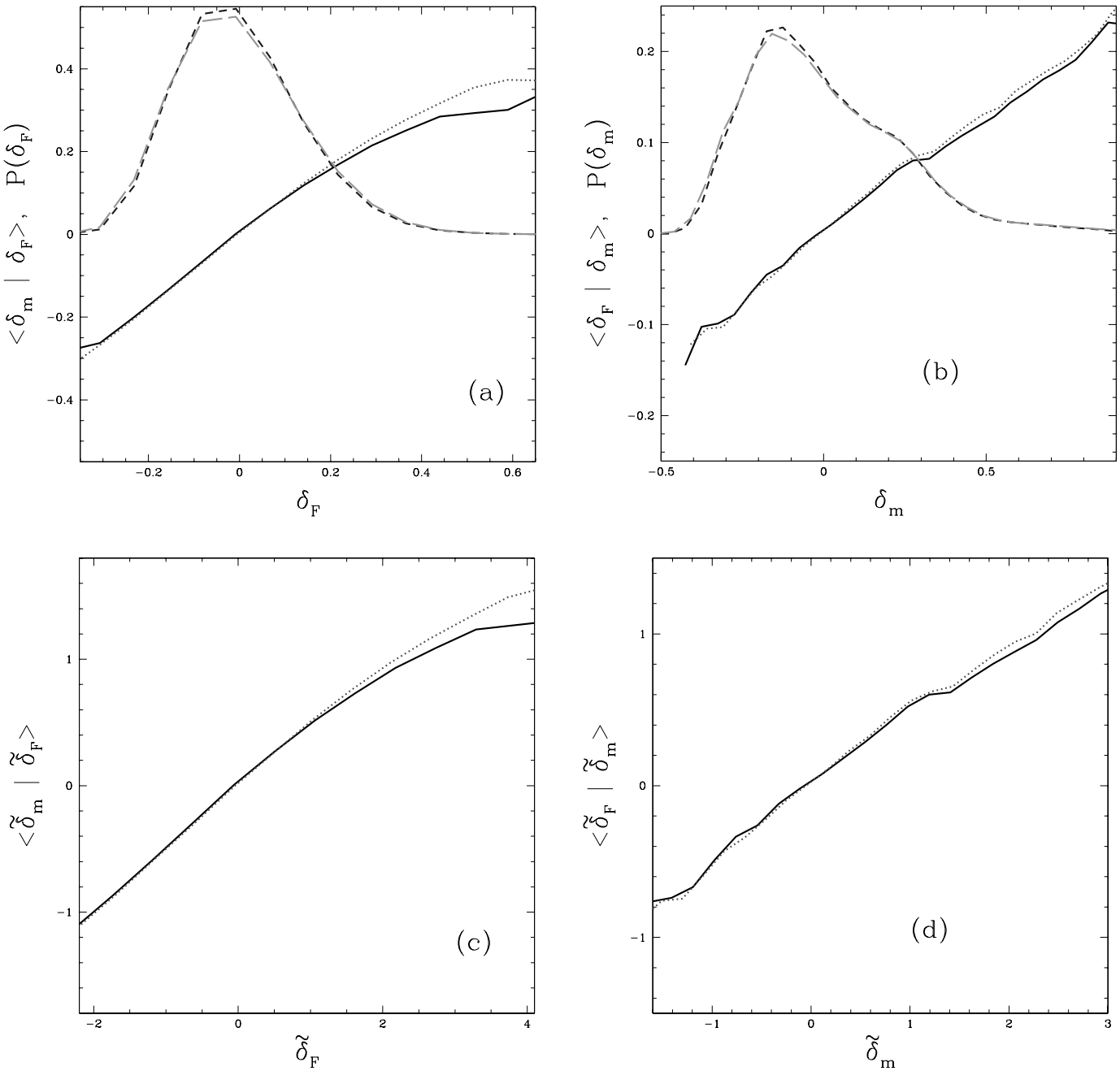

FIG. 4.—Test of the HPM approximation. *Solid lines*: HPM results of our standard statistics; *dotted lines*: fully hydrodynamic results for our standard statistics; *short-dashed lines*: PDFs in the HPM simulation; *long-dashed lines*: PDFs in the fully hydrodynamic simulation. The normalization of the PDFs is arbitrary. [*See the electronic edition of the Journal for a color version of this figure.*]

Figures 7*a*–7*d* show the four functions defined previously for $\bar{F} = 0.67$ (*solid lines*), $\bar{F} = 0.8$, appropriate for $z \simeq 2.5$ (*long-dashed lines*), $\bar{F} = 0.6$, appropriate for $z \simeq 3.5$ (*short-dashed lines*), and $\bar{F} = 0.7$ (*dotted lines*).

The functions $\langle \delta_m | \delta_F \rangle$ and $\langle \delta_F | \delta_m \rangle$ depend strongly on $\bar{F}$, whereas the functions with the normalized quantities $\langle \tilde{\delta}_m | \tilde{\delta}_F \rangle$ and $\langle \tilde{\delta}_F | \tilde{\delta}_m \rangle$ have a much weaker dependence. Essentially, the mean transmitted flux changes the effective "bias" in the Lyα forest that determines the value of δ_F that corresponds to a given mass fluctuation, and once a linear bias is eliminated the dependence on $\bar{F}$ is much smaller. Note that Figures 7*a*–7*b* reflect the fact that the Lyα forest bias is larger for higher mean flux decrements (see Fig. 10*c* of McDonald 2002; Croft et al. 1999; McDonald et al. 2000).

We give the results for $\bar{F} = 0.6$ and $\bar{F} = 0.8$ in Table 1, as rows 7sd and 7ld, respectively. The first row in the table already contains the results for $\bar{F} = 0.67$.

3.2. *Temperature-Density Relation*

The temperature-density relation has practically no effect on the functions that we are analyzing, as shown in Figures 8*a*–8*d*. The reason for this insensitivity is that the temperature-density relation does not affect the large-scale properties of the Lyα forest directly, and apparently

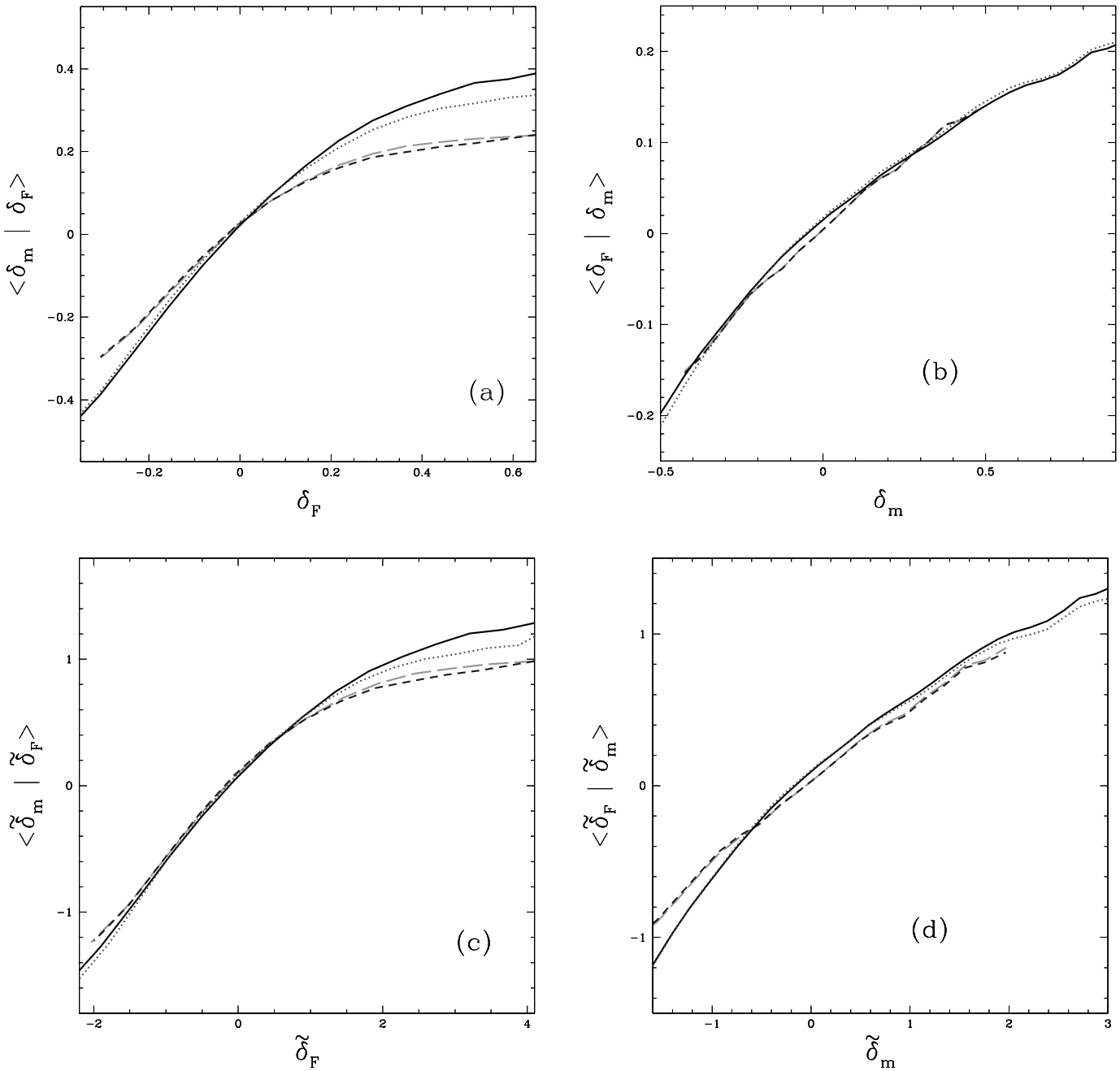

FIG. 5.—Resolution tests. *Long-dashed line*: 20 h^{-1} Mpc simulation box with 512^3 particles; *short-dashed line*: 20 h^{-1} Mpc, 256^3 particles; *solid line:* 40 h^{-1} Mpc, 512^3 particles; *dotted line*: 40 h^{-1} Mpc, 256^3 particles. (*a*) Mean mass as a function of transmitted flux. (*b*) Mean flux as a function of mass. (*c–d*) Same as (*a–b*), except that δ_F and δ_m are divided by their rms fluctuations. The mass is always computed in redshift space. [*See the electronic edition of the Journal for a color version of this figure.*]

its effect through the large-scale bias factor is quite small (see Fig. 10*b* in McDonald 2002).

3.3. *Amplitude of the Power Spectrum*

In Figure 9, we show the variations with the amplitude of the mass power spectrum by increasing the amplitude in our standard model, shown by the solid line, by 33%, to produce the dotted line. In practice, we do this by simply using an output of the same simulation at a scale factor 1.33 times larger, but keeping the mean transmitted flux in the Lyα spectrum constant (this is completely equivalent to running a new simulation with a different amplitude, except that the effective temperature changes because of the change in the relation between comoving distance and velocity, but the previous figures show that the temperature does not affect the results). The result reflects that the bias factor of the Lyα forest decreases with increasing power spectrum amplitude (see Fig. 10*a* of McDonald 2002). The difference is partially canceled when using the normalized variables $\tilde{\delta}_F$ and $\tilde{\delta}_m$, although less so for the function $\langle\tilde{\delta}_F|\tilde{\delta}_m\rangle$.

We also vary the power-law slope of the power spectrum in Figure 9, by changing n from 0.95 to 0.85 while holding the power fixed at wavenumber $k = 1$ $(h^{-1}$ Mpc$)^{-1}$. The

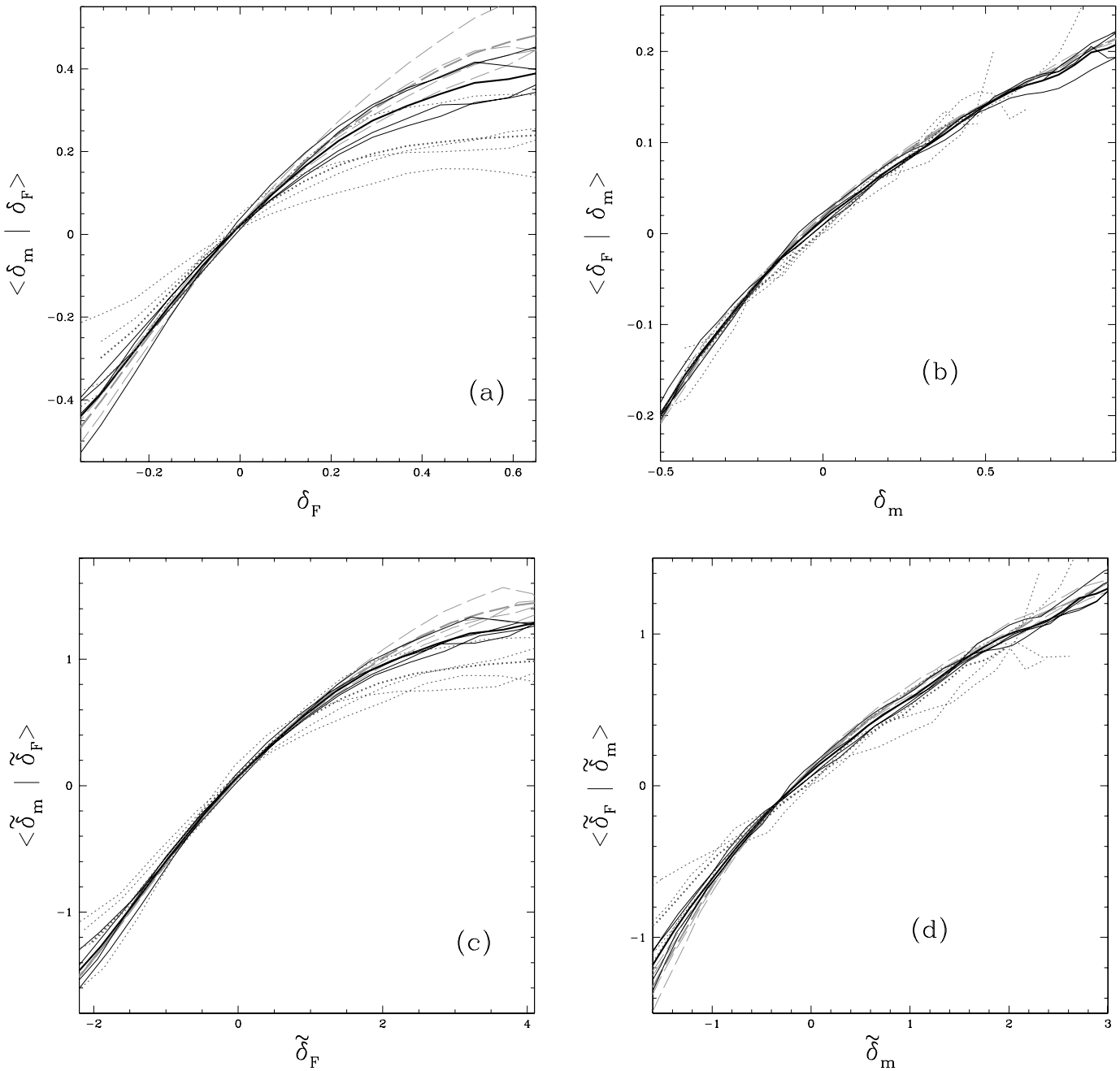

FIG. 6.—Box size and statistical error test. The thin curves are from four independent simulations; the thick curves are the average of the four. *Solid lines*: 40 h^{-1} Mpc simulations; *dotted lines*: 20 h^{-1} Mpc simulations; *dashed lines*: 80 h^{-1} Mpc simulations. All are with 512^3 particles (the 80 h^{-1} Mpc simulations are slightly underresolved). [*See the electronic edition of the Journal for a color version of this figure.*]

effect that we see is relatively small. Had we carefully chosen the value of k at which we hold the power fixed to minimize the variation with n, the effect might be even smaller.

We give the results for $\sigma_8 = 1$ and $n = 0.85$ in Table 1, as rows 9d and 9ld, respectively. The first row in the table is for the standard power spectrum.

4. DISCUSSION

We have presented predictions for the expected correlation of the mass and the Lyα forest transmitted flux, smoothed over a cube size of $\sim 10\ h^{-1}$ Mpc. Our most basic result is that the relation between $\langle \tilde{\delta}_m | \tilde{\delta}_F \rangle$ and $\tilde{\delta}_F$ is described by the formula $\langle \tilde{\delta}_m | \tilde{\delta}_F \rangle = 0.073 + 0.58\tilde{\delta}_F - 0.070\tilde{\delta}_F^2$, which is accurate to ~ 0.1 for the range $-2 \lesssim \tilde{\delta}_F \lesssim 2$. We have shown that this correlation is not sensitive to the temperature-density relation of the gas. There is a dependence on the mean transmitted flux and the amplitude of the mass power spectrum, but these quantities are already measured to reasonable accuracy (Croft et al. 1999, 2002; McDonald et al. 2000). We have also shown that the predictions do not suffer from uncertainties due to the resolution of the simulations. Calculations with even larger boxes than used here will be desirable, because the uncertainties in the predictions due to the variance in the simulations and the suppression

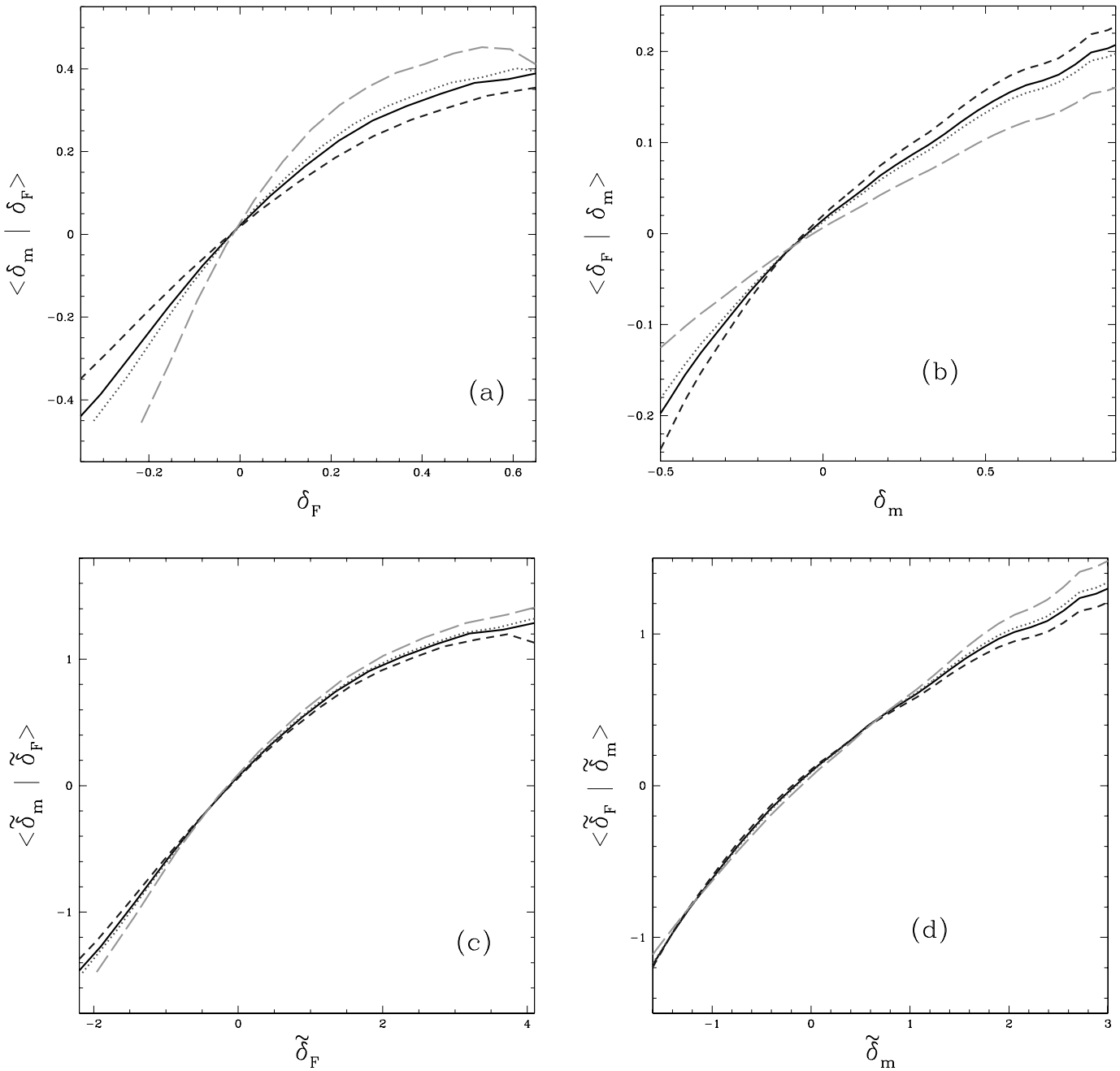

FIG. 7.—Effect of mean transmitted flux fraction. Shown are $\bar{F} = 0.67$ (*solid lines*), $\bar{F} = 0.8$ (*long-dashed lines*), $\bar{F} = 0.6$ (*short-dashed lines*), and $\bar{F} = 0.7$ (*dotted lines*), for the same four functions as previously. [*See the electronic edition of the Journal for a color version of this figure.*]

of the large-scale power may still be significant; however, this uncertainty is mostly isolated in the most rare, high-density regions ($\tilde{\delta}_F > 2$).

A similar relation holds between $\langle \tilde{\delta}_F | \tilde{\delta}_m \rangle$ and $\tilde{\delta}_m$, described by the formula $\langle \tilde{\delta}_F | \tilde{\delta}_m \rangle = 0.077 + 0.61\delta_m - 0.077\tilde{\delta}_m^2$ (accurate to $\sim$0.1 in the interval $-1.2 \lesssim \tilde{\delta}_m \lesssim 2$). We notice, however, that the observational determination of this other function will be affected by galaxy shot noise. The number of galaxies observed in a certain cube in redshift space containing a fixed mass is subject to shot noise, and this inevitably introduces a smoothing of the function $\langle \tilde{\delta}_F | \tilde{\delta}_m \rangle$, which must be taken into account before any comparisons to our theoretical results are made (in contrast, galaxy shot noise does not change the average $\langle \delta_m | \delta_F \rangle$, but it does alter the rms dispersion σ_m needed to compute $\tilde{\delta}_m$). Measuring the effects of galaxy shot noise can also teach us useful information about how galaxies form, because galaxy shot noise does not generally need to be strictly described by Poisson statistics. For example, for a fixed mass contained within a cube, once a galaxy is found in the cube the probability to find others may be lower because some mass has already been used up by that galaxy.

Comparing these theoretical predictions with observations allows for several tests of the basic theory of the Lyα forest and can reveal new information on the spatial distribution of the galaxies. The functions $\langle \tilde{\delta}_m | \tilde{\delta}_F \rangle$ and $\langle \tilde{\delta}_F | \tilde{\delta}_m \rangle$, which should not be affected by any linear galaxy bias,

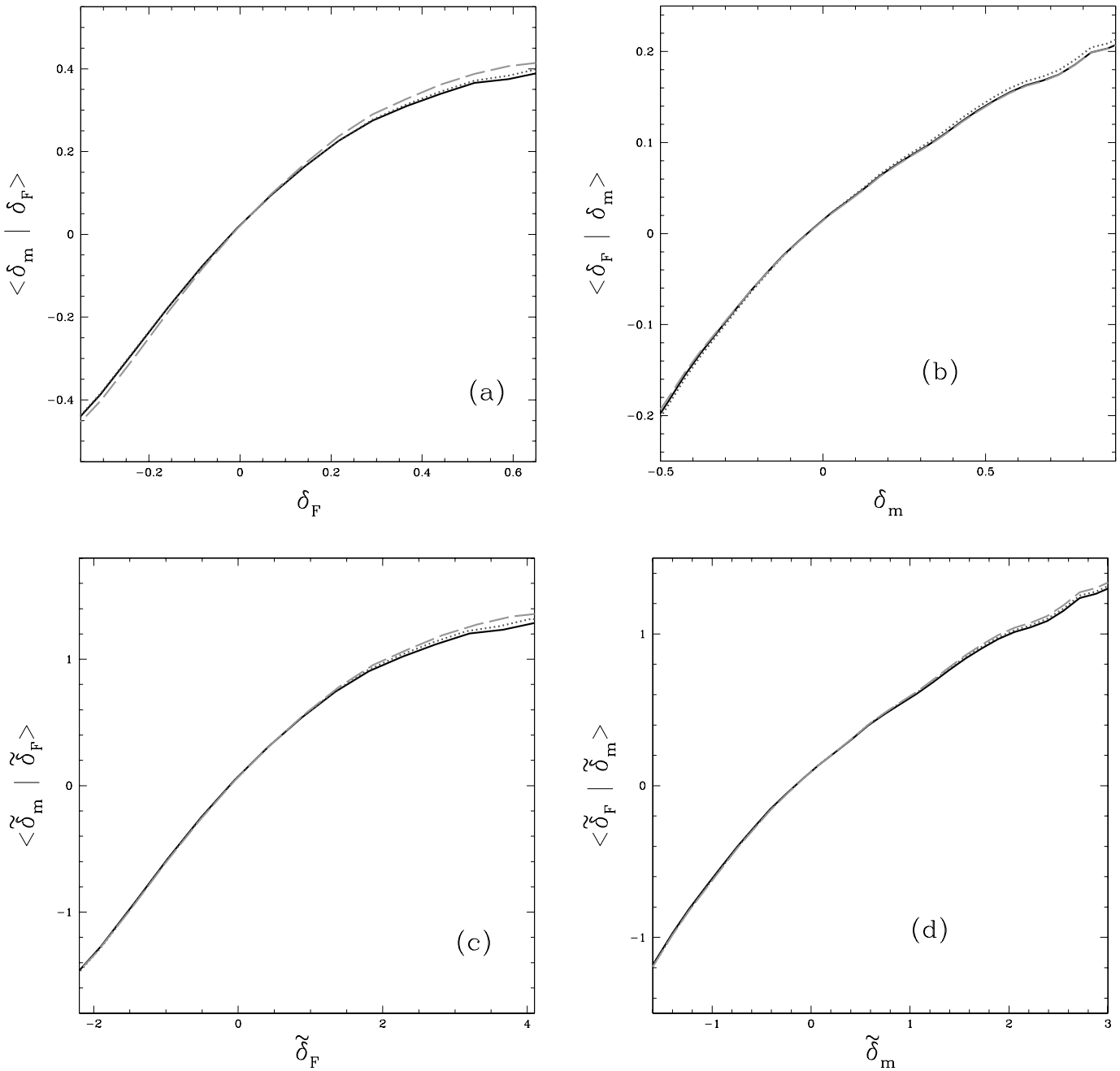

FIG. 8.—Effect of temperature-density relation. *Solid lines*: $T_{1.4} = 17,000$ K, $\gamma - 1 = 0.3$ (standard values); *dotted lines*: $T_{1.4} = 22,000$ K, $\gamma - 1 = 0.3$; *dashed lines*: $T_{1.4} = 17,000$ K, $\gamma - 1 = 0.6$. [*See the electronic edition of the Journal for a color version of this figure.*]

provide a powerful test of the assumption that the basic framework assumed here for the nature of the Lyα forest is correct and that galaxies trace the mass on large scales apart from linear bias. If the slope of $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle$ as a function of $\tilde{\delta}_F$ were found to be larger than predicted, this would imply that the Lyα forest is much more closely associated with galaxies than is expected from their common correlation with the mass distribution. If the observed slope were smaller than predicted, it would indicate that the Lyα-absorbing gas and the galaxies tend to avoid each other for some reason.

If the correlation of $\tilde{\delta}_m$ and $\tilde{\delta}_F$ is as expected, this will imply a strong confirmation of the basic model that we have for the Lyα forest and will justify using the Lyα forest as a predictor of the mass fluctuations. Comparing the predicted function $\langle\delta_m|\delta_F\rangle$ with the observed $\langle\delta_g|\delta_F\rangle$ will then yield the bias factor of any type of galaxies for which these observations can be made.

The correlation of galaxies and the Lyα forest can be measured as a function of scale. Our predicted value for the slope of $\langle\tilde{\delta}_m|\tilde{\delta}_F\rangle$ as a function of δ_F decreases with the smoothing scale, in a way that reflects the shape of the mass autocorrelation function. This will allow a precise test of the idea that the large-scale distributions of different types of galaxies differ only in a constant linear bias relative to the mass fluctuations.

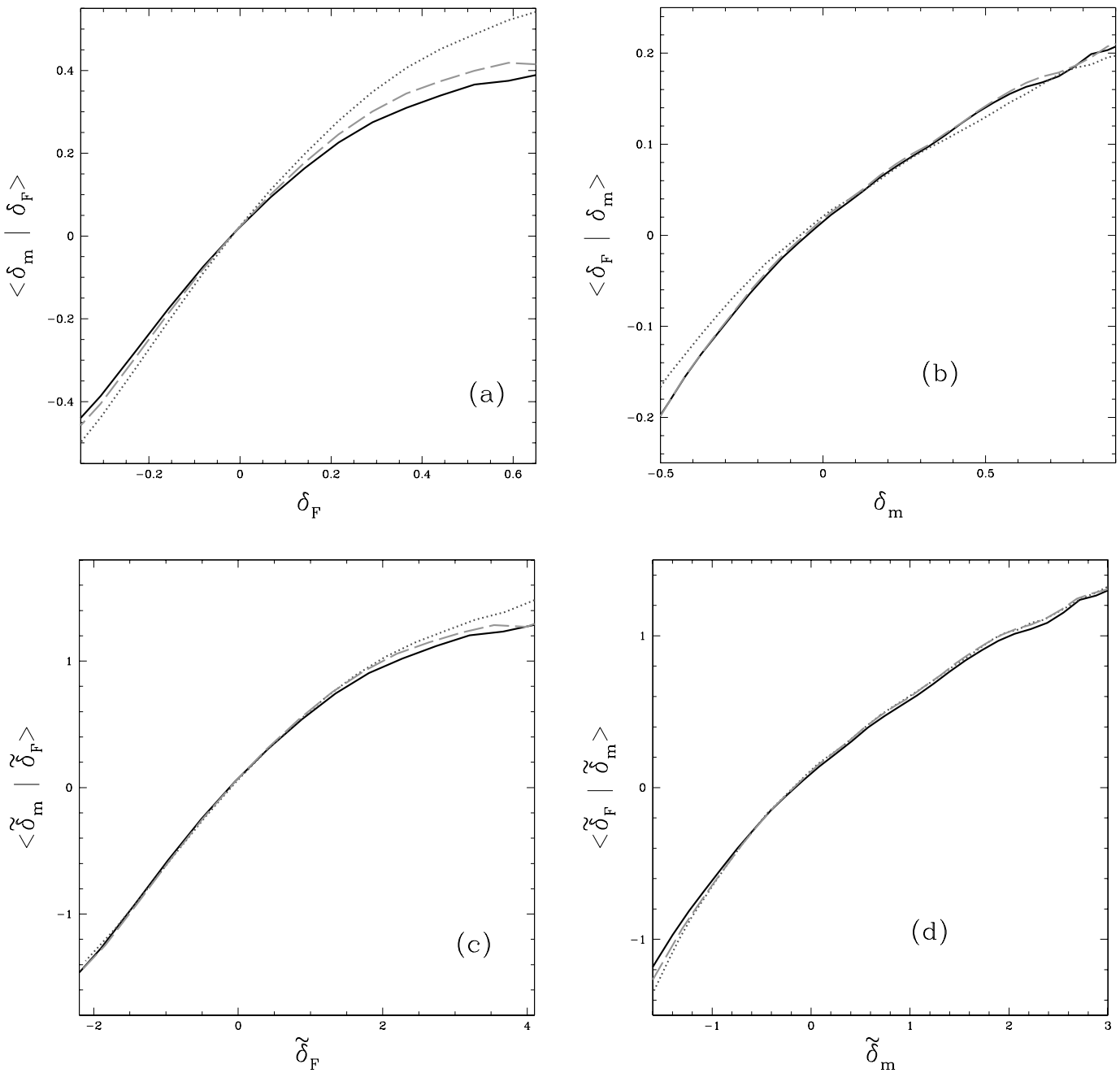

FIG. 9.—Effect of the mass power spectrum. *Solid lines*: Standard power spectrum; *dotted lines*: rms amplitude increased by 33%; *dashed lines*: $n = 0.85$ instead of $n = 0.95$, with fixed amplitude at $k = 1$ $(h^{-1}\ \mathrm{Mpc})^{-1}$. [*See the electronic edition of the Journal for a color version of this figure.*]

We acknowledge useful discussions with Kurt Adelberger, Charles Steidel, and David Weinberg. We thank Nick Gnedin for the HPM code. P. M. is supported by a grant from the Packard Foundation. R. C. is supported by grants AST 93-18185 and ASC 97-40300. J. M. is supported by grant NSF 00-98515. The HPM simulations were performed at the National Center for Supercomputing Applications.

REFERENCES

Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M. 2002, ApJ, submitted
Bechtold, J., Crotts, A. P. S., Duncan, R. C., & Fang, Y. 1994, ApJ, 437, L83
Bergeron, J., & Boissé, P. 1991, A&A, 243, 344
Bi, H. G. 1993, ApJ, 405, 479
Bi, H. G., & Davidsen, A. F. 1997, ApJ, 479, 523
Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9
Cen, R., Ostriker, J. P., Prochaska, J. X., & Wolfe, A. M. 2002, ApJ, submitted (astro-ph/0203524)
Chen, H.-W., Lanzetta, K. M., Webb, J. K., & Barcons, X. 2001, ApJ, 559, 654
Croft, R. A. C., Weinberg, D. H., Bolte, M., Burles, S., Hernquist, L., Katz, N., Kirkman, D., & Tytler, D. 2002, ApJ, in press
Croft, R. A. C., Weinberg, D. H., Pettini, M., Hernquist, L., & Katz, N. 1999, ApJ, 520, 1
Dinshaw, N., Impey, C. D., Foltz, C. B., Weymann, R. J., & Chaffee, F. H. 1994, ApJ, 437, L87
Dinshaw, N., Weymann, R. J., Impey, C. D., Foltz, C. B., Morris, S. L., & Ake, T. 1997, ApJ, 491, 45

Dolan, J. F., Michalitsianos, A. G., Nguyen, Q. T., & Hill, R. J. 2000, ApJ, 539, 111
Gnedin, N. Y., & Hui, L. 1998, MNRAS, 296, 44
Grogin, N. A., & Geller, M. J. 1998, ApJ, 505, 506
Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, ApJ, 457, L51
Lanzetta, K. M., Bowen, D. V., Tytler, D., & Webb, J. K. 1995, ApJ, 442, 538
López, S., Hagen, H.-J., & Reimers, D. 2000, A&A, 357, 37
McDonald, P. 2002, ApJ, submitted (astro-ph/0108064)
McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., & Cen, R. 2001, ApJ, 562, 52
McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., Cen, R., & Ostriker, J. P. 2000, ApJ, 543, 1
Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M. 1996, ApJ, 471, 582
Monier, E. M., Turnshek, D. A., & Hazard, C. 1999, ApJ, 522, 627
Penton, S. V., Stocke J. T., & Shull, J. M. 2002, ApJ, 565, 720
Petitjean, P., Surdej, J., Smette, A., Shaver, P., Muecket, J., & Remy, M. 1998, A&A, 334, L45
Primack, J. R. 2000, preprint (astro-ph/0007187)
Rauch, M., et al. 1997, ApJ, 489, 7
Schaye, J. 2001, ApJ, 559, 507
Steidel, C. C., Dickinson, M., & Persson, E. 1994, ApJ, 437, L75
Zhang, Y., Anninos, P., & Norman, M. L. 1995, ApJ, 453, L57
Zhang, Y., Meiksin, A., Anninos, P., & Norman, M. L. 1998, ApJ, 495, 63