Statistics of velocity centroids: effects of density–velocity correlations and non-Gaussianity


ABSTRACT

We continue with our previous work on statistics of velocity centroids, to retrieve information about the scaling properties of an underlying turbulent velocity field from spectroscopic observations. We use synthetic data sets with extreme effects of velocity–density correlations that we create artificially, which also have a non-Gaussian distribution of fluctuations. We confirm that centroids can be used to obtain the scaling properties of the turbulent velocity when the ratio of the density dispersion to the mean density is less than unity, regardless of velocity–density correlations and non-Gaussianity. It was found that extreme velocity–density correlations can distort the statistics of velocity centroids, impeding the recovery of the turbulent velocity spectral index from centroids. We show that such correlations introduce high-order moments to the maps of centroids, which we disregarded in previous work, but that they are only important when the density dispersion is large in comparison with the mean density. It was also found that non-Gaussian velocity and/or density distort the statistics of centroids too, but to a lower degree than extreme cross-correlations.

Key words: MHD – turbulence – ISM: general – ISM: structure – radio lines: ISM.

1 INTRODUCTION

The present picture of the interstellar medium (ISM) is that of a turbulent medium. This turbulence expands over several scales, that range from au to kpc (Larson 1992; Armstrong, Rickett & Spangler 1995; Deshpande, Dwarakanath & Goss 2000; Stanimirović & Lazarian 2001). It is of paramount importance for many physical processes that take place in the ISM, including star formation, cosmic ray and dust dynamics, magnetic field formation and evolution, and heat transport (for reviews see Vázquez-Semadeni et al. 2000; Cho & Lazarian 2005, and references therein). For those reasons there is an increasing interest in the astronomical community on turbulence. From an observational perspective, we can mention the studies of interstellar scintillations [for instance the work of Narayan & Goodman (1989) and Spangler & Gwinn (1990)]. Although this technique has been very successful, it has important limitations. It is restricted to the ionized media and it is only sensitive to density fluctuations.

Radio observations of spectral linewidths (Larson 1981, 1992; Scalo 1984, 1987) and velocity centroids (von Hoerner 1951; Münch 1958; Dickman & Kleiner 1985; Kleiner & Dickman 1985; O’Dell & Castañeada 1987; Miesch & Bally 1994) have been used to diagnose turbulence in molecular clouds and H II regions. The spectral lines contain important information on the turbulent velocity field, a quantity that is more directly related to the theories of turbulence compared to density fluctuations. Thus, they can provide us with an enormous amount of information that might help to comprehend interstellar turbulence. However, the shape of a spectral line depends not only on the velocity, but also on the density structure of the emitting material. The separation of the individual contribution of density and velocity has been a long-standing problem (see review by Lazarian 2006a).

From the theoretical side, there have been substantial advances in the understanding of the scaling properties of compressible magnetohydrodynamic (MHD) turbulence [see the reviews by Cho & Lazarian (2005) and Lazarian & Cho (2005) and references therein]. These scaling properties have been used, for instance, to find the rates of scattering of cosmic rays (Yan & Lazarian 2004) and the acceleration of cosmic dust (Yan, Lazarian & Draine 2004).

In addition, scaling laws provide a direct point of comparison between theoretical predictions and observations. They can be studied by analysing the statistics of spectral maps. Such an analysis is often done with statistical tools like power spectra, correlation, or...
structure functions and, recently, with more elaborated and robust methods, like wavelets. As an example of these, we can cite the \( \Delta \)-variance (Stutzki et al. 1998; Mac Low & Ossenkopf 2000; Bensch, Stutzki & Ossenkopf 2001; Ossenkopf & Mac Low 2002).

At the same time, there is an effort to develop new statistics to be analysed by those techniques, where the main concern is to trace or to isolate the velocity contribution from spectral line maps. Among such statistics we can mention the ‘Spectral Correlation Function’ (SCF; Rosolowsky et al. 1999; Padoan, Rosolowsky & Goodman 2001), ‘Velocity Channel Analysis’ (VCA; Lazarian & Pogosyan 2000; Lazarian et al. 2001; Lazarian, Pogosyan & Esquivel 2002; Esquivel et al. 2003; Lazarian & Pogosyan 2004), ‘Modified Velocity Centroids’ [MVCs; Lazarian & Esquivel (2003) and Esquivel & Lazarian (2005), hereafter LE03 and EL05, respectively], and the ‘Velocity Coordinate Spectrum’ (VCS; Lazarian 2004, Lazarian & Pogosyan 2006). The SCF and VCA are related techniques. While the SCF was developed to compare observations and numerical simulations, the VCA addressed the underlying velocity spectrum of turbulence. The difference between the SCF and VCA is, in some sense, a normalization issue. However, because of this normalization the analytical expressions derived for the VCA do not apply to the SCF. In other words, while the VCA can be used for the same purposes as the SCF, the opposite is not true.

LE03 and EL05 addressed how velocity centroids trace the statistics of turbulent velocity. There, the structure function of centroids was decomposed into four terms. The first contains information about density fluctuations and can be obtained directly from observations (subtraction of that term yielded the MVCs). The second contains the desired information about the scaling of velocity. The remaining two terms are: a ‘cross-term’ (a combination of density and velocity fluctuations), and a term that contains the ‘cross-correlations’ between density and velocity. These last two terms cannot be isolated from observations and, for strongly supersonic MHD turbulence simulations, they contaminate the statistics of the centroids, making them unreliable to trace the scaling properties of velocity.

In LE03, a necessary condition was proposed to discriminate when centroids of velocity trace the scaling of the actual 3D velocity. If such a condition is satisfied, and the cross-term as well as the density–velocity cross-correlations are negligible, then the centroids would scale as the underlying turbulent velocity field. When the necessary condition is marginally met, MVCs trace the velocity better than ordinary velocity centroids. Using synthetic observations from MHD simulations, EL05 showed that the cross-term and cross-correlations of density and velocity are indeed not important in subsonic turbulence; however, they become important as the sonic Mach number increases.

Using the \( \Delta \)-variance analysis on velocity centroids, Ossenkopf et al. (2005, hereafter OELS05) have provided an additional criterion to decide whether centroids trace reliably the statistics of velocity. In OELS05, it is shown that the dispersion of mass density to mean mass density ratio \( \sigma_\rho/\rho_0 \) is a more robust indicator than the criterion in LE03. When such a ratio is \( \lesssim 0.7 \), velocity centroids can be used to represent velocity statistics. The results in OELS05 are compatible with those in LE05, since a small \( \sigma_\rho/\rho_0 \) ratio can be identified with subsonic turbulence. However, the new criterion requires knowledge of \( \rho_0 \), that is, additional information that may not be easy to obtain. Moreover, the data sets used in that work did not take into account the correlations of density and velocity (by construction, the data sets were uncorrelated).

In this paper, our primary concern are the cross-term and cross-correlation terms. We perform a detailed test on how the correlations of velocity and density, and the \( \sigma_\rho/\rho_0 \) ratio affect those two components of the centroid statistics. We do this by constructing data that have different values of \( \sigma_\rho/\rho_0 \), and by introducing various levels of artificial cross-correlations.

The layout of this paper is as follows. In Section 2, we review the statistics of velocity centroids. In Section 3, we discuss the data sets, and the general methodology we use. In Section 4, we show how the different parameters affect the cross-term and cross-correlation term. A summary can be found in Section 5. We include an appendix with the definitions and derivations in terms of the \( \Delta \)-variance, which can be used instead of structure functions for the same type of analysis.

2 STATISTICS OF VELOCITY CENTROIDS

2.1 Basic definitions and tools

To characterize the scaling properties of the turbulent velocity field, we can use (second-order) correlation or structure functions, power spectra, or wavelets (for instance the \( \Delta \)-variance, see Appendix A). Our original work of velocity centroids was laid out in terms of structure functions, but the \( \Delta \)-variance can be used as well, with some advantages when applied to real observational data. In Appendix A, we provide the reader with an equivalent decomposition of the statistics of centroids in terms of \( \Delta \)-variance.

The second-order structure function of a quantity \( f(x) \) is defined as

\[
S_F(r) = \langle (f(x) - f(x + r))^2 \rangle,
\]

where \( \langle \ldots \rangle \) stands for an ensemble average over all space \( x \) and \( r \) is the so-called ‘lag’. The structure function can be expressed in terms of the correlation function \( C_F(r) = \langle f(x)f(x + r) \rangle \) as: \( S_F(r) = 2[C_F(0) - C_F(r)] \). An alternative description, very much used in turbulence studies, is the power spectrum. It is defined as the Fourier transform of the correlation function, \( P(k) = \mathcal{F}[C_F(r)] \); thus, it is closely related to the structure function as well. The power spectrum can be defined equivalently as the square of the Fourier components of \( f(r) \), and therefore it is a function of the spatial frequency or wavenumber \( k \). For isotropic fields, structure and correlation functions depend only on the magnitude of the lag \( r = |r| \), and similarly the power spectrum is only function of \( k = |k| \). For the sake of simplicity, we restricted ourselves to a model of isotropic, homogeneous turbulence. Although this model is an idealization, the analysis of MHD simulations in Esquivel et al. (2003) has shown that anisotropies hardly change the slope of the \( k \)-averaged power spectrum of observed fluctuations. Therefore, we believe that the model provides a valid approximation to the actual astrophysical turbulence.

Observations of spectral lines contain information on the velocity field, and therefore on the turbulence therein. However, the emission in a spectral line depends on the density of the emitting material as well. In what follows, we will consider an optically thin medium with an emissivity linearly proportional to the density (e.g. H\( I \)). One cannot directly observe the distribution of the emitting material in configuration space \( (x, y, z) \), but the Doppler shift gives access to its velocity distribution. In fact, we observe the density-convolved velocity component along the line of sight – hereafter LOS – at a given position in the sky. Thus, spectroscopic data are usually arranged in position–position–velocity arrays (PPV cubes for short), in which two of the coordinates denote the position in the plane of the sky and the third denotes the LOS velocity. We will denote the position in the sky by \( X = (x, y) \), and the LOS direction by \( z \). In
Statistics of velocity centroids

The centroids of velocity are usually defined as (von Hoerner 1951; Münch 1958):

\[ C(X) = \frac{\int v_i I_{\text{line}}(X, v_i) \, dv_i}{\int I_{\text{line}}(X, v_i) \, dv_i}, \]

(2)

where \( I_{\text{line}} \) is the line intensity at the sky position \( X \), and at an LOS velocity \( v_i \), the limits of integration are determined by the LOS velocity extent of the object. Under the assumption of optically thin media and an emissivity in linear proportion to the density, \( I_{\text{line}} \) can be identified with the column density per velocity interval. In that case, the line intensity is proportional to the density of emitting material at each velocity (see Lazarian & Pogosyan 2000; EL05), and we can replace the velocity integrals by integrals along the actual LOS (\( z \)):

\[ C(X) = \frac{\int v_i I_{\text{line}}(X, v_i) \, dv_i}{\int \rho(x) \, dz}, \]

(3)

where \( \rho \) is the mass density and the limits of integration are now determined by the spatial extent along the LOS. The denominator ('normalization') in this definition makes it very difficult to provide a direct analytical treatment of the statistics, and its inclusion provides no significant improvement (see Levrier 2004; EL05). For that reason, it is useful to consider 'unnormalized centroids' as proposed in LE03:

\[ S(X) = \int v_i I_{\text{line}}(X, v_i) \, dv_i = \alpha \int v_i(x) \rho(x) \, dz, \]

(4)

where \( \alpha \) is the proportionality constant that relates the emissivity of the medium to its density (\( \alpha \) cancels in equation 3 because it appears both in the numerator and in the denominator).

2.2 The structure function of velocity centroids

Replacing \( x_1 = x \) and \( x_2 = x + r \), and similarly \( X_1 = X \) and \( X_2 = X + R \), the structure function of centroids can be written as

\[ \langle |S(X_1) - S(X_2)|^2 \rangle = \alpha^2 \int \int d^2 z \int d^2 r \left[ D(r) - D(0) \right]_{|X_1 = X_2|}, \]

(5)

where

\[ D(r) = \langle \rho(x_1)v_i(x_1) - \rho(x_2)v_i(x_2) \rangle^2. \]

(6)

The notation \(|X_1 = X_2|\) indicates that the integral is to be computed for a zero distance between \( X_1 \) and \( X_2 \) for the second term (i.e. varying only in \( z \)).

Equations (5) and (6) are exact; the double integrals are limited by the size of the object in the LOS (\( z \)) direction. If the density and velocity fields are isotropic (which might not be true in the presence of a strong mean magnetic field), one can reduce the double integrals to a single integral by changing variables from \( z_1 \) and \( z_2 \) to \( z = (z_1 + z_2)/2 \) and \( z_2 = z_2 - z_1 \), for details see appendix B in EL05.

Presenting the density and velocity fields as a mean plus a fluctuating part (\( \rho = \rho_0 + \bar{\rho}, v = v_0 + \bar{v} \)), and approximating the fourth-order moments as a combination of second-order moments:

\[ \langle h_1 h_2 h_3 h_4 \rangle \approx \langle h_1 h_2 \rangle \langle h_3 h_4 \rangle + \langle h_1 h_3 \rangle \langle h_2 h_4 \rangle + \langle h_1 h_4 \rangle \langle h_2 h_3 \rangle, \]

one can further decompose the structure function of velocity centroids as (see LE03; EL05):

\[ \langle |S(X_1) - S(X_2)|^2 \rangle \approx I_1(R) + I_2(R) + I_3(R) + I_4(R). \]

(7)

with

\[ I_1(R) = \alpha^2 \langle v^2 \rangle \int \int d^2 z \int d^2 r \left[ \bar{\rho}(r) - \bar{\rho}(0) \right]_{|X_1 = X_2|}, \]

(8a)

\[ I_2(R) = \alpha^2 \langle \rho^2 \rangle \int \int d^2 z \int d^2 r \left[ \bar{\rho}(r) - \bar{\rho}(0) \right]_{|X_1 = X_2|}, \]

(8b)

\[ I_3(R) = -\frac{1}{2} \alpha^2 \int \int d^2 z \int d^2 r [\bar{\rho}(r) - \bar{\rho}(0)]_{|X_1 = X_2|}, \]

(8c)

\[ I_4(R) = \alpha^2 \int \int d^2 z \int d^2 r \left[ c(r) - c(0) \right]_{|X_1 = X_2|}. \]

(8d)

where we have made use of the 3D structure functions of the density and of the LOS velocity:

\[ \bar{\rho}(r) = \langle \bar{\rho}(x_1) - \bar{\rho}(x_2) \rangle^2, \]

(9a)

\[ \bar{v}(r) = \langle \bar{v}(x_1) - \bar{v}(x_2) \rangle^2. \]

(9b)

The remaining density–velocity cross-correlations have been grouped into

\[ c(r) = 2 \langle \bar{v}(x_1) \bar{\rho}(x_2) \rangle^2 - 4 \rho_0 \langle \bar{\rho}(x_1) \bar{v}(x_2) \rangle^2. \]

(10)

The double integrals in equations (7) and (8) have been evaluated in this paper using the isotropy condition (i.e. changing to variables \( z_1, z_2 \) as mentioned above). However, some of the terms, such as \( I_1 \) and \( I_2 \) can be computed independently from 2D maps:

\[ I_1(R) = \alpha^2 \langle v^2 \rangle \left[ \int \rho(x) \, dz - \int \rho(x) \, dz \right] \]

(11a)

\[ I_2(R) = \alpha^2 \langle \rho^2 \rangle \left[ \int v(x) \, dz - \int v(x) \, dz \right]. \]

(11b)

For all the data sets used in the following sections, we have calculated \( I_1 \) and \( I_2 \) from 3D structure functions (equations 8a and 8b), and also from 2D structure functions (equations 11a and 11b), finding always an excellent agreement between the two. This indicates that the LOS integration procedure of the 3D structure functions is not affected by numerical noise from the discretization or by any artificial anisotropy.

Note that all of the terms in equations (7) and (8) can be measured in synthetic data sets or numerical simulations. However, only the structure function of centroids \( \langle |S(X_1) - S(X_2)|^2 \rangle \) and \( I_1 \) are accessible in observational data (PPV cubes). The centroid map is computed from PPV data using equation (4). The structure function (or any other measure such as the correlation function, power spectra, or the \( \Delta \)-variance spectra) can be calculated from such a map.

To compute \( I_1 \), we can interchange the integration along the LOS (\( z \)) variable in equation (11a), with the LOS velocity (\( v_i \)) and obtain \( I_1 \) as the structure function of a 2D map of integrated line intensity, which corresponds to the structure function of column density in case of an optically thin medium with constant \( \alpha \). Actually, the scaling laws of the integrated line intensity give a very robust measure of the properties of the density fluctuations. If the medium is
optically thin, the scaling of the density fluctuations can be obtained from the fluctuations of the column density.

3 GENERAL METHODOLOGY

In previous works (LE03, EL05 and OELS05), we have studied when velocity centroids can be used to retrieve the scaling properties of the underlying turbulent velocity field from spectroscopic observations. The main focus of this paper is to investigate in more detail the impact of the cross-term (equations 8c, A4c and A6d) and density–velocity cross-correlations (equations 8d, A4d and A6e) on the statistics of velocity centroids.

For simplicity, we will concentrate on the structure function description, because $\Delta$-variance has some advantages when applied to observational data, while in this work we use idealized data sets. Expressions in terms of correlation functions and power spectra can be found in EL05, and in terms of $\Delta$-variance can be found in Appendix A.

Our general method and the data sets we use are explained below.

3.1 Data sets

3.1.1 Density data cubes

We model the underlying density by fractional Brownian motion (fBm) fractal structures, that can be characterized by a spectral index $n$ (the log–log slope of its power-law power spectrum) and random Fourier phases (see Stutzki et al. 1998; Bensch et al. 2001). Our data cubes have a resolution of $128^3$ pixels, unless specified otherwise. All density data cubes are constructed with $n = -11/3$, which corresponds to the Kolmogorov index and that falls within the range found in observations. As discussed in detail in OELS05, fBm data cubes have a Gaussian distribution with a mean of zero (thus, they include negative values). We will refer to this original data with zero mean as ‘DATA0’. However, since negative mass density is unphysical, it is necessary to modify the data, but at the expense of distorting the idealized scaling properties of the fBms, in particular their probability distribution function (PDF). We remove the negative densities in two ways: we either add a constant and discard (set to zero) all negative values, or we exponentiate the whole fBm data cube (see OELS05 for more details about the impact of this procedures). Throughout this paper, we will refer to the data in which the mean is shifted by adding a constant as ‘DATA1’, and the exponentiated data as ‘DATA2’. When constructing the density cubes, we keep a dispersion $\sigma_\rho = 1$, and control the ratio of the density dispersion to the mean density ($\sigma_\rho/\rho_0$). The value of this ratio determines the magnitude of the cross-term $I3$ with respect to $I2$ (see OELS05), but it changes as well the level of the density–velocity cross-correlations ($I4$). We explore three different values of that ratio: 0.1, 1.0 and 3.0.

In Fig. 1, we present examples of histograms for the different density data sets. It is clear that, for both DATA1 and DATA2, an increase in the $\sigma_\rho/\rho_0$ ratio translates in a larger departure of density from a Gaussian PDF.

Figure 1. Density histograms: in panel (a) for DATA1, created by shifting the original fBm that was used to derive the density (DATA0) and use it as if it would correspond to the LOS velocity. These velocity fields are maximally correlated with density in the sense that they have the same Fourier phases. We will refer to them as ‘MAX’. They have a Gaussian PDF.

3.1.2 Velocity data cubes

For the velocity fields, we devised two ways of introducing correlations with the density data cubes. The simplest approach is to take the original fBm that was used to derive the density (DATA0) and use it as if it would correspond to the LOS velocity. These velocity fields are maximally correlated with density in the sense that they have the same Fourier phases. We will refer to them as ‘MAX’. They have a Gaussian PDF.

In reality, however, the correlation of velocity and density comes through physical processes, such as gravity. In an attempt to construct correlations similar to those that self-gravity would induce we followed a simple procedure: we compute the gravitational pull that would be caused by each density data cube (DATA0, DATA1 and DATA2) by solving Poisson’s equation using spectral methods (fast Fourier transforms, hereafter FFTs). We then assign to each pixel a velocity that points in the direction of the gradient of the gravitational potential and whose magnitude is proportional to the density at that particular point. Finally, we use only the velocity component along the LOS. We will term this artificially correlated velocities ‘1ST’; their PDF is not Gaussian.

We recognize these procedures as very rudimentary and even ad hoc. However, they serve the purpose of quickly generating full ensembles of data in a systematic way, The level of cross-correlations achieved in these procedures can be later compared with numerical

---

2 For optically thick media, it was shown by Lazarian & Pogosyan (2004) that the integrated self-absorbing line profiles give the scaling of the density fluctuations, provided that the density spectrum is shallow (i.e. $d\rho \sim n^\gamma$, with $\gamma < 0$).
In this paper, we are less interested in obtaining the quantitative spectral index of centroids for different density and velocity scalings but we perform a qualitative study of the dominant contributions including the role of cross-correlations of density and velocity, testing some of the assumptions we have used before.

3.2 Calculation and conventions in the figures

We obtain all the structure and correlation functions necessary to calculate the terms in equation (7), and in consequence all terms in equations (8) using FFTs. As explained in more detail in EL05, the use of FFT is well justified for the synthetic data used here (Cartesian grid, periodic boundaries, no instrumental noise). However, for real observational data it is preferable to compute such functions directly in real space, or use for instance the Δ-variance instead (as in OELS05).

To improve our statistics, we have computed all the quantities for 10 different realizations (different seeds to generate the fBms). The average of the 10 realizations is what we will show in the figures, unless specified otherwise.

In order to facilitate the distinction of the different data sets to the reader, we put labels in the title of all individual plots, signalling the density and the velocity data types. For instance, ‘DATA1-0TH’ corresponds to the average of 10 realizations of data in which the density distribution had been shifted by addition of a constant and the negative values set to zero, combined with an uncorrelated Gaussian velocity field. We include also the value of $\sigma_p/\rho_0$ in the title of each plot. Since $I_3$ and $I_4$ can be negative, and negative numbers cannot be plotted in the log–log scale, we only plot their magnitude and, in order to discriminate the sign, we join the symbols with lines when the values are positive. The symbols remain disconnected when the quantity is negative.

The magnitude of the lag ($R$ for 2D statistics, or $r$ for 1D and 3D statistics) is given in units relative to the size of the computational box $L$ (unless we specify otherwise $L = 128$). For simplicity, we use the following convention to label the structure function of centroids: $\langle (S_1 - S_2)^2 \rangle = \langle (S(X_1) - S(X_2))^2 \rangle$.

4 RESULTS

4.1 Density–velocity cross-correlations in Gaussian data

To study the effects of density–velocity cross-correlations alone, that is, independent of possible consequences of the data not being Gaussian, we computed the centroids and all terms in equation (7) using the density set DATA0, combined with velocity fields 0TH and MAX. All these fields have a Gaussian distribution. As mentioned in the previous section, the velocity sets 0TH are uncorrelated with density, while MAX have extreme density–velocity correlations.

In Fig. 3, we plot the results. In each panel, we include the individual contribution of each term of equation (7) along with the result of adding $I_1 + I_2 + I_3 + I_4$ (dashed line with ‘x’ symbols). The difference of this curve and the structure function of centroids ($\langle (S_1 - S_2)^2 \rangle$ (solid line) can help estimate the overall error of the decomposition in equation (7). If the two curves deviate notably from each other, other terms, neglected here, must play a significant role (see Section 4.4). We must stress that the only quantities that can be obtained from observations are $\langle (S_1 - S_2)^2 \rangle$ and $I_1$.

In uncorrelated data shown in the first row in Fig. 3 (panels a, b and c), we have an excellent agreement of the decomposition of the structure function of velocity centroids ($I_1 + I_2 + I_3 + I_4$) and the structure function of centroids. For maximally correlated data...
Figure 3. Structure functions for Gaussian density (DATA0) with uncorrelated velocity (OTH in panels a, b and c), and highly correlated velocity (MAX – in panels d, e and f) fields. The $\sigma_\rho/\rho_0$ ratio increases from the left-hand to right-hand side, as indicated in the title of each plot.

4.1.1 Spurious correlations

For uncorrelated data, the two terms in $I_4$ (equation 10) should be each equal to zero: if no correlations between density and velocity are present, $(\bar{v}(x_1)\bar{\rho}(x_2))^2 = (\bar{v}(x_1))^2(\bar{\rho}(x_2))^2$, $(\bar{\rho}(x_1)\bar{v}(x_1)\bar{v}(x_2)) = (\bar{\rho}(x_1))(\bar{v}(x_1))\bar{v}(x_2)$, where by definition $(\bar{\rho}(x)) = \langle \rho(x) \rangle = 0$. Some level of incidental correlations is unavoidable when creating data sets using a random number generator. That is the reason why we decided from the beginning to average 10 individual realizations for each figure, to improve our statistics.

To investigate in more depth the contribution of cross-correlations, let us split the two types of cross-correlations that appear in $I_4$ as

$$cc1(r) = 2(\bar{v}(x_1)\bar{\rho}(x_2))^2,$$

$$cc3(r) = -4\rho_0(\bar{\rho}(x_1)\bar{v}(x_1)\bar{v}(x_2)),$$

and calculate them for a larger number of realizations. To illustrate how the statistics change with a much larger ensemble, we did the experiment for 300 independent realizations of DATA0 and uncorrelated (OTH) velocity, all with $\sigma_\rho/\rho_0 = 1$.

The results are presented in Fig. 4. In the first two panels (a and b), we present the correlation and structure functions before the LOS integration. The last panel (c) shows the resulting $I_1$, $I_2$ and $I_4$. In the figure, we present the quantities for each individual realization. Since $I_4$ can be negative, this time the vertical axis is linearly scaled here, not logarithmically. We can see from the figure that the scatter of $I_1$ and $I_2$, for all the different realizations is marginal (and the same holds for $I_3$, which we have calculated, but is not shown to avoid confusion). This is not the case for $I_4$, that shows a rather large dispersion.

An easier way to visualize the variability of $I_4$ is to reduce the problem to one dimension. For this experiment, we produce 300 1D fBm vectors (with $L = 1000$ pixels), and we kept the same $\sigma_\rho/\rho_0$ of unity and did not remove negative densities. In one dimension, the density and velocity structure functions, the cross-terms $cc1$ and $cc3$ depend on the separation $r$, but the integrated quantities ($I_1$, $I_2$ and $I_3$) are practically identical, i.e. they are independent of $r$. The case of $I_4$ is different, and it is more meaningful to look at 2D fBm vectors, and calculate the statistics in the 2D case.
The cross-correlation terms $cc_2$ and/or with $cc_3$ before integration (panels a and b) from 1D data with $\sigma_w$ the 3D counterparts of $f_1$ and $f_2$. The lines show the averages of 300 realizations of DATA0 with an uncorrelated velocity (0TH). On top of the averages, the individual results are plotted as scattered points to show the variation in $cc_1$ and $cc_3$. Evidently, these two terms are much more uncertain compared to the structure functions of $\rho$ and $v$. The averaging procedure reduces the random fluctuations. In panel (c), the resulting 2D quantities ($f_1$, $f_2$ and $f_4$) are plotted. It shows that each realization of $f_4$ has large uncertainty, but on the average it is close to zero.

Fig. 4 illustrates the result after averaging 30 individual realizations; Fig. 5(b) with 300 realizations. In the last panel (Fig. 5c), we have called $f_4$ the portion of $f_4$ that stems from $cc_1$, and $f_4$ the portion that stems from $cc_3$. The terms containing $cc_3$ decrease with larger ensembles, while $cc_1$ saturates at a non-zero number. The explanation is trivial: $cc_1$ is positive defined, therefore any incidental correlations add up instead of average out, as in the case of $cc_3$. It is also notable that the correlations $cc_1$ and $cc_3$ on their own are small compared to $d_v$ and $d_r$, but $f_4$ can become large relative to $f_1$ and $f_2$ due to the LOS integration. We have carefully checked the results for $f_1$ and $f_2$ from 2D and 3D structure functions, and found an excellent agreement. Thus, we must stress that the problem is not the algorithm of integration, but the fact that we are integrating at all.

Since it is not practical to produce such large ensembles of data sets for every combination of parameters, we stick to the average of 10 independent realizations for each combination of parameters. This can be partially justified by the fact that, except for the smallest scales, $f_4$ is typically smaller than the velocity term $f_2$. If they become comparable, $f_3$ is also large so that if $f_4$ would dominate the statistics of the centroids, those were already unreliable due to $f_3$. With these limitations in mind, we proceed to study how the centroids change with $\sigma_\rho/\rho_0$ and/or with the level of artificial correlations introduced when the data are not Gaussian.

4.2 Non-Gaussianity in uncorrelated data

We now investigate the impact of non-Gaussianity of the data on the statistics of centroids. Strictly speaking, we will study the effect of data with non-negligible third- and fourth-order moments of the density and velocity PDFs, but for simplicity we will refer to these data simply as non-Gaussian.

The non-Gaussianity can be introduced in either the density field, the velocity field, or both. On the one hand, the density is naturally non-Gaussian; in fact, one should expect something closer to a lognormal distribution, such as the DATA2 sets (see for instance Padoan, Jones & Nordlund 1997; Beresnyak, Lazarian & Cho 2005). On the other hand, from a simplistic view of ‘pure turbulence’ one can expect the velocity to be more or less Gaussian, although real turbulence exhibits departures from Gaussianity, for instance in the form of intermittency (see Lazarian 2006b). More departures from Gaussianity in nature are likely to occur with the aid of additional phenomena such as self-gravity. In fact, our artificially correlated velocity in the set 1ST, shown in Fig. 2(c) is a rather extreme case; even at scales when self-gravity becomes important the observed line profiles are often more Gaussian (Ossenkopf, Klessen & Heitsch 2001) than those obtained with our correlating procedure.
In Figs 6 and 7, we show the results of combining uncorrelated non-Gaussian density and velocity fields to construct spectral lines and their centroids.

Combinations of Gaussian velocity (0TH) and uncorrelated non-Gaussian density fields are shown in Fig. 6. In the first row (panels a–c), the density corresponds to DATA1 and in the second row (panels d–f) to DATA2. In all the cases, $\sigma_\rho/\rho_0$ increases from the left-hand to right-hand side. In all panels of Fig. 6, we find a very good agreement of the structure function of centroids ($\langle S_1 - S_2 \rangle^2$) and its decomposition in terms of $I^1 + I^2 + I^3 + I^4$. We see, however, that for $\sigma_\rho/\rho_0 = 3$ (Figs 6c and f) the slope of velocity ($I^2$) does not coincide with that of $\langle S_1 - S_2 \rangle^2$, because the cross-term $I^3$ dominates. In all cases, the cross-correlations $I^4$ are negligible compared with the rest of the quantities plotted. It is notable that the results with non-Gaussian density are slightly more favourable than those obtained with Gaussian density (Fig. 3). The reason is that the non-zero mean density gives more weight to the $I^2$ term because the factor $(\rho^2) = \rho_0^2 + \sigma_\rho^2$ in equation (8b) becomes larger (a more detailed discussion of this issue can be found in OELS05).

The results with non-Gaussian velocity fields (but uncorrelated with density), are presented in Fig. 7. Gaussian density (DATA0-IND) is used for the panels in the first row (panels a–c) of Fig. 7; the rest of panels correspond to non-Gaussian density [DATA1-IND (panels d–f) and DATA2-IND panels (g–i)].

For $\sigma_\rho/\rho_0 = 0.1$ (panels a, d and g), we find an excellent agreement between the structure function of centroids and the sum of the terms of their decomposition, which in this case is clearly dominated by $I^2$. Thus, the unnormalized velocity centroids (MVCs as well) trace directly the scaling of the turbulent velocity.

For $\sigma_\rho/\rho_0 = 1$ (panels b, e and h), we find a reasonably good agreement [although to a lesser degree in (h)] between $\langle (S_1 - S_2)^2 \rangle$ and $I^1 + I^2 + I^3 + I^4$. At small scales, we can also see that both $I^3$ and $I^4$ are smaller than the rest of the terms. However, $I^1$ is comparable in magnitude with $I^2$, meaning that it is possible to retrieve the scaling of velocity (using the fluctuations at small scales), but one would need to use MVCs instead of unnormalized centroids directly.

For $\sigma_\rho/\rho_0 = 3$ (panels c, f and i), the situation is less encouraging because of the discrepancy between the structure function of centroids and the sum of the terms in the decomposition. This difference is small in panels (c and f) (it can also be noted in panel h), but is quite prominent in panel (i). For $\sigma_\rho/\rho_0 = 3$, we also see that the cross-correlations remain small, but the cross-term $I^3$ is not negligible. Clearly, in panel (i) both unnormalized centroids and MVCs would fail to trace the velocity scaling properties. For panels (c) and (f), unnormalized centroids would also fail to trace velocity, but is not clear for MVCs.

### 4.3 Correlated non-Gaussian data

Finally, we perform the decomposition proposed in LE03 and EL05 with correlated density and velocity fields. For this experiment, we use the artificially correlated velocity data cubes 1ST, with the various density data cubes (DATA0, DATA1 and DATA2). The results are shown in Fig. 8.

Fig. 8 shows a similar behaviour to the other non-Gaussian combinations of parameters; however, with a much larger discrepancy between the structure function of centroids and $I^1 + I^2 + I^3 + I^4$ for the least-Gaussian cases (cf. Figs 8c, f, h and i). It is evident that the correlated velocity fields used here (1ST) produce a very prominent $I^4$ term, which distorts the slope of $\langle (S_1 - S_2)^2 \rangle$ relative to the velocity scaling in $I^2$, even for $\sigma_\rho/\rho_0 = 0.1$ (panels a, d and g) where $I^2$ is the dominant term. In this case, centroids (and MVCs) would trace the scaling of velocity with an acceptable accuracy. However, for the rest of the cases either all the terms in the decomposition ($I^1 + I^2 + I^3 + I^4$) are close in magnitude (panels b, c and e), or the error of the decomposition and $\langle (S_1 - S_2)^2 \rangle$ are too big (panels f, h and i), meaning that both centroids and MVCs would not be useful to recover the velocity spectral index.
4.4 Validity of the decomposition for highly non-Gaussian data

The large discrepancies observed for the centroids from the least-Gaussian data cubes can be worrisome. As mentioned earlier, the decomposition of the structure function (as well as the correlation function and the Δ-variance) presented in Section 2.2 is only as accurate as the fourth-order moments and can be split into second-order ones, which for Gaussian fluctuations is exact, but it is not the general case. We must stress that velocity fluctuations might have a Gaussian distribution (or at least close to it), but density fluctuations certainly do not.

The discrepancies between the structure function of centroids \( \langle (S_1 - S_2)^2 \rangle \) and \( I_1 + I_2 + I_3 + I_4 \) are only notable for \( \sigma_p/\rho_0 \geq 1 \). Such discrepancies are most recognizable in Figs 3(e), 3(f), 7(f), 8(f), 8(h) and 8(i).

It is notable that these differences are more evident for correlated density and velocity fields, and they are present even in the case of a density structure with Gaussian fluctuations (DATA0). We can compare for instance Figs 3(b) and (c), where density and velocity are uncorrelated, and no discrepancy is seen, with Figs 3(e) and (f) where the density and velocity are correlated (MAX) and the error is obvious. For uncorrelated density and velocity, the error is only seen for the least-Gaussian fields (DATA2-IND, see Fig. 7i). Even for a highly non-Gaussian density combined with an uncorrelated Gaussian velocity (Fig. 6), we have a very good agreement. The most pronounced errors in the decomposition of the centroid statistics is found for correlated, non-Gaussian data, such as in Figs 8(f), (h) and (i).

Thus, it is clear that at some point, mainly due to a combination of the cross-correlations and the non-Gaussianity of the velocity fields, the decomposition presented in equations (7) and (8) fails.

To investigate further on such a discrepancy, let us consider the ‘exact’ decomposition of the centroid structure function, that is, before splitting the fourth-order moments into second-order ones.
Figure 8. Decomposition of velocity centroids with correlated non-Gaussian velocity (1ST) with: Gaussian density (DATA0, panels a, b and c), and non-Gaussian density (DATA1, panels d–f) and DATA2 (panels g–i).

(see EL05):
\[
\langle [S(X_1) - S(X_2)]^2 \rangle = \alpha^2 \varrho_0^2 \int \! \! \int d\tilde{z}_1 d\tilde{z}_2 [d_{\rho}(r) - d_{\rho}(r)]_{X_1 = X_2}
\]
\[
+ \alpha^2 \rho_0^2 \int \! \! \int d\tilde{z}_1 d\tilde{z}_2 [d_{\rho}(r) - d_{\rho}(r)]_{X_1 = X_2}
\]
\[
+ \alpha^2 \varrho_0 \int \! \! \int d\tilde{z}_1 d\tilde{z}_2 \left[ cc_{3}(r) - cc_{3}(r) \right]_{X_1 = X_2}
\]
\[
- 2 \alpha^2 \int \! \! \int d\tilde{z}_1 d\tilde{z}_2 \left[ p(r) - p(r) \right]_{X_1 = X_2},
\]
\[\tag{13}\]
where
\[
p(r) = \langle \tilde{b}(x_1) \tilde{b}(x_2) \rangle \tilde{b}(x_2) \tilde{b}(x_2)
\]
\[\tag{14}\]
or, equivalently, setting \(v_0 = 0\) (it always exists in a reference frame in which this is true):
\[
\langle [S(X_1) - S(X_2)]^2 \rangle = \frac{\rho^2}{\rho_0^2} I2(R) + I43(R) + I5(R).
\]
\[\tag{15}\]
with
\[
I5(R) = \alpha^2 \left( v_1^2 \right) \int \! \! \int d\tilde{z}_1 d\tilde{z}_2 [p(r) - p(r)]_{X_1 = X_2}.
\]
\[\tag{16}\]
The last quantity introduced \((p)\) can only be obtained from the 3D statistics and not from projected maps. In other words, we can calculate it in our simulations, but not for observational data.

In Fig. 9, we present the exact decomposition for the example of correlated non-Gaussian data (the same data in the last column of Fig. 8, panels c, f and i, with \(\sigma_{\rho}/\rho_0 = 3.0\)). The solid line represents the structure function of the centroid map, the dashed line represents the sum in equation (15). We see immediately that the discrepancy of the two is marginal, and that the dominant term is \(I5\) in every case. This suggests that indeed the cause of the previous discrepancy is the lack of higher order moments in the decomposition as presented in LE03. However, this lack of higher order moments is only problematic when the \(I3\) term already contaminates the statistics of centroids (i.e. \(\sigma_{\rho}/\rho_0 \gtrsim 1\)).

5 SUMMARY

In this paper, we have made a follow-up to the work presented in LE03, EL05 and OELS05. With the goal to extract the spectral index of the turbulent velocity field from measured centroid maps, a decomposition of the structure function of centroids was presented in LE03 and studied in more detail in EL05. Such a decomposition was possible assuming that fourth-order moments could be split into second-order moments, and resulted basically into four different
terms: a contribution of the column density ($I_1$), a contribution of the integrated velocity ($I_2$, the quantity that one would have to isolate to get the velocity spectral index), a cross-term ($I_3$), and a term with density–velocity cross-correlations ($I_4$). Later in OELS05, it was proposed that $\sigma I/\rho_0$ was an important parameter that determines whether velocity centroids can retrieve the velocity scaling. In this paper, we combine fractal density cloud maps and velocity data cubes with ad hoc trial density correlation study to $I_3$ and $I_4$. We found the following.

(i) Centroids trace the scaling properties of turbulence if $\sigma I/\rho_0 \lesssim 1$. This is true in spite of cross-correlations and/or non-Gaussian nature of the fluctuations of density and velocity. Such small density dispersion can be identified with subsonic turbulence.

(ii) We found that in increasing order of importance the following effects make velocity centroids unreliable to recover the velocity scaling properties: cross-correlations of density and velocity (although we use extremely correlated fields), velocity non-Gaussianity, and density non-Gaussianity. All these effects are non-negligible for $\sigma I/\rho_0 > 1$.

(iii) $I_3$ was found to correlate with $\sigma I/\rho_0$ as expected by OELS05. For values of this ratio greater than unity, the $I_3$ term is comparable to $I_2$ contaminating the desired velocity statistics. The effect of this term has been explored in OELS05, and for power-law statistics its dependence on the spectral index of density (which can be measured straightforwardly) can be found in appendix E of EL05.

(iv) $I_4$ reflects the level of density–velocity cross-correlations. However, it was found to have a rather erratic behaviour. We attribute this mainly to our limited statistical sample and to our rudimentary generation of correlated fields. However, $I_4$ did not contaminate the velocity statistics when $\sigma I/\rho_0$ was small ($\sim 0.1$).

(v) Our results show a problem in retrieving the scaling properties of velocity from centroids (for $\sigma I/\rho_0 > 1$) that lies in high-order moments which are introduced by high-density–velocity cross-correlations. These high-order moments were not considered in our earlier work, in which fourth-order moments were approximated by a combination of second-order ones. For extremely correlated data (coinciding with the least-Gaussian data sets as well), we found that such an approximation is not accurate. However, in the cases that our proposed decomposition failed, the term that has useful information of velocity was already very small with respect to all other terms, including those we have studied in previous work.

All these findings are consistent with the criteria for deciding when centroids trace velocity fluctuations provided in previous works: $\langle \tilde{S}^2 \rangle / \langle \tilde{v}^2 \rangle / \langle I^2 \rangle \gg 1$ from LE03 and confirmed in EL05, or $\sigma I/\rho_0 < 1$ proposed in OELS05. Our results reinforce the notion that velocity centroids trace the velocity scaling for subsonic turbulence, but not for supersonic turbulence.

ACKNOWLEDGMENTS

AE acknowledges financial support from Mexico’s Consejo Nacional de Ciencia y Tecnología and the ‘Macroproyecto de Tecnologías para la Universidad de la Información y la Computación’ (Secretaría de Desarrollo Institucional de la UNAM, Programa Transdisciplinario en Investigación y Desarrollo para Facultades y Escuelas, Unidad de Apoyo a la Investigación en Facultades y Escuelas). AL acknowledges the NSF grant AST-0307869 and the support from the Centre for Magnetic Self-Organization in Laboratory and Astrophysical Plasmas. The whole project was supported through the Deutsche Forschungsgemeinschaft SFB 494A.

REFERENCES

Deshpande A. A., Dwarkanath K. S., Goss W. M., 2000, Apj, 543, 227
Kleiner S. C., Dickman R. L., 1985, Apj, 295, 466
Lazarian A., 2006a, AIPC, 874, 301
Lazarian A., 2006b, Int. J. Modern Physics D., 15, 1099
Lazarian A., Pogosyan D., 2000, Apj, 537, 72
Lazarian A., Pogosyan D., 2004, Apj, 616, 943
Lazarian A., Pogosyan D., 2006, Apj, 652, 1348
Lazarian A., Pogosyan D., Vázquez-Semadeni E., Pichardo B., 2001, Apj, 555, 130
Miesch M. S., Bally J., 1994, Apj, 429, 645

© 2007 The Authors. Journal compilation © 2007 RAS. MNRAS 381, 1733–1744

Downloaded from https://academic.oup.com/mnras/article-abstract/381/4/1733/972713 by University of Groningen user on 19 April 2018
A similar decomposition to that of equation (7), in terms of $\Delta$-variance, can be obtained considering the correlation function of un-normalized centroids (see Esquivel & Lazarian 2005; Levrier 2004):

$$\left\langle S(X_1)S(X_2) \right\rangle \approx \left\langle \rho^2 \left( v_2^2 \right) \right\rangle (\sigma_{z_{\alpha 2}})^2 + B1(R) + B2(R) + B3(R) + B4(R),$$  \hfill (A3)

with

$$B1(R) = \sigma^2 \rho_0^2 \int \int \langle \rho(x_1) \rho(x_2) \rangle \, dx_1 \, dx_2,$$  \hfill (A4a)

$$B2(R) = \sigma^2 \rho_0^2 \int \int \langle \tilde{v}(x_1) \tilde{v}(x_2) \rangle \, dx_1 \, dx_2,$$  \hfill (A4b)

$$B3(R) = \sigma^2 \int \int \langle \rho(x_1) \rho(x_2) \rangle \langle \tilde{v}(x_1) \tilde{v}(x_2) \rangle \, dx_1 \, dx_2,$$  \hfill (A4c)

$$B4(R) = -2 \sigma^2 \int \int \langle \rho(x_2) \tilde{v}(x_2) \rangle \, dx_1 \, dx_2$$

$$+ 2 \sigma^2 \rho_0 \int \int \left[ \langle \rho(x_1) \tilde{v}(x_1) \tilde{v}(x_2) \rangle \right] \, dx_1 \, dx_2,$$  \hfill (A4d)

yielding

$$\sigma_2^2(r) = A0 + A1(R) + A2(R) + A3(R) + A4(R),$$  \hfill (A5)

where

$$A0 = \left\langle \rho^2 \right\rangle \left\langle v_2^2 \right\rangle (\sigma_{z_{\alpha 2}})^2 \left( \bigodot_{\alpha}(0) \right)^2,$$  \hfill (A6a)

$$A1(R) = \rho_0^2 \int P_{2D,I}(k) \left( \bigodot_{\alpha}(k) \right)^2 \, d^2k,$$  \hfill (A6b)

$$A2(R) = \sigma^2 \rho_0^2 \int P_{2D,V}(k) \left( \bigodot_{\alpha}(k) \right)^2 \, d^2k,$$  \hfill (A6c)

$$A3(R) = \int J(B3(R)) \left( \bigodot_{\alpha}(k) \right)^2 \, d^2k,$$  \hfill (A6d)

and

$$A4(R) = \int J(B4(R)) \left( \bigodot_{\alpha}(k) \right)^2 \, d^2k.$$  \hfill (A6e)