ORIGINALS

Prakash Vedula · R. J. Adrian

Optimal solenoidal interpolation of turbulent vector fields: application to PTV and super-resolution PIV

Abstract A new approach for the interpolation of a filtered turbulence velocity field given random point samples of unfiltered turbulence velocity data is described. In this optimal interpolation method, the best possible value of the interpolated filtered field is obtained as a stochastic estimate of a conditional average, which minimizes the mean square error between the interpolated filtered velocity field and the true filtered velocity field. Besides its origins in approximation theory, the optimal interpolation method also has other advantages over more commonly used ad hoc interpolation methods (like the 'adaptive Gaussian window'). The optimal estimate of the filtered velocity field can be guaranteed to preserve the solenoidal nature of the filtered velocity field and also the underlying correlation structure of both the filtered and the unfiltered velocity fields. The a posteriori performance of the optimal interpolation method is evaluated using data obtained from high-resolution direct numerical simulation of isotropic turbulence. Our results show that for a given sample data density, there exists an optimal choice of the characteristic width of cut-off filter that gives the least possible relative mean square error between the true filtered velocity and the interpolated filtered velocity. The width of this 'optimal' filter and the corresponding minimum relative error appear to decrease with increase in sample data density. Errors due to the optimal interpolation method are observed to be quite low for appropriate choices of the data density and the characteristic width of the filter. The optimal interpolation method is also seen to outperform the 'adaptive Gaussian window', in representing the interpolated field given the data at random sample locations. The overall a posteriori performance of the optimal interpolation method was found to be quite good and hence makes a potential candidate for use in interpolation of PTV and super-resolution PIV data.

1 Introduction

Particle tracking velocimetry (PTV) measures the velocities of sparse, individual particles located at random points in the flow; super-resolution particle image velocimetry (PIV) (Keane et al. 1995) achieves the same type of measurements in higher concentrations of particles by first estimating the vector field with a multi-particle correlation algorithm. In each case, an important post-processing step is the interpolation of the random point samples of the velocity vector field onto a uniform grid, or onto a continuous function. Interpolation of randomly sampled, random data is interesting because it is not subject to the requirements of Nyquist's criterion, but it may introduce significant distortion of the signal (Adrian and Yao 1987; Mueller et al. 1998; Benedict et al. 2000). Discontinuities in the interpolating function can introduce wide-band noise, and missing information about the small-scale fluctuations between interpolating points leads to attenuation of the measured spectrum.

Perhaps the most common method for interpolating PTV data is the 'adaptive Gaussian window' or AGW (Koga 1986; Agui and Jimenez 1987; Spedding and Rignot 1993). Letting $\mathbf{v}_p$ denote the velocity of the article 'p' that resides at $\mathbf{x}_p$ at time t, and $\mathbf{u}(\mathbf{x},t)$ denote the Eulerian velocity field, the velocity of a particle that follows the fluid motion with negligible slip is equal to a random point sample of the Eulerian field:

$$\mathbf{u}(\mathbf{x}_p, t) = \mathbf{v}_p(t). \tag{1}$$

P. Vedula (✉) · R. J. Adrian
Department of Theoretical and Applied Mechanics,
University of Illinois, Urbana, IL 61801, USA
E-mail: pvedula@uiuc.edu
E-mail: rjadrian@uiuc.edu

The AGW algorithm uses the set of samples $\{\mathbf{u}_p\}$ contained in a spatial window $W(\mathbf{x})$ to form the interpolated field:

$$\bar{\mathbf{u}}(\mathbf{x}, t) = \frac{\sum_{\{\mathbf{x}_p\} \in W(\mathbf{x})} \mathbf{u}(\mathbf{x}_p, t) G(\mathbf{x}_p - \mathbf{x})}{\sum_{\{\mathbf{x}_p\} \in W(\mathbf{x})} G(\mathbf{x}_p - \mathbf{x})}, \tag{2}$$

where $G(\mathbf{x})$ is typically a Gaussian function. (By letting G be a second order tensor, one can form the interpolation in such a way that the interpolated field is solenoidal, cf. Zhong et al. 1991.) Since G vanishes for large distances from x, the summations can be extended over the entire space of particles; restricting the summations to the window is just a computational convenience. Some insight is gained by expressing Eq. 2 in terms of the random point sample function:

$$g(\xi, t) = \sum_{\forall p} \delta(\xi - \mathbf{x}_p). \tag{3}$$

The interpolated field can be expressed as

$$\bar{\mathbf{u}}(\mathbf{x}, t) = \frac{\int \mathbf{u}(\xi, t) G(\xi - \mathbf{x}) g(\xi, t) \mathrm{d}\xi}{\int G(\xi - \mathbf{x}) g(\xi, t) \mathrm{d}\xi}, \tag{4}$$

from which one sees that the interpolated field is linear in u, but decidedly non-linear in the random locations of the samples. This causes a problem in interpreting **u** which cannot be expressed simply as a filtered form of the Eulerian field, **u**, due to the appearance of g and x in both the numerator and the denominator. If we tried to write **u** as a convolution with a filter response function, Eq. 4 implies that the impulse response would depend upon g, i.e. the filter would be a random function of the random particle positions. A second issue with the AGW and similar variations is simply that it is ad hoc. This is common to most interpolation schemes, but unsatisfying nevertheless. Another undesirable feature resulting from the ad hoc nature of these interpolation methods is that the correlations between the interpolated field and the randomly sampled data differ from the corresponding correlations between the underlying field (which is interpolated) and the randomly sampled data.

In order to address these limitations, we consider an 'optimal' interpolator on the mathematical basis that it is required to yield the least mean square error between the velocity field and its estimate based on the actual velocity samples at random locations. We can further generalize our definition of the mean square error function to represent the error (in the mean square sense) between some known (and useful) function of the velocity field and its estimate based on the actual velocity samples at random locations. For instance, an optimal estimate of the filtered velocity field (given the actual velocity samples) can be obtained by minimizing the mean square error between the filtered velocity field and its estimate. It is plausible that for any given filter function operating on the velocity field, there exists an optimal value of filter width parameter that gives the best possible estimate of the filtered velocity field, in the sense that the mean-square error (between the filtered field and its estimate) relative to the filtered signal, is a minimum. Note that the unfiltered velocity field can be considered as a special case of the filtered velocity field, with filter width parameter being zero. We represent a Cartesian component (along the x_I coordinate axis) of the filtered version of the Eulerian velocity field by $\tilde{u}_i(\mathbf{x}, t)$, which is defined as

$$\tilde{u}_i(\mathbf{x}, t) = \int h_{ij}(\xi, \mathbf{x}) \, u_j(\xi, t) \mathrm{d}\xi, \tag{5}$$

where h_{ij} denotes the filter impulse response function and u_j denotes a Cartesian component (along the x_j coordinate axis) of the unfiltered Eulerian velocity field. This form allows for inhomogeneous filters that depend explicitly on x. To preserve the solenoidal property of the filtered field, it is necessary to have a second order tensor for the filter impulse response, h_{ij}. Then, it is possible to construct the impulse response so that $\partial \tilde{u}_i / \partial x_i = \int \partial h_{ij}(\xi, \mathbf{x}) / \partial x_i u_j(\xi, t) \mathrm{d}\xi = 0$. For instance, the convolution of a Fourier cut-off filter function (defined in the following section) and the actual Eulerian velocity field results in a low-pass filtered velocity field that is solenoidal. We would like for our optimal estimate of the low-pass filtered velocity field to also inherit the solenoidal property of the underlying filtered field. It may be noted that not all filter response functions convolving with the actual Eulerian velocity field result in solenoidal (filtered) fields, in spite of the actual Eulerian velocity field being solenoidal. An inhomogeneous filter response function convolving with the actual velocity field can result in filtered fields that are non-solenoidal. In this paper, we focus our attention on the optimal interpolation of solenoidal filtered fields alone; although a generalization to non-solenoidal filtered fields can be carried out quite easily within the framework of our formulation.

An optimal estimate of a filtered Eulerian velocity field based on the random point samples of the velocity field can be obtained using the theory of stochastic estimation. For simplicity of exposition, we consider a random variable Y, which is to be estimated based on some known information, represented as a vector of random event data **E**. Let $y(\mathbf{E})$ be an estimate of Y. Of all possible estimates, the best estimate $y_b(\mathbf{E})$ is the one that has minimum error between the estimate and the random variable in the mean-square sense. It can be shown (cf. Papoulis 1984; Adrian 1996) that this best estimate $y_b(\mathbf{E})$ (in the mean-square sense) is given by the conditional average $\langle Y | \mathbf{E} \rangle$. Often it may not be possible to determine the best estimate because this conditional average may be difficult to obtain. Hence, we use stochastic estimation to constrain $y(\mathbf{E})$ in some fashion and require that it approximate the conditional average $\langle Y | \mathbf{E} \rangle$ as closely as possible. For instance, for linear stochastic estimation, we insist that $y(\mathbf{E})$ be linear in **E**.

If $\varepsilon_1 \equiv \left\langle |\langle Y \mid \mathbf{E} \rangle - y(\mathbf{E})|^2 \right\rangle$ denotes the mean-square error involved in the approximating conditional average and $\varepsilon_2 \equiv \left\langle |Y - y(\mathbf{E})|^2 \right\rangle$ denotes the mean-square error between the random variable and the estimate, one can show that these errors are related by the following identity: $\varepsilon_2 \equiv \varepsilon_1 + \left\langle |\langle Y \mid \mathbf{E} \rangle - Y|^2 \right\rangle$. From this it follows that a stochastic estimate that minimizes ε_2 also minimizes ε_1. For our purpose of obtaining an optimal interpolator, we minimize ε_2 instead of directly minimizing ε_1. A linear stochastic estimate that minimizes the mean-square error, ε_2, is obtained when $\langle Y \rangle = \langle y \rangle$ and $\langle E_i'(Y' - y') \rangle = 0$. Note that the primes denote fluctuations about the mean values, for instance, $Y' \equiv Y - \langle Y \rangle$. Given the data $\{\mathbf{x}_p\}, \{\mathbf{u}_p\}$ we seek to determine an estimator $\tilde{\mathbf{u}}$ that represents the filtered vector field with least mean square error:

$$e_i = \left\langle \left(\breve{\tilde{u}}_i - \tilde{u}_i \right)^2 \,\middle|\, \{\mathbf{u}_p\}, \{\mathbf{x}_p\} \right\rangle, \quad i = 1,2,3. \tag{6}$$

The exact solution to this minimization problem can be shown to be

$$\breve{\tilde{u}}_i(\mathbf{x}, t) = \langle \tilde{u}_i(\mathbf{x}, t) \mid \{\mathbf{u}_p\}, \{\mathbf{x}_p\} \rangle, \tag{7}$$

the conditional average of the turbulent velocity field given the values of the random velocity samples and the points at which they were measured (cf. Papoulis 1984; Adrian 1996). Equation 7 averages over all possible velocity fields that are consistent with the measured data. Note that Eq. 7 is linear in u, but it may be non-linear in both the position data and the velocity data.

Direct use of the exact solution given by Eq. 7 is infeasible because the space of velocity-position data is huge. A typical two-dimensional PTV experiment might have of order 10^4 two-dimensional velocity samples corresponding to a 10^8 dimensional function. As in the optimal algorithm numerical method of Adrian (1977) for the turbulent Navier–Stokes equations, practical implementation of the optimal method requires approximation. The linear stochastic estimate of the conditional average in Eq. 7 is found by expanding the conditional average as a power series in the velocity data and truncating at first order:

$$\hat{\tilde{u}}_i = \text{linear estimate} \langle \tilde{u}_i(\mathbf{x}, t) \mid \{\mathbf{u}_p\}, \{\mathbf{x}_p\} \rangle$$
$$= A_i(\mathbf{x}, \{\mathbf{x}_p\}) + \sum_p B_{pij}(\mathbf{x}, \{\mathbf{x}_p\}) \, u_{pj}. \tag{8}$$

Substituting Eq. 8 into

$$e_i = \left\langle \left(\hat{\tilde{u}}_i - \tilde{u}_i \right)^2 \,\middle|\, \{\mathbf{u}_p\}, \{\mathbf{x}_p\} \right\rangle, \quad i = 1,2,3 \tag{9}$$

and minimizing with respect to variations of A_i and B_{pij}, one finds, after some manipulation:

$$\hat{\tilde{u}}_i = \langle \tilde{u}_i \rangle + \sum_p B_{pij} \left(u_{pj} - \langle u_{pj} \rangle \right), \tag{10}$$

where B_{pij} is found by solving the system of linear equations:

$$\sum_p B_{pij} \left\langle u_{pj}' u_{qk}' \right\rangle = \left\langle \tilde{u}_i' u_{qk}' \right\rangle. \tag{11}$$

Here, a prime denotes fluctuation with respect to the mean:

$$(\cdot)' = (\cdot) - \langle (\cdot) \rangle. \tag{12}$$

The linear stochastic estimate has been shown to be quite good for various types of conditional averages of turbulence flows (Adrian et al. 1989).

Implementing Eqs. 11 and 12 requires data input for $\langle u_i \rangle, \langle \tilde{u}_i \rangle, \left\langle u_{pj}' u_{qk}' \right\rangle, \left\langle \tilde{u}' u_{qk}' \right\rangle$ The mean of the unfiltered field can be found from the ensemble of experimental flow fields, provided that the number of samples is large enough to give a good average over a small area around the point in question. The maximum size of this area depends upon the rate at which the mean varies in space, but clearly its dimension can be at least as great as the displacement of the particles in the image fields. The mean of the filtered field can be found by filtering the mean field, since filtering and averaging commute.

The two-point spatial correlations between the velocity samples, and the velocity sample and the velocity at the arbitrary location x, can be approximated by assuming that the spatial separations lie in a range of distances corresponding to the inertial subrange. The correlations may be modeled using correlations derived from the isotropic viscous-inertial turbulence spectrum, as an approximation to the true spectrum of the flow. While this approximation is not always justified, either because of insufficiently large Reynolds number or strong inhomogeneity and/or anisotropy, the interpolating functions still have the virtue of being rooted in fluid dynamics, and consequently they are less ad hoc than those used in other methods. The physics of high wave-number turbulence is imbedded in the process that determines the interpolating functions. The standard form of the isotropic representation only requires the longitudinal two-point correlation as input, and it guarantees that the correlation tensor is solenoidal (Batchelor 1960). It can be shown that the optimally interpolated field given by Eq. 10 is then also solenoidal.

These attractive features of optimal interpolation may not, however, guarantee better interpolation results. To judge this aspect one must appeal to a posteriori analysis of the optimal interpolation method. In this paper, we analyze the performance of the optimal interpolation method using random point samples of data, obtained from a high-resolution direct numerical simulation (DNS) of turbulence (with micro-scale Reynolds number, $R_\lambda \sim 164$), that mimic PTV and super-resolution PIV results.

2 Simulation and error analysis

In order to test the a posteriori performance of the optimal interpolator, we consider known turbulence fields where the statistics of filtered and unfiltered velocities can be accurately determined (except for sampling errors). This enables us to provide accurate inputs for the two-point correlations needed for obtaining a linear stochastic estimate of the 'optimal' interpolator, thereby eliminating a priori errors. The turbulence fields used in the present analysis are obtained using a well-resolved 256^3 numerical simulation of incompressible, forced isotropic turbulence ($R_\lambda \sim 164$). Negative viscosity forcing on modes with wavenumber magnitude $k \leq 3$ was used to attain stationarity in time. The simulation is based on a pseudo-spectral algorithm (Rogallo 1981) and the time integration is done using a second-order Runga–Kutta method. The magnitudes of the minimum and maximum resolved wavenumbers are $k = 1$ and 121, respectively.

The effects of filtering are analyzed using a sharp Fourier cut-off filter, whose effect is to annihilate all Fourier modes of $|\mathbf{k}|$ greater than a cut-off wavenumber k_c and leave the rest of the modes unaltered. For each k_c, we associate a characteristic filter width $\Delta_c \equiv \pi/k_c$. The transfer function of this filter is $H(k_c - |\mathbf{k}|)$, where H denotes the Heaviside function. It can be shown that the filtered velocity field (Eq. 5) obtained using this filter is solenoidal. A sharp cut-off filter in Fourier space tends to introduce oscillations in the filtered velocity field (in physical space). These oscillations are inherent to the nature of the sharp cut-off filter and hence should be captured by our optimal estimates. In cases where these oscillations are significant, it is possible that our optimal interpolator may result in a poor estimate, as reflected by a large error in the mean square sense. The resulting optimal estimate is the one with the minimum possible error (in the mean square sense), but in such cases may be contaminated with large errors.

To obtain an estimate of a component of the filtered velocity field at any chosen point, given the three-dimensional vector information of velocities at N random points uniformly distributed over the entire cubic domain of side L, we need to solve $3N$ equations (Eq. 11). For high concentrations, the number of equations to be solved is very large and the linear system may include a number of equations that contribute little to the estimate owing to loss of correlation between the estimate and events for very large separations. This problem is addressed by dividing the entire cubic domain (with side L) into non-overlapping cubic sub-domains, each with side L_s ($< L$), and considering M ($\ll N$) random samples that are uniformly distributed within each sub-domain. For each grid point location within a sub-domain, the estimate is evaluated based only on the events (or velocity samples) within that sub-domain and hence requires fewer equations ($3M$) to solve. The length L_s characterizes the maximum allowable loss of (longitudinal or transverse) correlation between the estimate and the event-data. It may also be useful to think of the role of L_s as that of an additional cut-off filter.

We note that the partitioning of our domain into non-overlapping sub-domains is done only to reduce the computational costs involved. The estimates within each sub-domain are continuous and satisfy (local) divergence-free constraints. There may be jumps or discontinuities in the estimate across the boundaries of the non-overlapping sub-domains. But this is a result of a trade-off for computational efficiency. It is not clear that the use of overlapping sub-domains would alleviate the discontinuity problems at the boundaries of the sub-domains.

Given our choice of the filter (i.e., the Fourier cut-off filter), the performance of our optimal interpolator (Eq. 10) for filtered velocities depends, a priori, on the following non-dimensional parameters: (i) L_s/λ, (ii) $\psi \equiv M\lambda^3/L_s^3$, and (iii) Δ_c/λ, where λ denotes the Taylor micro-scale of the unfiltered velocity field and ψ represents the (normalized) data density. Hence, the normalized mean square error, $\phi \equiv e/\sigma_{\tilde{u}}^2$ may be expressed as: $\phi = f(L_s/\lambda, \psi, \Delta_c/\lambda)$, where $e \equiv (e_1^u + e_2 + e_3)/3$ denotes the component averaged mean square error between the true filtered velocity and the interpolated velocity (see Eq. 9) and $\sigma_{\tilde{u}}^2$ denotes the variance of the filtered velocity field.

The true filtered velocity at each of the (256^3) grid points in our computational domain can be obtained by the use of the transfer function on the true unfiltered velocity field. The evaluation of the interpolated filtered velocity requires the knowledge of the true velocities (i.e., event data) at each of the random sample points obtained from a uniform distribution. The true velocities at these randomly sampled points are computed using a spectral interpolation procedure. Besides satisfying the solenoidal constraint on the event data, the spectral interpolation procedure also ensures that there are no errors in the specification of event data.

The correlations among event data, appearing on the left-hand side of Eq. 11, are evaluated using longitudinal velocity correlations obtained from DNS, along with the use of constraints imposed by isotropy and incompressibility of the flow. The correlations $\left\langle \tilde{u}' u'_{qk} \right\rangle$, between filtered and unfiltered velocities, appearing on the right hand side of Eq. 11, can be shown to be equal to $\left\langle \tilde{u}' \tilde{u}'_{qk} \right\rangle$ for the case of a Fourier cut-off filter. Similar to the unfiltered correlation tensor, the latter correlation can be evaluated from the longitudinal filtered velocity correlations obtained from DNS. For more general use of the optimal interpolation method for a wide range of Reynolds numbers, we can use the three-dimensional model spectrum of Pope (2000) to derive all correlations appearing in Eq. 11.

3 Results and discussion

The basic simulation parameters of the turbulence field obtained from DNS (using 256^3 grid points) are summarized in Table 1 (in arbitrary simulation units). Note that k_{max} ($k_{max} = 121$) is the magnitude of the maximum resolved wavenumber in our DNS. The three-dimensional energy spectrum obtained from DNS is shown in Fig. 1, as a function of the wavenumber magnitude. We investigate the effects of filtering on the performance of the optimal interpolation procedure by considering sharp Fourier cut-off filters at wavenumber magnitudes $k_c = 121$, 64, 32, 16, 8 and 4. The case with $k_c = 121$ (or greater) corresponds to DNS (i.e., unfiltered velocity field). The dashed lines in Fig. 1 show the cut-off filter locations, each of which may be associated with a normalized characteristic filter width Δ_c/λ.

The correlation tensor that appears on the left hand side of Eq. 11 can be derived from a unique scalar function, i.e., the longitudinal correlation of filtered velocities, which is dependent on the cut-off wave number k_c. Figure 2a and b shows the dependence of the longitudinal and transverse correlation coefficients of filtered velocities on k_c and the separation distance (normalized by the Kolmogorov length scale, η). The transverse correlation function can also be derived from the longitudinal function (and vice versa), for the case of incompressible, isotropic turbulence. The longitudinal correlation coefficient ($\rho^{\parallel}$) of filtered velocities increases with a decrease in k_c (or increase in Δ_c). The magnitude of the transverse correlation coefficient ($\rho^{\perp}$) of filtered velocities also appears to increase with k_c for both small and large separations. However, the trend with k_c is not very clear in the zero-crossing region (around $r/\eta \sim 180$–200). The amount of kinetic energy of the filtered velocity field, relative to that of the unfiltered velocity field, for different k_c is shown in Fig. 3. The ratio of the variance of the filtered velocity field to that of the unfiltered velocity field decreases with increase in the characteristic width of the filter (or decrease in k_c), as expected. A quantitative measure of the variance of the filtered velocity field (shown in Fig. 3) is useful for two reasons: (a) the variance of the filtered velocity field is relevant to estimation equations (i.e., right–hand side of Eq. 11), and (b) it is also used in defining a normalized mean square error, ϕ, due to the interpolation method. The behavior of the normalized error ϕ, which is a true indicator of the a posteriori performance of our optimal interpolation method, is shown in Fig. 4a–c. In Fig. 4a, the dependence of ϕ on the (normalized) sample data density (ψ) and filter width is illustrated for a fixed $L_s/\lambda = 0.939$. For any given ψ, the normalized mean square error ϕ appears to have a minimum value (say ϕ_{min}). The value ϕ_{min} and the width of the cut-off filter where ϕ_{min} occurs are observed to decrease with increase in data density. Except for large bandwidth filters, the error ϕ is also seen to decrease with increase in data density for any given width of the cut-off filter. The data points for the lowest Δ_c/λ shown in Fig. 4a also represent the unfiltered DNS case. For the highest data density case chosen, the minimum mean square error between the true filtered velocity and the velocity obtained from the optimal interpolation method is about 0.9% of the filtered velocity variance. For the lowest ψ chosen here (for $L_s/\lambda = 0.939$), the minimum error is about 13.5% of the filtered velocity variance.

The mean square error obtained using the AGW is also shown in Fig. 4c (dashed line) for comparison. To enable a fair comparison, the same data density $\psi = 2.412$ (highest data density case) was used for both the AGW and our optimal interpolation method. The evaluation of the mean-square error in the AGW case is carried out by comparing the estimate from AGW with the actual unfiltered velocity field. The performance of the optimal interpolation method appears to be better than the AGW for this particular density (highest data density case chosen), with a relative error of only about 4% in the former case (for moderate filter widths) and
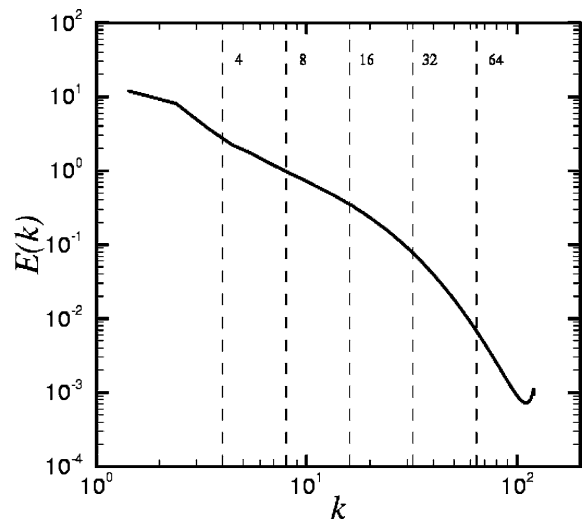
Table 1 Basic simulation parameters. N_0 (grid size), L (domain size), R_λ (micro-scale Reynolds number), $\langle \varepsilon \rangle$ (energy dissipation rate), ν (viscosity), σ_u^2 (velocity variance), Δ (grid spacing), η (Kolmogorov length scale), λ (Taylor micro-scale), L_{11} (longitudinal integral length scale) in arbitrary simulation units

N_0	256
L	2π
R_λ	164
$\langle \varepsilon \rangle$	62.9
ν	6.667×10^{-3}
σ_u^2	27.4
Δ/η	2.963
λ/η	25.228
L_{11}/η	121.331
k_{max}	121

Fig. 1 Three-dimensional energy spectrum $E(k)$ versus k. The *dashed lines* denote the several cases of k_c used in our analysis

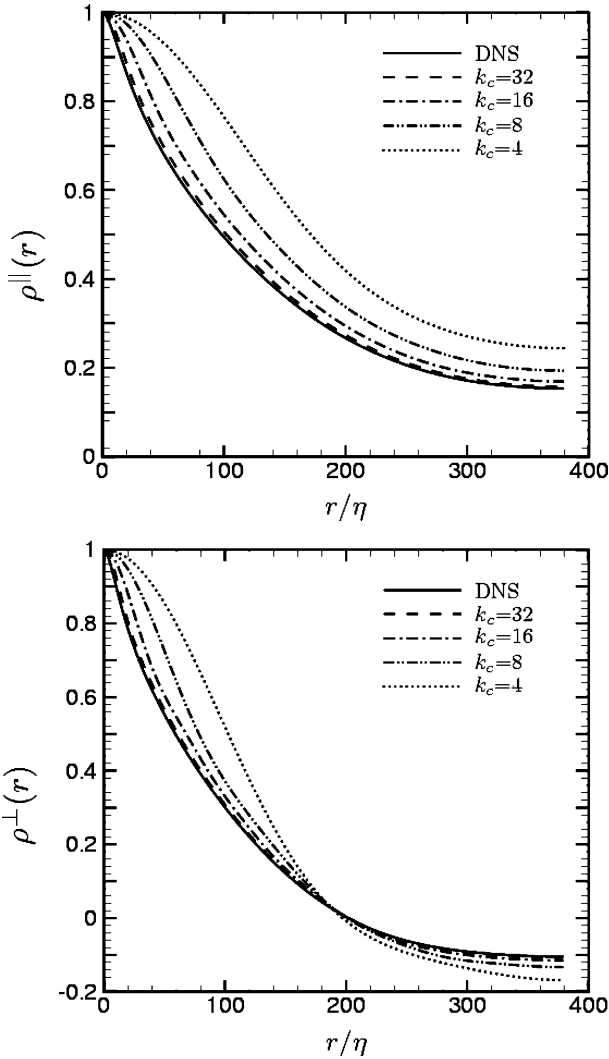Fig. 2 Correlation coefficients versus r/η for different k_c. **a** Longitudinal ($\rho^{\|}$) (*top*). **b** Transverse ($\rho^{\perp}$) (*bottom*). The correlation coefficients are defined as $\rho^{\|}(r) \equiv \langle \tilde{u}_{\|}(\mathbf{x})\tilde{u}_{\|}(\mathbf{x}+\mathbf{r})\rangle / \langle \tilde{u}_{\|}(\mathbf{x})\tilde{u}_{\|}(\mathbf{x})\rangle$ and $\rho^{\perp}(r) \equiv \langle \tilde{u}_{\perp}(\mathbf{x})\tilde{u}_{\perp}(\mathbf{x}+\mathbf{r})\rangle / \langle \tilde{u}_{\perp}(\mathbf{x})\tilde{u}_{\perp}(\mathbf{x})\rangle$

about 14% in the latter. Although not shown in the figure, our optimal interpolation method appears to have lower relative errors compared to the AGW for all data density cases considered. It may be noted that the AGW width selected in our computations is based on the optimum ratio $H/\delta = 1.24$ used in Agui and Jimenez (1987) and Spedding and Rignot (1993), where H and δ denote the Gaussian window width and mean nearest neighbor distance, respectively.

Figure 4b and c represents cases with L_s/λ fixed at twice and four times the value used in Fig. 4a. The effects of moderate and low data densities are also illustrated in Fig. 4b and c, respectively. The behavior of ϕ and the trends with change in ψ are similar to the ones observed for Fig. 4a. The results from Fig. 4a–c suggest that our optimal interpolation method gives the best possible results for interpolating a filtered velocity field from a random sample of unfiltered velocities, if we
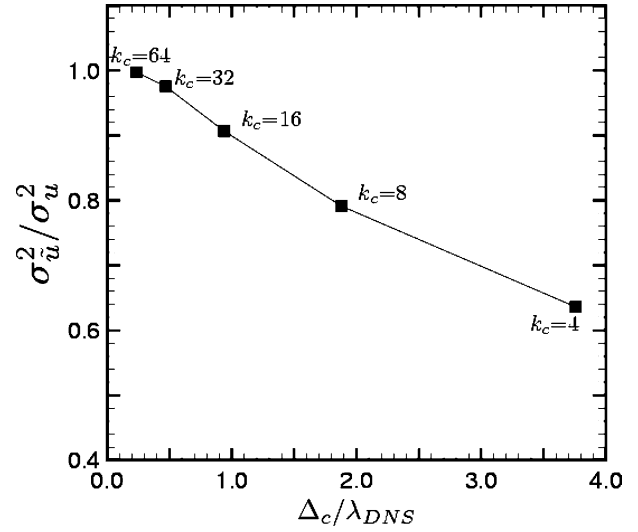

Fig. 3 Ratio of the variances of unfiltered to filtered velocities versus the characteristic filter width relative to the Taylor micro-scale (of unfiltered velocity field)

appropriately choose the width of the cut-off filter. For very small filter widths, our optimal interpolation method gives better results with increase in filter cut-off width (for fixed L_s/λ). The dependence of the error ϕ on the data density ψ and L_s/λ is better depicted in Fig. 5, which shows a plot of ϕ versus ψ for different L_s/λ. Two sets of curves shown here are for interpolation of (a) unfiltered velocity data and (b) filtered velocity data. For a given data density and filter width, the mean square error decreases with increase in L_s/λ. The figure also suggests that for low data densities ψ and large L_s/λ, optimal interpolation of a filtered velocity field yields a lower normalized error ϕ compared to the interpolation of an unfiltered velocity field (for small widths of the cut-off filter). On the contrary, optimal interpolation of unfiltered velocity fields performs better than that of the filtered case for high data densities and low L_s/λ. The error in the optimal interpolation of unfiltered velocity data appears to scale with the data density as $\phi \sim \psi^{-2/3}$. The effects of data density on the spectral performance of the optimal interpolation method and AGW are shown in Fig. 6a and b, respectively. For the optimal interpolation case, the agreement between the unfiltered interpolated energy spectrum and the true (DNS) energy spectrum appears to get better with increase in data density especially for low wavenumbers. However, for the AGW case, the trend with data density appears less certain and the errors in the low wave-number range are greater than those of the optimal interpolation method (for the highest data density case considered). The disparity between the two methods in the nature of convergence of energy spectra, with increase in data density, possibly has its origins in the very nature of the interpolation method. It can be shown that with increase in data density, the optimally interpolated signal converges to the true signal (since optimal esti-

Fig. 4 Normalized mean square error ϕ versus normalized widths of the cut-off filters for several values of the normalized data density ψ. **a** $L_s/\lambda = 0.939$ (*top*). **b** $L_s/\lambda = 1.878$ (*bottom left*). **c** $L_s/\lambda = 3.756$ (*bottom right*). The *dashed line* in **c** shows the error due to AGW for $\psi = 2.412$
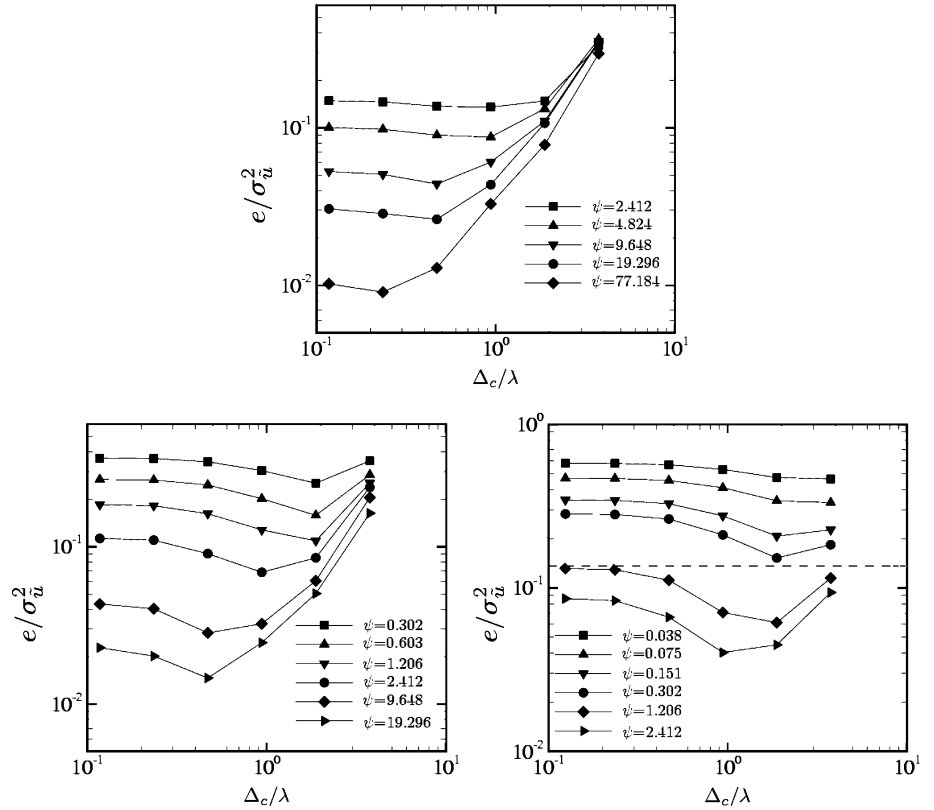
mation gives an exact result at sample points). The high wavenumber noise in the spectrum shown in Fig. 6b, for the AGW case, is a result of using non-overlapping sub-domains, whose significance was discussed in the previous section.

The spectral behavior of the error due to the optimal interpolation method at low wavenumbers is illustrated
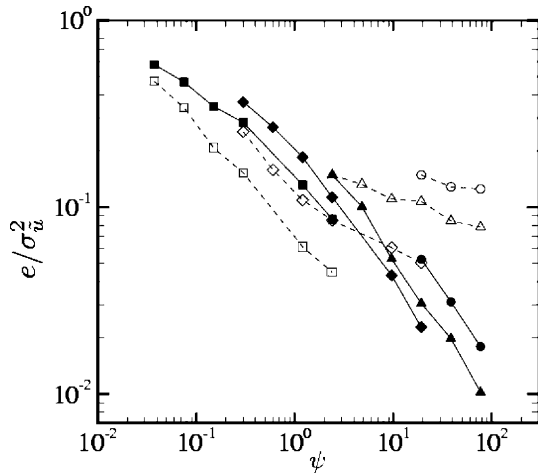
Fig. 5 Normalized mean square error $\phi \equiv e/\sigma_{\tilde{u}}^2$ versus normalized data density ψ. The *solid lines* correspond to the cases for unfiltered velocities, whereas the *dashed lines* correspond to those for the filtered field, with $k_c = 8$. The symbols, viz. *circles*, *triangles*, *diamonds* and *squares*, denote cases with $L_s/\lambda = 0.470, 0.939, 1.878$ and 3.756, respectively

in Fig. 7, wherein a plot of the relative error spectrum versus the wavenumber (normalized by the Kolmogorov scale) is shown for different data densities. The relative error spectrum is defined by

$$\theta^*(k) \equiv \frac{E_{\text{true}}(k) - E_{\text{est}}(k)}{E_{\text{true}}(k)}, \tag{13}$$

where $E_{\text{true}}(k)$ and $E_{\text{est}}(k)$ denote the true spectrum and the estimated spectrum due to the optimal interpolation method, respectively. The relative error clearly decreases with an increase in data density at low wavenumbers and relative errors of the order of 1% were obtained in the low wavenumber region (containing the most energetic modes) for the highest data density chosen ($\psi = 2.40$).

The behavior of this relative error due to the optimal interpolation method is compared to that of the AGW in Fig. 8 for the highest data density chosen ($\psi = 2.40$). We find that the optimal interpolation method performs better than the AGW in the low wave number range. Such a comparison would be fair for nearly identical bandwidths (i.e., wavenumber range where the relative error is less than 0.5) of the optimal interpolation and AGW filters, which is indeed the case shown in Fig. 8. The bandwidth (normalized by the Kolmogorov scale) is about 0.2 in this case.

A sample trace of the true unfiltered velocity signal along a line parallel to the x-axis is compared to the signal obtained from the optimal interpolation method (in Fig. 9a) and that obtained using the AGW (in Fig. 9b) for $(\Delta_c/\lambda, \psi, L_s/\lambda) = (0.117, 19.296, 1.878)$. The
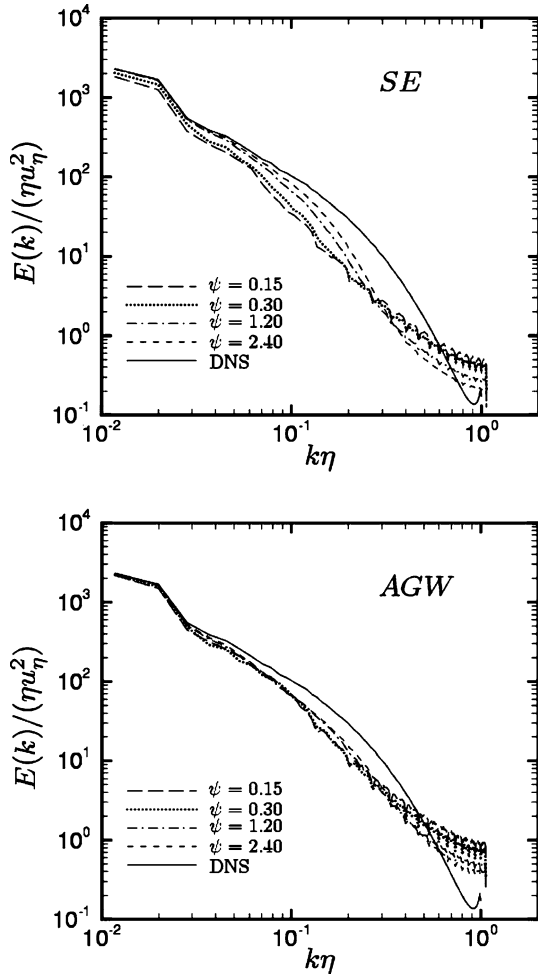
Fig. 6 Comparison of normalized energy spectra for different data densities using **a** the optimal interpolation (i.e., stochastic estimation) method (*top*), and **b** the adaptive Gaussian window (*bottom*) for $k_c = 128$ and $L_s/\lambda = 3.758$. The *solid line* shows the exact spectrum obtained from direct numerical simulation (DNS)
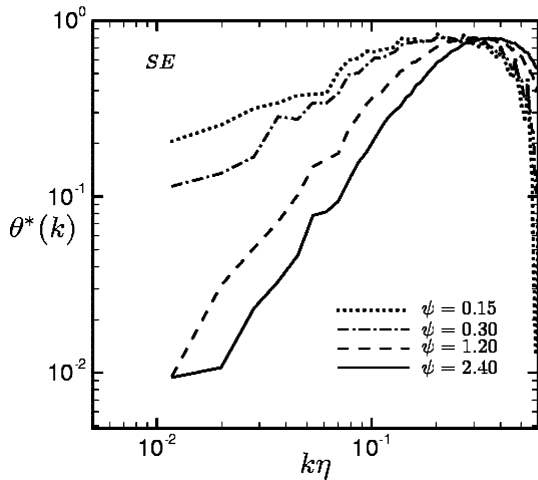

Fig. 7 Relative error spectra versus normalized wavenumber for different data densities using the optimal interpolation method ($k_c = 128$ and $L_s/\lambda = 3.758$)
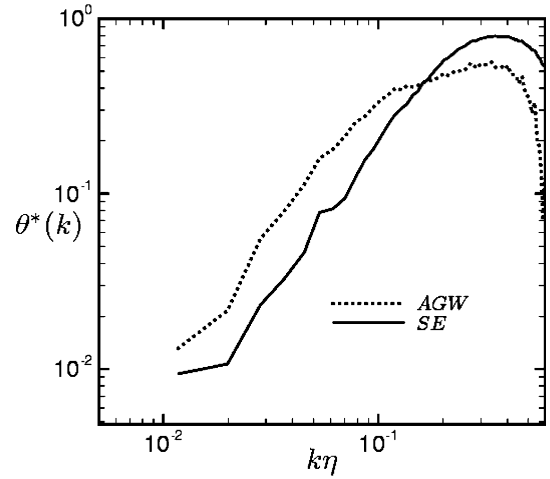

Fig. 8 Comparison of relative error spectra obtained using the optimal interpolation method (SE) and the adaptive Gaussian window (AGW) for $k_c = 128$, $\psi = 2.40$ and $L_s/\lambda = 3.758$

optimal interpolation method appears to perform better than the AGW case in approximating the true signal. The local extrema are found to be better represented by the optimal interpolation method, compared to the AGW case. For the cases shown in Fig. 9, the overall relative mean square error ϕ is about 2% for the optimal interpolation method compared to 5% for the AGW case, for the chosen parameter values.

4 Summary

The optimal interpolation process defines the best possible interpolator of the continuous turbulent velocity field given the available velocity data located at random points. Linear estimation of the conditional average is a method of proven accuracy in other contexts of turbulence representation. Application of the method to the interpolation problem yields a simple procedure that offers three attractive features: (1) the interpolated field is solenoidal, (2) the interpolation functions reflect approximate fluid dynamics of small scale turbulence, and (3) the interpolation allows for a readily interpretable filtered field defined independently of the sampling process.

Our results indicate that there exists an optimal choice of the width of the cut-off filter that gives the least possible error in the interpolation of the filtered velocity field for any given density of the random point samples. For low data densities, optimal interpolation of the filtered velocities was found to be more reliable than the optimal interpolation of unfiltered velocities. The most energetic modes, i.e., the low wavenumber modes, appear to be well approximated by the estimated signal using the optimal interpolation method. The a posteriori performance of the optimal interpolation method ranks better than the AGW case and can be improved with an increase in data density.
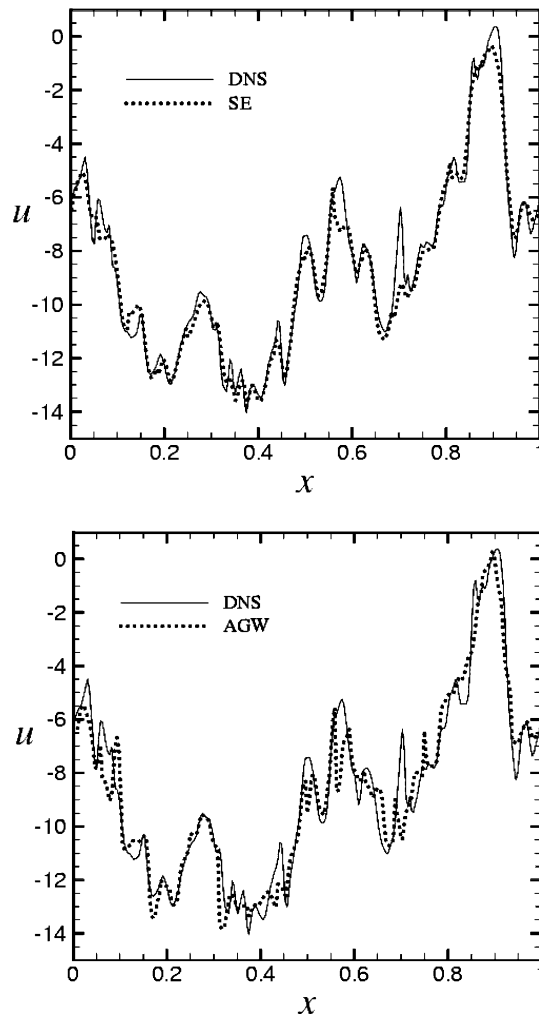
Fig. 9 Instantaneous velocity component u along a line parallel to the x-axis. Choice of parameters: $(\Delta_c/\lambda,\ \psi,\ L_s/\lambda) = (0.117,\ 19.296,\ 1.878)$. **a** The *solid* and *dotted lines* denote the true filtered velocity signal (obtained using DNS) and the interpolated signal obtained using the optimal interpolation method, respectively. **b** The *dotted line* corresponds to the signal obtained using AGW

In contrast to the ad hoc nature of AGW, the optimal interpolation method is based on a rational approach involving minimization of a mean square error function; this minimization is done by imbedding the known statistical properties of the flow. These advantages are perhaps some of the reasons for the better a posteriori performance of the optimal interpolation method, compared to AGW. In the case of high Reynolds number flow experiments in a laboratory, the input correlations may be approximated using classical models of energy spectra. The sensitivity of the a posteriori performance of the optimal interpolation method to a priori errors in the approximation of these input correlations needs to be investigated further. An alternative approach for specifying the input correlations (Eq. 11) in the case of more practical turbulent flows (i.e., not just isotropic turbulence) is to use a large ensemble of ran-domly sampled velocities for constructing the statistical correlation functions. For instance, using this alternative approach, the longitudinal and transverse correlation functions can be evaluated from a large ensemble of randomly sampled velocities in a self-consistent manner, thereby reducing the need for other simplifying modeling assumptions (on correlations or spectra).

The optimal interpolation method appears to be a promising approach applicable for approximating a filtered velocity field given the unfiltered velocities at random sample locations obtained from PTV or super-resolution PIV. This approach may also be extended for interpolating other physical quantities of interest in turbulent flows such as filtered versions of velocity gradients, vorticity, scalar concentrations and scalar gradients given the relevant randomly spaced event data.

References

Adrian RJ (1977) On the role of conditional averages in turbulence theory. In: Zakin JL, Patterson GK (eds) Turbulence in liquids. Science Press, Princeton, pp 323–332

Adrian RJ (1996) Stochastic estimation of the structure of turbulent fields. In: Bonnet JP (ed) Eddy structure identification. Springer, Berlin Heidelberg New York, pp 45–196

Adrian RJ, Jones BG, Chung MK, Hassan Y, Nithianandan CK, Tung AT-C (1989) Approximation of turbulent conditional averages by stochastic estimation. Phys Fluids A 1:992–998

Adrian RJ, Yao CS (1987) Power spectra of fluid velocities measured by laser Doppler velocimetry. Exp Fluids 5:17–28

Agui J, Jimenez J (1987) On the performance of particle tracking velocimetry. J Fluid Mech 185:447–468

Batchelor GK (1960) Homogeneous turbulence. Cambridge University Press, New York

Benedict LH, Nobach H, Tropea C (2000) Estimation of turbulent velocity spectra from laser Doppler data. Meas Sci Tech 11:1089–1104

Keane RD, Adrian RJ, Zhang Y (1995) Super-resolution particle-imaging velocimetry. Meas Sci Tech 6:754–768

Koga DJ (1986) An interpolation scheme for randomly spaced sparse velocity data, unpublished note. Mechanical Engineering Department, Stanford University, Stanford, CA

Mueller E, Nobach H, Tropea C (1998) Refined reconstruction-based correlation estimator for two-channel, non-coincidence laser Doppler anemometry. Meas Sci Tech 9:442–451

Papoulis A (1984) Probability, random variables and stochastic processes. McGraw-Hill, New York

Pope SB (2000) Turbulent flows. Cambridge University Press, Cambridge

Rogallo RS (1981) Numerical experiments in homogeneous turbulence. Technical Report TM-81315, NASA Ames

Spedding GR, Rignot EJM (1993) Performance analysis and application of grid interpolation techniques for fluid flows. Exp Fluids 15:417–430

Zhong JL (1995) Vector-valued multidimensional signal processing and analysis in the context of fluid flows. PhD Thesis, University of Illinois, Chicago, USA

Zhong JL, Weng JY, Huang TS (1991) Vector field interpolation in fluid flow. Digital signal processing 1991. In: International conference on DSP, Florence, Italy