Reconstructing the velocity field beyond the local universe

R. Johnston · D. Bacon · L. F. A. Teodoro · R. C. Nichol · M. S. Warren · C. Cress

Abstract  We present a maximum probability approach to reconstructing spatial maps of the peculiar velocity field at redshifts $z \sim 0.1$, where the velocities have been measured from distance indicators (DI) such as $D_n - \sigma$ relations or Tully–Fisher. With the large statistical uncertainties associated with DIs, our reconstruction method aims to recover the underlying true peculiar velocity field by reducing these errors with the use of two physically motivated filtering prior terms. The first constructs an estimate of the velocity field derived from the galaxy over-density $\delta_g$ and the second makes use of the matter linear density power spectrum $P_k$. Using N-body simulations we find, with an SDSS-like sample ($N_{\text{gal}} \simeq 33 \text{ per deg}^2$), an average correlation coefficient value of $r = 0.55 \pm 0.02$ between our reconstructed velocity field and that of the true velocity field from the simulation. However, with a suitably high number density of galaxies from the next generation surveys (e.g. $N_{\text{gal}} \simeq 140 \text{ per deg}^2$) we can achieve an average $r = 0.70 \pm 0.02$ out to moderate redshifts $z \sim 0.1$. This will prove...

R. Johnston (✉)
Department of Physics, University of Western Cape, Belville, Cape Town, South Africa
e-mail: rwi.johnston@gmail.com

D. Bacon · R. C. Nichol
Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Road, Portsmouth PO1 3FX, UK

L. F. A. Teodoro
BAER Institute, Space Science and Astrobiology Division, NASA Ames Research Center, MS: 245-3, Moffett Field, CA 94035-1000, USA

M. S. Warren
Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

C. Cress
Centre for High Performance Computing, 15 Lower Hope Rd, Cape Town 7700, South Africa
useful for future tests of gravity, as these relatively deep maps are complementary to weak lensing maps at the same redshift. LA-UR 12-24505.

**Keywords**  Methods: data analysis, numerical, statistical · Galaxies: distances and redshifts · Cosmology: large-scale structure of universe

1 Introduction

With the advent of all-sky redshift surveys during the 1990s [1–5], the study of peculiar velocities grew into a thriving industry with sophisticated velocity and density field reconstruction techniques being developed such as: POTENT [6,7], VELMOD [8,9], Wiener filtering [10,11] the unbiased minimal variance (UMV) method of [12] and the inverse Tully–Fisher (ITF) method [13,14]. These approaches allowed powerful constraints on the cosmological model (for a comprehensive review see [15]).

Central to the velocity field reconstruction methodology is the acquisition of peculiar velocity data, which is achieved through distance indicators-empirical relationships between two or more intrinsic properties of galaxies, which allow us to estimate a redshift-independent distance $d$ (e.g. the Tully–Fisher relation for spiral galaxies, or $D_n - \sigma$ relations for elliptical galaxies). In the simplest scenario, we can use this distance to estimate a peculiar velocity $u$ via the well known relation

\[ u = cz_{\text{obs}} - H_0 d, \]

where $c$ is the speed of light, $H_0$ is the Hubble constant and $z_{\text{obs}}$ is the observed redshift, measured spectroscopically. The scatter in distance indicator relations introduces a typical statistical error for a given galaxy of $\approx 20\%$ in distance at scales $\gtrsim 50$ Mpc (see e.g. [16–18]). These large uncertainties coupled with systematics associated with distance indicators have thus far limited their range to reconstructing spatially resolved peculiar velocities typically within $z \sim 0.015$ (e.g. [6–8,10,19–21]).

Recent surveys have provided large catalogues of measured peculiar velocities (e.g. the SFI ++ galaxy peculiar velocity survey [22], the 6dF galaxy survey [23] and the 2MASS selected Flat Galaxy Catalog 2MFGC [24]). There has also been ongoing work extending catalogues of inferred distances. These include Fundamental Plane studies by [25] which measured $\approx 50,000$ early type galaxies in SDSS; work by Magoulas et al. [26] who have published 10,000 near-infrared early types in the 6dF Galaxy Survey out to $z < 0.055$; and [27] have recently used $\approx 93,000$ from SDSS-DR8 to calibrate the FP out to $z \approx 0.2$.

However, we are approaching a new era in galaxy survey science where forthcoming telescopes such as SKA (and their precursors, MeerKAT and ASKAP through HI measurements), Euclid and LSST will measure properties of billions of objects to an unprecedented scale and depth. It is therefore of great interest to explore the possibility of extending velocity field reconstruction to redshifts that compliment other cosmological probes of the gravitational potential.
In recent years there has been a flurry of activity to develop new probes of gravity on cosmological scales (e.g. [28]). Several tests have emerged that involve cross-correlating observables from spectroscopic surveys, galaxy imaging and the CMB. It is useful to examine how one can constrain quantities in the perturbed FLRW metric. For instance, in the Newtonian gauge,

$$ds^2 = -(1 - 2\Psi)dt^2 + a^2(t)(1 + 2\Phi)dx^2,$$

where $\Psi$ and $\Phi$ are the Newtonian potentials that describe the temporal and spatial perturbations to the metric respectively, $a(t)$ is the expansion factor of the Universe and $x$ is in comoving coordinates (e.g. [29,30]). Galaxy clustering measurements are sensitive to the $\Psi$ component, the weak lensing shear is sensitive to the potentials $(\Psi + \Phi)$, and the CMB (via the Integrated Sachs–Wolf effect) is dependent on $d(\Psi + \Phi)/d\eta$, where the derivative is taken with respect to the the comoving time $\eta$ (see e.g. [31–34]). More recently, redshift distortions have been combined with other gravitational probes to further constrain dark energy and gravity models (e.g. [35–40]).

Our main goal in this paper is to explore the possibility of reconstructing spatially resolved maps of the peculiar velocity field beyond the local Universe and out to moderate redshifts ($z \sim 0.1$) that will be complementary to other orthogonal cosmological probes from e.g. weak lensing, as a new probe of gravity.

Given the large statistical uncertainties on the distances indicators ($\gtrsim 20\%$) at moderate redshifts, we will apply a maximum probability approach, making use of two physically motivated prior terms to weakly regularise and filter the reconstruction. The first uses an estimate of the velocity field derived from the galaxy over-density $\delta_g$ and the second makes use of the matter density power spectrum $P_\delta$. Through the use of $N$-body simulations we demonstrate that, with measurements with a realistically high signal-to-noise, we can successfully reconstruct the velocity and gravitational potential field out to $z \lesssim 0.1$.

The structure of this paper is as follows. In Sect. 2 we summarise our approach to creating the input velocity maps. Details of the velocity reconstruction algorithm are presented in Sect. 3 together with results of applying this reconstruction to our simulations. We provide a summary of these results in Sect. 4. Finally, in Sect. 5 we provide a summary and discussion.

2 Simulating the peculiar velocity field

Here we describe the procedure used to create simple models of noisy input velocity fields from mock catalogues, in order to demonstrate the concept of this paper.

---

1 In regions with negligible anisotropic stress, that is, where the stress-energy tensor is invariant under spatial rotations, or the three principal pressures are identical, the Einstein equations set $\Phi = \Psi$. 
2.1 Peculiar velocities from distance indicators

Estimating an accurate peculiar velocity from current distance indicators (DI) remains a challenging task. In this section we discuss some of the key sources of uncertainty associated with obtaining a redshift independent distance measurement. For low redshifts, the radial peculiar velocity component \( u(r) \) at the position \( r \) of a galaxy moving with velocity \( v(r) \) is given by the relation

\[
u(r) \equiv \left[ v(r) - v(r = 0) \right] \cdot \hat{r} = cz - H_0 r,
\]

where \( c \) is the speed of light, \( z \) is the observed redshift for the galaxy, \( H_0 \) is the Hubble constant, \( r \equiv |r| \) is the true distance to the object, \( v(r = 0) \) denotes the the observer’s velocity assumed to be at \( r = 0 \), and \( \hat{r} \) represents the unit vector along the object’s position vector \( r \). The last equality in the above equation is only valid for redshifts \( \lesssim 0.3 \) (see e.g. [41]), which is the regime considered in this paper. If we are to use the relationship above to estimate \( u \), we require not only the measurement of galaxy redshifts but also the inference of a redshift-independent galaxy distance, the true value of which is \( r \).

The inferred distance is denoted by \( d \) and is obtained by means of a DI; empirical relationships between two or more intrinsic properties of galaxies, some of which are distance-independent (e.g. velocity dispersion \( \sigma_v \), and Luminosity \( L \)) and some others that are distance dependent (e.g. diameter \( D \)). The observables used to characterise galaxies differ according to their galaxy type. Thus, distance indicators include the Tully–Fisher relation [42] for spiral galaxies, and the Fundamental Plane (or one of its variations, e.g. the \( D_n - \sigma \) relation), for elliptical galaxies [43,44].

The uncertainties due to the internal scattering in the DI, \( \sigma_d \), leads to large uncertainties in the galaxy distances of \( \sigma_d/d \approx 0.2 \). One uses Eq. 3 with the inferred distance \( d \), in order to calculate the peculiar velocity. Note that this requires us to adopt a particular cosmology (i.e. the cosmological parameters assumed must be either stated or marginalised over). Moreover, the uncertainty in \( d \) will propagate via Eq. 3 into a large uncertainty on the peculiar velocity for a given galaxy. From Eq. 3 the error on the peculiar velocity is the sum in quadrature of the uncertainties in \( cz \) and \( H_0 d \) since these two variables are measured independently:

\[
\sigma_u = \sqrt{\sigma_{cz}^2 + \sigma_{H_0 d}^2} \approx \sigma_{H_0 d} \approx H_0 \sigma_d = H_0 \times (0.2 \ d),
\]

assuming that typically \( \sigma_{cz} \) will be negligeble compared to \( \sigma_{H_0 d} \). Hence for instance, a galaxy at \( d = 150 \text{ Mpc h}^{-1} \) has \( \sigma_u \approx 3,000 \text{ km s}^{-1} \) while its typical peculiar velocity is \( u \approx 300 \text{ km s}^{-1} \). It is therefore necessary to average over many galaxies to obtain reasonable signal-to-noise for a mean velocity signal. However, peculiar velocity work has long been plagued by biases stemming from both systematics from distance indicators in the form of Malmquist bias (e.g. [15,45–47]) and calibration bias of distance indicators. In this work we ask, if systematics are successfully corrected for, would it be possible to then overcome the large statistical errors that would result from studying velocities out to moderate redshifts? To answer this we explore whether...
it is possible to recover the velocity field from data with realistic noise levels, assuming that sources of systematics are well understood and accounted for.

2.2 The observed velocity field

This study is concerned with extracting the peculiar velocity signal from a statistically noisy measurement that would be measured from e.g. a $D_n - \sigma$ DI at redshifts of order $z \sim 0.1$. We are interested in the typical signal-to-noise one needs to successfully reconstruct a peculiar velocity map; we therefore consider two cases where one has either measured a high or low number density of peculiar velocities.

To do this we use a simulated set of galaxies within a CDM cosmological cube that is $384 h^{-1}$ Mpc on the side [48]. We define $r$ as the true (e.g. co-moving) distance of the object from the observer, and the line of sight peculiar velocity for an individual object is referred to as the true velocity $v_{z \text{true}}$. This set of true velocities represents the ideal peculiar velocity since no statistical noise has yet been added.

To create a catalogue of simulated observed peculiar velocities (in the absence of selection effects) we firstly compute the distribution of inferred distances $d$ given a set of true distances $r$ within each realisation of the mock catalogues. If we have Gaussian errors in the magnitudes obtained from our DI, which are used to compute the inferred distances, then the errors in the inferred distances are lognormal. The conditional probability density distribution that the inferred distance is between $d$ and $d + dd$ given the true distance $r$, is given by

$$Pr(d|r) = \frac{d(\ln d)}{\sqrt{2\pi} \sigma_d^2} \exp \left( -\frac{[\ln r - \ln d]^2}{2\sigma_d^2} \right),$$

where $\sigma_d$ is a measure of the fractional distance uncertainty of the distance indicator [15,49]. Throughout our analysis we assume $\sigma_d/d=0.3$ to propagate into our mock observed velocity field. For a given galaxy in our simulation we can define an observed redshift $z_{\text{obs}}$ to be

$$z_{\text{obs}} = \frac{H_0 r + v_z^{\text{pec}}}{c} \tag{6}$$

where, $H_0/c = z_{\text{cos}}$, the cosmological redshift.

We now wish to average the data on a particular scale. This is achieved by pixelising the cube equally in $x$, $y$ and redshift $z_{\text{obs}}$ with a transverse pixel scale of $(24 h^{-1} \text{ Mpc})^2$. Thus we have $16 \times 16 \times 16 = 4,096$ pixels in total. This scale is sufficiently large to smooth out non-linearities in the velocities. We create two galaxy samples which both cover a range $0.08 < z \leq 0.18$ but which have different number densities of objects. For the high S/N case we have a total of 300,000 galaxies which equates to a number density of $N_{\text{gal}} \simeq 140 \text{ per deg}^2$. In the lower S/N sample we use the estimated number density from the SDSS Legacy survey, as detailed in [50], with a total number of 70,000 objects ($N_{\text{gal}} \simeq 33 \text{ per deg}^2$). Throughout we will refer to the high and low S/N catalogues as Mock 1 and Mock 2 respectively. For our method we use the observed redshift as a proxy for distance and so bin the data in observed
redshift space. Therefore, for each slice in $z_{\text{obs}}$, we evaluate the average velocity $V_{ijk}$ for the $(i, j)$th pixel in the $k$th redshift slice. We denote the true and observed velocity fields respectively as $V^{\text{true}}$ and $V^{\text{obs}}$.

3 The velocity reconstruction method

In this section we describe our velocity reconstruction approach. The task of a reconstruction algorithm is to infer a good estimate of the underlying true quantity $T$. The probability of a particular hypothesis for the inferred quantity, $\hat{T}$, is to be calculated taking into account the noise of the observations, $N$, and the data $D$. A common approach is to parameterize $\hat{T}$ and try to extract such parameters from the observed data, $D$, via Bayes’ theorem

$$\Pr(\hat{T}|D) \equiv \frac{\Pr(D|\hat{T}) \times \Pr(\hat{T})}{\Pr(D)}.$$ \hspace{1cm} (7)

$\Pr(D|\hat{T})$ is the likelihood while $\Pr(\hat{T})$ denotes the prior, which contains any previous information or prejudice we have regarding the value of $\hat{T}$. Since $D$ is measured before it is used to infer $\hat{T}$, $\Pr(D)$ is a constant. In our algorithm, the relevant data are the noisy measurements of the radial velocity field $V^{\text{obs}}$ and the quantity we seek to infer is the reconstructed radial velocity field $\hat{V}^{\text{rec}}$. Finding the most probable $\hat{V}^{\text{rec}}$ is an application of maximum probability reconstruction to velocity distance measures; much work has been done in applying this methodology in lensing e.g. [51–54] and to large-scale structure e.g. [55,56]. In our approach the prior will be crucial to reduce noise in the reconstruction. With a particular choice of prior which we will motivate below, the previous equation can be cast as

$$\log[\Pr(\hat{V}^{\text{rec}}|V^{\text{obs}})] = \log[\Pr(V^{\text{obs}}|\hat{V}^{\text{rec}})]$$
$$+ \log[\Pr(V|V_\delta)]$$
$$+ \log[\Pr(P_V(k)|P_V^{\text{th}}(k))] - \text{const.},$$ \hspace{1cm} (8)

where the quantity $V_\delta$ is an expected velocity field derived from the galaxy overdensity. $P_V(k)$ is the measured velocity power spectrum and $P_V^{\text{th}}(k)$ is its theoretical counterpart. In broad terms, the $V_\delta$ prior term is stating that we (weakly) prefer velocity fields that are in keeping with the velocity field expected from linear theory given a density field. In addition, the $P_V(k)$ prior term acts to prefer velocity fields which have a power spectrum reasonably close to theoretical predictions for the velocity power. Together, these terms state that the prior probability of a reconstructed field is treated as proportional to an assigned probability that this field would deviate from the linear theory prediction, multiplied by the probability that the field’s behaviour would deviate from our preferred power spectrum. Since the permitted deviations in each case are very large, these two terms are treated as independent. The meaning and construction of the prior terms will be discussed in detail in the following subsections. The second and third terms of the previous equation are the logarithm of the prior while the first one is the logarithm of the likelihood:
Fig. 1 Covariance of the $V_{\text{obs}}$ field as calculated from 10,000 realisations of the distance errors on the catalogue galaxies

\[
-2 \log[\Pr(V_{\text{obs}} | \hat{V}_{\text{rec}})] \equiv \chi^2_k = \sum_{i,j=1}^{N_{\text{pix},k}} \frac{(\hat{V}_{\text{rec}}_{ijk} - V_{\text{obs}}_{ijk})^2}{\sigma^2_{v,ijk}},
\]

where $N_{\text{pix},k}$ is the number of pixels in the $k$th $z_C$ slice. In the likelihood term, we find the covariance to be strongly diagonal due to the dominance of noise. To confirm this, we took one of our simulated galaxy cubes at $0.08 < z < 0.18$ and generated 10,000 random realisations. For each realisation we add a new set of random errors to the distances, with $\sigma_d/d = 0.3$. The resulting covariance is shown below in Fig. 1; as expected, the matrix is strongly diagonal. We do not at this stage (as one could, as in e.g. the VELMOD approach) include systematic effects and biases into our likelihood analysis; here we are interested in a proof-of-concept approach to probe the high redshift regime.

3.1 Approximate form for likelihood

We note that in its current form, Eq. 9 is inadequate: the $\chi^2$ distribution is correct for data with Gaussian errors; however, as shown in expression 5 our inferred distances $d$ are drawn from a lognormal distribution which is subsequently propagated into the observed velocity field $V_{\text{obs}}$. Since there are many galaxies in each map pixel, the error on the mean velocity in the pixel will be Gaussian, but the lognormal distribution for each galaxy introduces an offset in the mean velocity. We therefore introduce a correction to account for this mean offset. Recall that for each object $i$ we have
defined an inferred distance $d_i$ and a true distance $r_i$ which are then converted into effective redshifts $z_{\cos,i}$ and $z_{\text{obs},i}$ as $H_0 d_i/c$ and $H_0 r_i/c$, respectively. We compute $A_{\text{corr}}$ by generating 100 different catalogues of observed distances, and computing $\langle z_{\text{obs}} - z_{\cos} \rangle |_{z_{\text{obs}}}$ and the dispersion on this for each realisation. Then we have as our corrected likelihood term

$$
\chi^2_k = \sum_{i,j=1}^{N_{\text{pix}},k} \left( \frac{\hat{V}_{ij} - V_{ij}^\text{obs} + A_{\text{corr},k}}{\sigma_{V,ijk}^2} \right)^2.
$$

The success of this approximation will be demonstrated in our results section Sect. 4 below. Having obtained a form for the likelihood, we now turn to the prior terms.

3.2 The $V_\delta$ field prior term

In this subsection we discuss the first prior term used in the reconstruction. We estimate the velocity field derived from the observed over-density $\delta_g$ of the galaxy distribution. The motivation for such a prior term is to express a modest preference for velocity fields that are consistent with linear gravitational infall given a galaxy density field.

From this starting point we can construct an estimate of the gravitational potential field $\Psi$ via the Poisson equation (assuming a linear bias on these scales; we currently set $b = 1$, but this could be altered for a particular galaxy sample). We include the required modification to take into account the effects of redshift space distortions; in Fourier space the gravitational potential can be determined by

$$
\Psi_{\text{gal}}^\delta(k) = \frac{3 \Omega_m H_0^2 \delta_z(k)}{2a(k^2 + \beta k_z^2)},
$$

where $\Omega_m$ is the dimensionless matter-density parameter, $H_0$ is the Hubble constant and $k$ is the wave-vector. We have assumed a value of $\beta = 0.309$ as constrained in [57]. The scale factor $a$ is computed at the midpoint of each redshift slice, given by $a = 1/(1+z)$. Throughout we assume $\Omega_m = 0.3$ and $H_0 = 72 \text{ km s}^{-1} \text{Mpc}^{-1}$.

The estimated velocity field is obtained using $V_\delta^\text{gal} = 2 f \partial_z \Psi_{\text{gal}}^\delta / 3 \Omega_m a \simeq \partial_z \Psi_{\text{gal}}$.

In Fourier space this is computed by

$$
FT \left( V_\delta^\text{gal} \right) = i \Psi_{\text{gal}}^\delta(k) k_z
$$

Here $k_z$ is the $z$-component of $k$. Now we can write the $V_\delta$ prior term in the form

$$
\chi^2_{V_\delta} = \sum_{i=1}^{N_{\text{pix}}} \left( \hat{V}_{\delta}^{\text{rec}} - V_{\delta}^{\text{gal}} \right)^2 / \sigma_{V_\delta}
$$

The quantity $\sigma_{V_\delta}$ is a parameter determining how much this prior term will dominate in the reconstruction. The true covariance of $V_\delta^{\text{gal}}$ will not be diagonal, as even on
degree scales the $\delta$ field is correlated between pixels. Hence if we wished to make a joint reconstruction between the velocity and overdensity fields, we should include this non-diagonal covariance. However, this joint reconstruction would strongly weight the information from the less noisy $V_{\delta}^{\text{gal}}$ field over the more noisy velocity field, which is not the point of this paper—we are interested in whether we can use the $V_{\delta}^{\text{gal}}$ field as a weakly informative prior for extracting information from the velocity field. That is, an important question is whether the final reconstruction is more correlated with the $V_{\delta}$ field or the true velocity field $V^{\text{true}}$ and, therefore, whether we are dominated by the prior. It is for this reason that we choose a large value for $\sigma_{\delta v}$; the value on the diagonal, for a joint reconstruction, would be the error one makes in estimating $V_{\delta}^{\text{gal}}$ within a pixel ($\simeq 50 \text{ kms}^{-1}$) whereas we choose $\sigma_{\delta v} = 1,000 \text{ kms}^{-1}$; we will show that this satisfies the condition that the prior does not dominate the reconstruction.

3.3 The peculiar velocity power spectrum prior term

The second prior term makes use of a theoretical velocity power spectrum to regularise the $\hat{V}^{\text{rec}}$ field. Using this information in reconstructions is not a new concept; [58] demonstrated the usefulness of adopting a power spectrum prior for regularising non-parametric fits in their work on density field reconstruction from peculiar velocities.

The $V_{\delta}$ prior from the previous section compares the reconstructed and position-derived velocity fields on a pixel by pixel basis. As we shall see, the introduction of the $P_k$ prior term complements this by regulating the reconstruction on a variety of scales.

In order to use this prior, we need to obtain a suitable theoretical velocity power spectrum; here we use the power spectrum of the true velocity field $V^{\text{true}}$ in the simulations, but in principle we could use a range of LCDM power spectrum predictions and marginalise our results over these. We compare the theoretical velocity power spectrum $P_v^{\text{true}}(k)$ with each trial reconstruction power spectrum $\hat{P}_v^{\text{rec}}(k)$ so that the $P_k$ prior term takes the form

$$\chi^2_{P_k} = \sum_{i=1}^{k_{\text{tot}}} \frac{[\hat{P}_v(k)^{\text{rec}} - P_v(k)^{\text{theory}}]^2}{\sigma_{P_v(k)}^2},$$

where $k_{\text{tot}}$ is the total number of wave numbers and we can choose

$$\sigma_{P_v(k)} = a P_v(k)^{\text{theory}}.$$  \hspace{1cm} (15)

The parameter $a$ in Eq. 15 allows us to modify the strictness of this prior term in the reconstruction. We have found that the off-diagonal terms are small and therefore can be neglected in the above prior term. Note that Eq. 14 assumes the power spectrum has Gaussian errors; this is valid except on the largest scales.
3.4 Implementation

In this section we briefly outline how we implement our algorithm.

1. The likelihood is maximised iteratively by beginning with $N_t$ initial trial velocity fields (we choose $N_t = 15$). For each iteration of the reconstruction we allow a random walk for $n = 2$, $N_t$ trial reconstructions by adding an amount of uniform random noise with of magnitude given by $\sqrt{P_{\text{theory}}(k)}$ to each mode. The $n = 1$ field is not perturbed [see step (ii)].

2. We compute the log probabilities for each $n$th trial reconstruction. We determine which of the trials has the minimum log probability. If the trial with minimum log probability corresponds to $n = 1$, we resample from the parent sample in the next iteration of the code and allow the pixel variation to reduce for the new set of trial fields. However, if the minimum log probability corresponds to $n \neq 1$, we set the corresponding trial $\hat{V}_{\text{rec}}$ to be the new template and feed this back in to Step (i), where a new set of trial velocity fields are generated based on this new template. This process is repeated until convergence is reached.

4 Results

We have applied our reconstruction technique to two simulated galaxy cubes (Mocks 1 and 2) over a redshift range $0.09 < z \leq 0.18$; the galaxy cubes have average number density of galaxies of $N_{\text{gal}} \simeq 140$ and $\simeq 33$ per deg$^2$ $N_{\text{gal}}$ respectively. Figure 2 shows velocity maps for a selected redshift slice at $0.118 < z < 0.124$. The top row represents the results for Mock 1, while the bottom row shows results for Mock 2. From

![Fig. 2 Velocity reconstruction maps. The top row shows results for a mock sample with a number density of galaxies $N_{\text{gal}} \simeq 140$ per deg$^2$ while the bottom row is for a mock sample with a number density of $N_{\text{gal}} \simeq 33$ per deg$^2$. For this example we show the results for a redshift slice of $0.118 < z < 0.124$. For each column going from left to right we show the input noise velocity maps, the true radial velocity field $V_{\text{true}}$, the reconstructed velocity map, and the resulting correlation between the two maps $r[V_{\text{true}}, V_{\text{rec}}]$ respectively. The dashed line is a one-to-one relation (colour figure online).](image-url)
Fig. 3  Left Correlation coefficient values for the final reconstruction of the peculiar velocity field. The bottom panels show the correlation between $r[V^{\text{obs}}, V^{\text{true}}]$ (red points) for each redshift slice Mock 1 (left) and Mock 2 (right). The black points show the mean correlation value for the 1,000 MC samples and the true velocity field i.e. $r[V^{\text{rec}}, V^{\text{true}}]$. The top panels show the respective correlation coefficient between the residuals $\epsilon$ (where $\epsilon = V^{\text{true}} - V^{\text{rec}}$) and $V^{\text{true}}$ (black points) and $V^{\text{rec}}$ (blue points). Right the error power spectrum compared for the various velocity fields. The red line shows the power spectrum of the true velocity field $V^{\text{true}}$, the blue line is the power spectrum on the $V^{\delta}_{\text{gal}}$ field (computed from the galaxy over density) and the green line is the power spectrum of the observed noisy input velocity map $V^{\text{obs}}$. The dashed black line shows the residual power of $(V^{\delta}_{\text{gal}} - V^{\text{rec}})$, where $V^{\text{rec}}$ is the reconstructed velocity field for the first $n=1$ iteration of the algorithm. The solid black line is then the residual power for the final iteration of the algorithm at $n=20,000$ (colour figure online).

In both mocks we can see that, due to the observational uncertainty $\sigma_d/d = 0.3$ we propagate into the observed velocity field $V^{\text{obs}}$, the observations are completely noise dominated. Comparing both maps to the input true velocity field we find little correlation between them. By applying our method including the two terms contributing to our prior, we can see that the final reconstructed fields correlate rather well with the $V^{\text{true}}$. Mock 1 ($N_{\text{gal}} \simeq 140$ per deg$^2$) shows an overall better reconstruction ($r = 0.71 \pm 0.05$) compared to Mock 2 which has considerably lower number density of objects ($r = 0.67 \pm 0.08$). However, in both cases with this example we can see that the main large scale structures are recovered well.

In Fig. 3, left panel we show the correlation between $\hat{V}^{\text{rec}}$ and $V^{\text{true}}$ for each redshift slice in the mocks. In the lower part of this panel we show correlation coefficient values $r$ between $[V^{\text{true}}, V^{\text{obs}}]$ (red points) and $[V^{\text{true}}, V^{\text{rec}}]$ (black points) for Mock 1 (bottom left) and Mock 2 (bottom right). In the upper part of the panel we show the corresponding correlation coefficient values between the residuals $\epsilon$ and $V^{\text{true}}$ (black points) and $V^{\text{rec}}$ (blue points). We observe that the overall correlation with the input observed velocity map $\hat{V}^{\text{rec}}$ across the mocks is predictably low, with a maximum $r = 0.23 \pm 0.11$ and a minimum $r = -0.008 \pm 0.122$. This is due to the addition of our large $\sigma_d/d = 0.3$ error on the distances at these relatively high redshifts, which can be seen visually in the noise map from Fig. 2.
In contrast, the black points on the lower part of the left panel of Fig. 3 demonstrate that we achieve a good reconstruction within each redshift slice. By comparing to the underlying true velocity map, we find a correlation of \( r = 0.69 \pm 0.01 \) in Mock 1 (the high S/N case). We can see that the result for this lower number density of objects (Mock 2) in the right hand upper panel follows a similar trend to the Mock 1 case. However, as expected there is a general systematic shift toward lower correlation values, with an average value of \( r = 0.55 \pm 0.02 \). In particular, we note an excessive dip in correlation in the Mock 2 with values of \( r = 0.23 \pm 0.12 \) and \( r = 0.27 \pm 0.12 \) \( z \approx 0.104 \) and \( z \approx 0.143 \) respectively. On closer inspection we find that this is due to a few outlying pixels with extremely high velocities of the order \( \sim 6,000 \text{ km/s} \) which the algorithm has failed to constrain effectively. Finally we note that in both mocks the correlations between the residuals \( \epsilon \) with \( V^{\text{true}} \) and \( V^{\text{rec}} \) are not consistent with zero, but show a systematic offset; this will need to be improved upon in future studies. Nevertheless, this is an encouraging result demonstrating that it may be possible to apply our technique to existing data from e.g. SDSS. At the very least, forthcoming surveys should provide the required signal-to-noise level to reconstruct the velocity field at moderate redshifts.

4.1 The residual power spectrum

A useful statistic with which to evaluate the reconstruction is the power spectrum itself. On the right hand panel of Fig. 3 we have plotted the power spectrum as measured from the observed velocity field \( V^{\text{true}} \) shown in green, the true velocity field shown in red, the \( V^{\delta \text{gal}} \) field shown in blue, and finally the residual power \( \epsilon P_k = P_k (V^{\delta \text{gal}} - \hat{V}^{\text{rec}}) \) after the first iteration of the code (dashed black line) and the final iteration.

The difference between \( P^{\text{obs}}(k) \) and \( P^{\text{true}}(k) \) is substantial, with the observed power being flat across most scales due to the dominant measurement noise. The power spectrum of the true velocity field, \( P^{\text{true}}(k) \) and that of the \( V^{\delta \text{gal}} \) field \( P^{\delta \Psi}(k) \) are rather similar, showing large power on large scales. The residual power \( \epsilon P_k \) reveals important information on how the reconstruction evolves; if we look at \( \epsilon P_k \) after the first iteration of the reconstruction then we can see that already the shape of its power spectrum is similar to that of \( P^{\text{true}}(k) \) and \( P^{\delta \Psi}(k) \) but with a comparative excess on all scales. By the final iteration of the code \( n = 20,000 \), \( \epsilon P_k \) (black solid line) now has similar power to \( P^{\delta \Psi}(k) \) on small scales \( k \gtrsim 0.06 \), but less on scales larger than this. At \( k \lesssim 0.05 \) we observe a similar trend when compared to \( P^{\text{true}}(k) \). This is a promising result as it shows that the residual power is smaller than the true fluctuations on these scales.

4.2 Exploring the impact of the priors

The results presented in the previous section use the power spectrum \( P(k) \) prior term and the \( V^{\delta} \) prior term in combination. In this section we examine how each prior is affecting the reconstruction process individually. In Fig. 4 we show four scenarios of possible reconstruction processes within \( 0.101 < z \leq 0.107 \). In all cases we show the \([V^{\text{rec}}, V^{\text{true}}]\) distributions with their respective correlation coefficient values \( r \) and the slope of their respective slopes \( a \). Firstly, we test the reconstruction method where no
Fig. 4 An example of the effect of the different prior terms on the reconstruction at a given redshift slice $0.101 < z \leq 0.107$. The left panel shows the final reconstruction of the velocity field where no prior terms have been activated. The second panel shows the final reconstruction when only the $P_k$ power spectrum prior term is applied with $\sigma^{P}_{v}(k) = 0.1 P^\text{true}_v(k)$. The third shows the reconstruction with the baryon $V_\delta$ prior term only with $\sigma_{V_\delta} = 1,000.0 \, \text{kms}^{-1}$. Finally, the fourth shows the combination of the two prior terms.

priors are applied as shown in the first panel (left hand side) of the figure. This serves as a null test, where we are essentially fitting to the observed velocity data which is a convolution of the true velocity information and the lognormal noise derived from the simulated observed distances. As expected, the plot shows a very noisy reconstruction with a correlation of $r = 0.24 \pm 0.11$ and a very broad distribution with velocities in excess of $\sim 3,000 \, \text{kms}^{-1}$ at the tails.

In the second panel we show the reconstruction where only the $P_k$ prior has been applied (with the same value for $\sigma^{V}_{P(k)}$ as used in the previous section). What is immediately clear is a much improved tighter relationship between $V_{\text{true}}$ and $\hat{V}_{\text{rec}}$ with a correlation of $r = 0.59 \pm 0.08$. We also find an improved slope $a$ of the correlation, $a = 0.44 \pm 0.04$.

The third panel of Fig. 4 shows the effect of the $V_\delta$ prior which compares the $\hat{V}_{\text{rec}}$ field with $V_\delta$ field derived from the galaxy over density. We adopt the same $\sigma_{\delta v} = 1,000 \, \text{kms}^{-1}$ value as in our previous analysis. The motivation for choosing such a conservative error on the prior is to limit its effect on regularising the reconstruction and avoid us simply fitting the $V_\delta$ field. As we can see, the final reconstruction with the $V_\delta$ prior alone shows an improved but still modest correlation of $r = 0.52 \pm 0.08$ and a slope of $a = 0.23 \pm 0.02$.

Finally, by combining the two priors as shown in the fourth panel, we find a much improved reconstruction with $r = 0.73 \pm 0.05$ and $a = 0.58 \pm 0.03$, apparently indicating the usefulness of including both priors. We can compare this correlation with the correlation between the reconstruction and the $V_\delta$ prior within the same redshift slice, which is $r = 0.63 \pm 0.08$; the reconstruction is closer to the truth than it is to the prior, which is an important criterion for the success of our method. To understand further whether the reconstruction is a faithful one, it is important to consider the effect of observational biases.

4.3 Exploring observational bias in the priors

Applying our method to real survey data will naturally be subject to various selection effects and biases. Understanding how this might impact the $V_\delta$ prior, and con-
Fig. 5  The left hand panel set shows: the $V_\delta$ field at $0.135 < z \leq 0.141$ (top left) and the same field with an artificial mass embedded; this can be seen as the dark blue region in the velocity field. The bottom left panel shows the [$V_{\text{true}}, V_\delta$] distribution, where $V_\delta$ relates to the map directly above. The bottom right panel is the [$V_{\text{true}}, V_\delta$] distribution where $V_\delta$ relates to the map with the overdensity embedded in the field. The right hand panel set shows the reconstructed velocity fields (top row) for the same redshift range—the left hand plots show the case where no artificial mass has been added to the $V_\delta$ prior and the right shows the result where it was been added (colour figure online).

sequently our velocity reconstruction, is necessary if the method is to be successfully applied to future data-sets. We have performed further tests of how much the $V_\delta$ prior will bias the reconstruction if in some region mass does not follow galaxy counts, i.e. if there is a baryonic feature present which is not proportionally represented in the underlying gravitational potential field. We test this by artificially embedding into the gravitational potential field a ‘fake’ over-dense region which has no counterpart in the true velocity field. We calculate the resulting $V_\delta$ field and hence the $V_\delta$ prior term. The purpose of this test is to observe if such effects will substantially contaminate the final reconstruction of the velocity field.

In Fig. 5 we consider the scenario where a large mass would be inferred to be present from the galaxy clustering, modelled by a gaussian embedded into the $\Psi_{\text{gal}}$ field over several redshift slices. In the left hand panel set of the figure, we show the redshift slice where this false mass is most prominent at $0.135 < z \leq 0.141$ and is then propagated into the $V_\delta$ field (top right). To distinguish between the two velocity fields we will refer to the altered prior information with a “hat” above the observable e.g. $\hat{V}_\delta$. The top map row in this panel set shows the two $V_\delta$ maps; we observe a negative excess in the $\hat{V}_\delta$ velocity potential extending to $\sim 1,000 \text{ kms}^{-1}$ centred at $[13,13]$ on the $[X, Y]$ pixel plane (top right). The difference between $V_\delta$ and $\hat{V}_\delta$ is perhaps more clearly shown in the bottom rows where we plot the distributions of both [$V_\delta, V_{\text{true}}$] (left) and [$\hat{V}_\delta, V_{\text{true}}$] (right). In the right-hand panel of this row we now see a skewed distribution of velocity pixels toward the bottom left part of the plot, compared to the $V_\delta$ on the left where it remains a relatively tight distribution. Comparing their relative correlations we can see that with the added mass we observe a reduced correlation coefficient from $r[V_{\text{true}}, V_\delta]=0.74\pm0.06$ to an $r[\hat{V}_{\text{true}}, \hat{V}_\delta]=0.59\pm0.09$ corresponding to a $\Delta r = 0.15$.

Our new $\hat{V}_\delta$ field provides a modified $V_\delta$ prior term of the form
\[ \frac{(\hat{V}_{\text{rec}} - \hat{V}_\delta)^2}{\sigma_{\hat{V}_\delta}^2}. \]  

(16)

We then carry out our reconstruction method, afflicted with this poor prior. Comparing the two distributions in the right hand panel set of Fig. 5, we can clearly see that there is very little residual bias in the final reconstruction of the velocity field despite the prominent artificial velocities present in the \( \hat{V}_\delta \) field. By comparing the respective correlations we now find for the case where no mass has been added 
\[ r[V_{\text{true}}, \hat{V}_{\text{rec}}] = 0.63 \pm 0.08 \]  
and the case where the mass has been added to the \( \hat{V}_\delta \) field 
\[ r[V_{\text{true}}, \hat{V}_{\text{rec}}] = 0.58 \pm 0.09 \]  
(corresponding a \( \Delta r = 0.05 \)). This further demonstrates how the \( V_\delta \) prior term does not dominate the reconstruction; we are extracting useful corrective information from the observed velocity field.

5 Discussion and considerations for the future

In this study we have carried out a first step towards a reconstruction the velocity field on cosmological scales, through the use of measured peculiar velocities as derived from current distance indicators. Previous methods have focused on velocity reconstruction within the local Universe. Our motivation is to look toward the possibility of pushing to cosmological depths where such maps of the velocity and gravity field could be used as a cosmological probe of GR.

To begin developing an approach that can reconstruct resolved maps of the velocity field out to \( z \lesssim 0.1 \), we created two simulated mock catalogues. Both catalogues were over the range \( 0.09 < z < 0.1 \) with a scale of 384 h\(^{-1}\) Mpc on the side. However, the two mocks had different number densities of objects; the first mock had a total of 300,000 objects equating to a number density of \( N_{\text{gal}} \approx 140 \) per deg\(^2\), while the second contained 70,000 objects (\( N_{\text{gal}} \approx 33 \) per deg\(^2\)). The mocks were pixelised to create a smoothed true velocity field \( V_{\text{true}} \) and a velocity field that represents the noisy observed data, \( \hat{V}_{\text{rec}} \), with a transverse pixel scale of \((24 h^{-1} \text{Mpc})^2\). We have modelled a volume-limited sample with a roughly uniform redshift distribution within the redshift range of interest, allowing us to bin the data in equal size redshift bins, in redshift \( z_{\text{obs}} \) space along the line of sight. In reality one would prefer to maximise the amount of peculiar velocity data available, and a volume limited selection is certainly not optimal.

In this work we have focused only on the contribution of the statistical noise from distance indicator measurements, which then propagate into the observed velocity field. We chose a conservative estimate for the statistical error on the distances of \( \sigma_d/d = 0.3 \), which is greater than current error estimates from distance indicators. However, we do note that a future development of our approach should include other observational effects/biases into our likelihood analysis in a similar way to that by Willick and Strauss VELMOD approach. Moreover, it will be useful to examine specifically the compounding effects due to DI calibration bias as described in detail in [15] by including this in our mocks as well as applying our approach to existing observed data.

Key to our method is the inclusion in our prior of two terms which guide the velocity reconstruction. The first prior term, featuring \( P_k \), regulates the reconstruction...
on various scales by comparing a theoretical velocity power spectrum $P^\text{true}_v(k)$ to the one computed for the trial reconstruction $\hat{P}^\text{rec}_v(k)$. The second prior term, featuring $V_\delta$, compares the velocity field $V_\delta$ derived from the galaxy over-density $\delta_g$ with the trial reconstructed velocity field $\hat{V}^\text{rec}$.

The results show that overall we can recover the velocity field to a high degree of correlation with $V^\text{true}$. For the cube at the high number density mock, we found that we could achieve a consistently high correlation coefficient value of $r[V^\text{true},\hat{V}^\text{rec}] \simeq 0.7$ for a given slice in redshift. For the mock with a number density closer to that of an SDSS type survey we found we could still reconstruct the main structures but with a lower overall correlation of $r = 0.55 \pm 0.02$.

As a next step, it will be of interest to modify our approach to use the peculiar velocity field to reconstruct the gravitational potential itself at intermediate redshifts of $z \sim 0.1$, which will complement gravity maps from other probes.

Acknowledgments We would like to extend special thanks to Enzo Branchini and Adi Nusser for stimulating discussions, their comments and reading several drafts of this paper. We also would like to thank Martin Hendry, Mat Smith, Andreas Faltenbacher, Roy Maartens, Daniele Bertacca, Rafal Szepeitkowski, Yong-Seon Song, Kazuya Koyama, Prina Patel, Emma Beynon, Robert Crittenden and Philip Marshall for useful discussions. RJ acknowledges the support of the SKA-South Africa and the National Research Foundation (NRF), as well as the hospitality of the Institute of Cosmology and Gravitation (ICG) at the University of Portsmouth where some of this work was carried out. DB and RN are supported by the UK Science and Technology Facilities Council (Grant Nos. ST/H002774/1 and ST/K0090X/1). The analysis was performed with the SCIAMA High Performance Computing cluster supported by the ICG, at the University of Portsmouth. We would like to thank the SCIAMA cluster administrator, Gary Burton, for all his help. Please contact the authors to request access to research materials discussed in this paper.

References

1. Fisher, K.B., Huchra, J.P., Strauss, M.A., Davis, M., Yahil, A., Schlegel, D.: Astrophys. J. Suppl. 100, 69 (1995). doi: 10.1086/192208
2. Sutherland, W.J., Maddox, S.J., Keeble, O., Oliver, S.J., Rowan-Robinson, M., McMahon, R.G., Efstathiou, G., et al.: MNRAS 317, 55 (2000). doi: 10.1046/j.1365-8711.2000.03528.x
3. Willick, J.A., Courteau, S., Faber, S.M., Burstein, D., Dekel, A., Kolatt, T.: VizieR Online Data Catalog 7198, 139 (1996)
4. Willick, J.A., Courteau, S., Faber, S.M., Burstein, D., Dekel, A., Strauss, M.A.: Astrophys. J. Suppl. 109, 333 (1997). doi: 10.1086/312983
5. Giovannelli, R., Haynes, M.P., Freudling, W., da Costa, L.N., Salzer, J.J., Wegner, G.: Astrophys. J. Lett. 505, L91 (1998). doi: 10.1086/311621
6. Bertschinger, E., Dekel, A.: Astrophys. J. Lett. 336, L5 (1989). doi: 10.1086/185348
7. Dekel, A., Eldar, A., Kolatt, T., Yahil, A., Willick, J.A., Faber, S.M., Courteau, S., Burstein, D.: Astrophys. J. 522, 1 (1999). doi: 10.1086/307636
8. Willick, J.A., Strauss, M.A., Dekel, A., Kolatt, T.: Astrophys. J. 486, 629 (1997). doi: 10.1086/304551
9. Willick, J.A., Strauss, M.A.: Astrophys. J. 507, 64 (1998). doi: 10.1086/306314
10. Fisher, K.B., Lahav, O., Hoffman, Y., Lynden-Bell, D., Zaroubi, S.: MNRAS 272, 885 (1995)
11. Zaroubi, S., Hoffman, Y., Fisher, K.B., Lahav, O.: Astrophys. J. 449, 446 (1995). doi: 10.1086/176070
12. Zaroubi, S.: MNRAS 331, 901 (2002). doi: 10.1046/j.1365-8711.2002.05229.x
13. Nusser, A., Davis, M.: MNRAS 276, 1391 (1995)
14. da Costa, L.N., Nusser, A., Freudling, W., Giovannelli, R., Haynes, M.P., Salzer, J.J., Wegner, G.: MNRAS 299, 425 (1998). doi: 10.1046/j.1365-8711.1998.01802.x
15. Strauss, M.A., Willick, J.A.: Phys. Rep. 261, 271 (1995). doi: 10.1016/0370-1573(95)00013-7
16. Bernardi, M., Sheth, R.K., Annis, J., Burles, S., Eisenstein, D.J., Finkbeiner, D.P., Hogg, D.W., Lupton, R.H., Schlegel, D.J., SubbaRao, M., Bahcall, N.A., Blakeslee, J.P., Brinkmann, J., Castander, F.J.
49. Landy, S.D., Szalay, A.S.: Astrophys. J. 391, 494 (1992). doi:10.1086/171365
50. Ahn, C.P., Alexandroff, R., Allende Prieto, C., Anderson, S.F., Anderton, T., Andrews, B.H., Aubourg, É., Bailey, S., Balbinot, E., Barnes, R., et al.: arXiv e-prints (2012)
51. Seitz, S., Schneider, P., Bartelmann, M.: A&A 337, 325 (1998)
52. Bridle, S.L., Hobson, M.P., Saunders, R., Lasenby, A.N.: Astronomical Society of the Pacific Conference Series. Gravitational Lensing: Recent Progress and Future Go, vol. 237, p. 307 (2001)
53. Marshall, P.J., Hobson, M.P., Gull, S.F., Bridle, S.L.: MNRAS 335, 1037 (2002). doi:10.1046/j.1365-8711.2002.05685.x
54. Szepietowski, R.M., Bacon, D.J., Dietrich, J.P., Busha, M., Wechsler, R., Melchior, P.: MNRAS 440, 2191–2200 (2014)
55. Kitaura, F.S., Jasche, J., Li, C., Enßlin, T.A., Metcalf, R.B., Wandelt, B.D., Lemson, G., White, S.D.M.: MNRAS 400, 183 (2009). doi:10.1111/j.1365-2966.2009.15470.x
56. Jasche, J., Kitaura, F.S., Li, C., Enßlin, T.A.: MNRAS 409, 355 (2010). doi:10.1111/j.1365-2966.2010.17313.x
57. Tegmark, M., Eisenstein, D.J., Strauss, M.A., Weinberg, D.H., Blanton, M.R., Frieman, J.A., Fukugita, M., Gunn, J.E., et al.: Phys. Rev. D 74(12), 123507 (2006). doi:10.1103/PhysRevD.74.123507
58. Kaiser, N., Stebbins, A.: In: Latham, D.W., da Costa, L.A.N. (eds.) Large-Scale Structures and Peculiar Motions in the Universe, Astronomical Society of the Pacific Conference Series, Astronomical Society of the Pacific Conference Series, vol. 15, p. 111 (1991)