

# Estimating cosmic velocity fields from density fields and tidal tensors

Francisco-Shu Kitaura<sup>1,2\*</sup>, Raul E. Angulo<sup>2</sup>, Yehuda Hoffman<sup>3</sup> and Stefan Gottlöber<sup>1</sup>

<sup>1</sup> *Leibniz-Institut für Astrophysik (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*

<sup>2</sup> *Max-Planck Institut für Astrophysik (MPA), Karl-Schwarzschildstr. 1, D-85748 Garching, Germany*

<sup>3</sup> *Racah Institute of Physics, Hebrew University, Jerusalem, Israel*

30 November 2011

## ABSTRACT

In this work we investigate the nonlinear and nonlocal relation between cosmological density and peculiar velocity fields. Our goal is to provide an algorithm for the reconstruction of the nonlinear velocity field from the fully nonlinear density. We find that including the gravitational tidal field tensor using second order Lagrangian perturbation theory (2LPT) based upon an estimate of the linear component of the nonlinear density field significantly improves the estimate of the cosmic flow in comparison to linear theory not only in the low density, but also and more dramatically in the high density regions. In particular we test two estimates of the linear component: the log-normal model and the iterative Lagrangian linearisation. The present approach relies on a rigorous higher order Lagrangian perturbation theory analysis which incorporates a nonlocal relation. It does not require additional fitting from simulations being in this sense parameter free, it is independent of statistical-geometrical optimisation and it is straightforward and efficient to compute. The method is demonstrated to yield an unbiased estimator of the velocity field on scales  $\gtrsim 5 h^{-1}$  Mpc with closely Gaussian distributed errors. Moreover, the statistics of the divergence of the peculiar velocity field is extremely well recovered showing a good agreement with the true one from  $N$ -body simulations. The typical errors of about 10 km/s (1 sigma confidence intervals) are reduced by more than 80% with respect to linear theory in the scale range between 5 and 10  $h^{-1}$  Mpc in high density regions ( $\delta > 2$ ). We also find that iterative Lagrangian linearisation is significantly superior in the low density regime with respect to the lognormal model.

**Key words:** (cosmology:) large-scale structure of Universe – galaxies: clusters: general – catalogues – galaxies: statistics

## 1 INTRODUCTION

Gravitational instability is one of the key ingredients to explain the rich hierarchy of structures we observe today in the Universe. It has amplified small mass fluctuations produced after inflation to give rise from large galaxy clusters to huge voids. Simultaneously, the same local gravitational field imprinted “peculiar” velocities in galaxies; deviations from the overall expansion of the Universe which is responsible for the Hubble flow.

The peculiar velocity of galaxies is a valuable quantity in cosmology since it contains similar but complementary information to that enclosed in the galaxies position. For instance, by requiring isotropy in the measured galaxy clustering, the cosmological mass density parameter and even

the nature of gravity can be constrained (see e.g. Davis et al. 1996; Willick & Strauss 1998; Branchini et al. 2001; Guzzo et al. 2008). In addition, these motions can be used to reconstruct the properties of the universe at a earlier time, in principle, even at recombination where perturbations were completely linear (Nusser & Dekel 1992; Gramann 1993a; Croft & Gaztanaga 1997; Narayanan & Weinberg 1998; Monaco & Efstathiou 1999; Frisch et al. 2002).

A method able to accurately determine the peculiar velocity field can be used in many different applications; ranging from bias studies combining galaxy redshift surveys with measured peculiar velocities (see e.g. Fisher et al. 1995; Zaroubi et al. 1999; Courtois et al. 2011), Baryon Acoustic Oscillations reconstructions (see e.g. Eisenstein et al. 2007), determination of the growth factor, to estimates of the initial conditions of the Universe which in turn can be used for constrained simulations (see e.g. Gottloeber et al. 2010; Klypin

\* E-mail: kitaura@aip.de, Karl-Schwarzschild fellow

Works	$\theta$ - $\delta$ relation	parameters	methodology
Nusser et al. (1991)	$\theta = \delta$	$\{\mathbf{p}\}$	linear theory: LPT
Bernardeau (1992)	$\theta = \delta/(1 + \alpha\delta)$	$\alpha, \{\mathbf{p}\}$	empirical approximation
Gramann (1993b)	$\theta = \alpha \left[ (1 + \delta)^\beta - 1 \right]$	$\alpha, \beta, \{\mathbf{p}\}$	PT
Willick et al. (1997)	$\theta = \delta - \alpha D_1^2 \mu^{(2)}(\phi_g)$	$\alpha, \{\mathbf{p}\}$	approx. 2LPT
Chodorowski et al. (1998)	$\theta = [(1 + \alpha^2 \sigma_\delta^2) \delta + \alpha \sigma_\delta^2] / (1 + \alpha \delta)$	$\alpha, \sigma_\delta^2, \{\mathbf{p}\}$	empirical approximation
Bernardeau et al. (1999)	$\theta = [\delta - \alpha (\delta^2 - \sigma_\delta^2) + \beta \delta^3]$	$\alpha, \beta, \sigma_\delta^2, \{\mathbf{p}\}$	PT+N-body
Kudlicki et al. (2000)	$\theta = [\gamma \delta + \alpha (\delta^2 - \sigma_\delta^2) + \beta \delta^3]$	$\alpha, \beta, \gamma, \sigma_\delta^2, \{\mathbf{p}\}$	PT+N-body
Bilicki & Chodorowski (2008)	$\theta = \alpha \left[ (1 + \delta)^{1/\alpha} - 1 \right] + (\alpha - 1)/(2\alpha) \sigma_\delta^2$	$\alpha, \sigma_\delta^2, \{\mathbf{p}\}$	PT+Eulerian grid-based code
this work	$\theta = \gamma \left[ (1 + \delta)^{1/\alpha} - (1 + \delta)^\beta \right]$	$\alpha, \beta, \gamma, \{\mathbf{p}\}$	spherical collapse model
	$\theta = D_1 \delta^{(1)} - D_2 f_2 / f_1 \mu^{(2)}(\phi^{(1)})$	$\{\mathbf{p}\}$	2LPT

**Table 1.** The parameters  $\alpha, \beta, \gamma$  are either derived from first principles (Bernardeau 1992; Gramann 1993b; Bilicki & Chodorowski 2008, this work) or from fitting to simulations being different for each case. These parameters also depend on the scale of interest. The parameters which are in parenthesis are fully determined by the chosen cosmology and the theoretical model. The variance of the density field is given by  $\sigma_\delta^2 = \langle \delta^2 \rangle$ . PT stands for perturbation theory and is applied only in the univariate case (local relation). 2LPT stands for second order Lagrangian perturbation theory and yields the only nonlocal (and nonlinear) relation from the list. Hivon et al. (1995); Monaco & Efstathiou (1999) proposed 2LPT to correct for redshift distortions. Let us mention here the least-action principle methods to determine the peculiar velocities from mass tracer objects (galaxies) introduced by Peebles (1989) and further extended by Nusser & Branchini (2000); Branchini et al. (2002); Lavaux et al. (2008).

et al. 2003). A particularly well suited application regards the topological methods to detect the kinematic Sunyaev-Zeldovich effect.

There is in addition recent interest in the measurement of large-scale flows. Some authors claim to have detected a so-called “dark flow” caused by super-horizon perturbations (see e.g. Kashlinsky et al. 2009, 2011). Others discuss such flows as a challenge to the standard cosmological model as a whole (see e.g. Watkins et al. 2009).

Unfortunately, the apparent shift in spectral features of a galaxy is also affected by the expansion of the universe, therefore it is not possible to directly measure the peculiar velocity field in spectroscopic surveys. For this reason, one has to resort to indirectly infer it from the mass density fluctuations (but see Nusser et al. 2011, for a recent alternative method). However, this is not a trivial procedure due to the highly nonlinear state of the density fluctuations today and due to its nonlocal relationship with the velocity field.

The simplest approach is given by the linear continuity equation, which is routinely used in clustering studies. However, it has a range of applicability only limited to very large scales (e.g. Angulo et al. 2008). Alternative methods devised to improve upon linear theory can be separated into three areas. The first one consists on developing nonlinear relationships with higher-order perturbation theory (Bernardeau 1992; Chodorowski et al. 1998; Bernardeau et al. 1999; Kudlicki et al. 2000), with spherical collapse models (Bilicki & Chodorowski 2008) or based on empirical relations found in cosmological N-body simulations (Nusser et al. 1991; Willick et al. 1997).

A second idea is to solve the boundary problem of finding the initial positions of a set of particles governed by the Eulerian equation of motion and gravity with the least action principle (see Peebles 1989; Nusser & Branchini 2000; Branchini et al. 2002). A similar approach consists on relating the observed positions of matter tracers (e.g. galaxies) in a geometrical way to a homogeneous distribution by minimizing a cost function, which combined with the Zeldovich approximation (Zel’dovich 1970) provides an estimation of

the velocity field (see Mohayaee & Tully 2005; Lavaux et al. 2008).

The diversity of strategies and approximations for obtaining the velocity from the density field hint at the difficulty of the problem. Some approaches are simply not accurate enough and some are computationally very expensive. This sets the agenda for an improvement in the field. Any new method should be accurate, unbiased, computationally efficient and applicable to observational data.

A further shortcoming of the existing approaches is that they are mostly particle-based, which is not applicable for more general matter tracers like the Lyman-alpha forest or the 21-cm line, nor they can be combined with optimal density field estimators (see Kitaura et al. 2010; Jasche & Kitaura 2009; Kitaura et al. 2010).

In this paper we investigate a different approach based on higher order Lagrangian perturbation theory, and it is motivated by the pioneering work of Gramann (1993b) and further extended by Hivon et al. (1995); Monaco & Efstathiou (1999). The theoretical basis for LPT was carefully worked out by Moutarde et al. (1991); Buchert & Ehlers (1993); Buchert (1994); Bouchet et al. (1995); Catalan (1995) (for further references see Bernardeau et al. 2002).

Contrary to classic applications of LPT, in which the properties of an evolved distribution are predicted from a linear density field in Lagrangian coordinates (e.g. in the generation of initial conditions for N-body simulations or of galaxy mock catalogues: Jenkins (2010); Scoccimarro & Sheth (2002), our starting point is an evolved field in Eulerian coordinates (e.g. the present-day galaxy distribution). The key realisation of our approach is that it is possible to decompose a nonlinear density field into a Gaussian and Non-Gaussian term, which are related to each other through LPT (see similar approaches Gramann 1993a,b; Monaco & Efstathiou 1999). In other words, it is possible to find a closely Gaussian field which would evolve, under the assumption of LPT, into the measured density field today. This Gaussian field can then naturally be used to predict the corresponding velocity field today in LPT.

Our method combines i) the equations of motion and

continuity for a fluid under self gravity in 2-nd (3-rd) order Lagrangian Perturbation theory (LPT) with ii) the idea that the present-day galaxy distribution can be expanded into a closely linear-Gaussian field and a highly skewed nonlinear component consistent with 2LPT (or 3LPT) (see Kitaura & Angulo 2011). The former aspect makes our approach physically motivated and also captures the nonlocal nature of the density-velocity relation via the gravitational tidal field tensor. The latter aspect self-consistently minimises the impact of the approximations of a 2nd order formulation of gravitational evolution, but more importantly, it enables the application of LPT to nonlinear fields.

We note, that the lognormal transformation including the subtraction of a mean field to obtain an estimate of the linear field was proposed by Kitaura et al. (2010) and has been recently confirmed to give already a good estimate of the divergence of the displacement field (Falck et al. 2011). To estimate the velocity field one needs, however, to go to higher order perturbation theory as we show in this work.

In the next section we recap Lagrangian perturbation theory and derive the velocity-density relation to second and third order. In section 2.2 we will present the method to compute the peculiar velocity field from the nonlinear density field. In section 3 we will present our numerical tests based on the Millennium Run simulation. Here we will analyze the goodness of the recovered velocity divergence as compared to the true one and the same for each velocity component. A study of the errors in our method is also presented. Finally we present our conclusions.

## 2 VELOCITY–DENSITY RELATION

The first part of this section presents the relation between density and velocity fields in 2LPT, and how it can be applied to an evolved field. In the second part, we outline a practical implementation of this method.

### 2.1 Second order Lagrangian perturbation theory

The basic idea in Lagrangian perturbation theory is that an initial homogeneous field expressed in Lagrangian coordinates  $\mathbf{q}$  can be related to a final field in Eulerian coordinates  $\mathbf{x}$  through a unique mapping:  $\mathbf{x} = \mathbf{q} + \mathbf{\Psi}(\mathbf{q})$  determined by a displacement field  $\mathbf{\Psi}$  (see e.g. Bernardeau et al. 2002).

The linear solution for this expression is the well-known Zeldovich approximation, in which the displacement field is given by the Laplacian of the gravitational potential at  $\mathbf{q}$ . This result has been successfully applied to many aspects of cosmology, but it fails to describe the dynamics of a nonlinear field where shell crossings and curved trajectories are common.

An improvement is found by expanding the displacement field and considering higher order terms (see e.g. Buchert et al. 1994; Melott et al. 1995; Bouchet et al. 1995). For instance, the displacement field to second order is given by

$$\mathbf{\Psi} = -D_1 \nabla \phi^{(1)} + D_2 \nabla \phi^{(2)}, \quad (1)$$

where  $D_1$  is the linear growth factor normalised to unity today,  $D_2$  the second order growth factor, which is given by

$D_2 \simeq -3/7 \Omega^{-1/143} D_1^2$  (see Buchert & Ehlers 1993; Bouchet et al. 1995). The potentials  $\phi^{(1)}$  and  $\phi^{(2)}$  are obtained by solving a pair of Poisson equations:  $\nabla^2 \phi^{(1)} = \delta^{(1)}$ , where  $\delta^{(1)}$  is the linear over-density, and  $\nabla^2 \phi^{(2)} = \delta^{(2)}$ .

It is important to realise that these terms are not independent of each other. The second order nonlinear term  $\delta^{(2)}$  is fully determined by the linear over-density field  $\delta^{(1)}$  through the following quadratic expression

$$\delta^{(2)} \equiv \mu^{(2)}(\phi^{(1)}) = \sum_{i>j} \left( \phi_{,ii}^{(1)} \phi_{,jj}^{(1)} - [\phi_{,ij}^{(1)}]^2 \right), \quad (2)$$

where we use the following notation  $\phi_{,ij} \equiv \partial^2 \phi / \partial q_i \partial q_j$ , and the indices  $i, j$  run over the three Cartesian coordinates.

We note that going to third order in Lagrangian perturbation theory provides modest improvements (see Buchert & Ehlers 1993; Melott et al. 1995; Catelan 1995; Bouchet et al. 1995; Bernardeau et al. 2002).

Under this framework, and expanding the inverse of the Jacobian of the transformation between Lagrangian and Eulerian coordinates (we follow here the approach by Gramann 1993b, extending the Zeldovich approximation to 2LPT), the final density field can be written in terms of the linear and nonlinear fields (Kitaura & Angulo 2011);

$$\begin{aligned} \delta_M &= -\nabla \cdot \mathbf{\Psi} + \mu^{(2)}(\Theta) + \mu^{(3)}(\Theta) \\ &= D_1 \delta^{(1)} - D_2 \mu^{(2)}(\phi^{(1)}) + \mu^{(2)}(\Theta) + \mu^{(3)}(\Theta) \\ &= \delta^L + \delta^{\text{NL}} \end{aligned} \quad (3)$$

with  $\mathbf{\Psi} = -\nabla \Theta$ ,  $\delta^L = D_1 \delta^{(1)}$  and  $\delta^{\text{NL}} = -D_2 \mu^{(2)}(\phi^{(1)}) + \mu^{(2)}(\Theta) + \mu^{(3)}(\Theta)$ . Note that the third order term in the Jacobian expansion is given by:

$$\mu^{(3)}(\Theta) = \det(\Theta_{,ij}). \quad (4)$$

Similarly, the particle co-moving velocities  $\mathbf{v}$  are given to second order by:

$$\mathbf{v} = -D_1 f_1 H \nabla \phi^{(1)} + D_2 f_2 H \nabla \phi^{(2)}, \quad (5)$$

where  $f_i = d \ln D_i / d \ln a$ ,  $H$  is the Hubble constant and  $a$  is the scale factor. For flat models with a non-zero cosmological constant, the following relations apply  $f_1 \approx \Omega^{5/9}$ ,  $f_2 \approx 2\Omega^{6/11}$  (see Bouchet et al. 1995), where  $\Omega(z)$  is the matter density at a redshift  $z$ .

Assuming that any primordial vorticity has no growing modes associated (the first vorticity terms appear in 3rd order Lagrangian perturbation theory, see appendix), implies that the velocity field today is fully characterised by its divergence ( $\nabla \cdot \mathbf{v}$ ), or, for convenience, by the scaled velocity divergence, defined as:

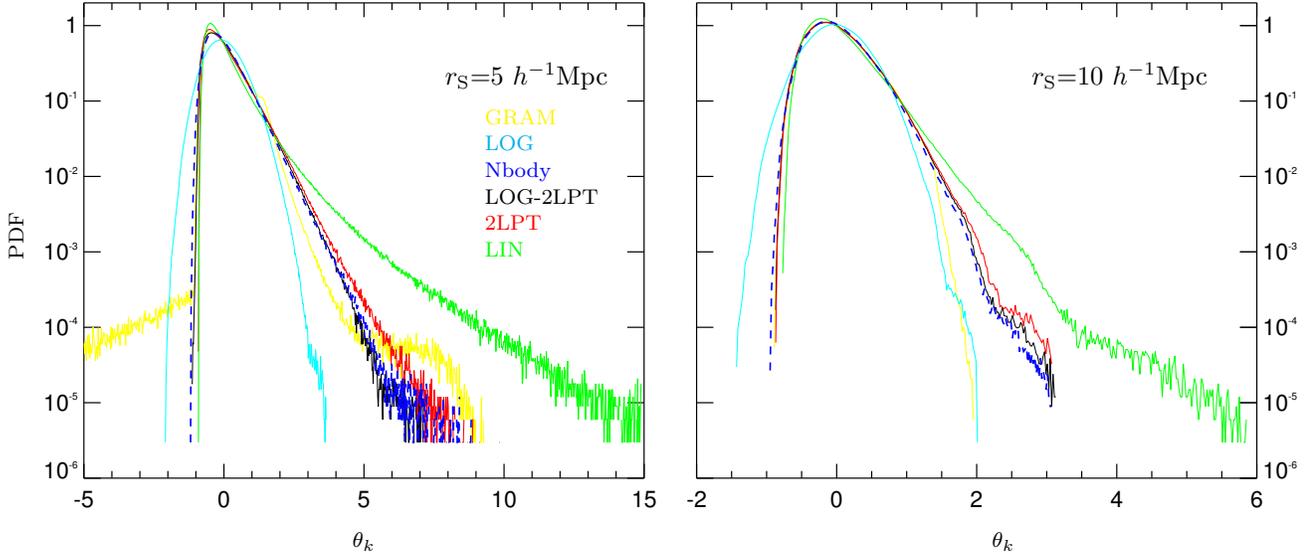
$$\theta \equiv -f(\Omega, \Lambda, z)^{-1} \nabla \cdot \mathbf{v}, \quad (6)$$

with  $f(\Omega, \Lambda, z) \equiv \dot{D}_1 / D_1 = f_1(\Omega, \Lambda, z) H(z)$ .

Combining Eqs. 3 and 5 truncated to quadratic terms in  $\Phi^{(1)}$ :  $\delta^{\text{NL}} = (D_1^2 - D_2) \delta^{(2)}$ , one gets

$$\theta = \delta_M - \left[ D_1^2 + \left( \frac{f_2}{f_1} - 1 \right) D_2 \right] \delta^{(2)}. \quad (7)$$

This expression is very similar to the one found by Gramann (1993b) (see Tab. 1). Note, however, that using directly the evolved field as the source for the second order



**Figure 1.** Statistics of the scaled velocity divergence ( $\theta_k$ : true from the simulation  $\theta^{\text{Nbody}}$  and different reconstructed ones  $\theta^{\text{rec}}$  as explained below) with different smoothing scales: Left:  $r_S = 5 h^{-1}\text{Mpc}$ , Right:  $10 h^{-1}\text{Mpc}$ . Curves from left to right in the order of appearance: yellow: approximate 2LPT solution from the nonlinear density field (GRAM) (see G93 in table 1), cyan: logarithm of the density field (LOG), blue dashed: true scaled velocity divergence at  $z = 0$  the Millennium Run (Nbody), black: 2LPT estimate from the logarithm of the density field (LOG-2LPT) (Eq. 8), red: 2LPT estimate from the iterative solution (2LPT) (see §2.2), green: linear theory (LIN) (density field).

term  $\delta^{(2)}$  is a good approximation for the true velocity field only on very large scales (where  $\delta_M$  is close to unity), as we will see in §3 and Fig. 1, breaking down on scales of even  $10 h^{-1}\text{Mpc}$  for both estimations of the nonlinear field and of the velocity divergence.

In this paper we follow a different approach, and solve iteratively the following equation;

$$\theta = D_1 \delta^{(1)} - D_2 \frac{f_2}{f_1} \delta^{(2)}. \quad (8)$$

which results from taking the divergence of Eq. 5 (for a similar approach see Monaco & Efstathiou 1999). For this, we rely on an estimation of the linear component of the present-day density field, which in turn can be calculated by solving iteratively Eq. (3). Note that we expand the inverse of the Jacobian and avoid in this way problems caused by low density regions as reported in Monaco & Efstathiou (1999) (see also Kitaura & Angulo 2011). In practice, a good approximation for such term is simply given by the logarithm of the density field;  $\delta^L = D_1 \delta^{(1)} = \ln(1 + \delta_M) - \mu$ , with  $\mu = \langle \ln(1 + \delta_M) \rangle$ , as shown by (Neyrinck et al. 2009; Kitaura & Angulo 2011). Note that this expression is essentially the lognormal approximation for the matter field (Coles & Jones 1991). This local transformation has the advantage of being computationally inexpensive.

In summary, what our approach does is to find a linear field which when plugged into 2LPT expressions produces the observed matter field (or third order, see appendix). If gravity worked only at a second order level, then this linear field would be identical to the actual linear field that originated  $\delta_M$ , but naturally in reality it is just an approximation. Then, our estimation for the velocity field is that predicted by this linear density field.

## 2.2 Method

The method to estimate the peculiar velocity field from the nonlinear density field is straightforward and fast to compute. We now outline the steps to be followed for a practical implementation. Here we have assumed that there is an unbiased and complete estimation of the matter field  $\delta_M$ . The extra layer of complication introduced by shot noise, a survey mask, biasing and redshift space distortions is out of the scope of this paper and will be addressed in a future publication.

### (i) Linear density field

One starts computing an estimate of the linear component of the density field. Here we propose two alternatives:

#### (a) Lognormal model

$$\delta_{\text{LOG}}^L = D_1 \delta^{(1)} = \ln(1 + \delta_M) - \mu \quad (9)$$

#### (b) Gaussianisation with LPT (Kitaura & Angulo 2011)

$$\delta_{\text{LPT}}^L = D_1 \delta^{(1)} = \delta_M - \delta^{\text{NL}} \quad (10)$$

### (ii) Linear potential

Then the Poisson equation is solved to obtain the linear potential:

$$\phi^{(1)} = \nabla^{-2} \delta^{(1)} \quad (11)$$

### (iii) Nonlinear second order density field

The tidal field contribution to second order is calculated from the linear potential:

$$\delta^{(2)}[\phi^{(1)}] = \sum_{i>j} (\phi_{,ii}^{(1)}\phi_{,jj}^{(1)} - (\phi_{,ij}^{(1)})^2), \quad (12)$$

#### (iv) Scaled velocity divergence

One can now construct the 2nd order divergence of the velocity field taking both the linear and the second order contribution:

$$\theta = D_1\delta^{(1)} - D_2\frac{f_2}{f_1}\delta^{(2)}, \quad (13)$$

#### (v) Peculiar velocity field

Finally, one obtains the 3D velocity field:

$$\mathbf{v} = -f(\Omega, \Lambda, z)\nabla\nabla^{-2}\theta. \quad (14)$$

Please note that the equations in steps (ib), (ii), (iii) and (v) can be solved with FFTs. Details of the Gaussianisation step with LPT can be found in Kitaura & Angulo (2011).

### 3 TESTING THE METHOD WITH *N*-BODY SIMULATIONS

In this section we test the performance of the method outlined above by comparing the velocity field directly extracted from a *N*-body simulation with our estimation based on the respective nonlinear density field.

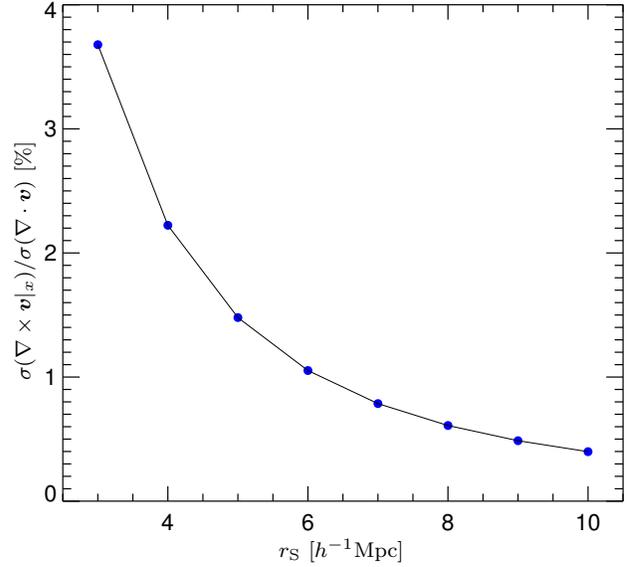
With this purpose, we employ the Millennium simulation which tracks the nonlinear evolution of more than 10 billion particles in a box of comoving side-length  $500 h^{-1}\text{Mpc}$  (Springel et al. 2005). In particular, we use the output corresponding to redshift 0, which is where the most nonlinear structures are present.

At such output we first compute the velocity and density field by mapping the information of dark matter particles onto a grid of  $256^3$  cells using the nearest-grid-point (NGP) assignment scheme, which gives a spatial resolution of about  $2 h^{-1}\text{Mpc}$ . We then apply the algorithm presented in §2.2 to infer the velocity divergence on a grid of identical dimensions. In the next two subsections we present our results and explore the accuracy when applied on two scales;  $10 h^{-1}\text{Mpc}$  on which linear theory is usually assumed to perform well, and  $5 h^{-1}\text{Mpc}$  which enters into the mildly nonlinear regime.

#### 3.1 The velocity divergence

The first ingredient in our algorithm is the linear component of the density field. We stress that this field is not the “initial conditions” of the universe, since structures have moved from its Lagrangian position to the Eulerian ones at  $z=0$ . Contrarily to the linear term, the nonlinear component is highly skewed.

Fig. 1 compares the PDF of the velocity divergence  $\theta$  as given by different estimators, with the one directly measured from the simulation (blue dashed line). The left panel shows the fields smoothed with a  $5 h^{-1}\text{Mpc}$  whereas the right panels do so with a larger smoothing of  $10 h^{-1}\text{Mpc}$ . In both cases, the predictions of linear perturbation theory (i.e.  $\theta = \delta$ ) displays the worst performance of all – it overestimates systematically the number of volume elements with



**Figure 5.** Standard deviation of the *x*-component of the curl  $\sigma(\nabla \times \mathbf{v}|_x)$  divided by the standard deviation of the velocity divergence  $\sigma(\nabla \cdot \mathbf{v})$  in % as a function of the scale  $r_S$ . The mean of both the curl and the divergence of the peculiar field are very close to zero.

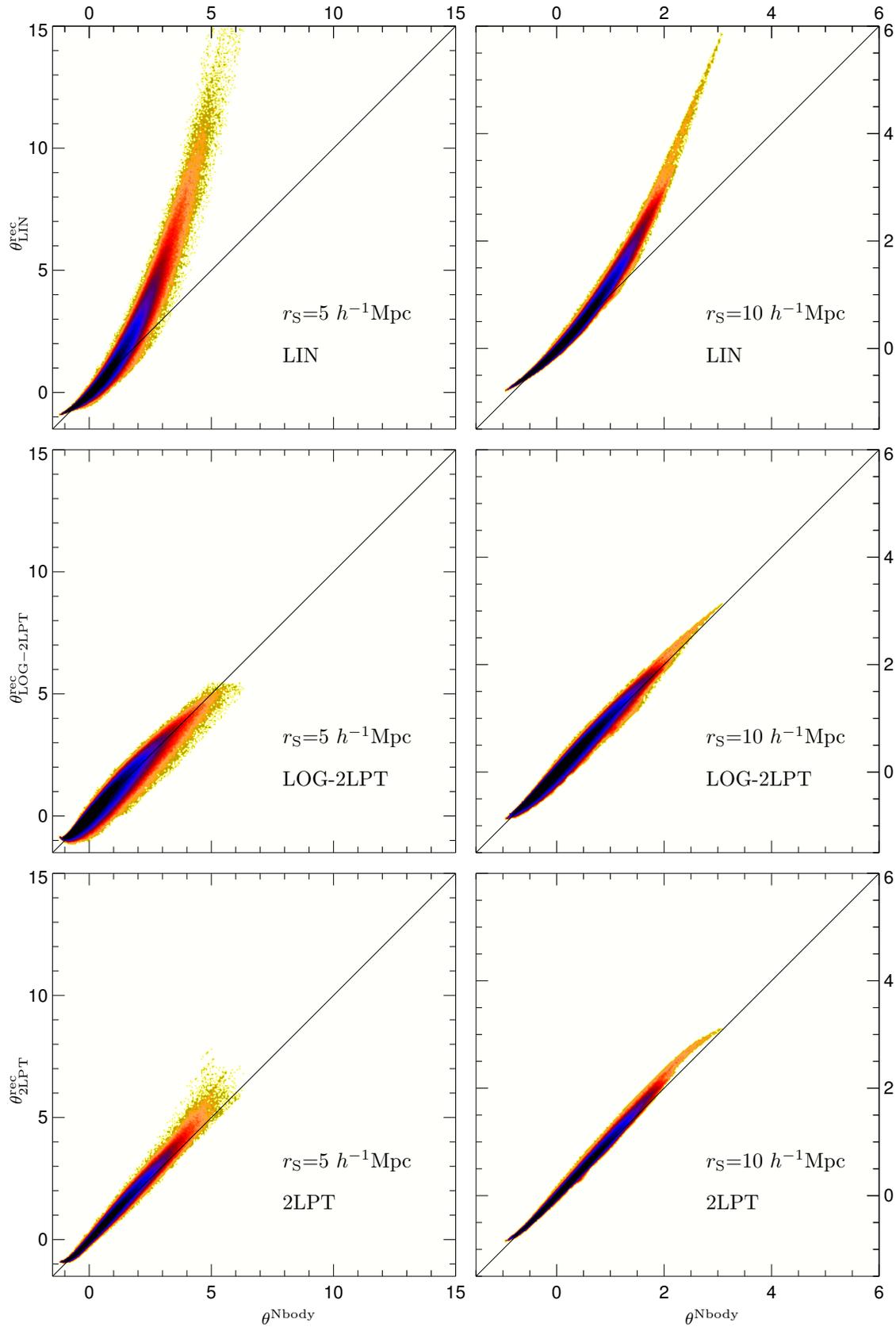
large value of  $\theta$  and underestimates the ones with low values of  $\theta$  (green curve). This is a consequence of linear theory breaking down in even mildly under- or overdense regions.

Using linear theory together with the linear component of the nonlinear density field given by the logarithm of the field ( $\ln(1 + \delta_M) - \mu$ , see cyan curve in Fig. 1) is still a poor description of the PDF, suggesting the need of higher order corrections.

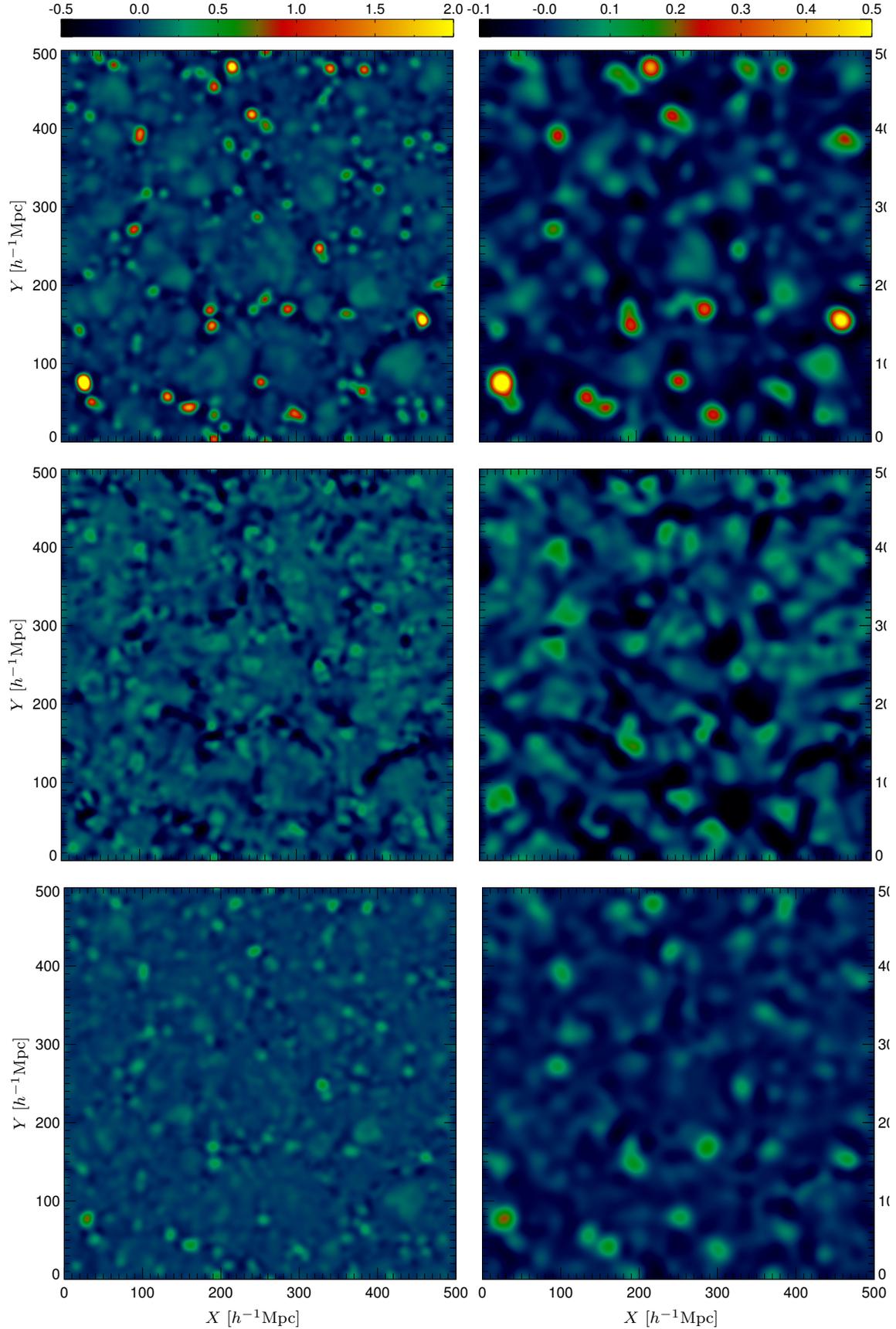
The first second order estimation we consider is that given by Eq. (7) which is closely related to the one proposed by Gramann and uses the nonlinear field as a proxy for the linear density field (yellow curve). In the regime where this approximation is valid ( $\delta \sim 0$ ) this approach performs remarkably well. However, there is a clear and rapid degradation for volumes with larger deviations of homogeneity. For instance, this solution yields to values even of  $\theta = -140$  at scales of  $5 h^{-1}\text{Mpc}$ ! This behaviour is due to a complete misestimation of the nonlinear term  $\delta^{(2)}$  and therefore of the nonlinear corrections to the velocity field.

Finally, black and red lines indicate the results of the method proposed in this paper with Lagrangian perturbation theory, based upon an estimate of the linear component of the density field using the lognormal model (LOG-2LPT: black) or the iterative Lagrangian linearisation (2LPT: red). On both panels, the predicted PDF follows very closely that measured in the Millennium simulation, even on the extreme tails (especially with LOG-2LPT). The only appreciable difference with the lognormal model is a slight overestimation for low values of  $\theta$  (static regions), we note however, that this could be potentially improved by higher order expressions. Indeed, 3rd order PT appears to perform better than 2LPT for underdense regions (see appendix). In spite of this, on both  $5$  and  $10 h^{-1}\text{Mpc}$ <sup>1</sup> our method is clearly superior to

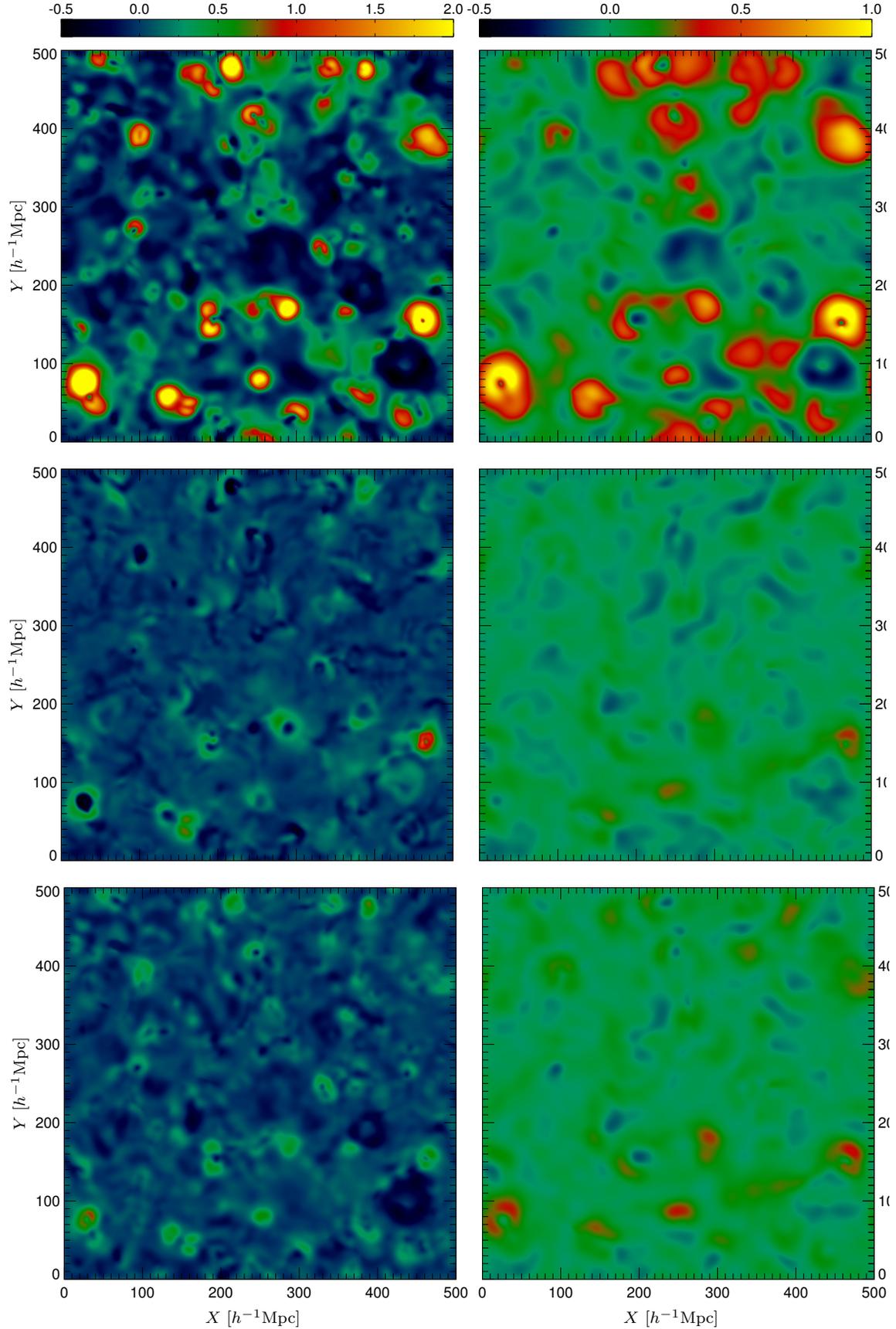
<sup>1</sup> We have also checked that this is true on  $3, 4, 6, 7$  and  $8 h^{-1}\text{Mpc}$



**Figure 2.** Cell-to-cell correlation between the *true* and the reconstructed scaled velocity divergence,  $\theta^{\text{Nbody}}$  and  $\theta^{\text{rec}}$  respectively. Left panels on scales of  $5 h^{-1} \text{Mpc}$  and right panels on scales of  $10 h^{-1} \text{Mpc}$ . Upper panels: LIN, middle panels: LOG-2LPT, lower panels: 2LPT.



**Figure 3.** Difference between the true and the estimated scaled velocity divergence. On scales of  $r_S = 5 h^{-1} \text{Mpc}$  (left panels) and  $10 h^{-1} \text{Mpc}$  (right panels) with different colour ranges. Upper panels: LIN, middle panels: LOG-2LPT, lower panels: 2LPT.



**Figure 4.** Difference between the true and the estimated speeds (velocity magnitude). On scales of  $r_s = 5 h^{-1} \text{Mpc}$  (left panels) and  $10 h^{-1} \text{Mpc}$  (right panels) with different colour ranges. Upper panels: LIN, middle panels: LOG-2LPT, lower panels: 2LPT. The color-code is in units of  $10^{-2} \text{s}^{-1} \text{km}$ .

any of the other methods we investigated here, as far as the PDF is concerned and for any value of  $\theta$ . The iterative 2LPT solution yields a moderate overestimation for high values of  $\theta$  due to the laminar flow approximation used in Lagrangian perturbation theory which does not fully capture nonlinear structure formation. Nevertheless, the PDF of  $\theta$  using this solution is clearly superior to the linear approximation.

We now continue with a more detailed testing of our method. In Fig. 2 we plot the predicted velocity divergence,  $\theta_{\text{LOG-2LPT}}^{\text{rec}}$  and  $\theta_{\text{2LPT}}^{\text{rec}}$  (based on  $\delta_{\text{LOG}}^{\text{L}}$  and  $\delta_{\text{LPT}}^{\text{L}}$ , respectively), for each of the  $256^3$  cells in our mesh, as a function of the value measured in the simulation. As in the previous plot, we display results on two different scales;  $5 h^{-1}\text{Mpc}$  on the top panel and  $10 h^{-1}\text{Mpc}$  on the bottom panel. For comparison we also provide the results using linear perturbation theory  $\theta_{\text{LIN}}^{\text{rec}} = \delta$ .

The values of  $\theta$  are remarkably well predicted by Lagrangian perturbation theory. In fact, measurements lie around the 1:1 line in the LOG-2LPT case, implying that there are no appreciable biases in our estimation over all the range probed by the Millennium simulation (with the exception of the low values of  $\theta$ , for an improvement on this see appendix). In contrast, the linear theory prediction presents overestimations of up to a factor of 3 for the  $5 h^{-1}\text{Mpc}$  smoothing and of 2 for  $10 h^{-1}\text{Mpc}$ .

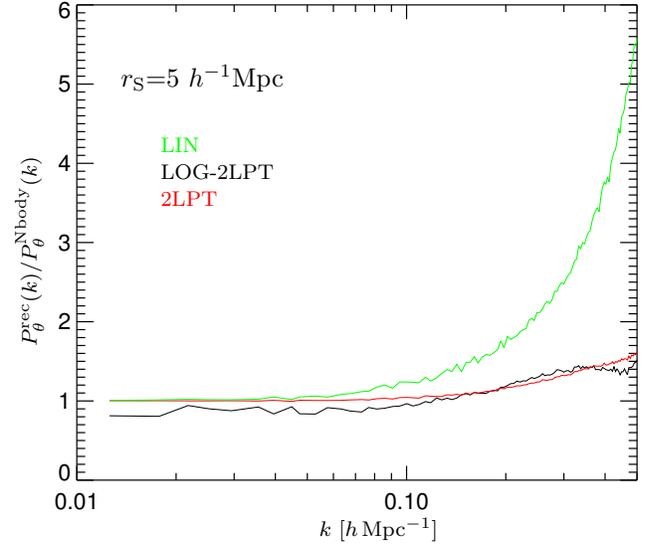
The iterative solution (2LPT) produces smaller dispersions but also a slight overestimation of  $\theta$  for high values, as we already mentioned before.

The distribution of differences in our method is well approximated by a Gaussian function, whereas in linear theory there are significant extended tails, we will return to this in more detail in §3.2. Overall, this plot suggest that our method not only performs adequately on a statistical basis, but also on predicting the actual average value of  $\theta$  in a given volume element.

Although not displayed by the figure, our method also performs better than the other methods shown in Fig. 1. In particular, the classic application of 2LPT recovers  $\theta$  quite well for the range  $-0.5 < \theta < 2$ , as shown in Gramann (1993b). However, outside the range it displays an erratic behaviour as it could have been anticipated from Fig. 1.

A complementary visual assessment of our method is provided in Fig. 4. In these images we display the relative difference  $\theta^{\text{rec}} - \theta^{\text{Nbody}}$  ( $\theta_{\text{LIN}}^{\text{rec}} - \theta^{\text{Nbody}}$  in the upper panels,  $\theta_{\text{2LPT}}^{\text{rec}} - \theta^{\text{Nbody}}$  in the central panels and  $\theta_{\text{LOG-2LPT}}^{\text{rec}} - \theta^{\text{Nbody}}$  in the lower ones) projected on a slice  $2 h^{-1}\text{Mpc}$  thick. As previously, we explore  $5$  and  $10 h^{-1}\text{Mpc}$  and also display the linear theory predictions for comparison. We see that this difference field is not uniformly distributed across the simulation but there are well defined regions in which the predictions are very accurate and others where the prediction is somewhat worse. Not surprisingly the latter coincide with high density regions. Nevertheless, and consistently with previous plots, we see that areas where linear theory fails dramatically are much better handled in our approach.

As a final crucial check, we compute the power-spectra of the scaled velocity divergence according to the  $N$ -body simulation and our reconstruction with both the LPT and the lognormal linearisation approaches. This is shown in Fig. 6. Linear theory, as expected, increasingly overestimates the power towards small scales, whereas the 2LPT solutions perform remarkably well in a wide range going down to scales



**Figure 6.** Ratio of the true  $P_{\theta}^{\text{Nbody}}(k)$  and reconstructed  $P_{\theta}^{\text{rec}}(k)$  power-spectra of the scaled velocity divergence  $\theta$  for green: linear theory (LIN), black: 2LPT estimate from the logarithm of the density field (LOG-2LPT) (Eq. 8), red: 2LPT estimate from the iterative solution (2LPT) (see §2.2)

of  $k \simeq 0.5 h/\text{Mpc}$ . We can also see that there is a systematic deviation originated by the lognormal transformation. The LPT estimate of the linear component corrects these and the results are extremely close to the actual power-spectrum over most of the  $k$ -range shown.

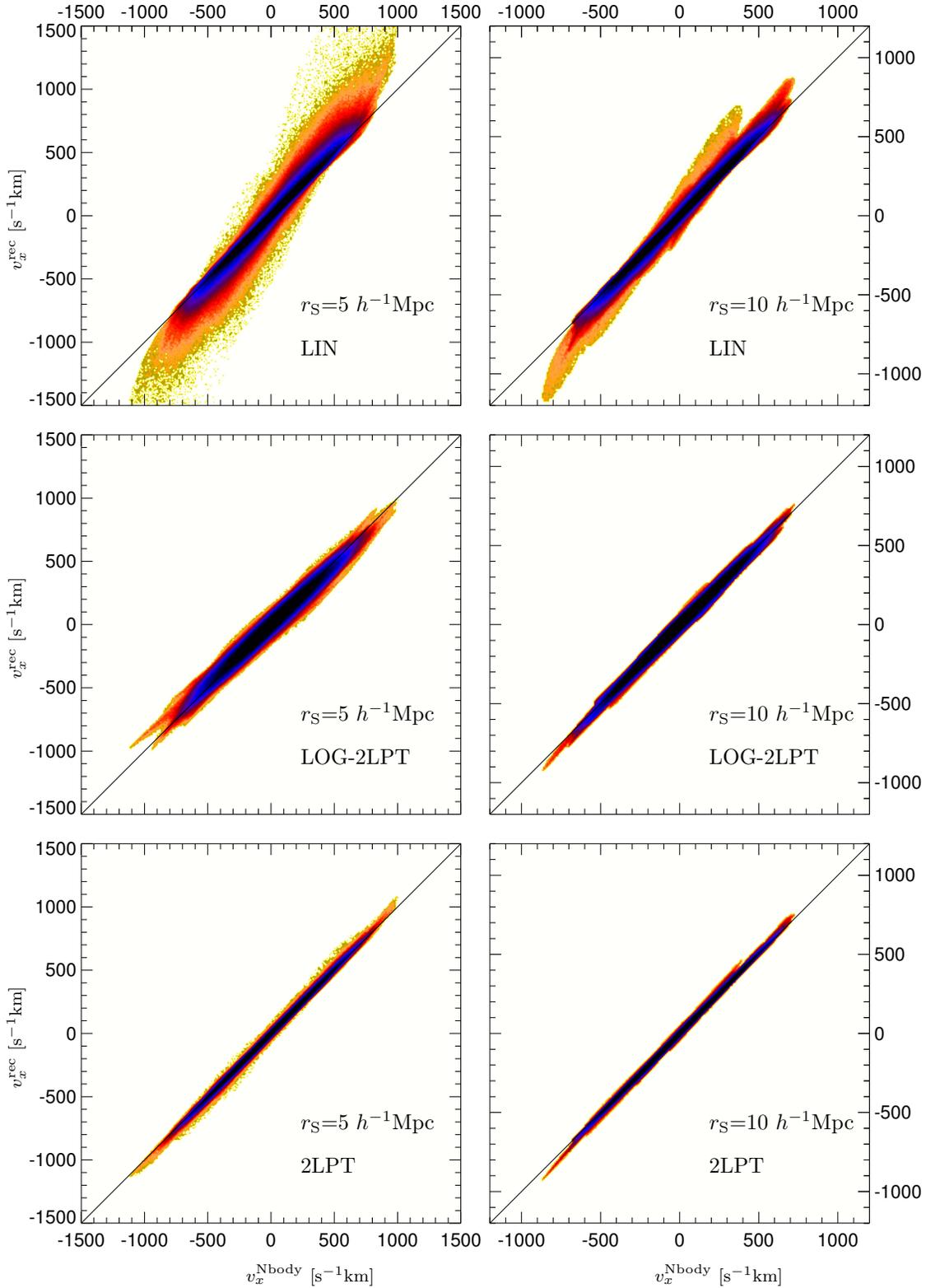
### 3.2 The full 3D peculiar velocity field

We have calculated each component of the peculiar velocity field ( $v_x, v_y, v_z$ ) from the inferred velocity divergence assuming  $\nabla \times \theta = 0$ . In this case, the Fourier transform of the velocity along a direction is given by  $\hat{v} = \mathbf{k} \cdot \hat{\theta}$  with  $\mathbf{k}$  being the  $k$ -vector. The approximation that the velocity field is irrotational is actually a good one for the scales and redshifts we consider here. In fact the curl of the velocity fields is on average more than 25 times smaller than its divergence in the  $z=0$  output of the Millennium simulation on scales of  $3 h^{-1}\text{Mpc}$  and even smaller on larger scales (see Fig. 5).

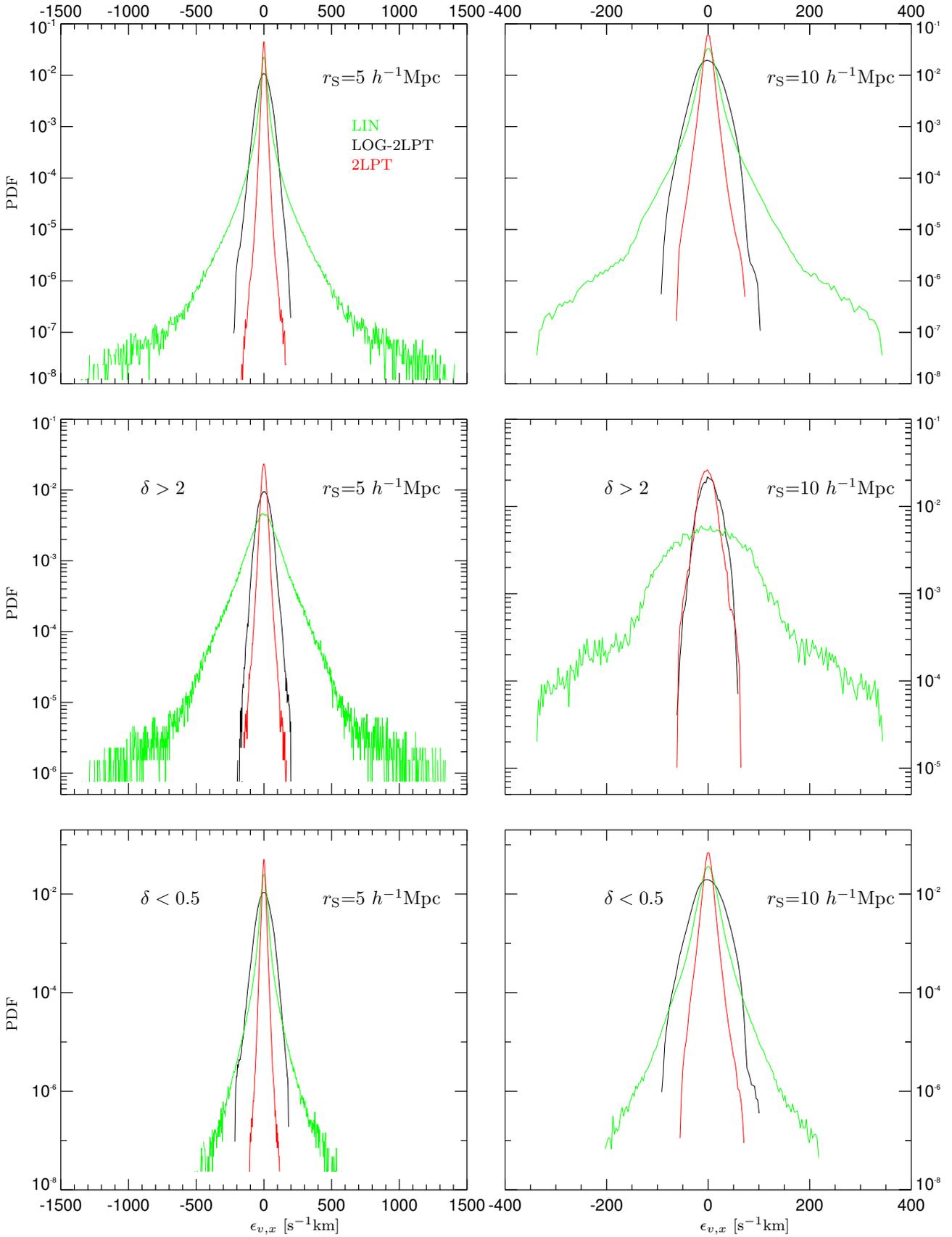
In Fig. 7 we compare the velocity field along the  $x$ -axis in the simulation and in our method (the two other axis show essentially the same features). The contours show the variation of number of cells (the darker colours represent higher numbers). By comparing both columns, it is clear that linear theory presents biases for large speeds, which are removed by our 2LPT approach, on both  $5$  and  $10 h^{-1}\text{Mpc}$ . We can see that the distribution gets sharper with 2LPT showing a clear decrease in the number of outliers.

We quantified the uncertainty in the recovered velocity in Fig. 8. The  $x$ -axis corresponds to the errors in the reconstruction, defined as  $\epsilon_{v,x} \equiv v_x^{\text{rec}} - v_x$ . As in the previous plot, we show only results for the  $x$ -axis since the other three Cartesian coordinates provide consistent results.

We find that the errors in our method are closely Gaussian distributed – the skewness and kurtosis are dramatically smaller than for linear theory. This property is very



**Figure 7.** Cell-to-cell correlation between the true velocity  $v_x^{\text{Nbody}}$  and the reconstructed one  $v_x^{\text{rec}}$  for the  $x$ -component. Left panels on scales of  $5 h^{-1} \text{Mpc}$  and right panels on scales of  $10 h^{-1} \text{Mpc}$ . Upper panels: LIN, middle panels: LOG-2LPT, lower panels: 2LPT (see §2.2).



**Figure 8.** Statistics of the errors in the reconstructed velocity for the  $x$ -component (green: LIN; black: LOG-LPT, red: 2LPT) with different smoothing scales  $r_S = 5$  (left panels) and  $10$  (right panels)  $h^{-1}$ Mpc. Upper panels in the entire  $\delta$  range, middle panels: for the range  $\delta > 2$ , lower panels: for the range  $\delta < 0.5$ .

important when applying the method to real data, since the observational uncertainties can be added to the model uncertainties within a Gaussian likelihood without the need of introducing complex error models. The typical errors are also significantly reduced with lower standard deviations. The errors in the reconstruction using linear theory have very long tails. Such outliers are not present in our method (see Fig. 8).

As we have discussed along this paper, the larger improvement of our method relates to velocities in the high density regions. For regions with  $\delta > 1$  and  $\delta > 2$  we find significant differences between linear and 2LPT. At  $10 h^{-1}\text{Mpc}$  and cells with  $\delta > 2$  the 1 sigma errors with linear theory are about 70 km/s and are reduced with 2LPT to  $\sim 13$  km/s, i.e. by about 81%. The corresponding 2 sigma confidence intervals are about 160 and 28 km/s, respectively, i.e. an error reduction of about 83%. One can see that the 2 sigma confidence intervals are about double as large as the 1 sigma confidence intervals for the 2LPT estimation. However, this is not the case for the linear estimates as these are not Gaussian distributed.

#### 4 CONCLUSIONS

In this paper we presented an improved method to reconstruct the peculiar velocity field from the density field. It builds from 2nd order Lagrangian perturbation theory and the realisation that the density field can be split into a linear plus a nonlinear term. The latter is the key concept, which enables the application of Lagrangian perturbation theory to an evolved field in Eulerian coordinates. This in turn, creates an approach that is nonlinear and nonlocal by including the information of the gravitational tidal field tensor.

We have shown that this approach is efficient and accurate over the dynamical range probed by the Millennium simulation. When the reconstruction is carried out on  $5 h^{-1}\text{Mpc}$ , each component of the velocity field can be recovered to an accuracy of about 10 kms/sec. On  $10 h^{-1}\text{Mpc}$  this figure is reduced to about 7 kms/sec. If we consider high density regions, the typical uncertainty is of 13, which improves dramatic over linear perturbation theory; 81%. An accurate description of the velocity divergence, both in terms of its PDF and on a point-by-point basis, is also achieved. In addition, we have shown that the statistics of the scaled velocity divergence is extremely well recovered, being almost indistinguishable from the true one, whereas linear theory dramatically over-estimates the velocity divergence. This especially on the mildly nonlinear scale of  $5 h^{-1}\text{Mpc}$  where our method shows more clearly its advantages. Finally, highlight that our method does not require calibration nor free parameters to derive the velocity divergence.

There are a number of simplifications and assumptions that we have adopted throughout our paper. First, our analyses focused on the peculiar velocity averaged over a volume. Another aspect is that we have neglected the rotational component of the velocity field. This however, does not seem to be relevant at the scales we have investigated (larger than  $5 h^{-1}\text{Mpc}$ ).

We have also performed our comparison assuming that there exist an unbiased estimator of the underlying real-space density field. But, of course, densities measured in

a survey are in redshift-space. The transformation can be done, but not trivially. One alternative is to apply the Gibbs-sampling method suggested by Kitaura & Enßlin (2008) to correct for redshift-space distortions. In this, the Gaussian distribution of errors in our method is highly convenient, since it permits to model the uncertainties in the peculiar velocity field including observational errors in a straightforward way.

We would like to note that our comparison and the uncertainties quoted here, were based on the present-day output of the Millennium Run. Such simulation was carried out with a value for  $\sigma_8$  about 10% higher than the current best estimations (see Angulo & White 2010, for a method to correct for this). Therefore, our uncertainties should be regarded as an upper limit of the reachable uncertainties for a hypothetical spectroscopic survey, which should contain a less nonlinear underlying dark matter distribution.

We anticipate that our method can be used in many different applications; ranging from bias studies combining galaxy redshift surveys with measured peculiar velocities, Baryon Acoustic Oscillation reconstructions, determination of the growth factor, to estimates of the initial conditions of the Universe.

#### ACKNOWLEDGEMENTS

FSK thanks Simon White for encouraging conversations. The authors appreciate Francisco Prada's comments to the manuscript. We are indebted to the German Astrophysical Virtual Observatory (GAVO) and the MPA facilities for providing us the Millennium Run simulation data. The work of REA was supported by Advanced Grant 246797 "GALFORMOD" from the European Research Council. YH has been partially supported by the ISF (13/08).

#### REFERENCES

- Angulo R. E., Baugh C. M., Frenk C. S., Lacey C. G., 2008, *MNRAS*, 383, 755
- Angulo R. E., White S. D. M., 2010, *MNRAS*, 405, 143
- Bernardeau F., 1992, *ApJ*, 390, L61
- Bernardeau F., Chodorowski M. J., Lokas E. L., Stompor R., Kudlicki A., 1999, *MNRAS*, 309, 543
- Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, *Phys.Rep.*, 367, 1
- Bilicki M., Chodorowski M. J., 2008, *MNRAS*, 391, 1796
- Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, *Astr.Astroph.*, 296, 575
- Branchini E., Eldar A., Nusser A., 2002, *MNRAS*, 335, 53
- Branchini E., Freudling W., Da Costa L. N., Frenk C. S., Giovanelli R., Haynes M. P., Salzer J. J., Wegner G., Zehavi I., 2001, *MNRAS*, 326, 1191
- Buchert T., 1994, *MNRAS*, 267, 811
- Buchert T., Ehlers J., 1993, *MNRAS*, 264, 375
- Buchert T., Melott A. L., Weiss A. G., 1994, *Astr.Astroph.*, 288, 349
- Catelan P., 1995, *MNRAS*, 276, 115
- Chodorowski M. J., Lokas E. L., Pollo A., Nusser A., 1998, *MNRAS*, 300, 1027
- Coles P., Jones B., 1991, *MNRAS*, 248, 1
- Courtois H. M., Hoffman Y., Tully R. B., Gottlober S., 2011, *ArXiv e-prints*
- Croft R. A. C., Gaztanaga E., 1997, *MNRAS*, 285, 793

Davis M., Nusser A., Willick J. A., 1996, ApJ, 473, 22  
Eisenstein D. J., Seo H.-J., Sirko E., Spergel D. N., 2007, ApJ, 664, 675  
Falck B. L., Neyrinck M. C., Aragon-Calvo M. A., Lavaux G., Szalay A. S., 2011, ArXiv e-prints  
Fisher K. B., Lahav O., Hoffman Y., Lynden-Bell D., Zaroubi S., 1995, MNRAS, 272, 885  
Frisch U., Matarrese S., Mohayaee R., Sobolevski A., 2002, Nature, 417, 260  
Gottloeber S., Hoffman Y., Yepes G., 2010, ArXiv e-prints  
Gramann M., 1993a, ApJ, 405, 449  
Gramann M., 1993b, ApJ, 405, L47  
Guzzo L., Pierleoni M., Meneux B., Branchini E., Le Fèvre O., Marinoni C., Garilli B., et al 2008, Nature, 451, 541  
Hivon E., Bouchet F. R., Colombi S., Juszkiewicz R., 1995, Astr.Astroph., 298, 643  
Jasche J., Kitaura F. S., 2009, ArXiv e-prints  
Jenkins A., 2010, MNRAS, 403, 1859  
Kashlinsky A., Atrio-Barandela F., Ebeling H., 2011, ApJ, 732, 1  
Kashlinsky A., Atrio-Barandela F., Kocevski D., Ebeling H., 2009, ApJ, 691, 1479  
Kitaura F., Jasche J., Metcalf R. B., 2010, MNRAS, 403, 589  
Kitaura F.-S., Angulo R., , 2011, in preparation  
Kitaura F. S., Enßlin T. A., 2008, MNRAS, 389, 497  
Kitaura F.-S., Gallerani S., Ferrara A., 2010, ArXiv e-prints  
Klypin A., Hoffman Y., Kravtsov A. V., Gottlöber S., 2003, ApJ, 596, 19  
Kudlicki A., Chodorowski M., Plewa T., Różyczka M., 2000, MNRAS, 316, 464  
Lavaux G., Mohayaee R., Colombi S., Tully R. B., Bernardeau F., Silk J., 2008, MNRAS, 383, 1292  
Melott A. L., Buchert T., Weib A. G., 1995, Astr.Astroph., 294, 345  
Mohayaee R., Tully R. B., 2005, ApJ, 635, L113  
Monaco P., Efstathiou G., 1999, MNRAS, 308, 763  
Moutarde F., Alimi J.-M., Bouchet F. R., Pellat R., Ramani A., 1991, ApJ, 382, 377  
Narayanan V. K., Weinberg D. H., 1998, ApJ, 508, 440  
Neyrinck M. C., Szapudi I., Szalay A. S., 2009, ApJ, 698, L90  
Nusser A., Branchini E., 2000, MNRAS, 313, 587  
Nusser A., Branchini E., Davis M., 2011, ArXiv e-prints  
Nusser A., Dekel A., 1992, ApJ, 391, 443  
Nusser A., Dekel A., Bertschinger E., Blumenthal G. R., 1991, ApJ, 379, 6  
Peebles P. J. E., 1989, ApJ, 344, L53  
Scoccimarro R., Sheth R. K., 2002, MNRAS, 329, 629  
Springel V., White S. D. M., Jenkins A., Frenk C. S., Yoshida N., Gao L., Navarro J., Thacker R., Croton D., Helly J., Peacock J. A., Cole S., Thomas P., Couchman H., Evrard A., Colberg J., Pearce F., 2005, Nature, 435, 629  
Watkins R., Feldman H. A., Hudson M. J., 2009, MNRAS, 392, 743  
Willick J. A., Strauss M. A., 1998, ApJ, 507, 64  
Willick J. A., Strauss M. A., Dekel A., Kolatt T., 1997, ApJ, 486, 629  
Zaroubi S., Hoffman Y., Dekel A., 1999, ApJ, 520, 413  
Zel'dovich Y. B., 1970, Astr.Astroph., 5, 84

## APPENDIX A: THIRD ORDER LAGRANGIAN PERTURBATION THEORY (3LPT)

For the sake of completeness we investigate here third order LPT. Following Buchert (1994) and Catelan (1995) one can write the displacement field to third order as

$$\Psi = -D_1 \nabla \phi^{(1)} + D_2 \nabla \phi^{(2)} + D_{3a} \nabla \phi_a^{(3)} + D_{3b} \nabla \phi_b^{(3)} + D_{3c} \nabla \times \mathbf{A}^{(3)} \quad (\text{A1})$$

where  $\{D_{3a}, D_{3b}, D_{3c}\}$  are the 3rd order growth factors corresponding to the gradient of two scalar potentials  $(\phi_a^{(3)}, \phi_b^{(3)})$  and the curl of a vector potential  $(\mathbf{A}^{(3)})$ . Particular expressions for the irrotational 3rd order growth factors  $(D_{3a}, D_{3b})$  can be found in Bouchet et al. (1995), the growth factor of the rotational term  $(D_{3c})$  is given in Catelan (1995).

We assume that the fields are curl-free on scales  $\gtrsim 5 h^{-1} \text{Mpc}$  (see Bouchet et al. 1995, and the comparison between the velocity divergence and the curl in the Millennium Run in Fig. 5). We therefore consider only the scalar potential terms  $\phi_a^{(3)}$  and  $\phi_b^{(3)}$ .

The first term is cubic in the linear potential

$$\delta_a^{(3)} \equiv \mu^{(3)}(\phi^{(1)}) = \det \left( \phi_{,ij}^{(1)} \right), \quad (\text{A2})$$

and the second term is the interaction term between the first- and the second-order potentials:

$$\delta_b^{(3)} \equiv \mu^{(2)}(\phi^{(1)}, \phi^{(2)}) = \frac{1}{2} \sum_{i \neq j} \left( \phi_{,ii}^{(2)} \phi_{,jj}^{(1)} - \phi_{,ij}^{(2)} \phi_{,ji}^{(1)} \right), \quad (\text{A3})$$

(see Buchert 1994; Bouchet et al. 1995; Catelan 1995).

In order to ensure that the term  $\delta_b^{(3)}$  is curl-free one has to introduce some proper weights in the expression A3 (see Buchert 1994; Catelan 1995).

From the displacement field one can derive the expression for the velocity field

$$\mathbf{v} = -D_1 f_1 H \nabla \phi^{(1)} + D_2 f_2 H \nabla \phi^{(2)} + D_{3a} f_{3a} H \nabla \phi_a^{(3)} + D_{3b} f_{3b} H \nabla \phi_b^{(3)}, \quad (\text{A4})$$

with  $f_i = d \ln D_i / d \ln a$  (particular expressions for  $f_{3a}$  and  $f_{3b}$  can be found in Bouchet et al. 1995). As we can see from Eq. A4 one can construct all components from the linear potential  $\phi^{(1)}$ .

We consider here two models for the linear potential. The first model relies on the local lognormal estimate (see §2.2) from which the scaled velocity divergence can be calculated in the following way to 3rd order LPT

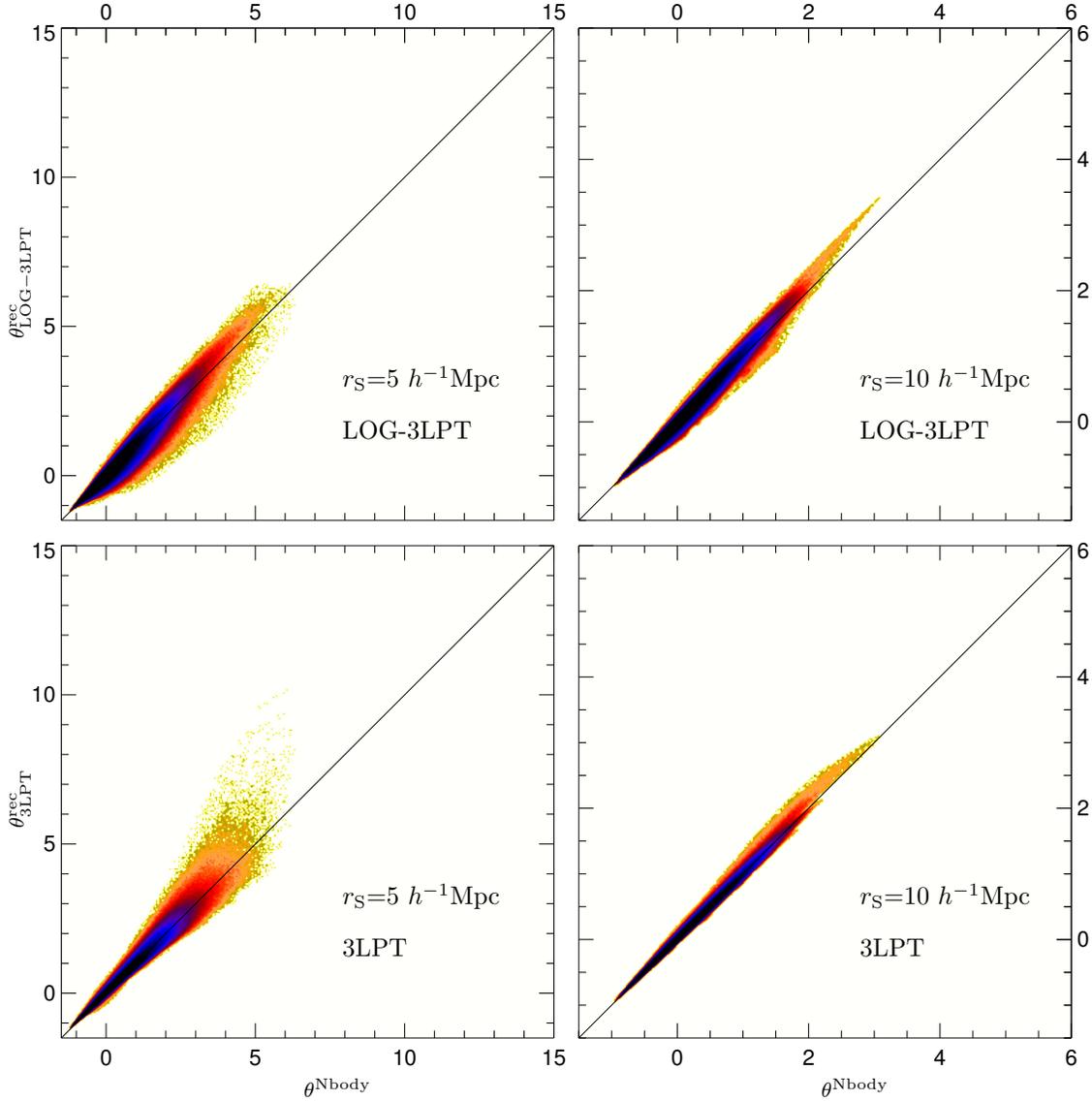
$$\theta = D_1 \delta^{(1)} - \frac{f_2}{f_1} D_2 \mu^{(2)}(\phi^{(1)}) - \frac{f_{3a}}{f_1} D_{3a} \mu^{(3)}(\phi^{(1)}) - \frac{f_{3b}}{f_1} D_{3b} \mu^{(2)}(\phi^{(1)}, \phi^{(2)}). \quad (\text{A5})$$

The second model yields a non-local estimate of  $\phi^{(1)}$  to 3rd order given by

$$\delta^L = \delta - \delta^{\text{NL}} \quad (\text{A6})$$

with  $\delta^{\text{NL}} = -D_2 \mu^{(2)}(\phi^{(1)}) - D_{3a} \mu^{(3)}(\phi^{(1)}) - D_{3b} \mu^{(2)}(\phi^{(1)}, \phi^{(2)}) + \mu^{(2)}(\Theta) + \mu^{(3)}(\Theta)$ . Note that the potential  $\Theta$  is also different and is determined by Eq. A1 recalling the relation to the displacement field  $\Psi = -\nabla \Theta$ .

Using the latter expression one can write the  $\theta$ - $\delta$  relation to 3rd order in LPT as



**Figure A1.** Cell-to-cell correlation between the *true* and the reconstructed scaled velocity divergence using 3LPT. Left panels on scales of  $5 h^{-1}\text{Mpc}$  and right panels on scales of  $10 h^{-1}\text{Mpc}$ . Upper panels: LOG-3LPT (based on the truncated 3LPT expansion computing Eq. A5 with the local model for the linear component), lower panels: 3LPT (based on the truncated 3LPT expansion computing Eq. A7 with the non-local model for the linear component).

$$\begin{aligned}
 \theta = & \delta - \left( \frac{f_2}{f_1} - 1 \right) D_2 \mu^{(2)}(\phi^{(1)}) \\
 & - \left( \frac{f_{3a}}{f_1} - 1 \right) D_{3a} \mu^{(3)}(\phi^{(1)}) - \left( \frac{f_{3b}}{f_1} - 1 \right) D_{3b} \mu^{(2)}(\phi^{(1)}, \phi^{(2)}) \\
 & - \mu^{(2)}(\Theta) - \mu^{(3)}(\Theta). \tag{A7}
 \end{aligned}$$

Note that Eq. A6 should be solved iteratively as we do in the 2LPT case (see §2.2 and Kitaura & Angulo 2011). Nevertheless to find a fast solution we plug-in the lognormal model for  $\phi^{(1)}$  into Eq. A5 yielding a non-local estimate of the linear field in one iteration. To minimize the deviation between LPT and the full nonlinear evolution we have additionally smoothed the density  $\delta$  in Eq. A7 with a Gaussian kernel of  $2 h^{-1}\text{Mpc}$  radius.

We do not find an improvement in the determination of

the velocity divergence with respect to 2LPT including both curl-free terms using any of the estimates of the linear component. This could be due to an inaccurate determination of the interaction term  $\phi_b^{(3)}$ , as uncertainties in the linear component  $\phi^{(1)}$  propagate more dramatically than in terms which depend only on the initial conditions ( $\phi^{(2)}, \phi_a^{(3)}$ ).

One may neglect the interaction term  $\delta_b^{(3)}$  and consider only the cubic term  $\delta_a^{(3)}$  to 3rd order. Such a truncated 3LPT model includes the *main body* of the perturbation sequence with the rest of the sequence being made up of interaction terms (Buchert 1994).

Our calculations show in this case better results than 2LPT for low values of the velocity divergence (see Fig. A1). However, for large values of  $\nabla \cdot \mathbf{v}$  we obtain larger dispersions which could be also due to numerical errors in the estimate of the linear density component  $\delta^L$ . The errors in the velocity

estimation are only moderately reduced with respect to the 2LPT case (see §3.2).