Data assimilation method to de-noise and de-filter particle image velocimetry data

Jurriaan J. J. Gillissen[1,†], Roland Bouffanais[2] and Dick K. P. Yue[3]

[1]Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK

[2]Singapore University of Technology and Design, 8 Somapah Road, 487372, Singapore

[3]Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

We present a variational data assimilation method in order to improve the accuracy of velocity fields $\tilde{v}$, that are measured using particle image velocimetry (PIV). The method minimises the space–time integral of the difference between the reconstruction u and $\tilde{v}$, under the constraint, that u satisfies conservation of mass and momentum. We apply the method to synthetic velocimetry data, in a two-dimensional turbulent flow, where realistic PIV noise is generated by computationally mimicking the PIV measurement process. The method performs optimally when the assimilation integration time is of the order of the flow correlation time. We interpret these results by comparing them to one-dimensional diffusion and advection problems, for which we derive analytical expressions for the reconstruction error.

---

1. Introduction

The particle image velocimetry (PIV) technique measures a fluid velocity field by seeding the fluid with micron sized particles, and imaging their positions in a plane, using a laser sheet (e.g. Adrian & Westerweel 2011). The measured fluid velocity $\tilde{v}(x, t)$ at time t and position x is obtained from two consecutive images $\mathscr{I}(y, t - \frac{1}{2}\Delta t_{PIV})$ and $\mathscr{I}(y, t + \frac{1}{2}\Delta t_{PIV})$, with time difference Δt_{PIV}, by maximising for each x the correlation function,

$$\mathscr{C}(x) = \int f_\ell(|x - y|)\mathscr{I}(y, t - \tfrac{1}{2}\Delta t_{PIV})\mathscr{I}(y + \tilde{v}(x, t)\Delta t_{PIV}, t + \tfrac{1}{2}\Delta t_{PIV})\, d^2 y, \quad (1.1)$$

where $f_\ell(x)$ is a (two-dimensional) mask function, whose size ℓ^2 is referred to as the interrogation window. As an effect, PIV does not resolve the eddies that are smaller than ℓ, i.e. the measured velocity field $\tilde{v}$ is a filtered version of the actual field v. Besides filtering, PIV introduces additional noise, which depends in a complicated way on the structure of the unresolved eddies, and on the out of plane velocity component

of $\boldsymbol{v}$, as well as on the particle seeding density, particle image diameter, non-uniform laser illumination and camera limitations, e.g. pixel locking and resolution.

Various methods have been proposed to mitigate the PIV noise. Some methods are based on the data assimilation technique. This technique reconstructs a velocity field $\boldsymbol{u}$ by minimising the difference between $\boldsymbol{u}$ and $\tilde{\boldsymbol{v}}$, under the constraint, that $\boldsymbol{u}$ satisfies conservation of mass and/or momentum. There are principally two distinct data assimilation methods, being the space integral method (SIM) (e.g. Humphrey 1993; Ruhnau, Stahl & Schnörr 2007; Suzuki 2012; Foures *et al.* 2014; Symon *et al.* 2017), and the space–time integral method (STIM) (e.g. Talagrand & Courtier 1987; Gronskis, Heitz & Mémin 2013; Mons *et al.* 2016; Gillissen *et al.* 2018). In both methods the mismatch between the reconstructed field $\boldsymbol{u}$ and the observed field $\tilde{\boldsymbol{v}}$ is translated into a force field, that drives $\boldsymbol{u}$ towards $\tilde{\boldsymbol{v}}$. In SIM the force at time $t = t_1$ depends only on measurements $\tilde{\boldsymbol{v}}$ at t_1, while in STIM the force at t_1 depends on $\tilde{\boldsymbol{v}}$ in a time window $t_1 < t < t_1 + \tau$, where τ is referred to as the assimilation time. In this approach the observations $\tilde{\boldsymbol{v}}$ from $t > t_1$ are propagated backward in time to t_1, to update the initial conditions of $\boldsymbol{u}$ at t_1. As an effect STIM is more computationally intensive and more accurate than SIM.

To our knowledge STIM has not yet been applied to improve PIV data. In this work we fill this gap, and apply STIM to improve the accuracy of synthetic PIV data for a two-dimensional (2-D) turbulent flow. In particular we study, how the reconstruction error ϵ depends on the assimilation time τ, and the PIV measurement error, where the latter is controlled by the size of the PIV interrogation window ℓ^2. We interpret these numerical results, by comparing them to one-dimensional diffusion and advection problems, for which we derive analytical relationships between ϵ and the above mentioned parameters.

2. Data assimilation methods

In this work, we study the performance of a STIM to improve the accuracy of synthetic PIV data for a 2-D turbulent flow. In order to evaluate its performance, we compare the STIM to a SIM, which is less computationally intensive, but also less effective in reducing the PIV measurement error. We derive the STIM and SIM in §§ 2.1 and 2.2 below.

2.1. *Space–time integral method*

We consider a PIV velocity field $\tilde{\boldsymbol{v}}$, which is a measurement of an actual velocity field $\boldsymbol{v}$, at evenly distributed time instances $t_i = i\Delta t$, where $i = 1, 2, 3, \ldots$, and Δt is the sampling time. We construct an improved velocity field $\boldsymbol{u}$, by minimising the difference between $\boldsymbol{u}$ and $\tilde{\boldsymbol{v}}$ under the constraint that $\boldsymbol{u}$ satisfies the Navier–Stokes equations,

$$\boldsymbol{R}(\boldsymbol{w}) = \begin{pmatrix} \partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + \nabla p - \nu \nabla^2 \boldsymbol{u} \\ \nabla \cdot \boldsymbol{u} \end{pmatrix} = \boldsymbol{0}, \tag{2.1}$$

where $\boldsymbol{w} = (\boldsymbol{u}, p)$ is referred to as the 'state variable', p is the fluid pressure and ν is the fluid kinematic viscosity. We therefore minimise a constrained cost functional, referred to as the Lagrangian $\mathscr{L}$,

$$\mathscr{L} = \int_{t_1}^{t_P} \left(\frac{1}{2} \|\boldsymbol{u} - \tilde{\boldsymbol{v}}\|^2 \sum_{i=1}^{P} \delta(t - t_i) + \langle \hat{\boldsymbol{w}}, \boldsymbol{R}(\boldsymbol{w}) \rangle \right) \mathrm{d}t, \tag{2.2}$$

which involves a set of P measurements, taken from an assimilation window $t_1 < t < t_P$, where $\tau = t_P - t_1$ is the assimilation time. The Lagrange multiplier $\hat{\boldsymbol{w}} = (\hat{\boldsymbol{u}}, \hat{p})$ is introduced to enforce the constraint (2.1). Note, that a hat $\hat{\ }$ above a field variable denotes, that the field variable is a Lagrange multiplier. In (2.2) time integration is performed from the first to the last (Pth) measurement, $\delta(\cdot)$ denotes the Dirac delta function and $\|\cdot\|$ is the norm based on the standard inner product $\langle\cdot, \cdot\rangle$, which, when applied to two complex vector fields $\boldsymbol{a}$ and $\boldsymbol{b}$, reads

$$\langle\boldsymbol{a}, \boldsymbol{b}\rangle = \int \mathrm{d}V \boldsymbol{a}^* \cdot \boldsymbol{b}. \tag{2.3}$$

Here V is the spatial domain, on which $\boldsymbol{a}$ and $\boldsymbol{b}$ are defined, and the superscript * denotes complex conjugation.

The goal is to find $\boldsymbol{u}$, that minimises $\mathscr{L}$ (2.2). Since $\boldsymbol{u}$ is determined by its initial conditions $\boldsymbol{u}_1$, we determine $\boldsymbol{u}$ by minimising $\mathscr{L}$ with respect to $\boldsymbol{u}_1$. It is noted, that subscript indices refer to time instances, e.g. $\boldsymbol{u}_1 = \boldsymbol{u}(t_1)$. To minimise $\mathscr{L}$ with respect to $\boldsymbol{u}_1$, we need the gradient of $\mathscr{L}$ with respect to $\boldsymbol{u}_1$, i.e. $\delta\mathscr{L}/\delta\boldsymbol{u}_1$. To derive an expression for $\delta\mathscr{L}/\delta\boldsymbol{u}_1$ we start by writing the variation of the cost functional $\delta\mathscr{L}$ due to an infinitesimal variation of the state variable $\delta\boldsymbol{w} = (\delta\boldsymbol{u}, \delta p)$,

$$\delta\mathscr{L} = \int_{t_1}^{t_P} \left(\langle\boldsymbol{u} - \tilde{\boldsymbol{v}}, \delta\boldsymbol{u}\rangle \sum_{i=1}^{P} \delta(t - t_i) + \langle\hat{\boldsymbol{w}}, \delta\boldsymbol{R}(\boldsymbol{w})\rangle \right) \mathrm{d}t. \tag{2.4}$$

We rewrite the Lagrange multiplier term in (2.4) using integration by parts (e.g. Gunzburger 2003),

$$\delta\mathscr{L} = \int_{t_1}^{t_P} \left(\langle\boldsymbol{u} - \tilde{\boldsymbol{v}}, \delta\boldsymbol{u}\rangle \sum_{i=1}^{P} \delta(t - t_i) + \langle\hat{\boldsymbol{R}}(\boldsymbol{w}, \hat{\boldsymbol{w}}), \delta\boldsymbol{w}\rangle \right) \mathrm{d}t + \langle\hat{\boldsymbol{u}}_P, \delta\boldsymbol{u}_P\rangle - \langle\hat{\boldsymbol{u}}_1, \delta\boldsymbol{u}_1\rangle, \tag{2.5}$$

where $\hat{\boldsymbol{R}}$ is the adjoint of the linearised Navier–Stokes equations $\boldsymbol{R}$ (2.1),

$$\hat{\boldsymbol{R}}(\boldsymbol{w}, \hat{\boldsymbol{w}}) = \begin{pmatrix} -\partial_t\hat{\boldsymbol{u}} - \boldsymbol{u} \cdot (\nabla\hat{\boldsymbol{u}} + \nabla\hat{\boldsymbol{u}}^{\mathrm{T}}) - \nabla\hat{p} - \nu\nabla^2\hat{\boldsymbol{u}} \\ -\nabla \cdot \hat{\boldsymbol{u}} \end{pmatrix}, \tag{2.6}$$

where a hat $\hat{\ }$ above an operator denotes the adjoint of the linearised version of that operator, and $\nabla\hat{\boldsymbol{u}}^{\mathrm{T}}$ is the transpose of $\nabla\hat{\boldsymbol{u}}$. In (2.5) the terms involving $\hat{\boldsymbol{u}}_P$ and $\hat{\boldsymbol{u}}_1$ are time-boundary terms, that are obtained by integrating by parts the time derivative term in $\langle\hat{\boldsymbol{w}}, \delta\boldsymbol{R}(\boldsymbol{w})\rangle$ in (2.4). These terms determine the initial and final conditions for $\hat{\boldsymbol{u}}$ at $t = t_P$ and at $t = t_1$, respectively. We do not consider similar spatial boundary terms, as we assume spatial periodicity, for simplicity.

We find the initial value $\boldsymbol{u}_1$ that minimises the cost functional $\mathscr{L}$ (2.2), using a conjugate gradient method, by iteratively updating $\boldsymbol{u}_1$ in a search direction, that is related to the functional derivative of $\mathscr{L}$ with respect to $\boldsymbol{u}_1$. From (2.5) we find that

$$\frac{\delta\mathscr{L}}{\delta\boldsymbol{u}_1} = -\hat{\boldsymbol{u}}_1. \tag{2.7}$$

To find the Lagrange multiplier $\hat{\boldsymbol{u}}_1$ at instant t_1, we use the equation of motion for $\hat{\boldsymbol{u}}$. This equation is derived from (2.5), by demanding that the Lagrangian $\mathscr{L}$ is at

an extremum with respect to the state variable w, i.e. $\delta\mathscr{L}$ is zero under infinitesimal variation δw,

$$\left.\begin{aligned}-\partial_t\hat{\boldsymbol{u}} - \boldsymbol{u}\cdot(\boldsymbol{\nabla}\hat{\boldsymbol{u}} + \boldsymbol{\nabla}\hat{\boldsymbol{u}}^{\mathrm{T}}) - \boldsymbol{\nabla}\hat{p} - \nu\nabla^2\hat{\boldsymbol{u}} + (\boldsymbol{u} - \tilde{\boldsymbol{v}})\sum_{i=1}^{P}\delta(t - t_i) = \boldsymbol{0},\\[4pt]-\boldsymbol{\nabla}\cdot\hat{\boldsymbol{u}} = \boldsymbol{0}.\end{aligned}\right\}\tag{2.8a}$$

According to (2.8a), the Lagrange multiplier $\hat{\boldsymbol{u}}$ is incompressible and is advected by $\boldsymbol{u}$ and is subjected to a (negative) viscosity $-\nu$ and is driven by a source term, which is the difference between the measurement $\tilde{\boldsymbol{v}}$ and the reconstruction $\boldsymbol{u}$, at discrete sampling times t_i. Due to the negative viscosity, the equation is integrated backward in time. The corresponding 'initial' conditions at time t_P and 'final' conditions at time t_1 are found from (2.5), by demanding that $\delta\mathscr{L}$ is zero under infinitesimal variations $\delta\boldsymbol{u}_P$ and $\delta\boldsymbol{u}_1$, giving

$$\hat{\boldsymbol{u}}_P = \boldsymbol{0},\tag{2.8b}$$

and

$$\hat{\boldsymbol{u}}_1 = \boldsymbol{0}.\tag{2.8c}$$

The conjugate gradient update direction $\hat{\boldsymbol{u}}_1$ (2.7) is found by integrating (2.8a) backward in time, using 'initial' conditions given by (2.8b). After updating the initial conditions for the forward problem $\boldsymbol{u}_1$, equation (2.1) is integrated (forward in time), which gives a new source term for the Lagrange multiplier equation (2.8a). This procedure is repeated until $\mathscr{L}$ is at an extremum, with respect to $\boldsymbol{u}_1$, which, according to (2.7), corresponds to (2.8c) being satisfied, to a certain degree of accuracy.

2.2. *Space integral method*

In this work we compare the performance of the STIM (2.2) to a SIM, which is similar to previously used methods, to improve PIV data (e.g. Ruhnau *et al.* 2005). The SIM minimises, at each instant $t = t_i$, the following cost functional, which penalises the difference between the reconstruction $\boldsymbol{u}$ and the measurement $\tilde{\boldsymbol{v}}$, under the constraint that $\boldsymbol{u}$ is divergence free

$$\mathscr{L} = \tfrac{1}{2}\|\boldsymbol{u} - \tilde{\boldsymbol{v}}\|^2 + \langle\hat{p}, \boldsymbol{\nabla}\cdot\boldsymbol{u}\rangle + \tfrac{1}{2}\kappa\|\nabla^2\boldsymbol{u}\|^2,\tag{2.9}$$

and where a regularisation term $\kappa\|\nabla^2\boldsymbol{u}\|^2$ is added to suppress the noise at large wavenumbers (e.g. Tikhonov & Arsenin 1977). Here κ is referred to as the regularisation strength. Minimising $\mathscr{L}$ with respect to $\boldsymbol{u}$ involves computing the variation of $\mathscr{L}$ due to a variation in $\boldsymbol{u}$

$$\delta\mathscr{L} = \langle\boldsymbol{u} - \tilde{\boldsymbol{v}}, \delta\boldsymbol{u}\rangle - \langle\boldsymbol{\nabla}\hat{p}, \delta\boldsymbol{u}\rangle + \langle\kappa\nabla^4\boldsymbol{u}, \delta\boldsymbol{u}\rangle,\tag{2.10}$$

where we have applied integration by parts. To find the reconstructed velocity field $\boldsymbol{u}$, we demand that $\delta\mathscr{L}$ is zero under infinitesimal variation $\delta\boldsymbol{u}$

$$(1 + \kappa\nabla^4)\boldsymbol{u} = \tilde{\boldsymbol{v}} + \boldsymbol{\nabla}\hat{p}.\tag{2.11}$$

Here the term $\boldsymbol{\nabla}\hat{p}$ ensures that $\boldsymbol{u}$ is divergence free

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = \boldsymbol{0}.\tag{2.12}$$

The solution to (2.11), (2.12) in Fourier space reads

$$u^k = \frac{\tilde{v}^k}{1 + \kappa k^4} \cdot \left(\delta - \frac{kk}{k^2} \right), \tag{2.13}$$

where the superscript indices refer to the wavevectors of the spatial Fourier modes. Equation (2.13) shows, that the regularisation term in the cost functional (2.9) effectively applies hyper diffusion to the measurement field $\tilde{v}$. The order of the hyper diffusion, which in this case equals four, depends on the exponent on the velocity gradient in the κ-term in (2.9), which in this case equals two. For a unit exponent we would have recovered normal, second-order diffusion (see e.g. Ruhnau *et al.* 2005). Fourth-order diffusion is chosen above second-order diffusion however, since the former affects more selectively the large wavenumbers, while leaving the small wavenumbers intact.

3. Linear problems

3.1. *General considerations*

In this work we study the performance of the STIM (2.2) to improve the accuracy of synthetic PIV data. Before considering 2-D turbulent flow cases in §4, we start in this section by considering one-dimensional (1-D) linear problems, for which analytical expressions can be derived for the method performance. In general, the linear equation of motion reads

$$\partial_t u - S(u) = 0, \tag{3.1}$$

where S is a linear spatial operator. We decompose the spatial part of u into the Fourier modes $\exp(\mathbb{i}kx)$, where k is the wavenumber and where $\mathbb{i} = \sqrt{-1}$,

$$u = u^k(t) \exp(\mathbb{i}kx). \tag{3.2}$$

Here $u^k(t)$ are the time-dependent Fourier coefficients, and summation of repeated indices is assumed. It is noted, that superscript indices refer to Fourier coefficients. Inserting (3.2) into (3.1) and taking the inner product (2.3) of the result with $\exp(\mathbb{i}lx)$ and using the orthogonality property of $\exp(\mathbb{i}kx)$, gives the equation of motion of u^l

$$\partial_t u^l - M^{lk} u^k = 0, \tag{3.3}$$

where $M^{lk} = \langle \exp(\mathbb{i}lx), S \exp(\mathbb{i}kx) \rangle$ is the Fourier representation of S. Similarly we decompose the Lagrange multiplier into $\hat{u} = \hat{u}^k(t) \exp(\mathbb{i}kx)$, where it is recalled, that a hat $\hat{}$ above a field variable denotes that the field variable is a Lagrange multiplier. The evolution equation for the Fourier coefficients $\hat{u}^l$ is derived, following similar lines as for (2.8), and is given by

$$- \partial_t \hat{u}^l - M^{*lk} \hat{u}^k + (u^l - \tilde{v}^l) \sum_{i=1}^{P} \delta(t - t_i) = 0, \tag{3.4a}$$

where M^{*lk} is the complex conjugate of M^{lk}. The initial and final conditions for (3.4a) read

$$\hat{u}_1^l = 0, \tag{3.4b}$$

and

$$\hat{u}_P^l = 0. \tag{3.4c}$$

To allow for an analytical treatment of the problem, we assume, in this section, that the PIV measured velocity field $\tilde{v}$ is related to the actual velocity field v as follows:

$$\tilde{v} = \overline{v + v'}, \tag{3.5a}$$

where v' is white noise with standard deviation σ_v, and $\overline{(\cdot)}$ is the following spatial filter:

$$\overline{v}(x) = \int f_\ell(|x - y|) v(y) \, \mathrm{d}y. \tag{3.5b}$$

Here $f_\ell(x)$ is a mask function of width ℓ. In Fourier space, the PIV measurement operation (3.5) reads

$$\tilde{v}^l = F^{lk}(v^k + v'^k), \tag{3.6}$$

where F^{lk} is the Fourier representation of the filter operator (3.5b), which is assumed to be a sharp cutoff filter, with a width of $k_{filter} = 2\pi/\ell$

$$F^{kl} = \delta^{kl} \begin{cases} 1, & \text{if } k \leqslant k_{filter}, \\ 0, & \text{if } k > k_{filter}. \end{cases} \tag{3.7}$$

It is noted, that actual PIV noise is more complex than white noise. Assuming white noise is necessary however to make analytical progress. In §4.3 we consider more realistic noise when applying the method to 2-D turbulent flow cases.

Without loss of generality, we assume that $t_1 = 0$ and we introduce the following notation for the forward integration, from $t = t_1$ to $t = t_i$, of the linear equation (3.3)

$$u_i^l = G_i^{lk} u_1^k, \tag{3.8a}$$

where G_i^{lk} is given by

$$G_i^{lk} = \exp(M^{lk} t_i). \tag{3.8b}$$

It is recalled, that subscript indices refer to time instances. With this notation, the backward in time integration of (3.4a), from $t = t_P$ to $t = t_1$, starting at $t = t_P$ with $\hat{u}_P^k = 0$, is written as

$$\hat{u}_1^l = \sum_{i=1}^{P} G_i^{*lk} (u_i^k - \tilde{v}_i^k), \tag{3.9}$$

Using (3.6) and (3.8a)

$$\tilde{v}_i^l = F^{lk}(G_i^{km} v_1^m + v_i'^k), \tag{3.10}$$

we rewrite (3.9) into

$$\hat{u}_1^l = \sum_{i=1}^{P} G_i^{*lm} [G_i^{mk} u_1^k - F^{mn}(G_i^{nk} v_1^k + v_i'^n)]. \tag{3.11}$$

It is recalled that superscript indices refer to spatial Fourier modes. For the ease of notation, we replace these superscript indices with vector notation, such that (3.11) reads

$$\hat{\boldsymbol{u}}_1 = \sum_{i=1}^{P} \boldsymbol{G}^*_i \cdot [\boldsymbol{G}_i \cdot \boldsymbol{u}_1 - \boldsymbol{F} \cdot (\boldsymbol{G}_i \cdot \boldsymbol{v}_1 + \boldsymbol{v}'_i)]. \tag{3.12}$$

Minimisation of the cost functional corresponds to $\hat{\boldsymbol{u}}_1 = \boldsymbol{0}$, which results in the following reconstruction error:

$$\boldsymbol{u}_1 - \boldsymbol{v}_1 = \left[\left(\sum_{i=1}^{P} \boldsymbol{G}^*_i \cdot \boldsymbol{G}_i \right)^{-1} \cdot \sum_{i=1}^{P} \boldsymbol{G}^*_i \cdot \boldsymbol{F} \cdot \boldsymbol{G}_i - \boldsymbol{\delta} \right] \cdot \boldsymbol{v}_1$$
$$+ \left(\sum_{i=1}^{P} \boldsymbol{G}^*_i \cdot \boldsymbol{G}_i \right)^{-1} \cdot \sum_{i=1}^{P} \boldsymbol{G}^*_i \cdot \boldsymbol{F} \cdot \boldsymbol{v}'_i. \tag{3.13}$$

The error (3.13) consists of two terms. The first term is referred to as the filter error, and the second term is referred to as the noise error. In the next subsections, we analyse the behaviour of these two terms, as a function of the assimilation time τ, for the cases of diffusion and advection.

3.2. Diffusion

We consider scalar diffusion with diffusivity ν on a domain of size 2π. The Fourier representation of the corresponding operator $S(u) = \nu \partial_x^2 u$ reads

$$M^{kl} = -\nu k^2 \delta^{kl}, \tag{3.14}$$

which is diagonal, meaning that the Fourier modes evolve independently. Inserting $G_i^{kl} = \exp(-k^2 \nu t_i) \delta^{kl}$ and (3.7) into (3.13), gives that for a single mode with wavenumber $k \leqslant k_{filter}$, the filter error is zero (first term on right-hand side of (3.13)), while the noise error is non-zero (second term on right-hand side of (3.13)), and is given by

$$u_1^k - v_1^k = \frac{\displaystyle\sum_{i=1}^{P} \exp(-k^2 \nu t_i)}{\displaystyle\sum_{j=1}^{P} \exp(-2k^2 \nu t_j)} v_i'^k. \tag{3.15}$$

Assuming that $v_i'^k$ has a zero mean and a non-zero standard deviation σ_v, the mean of the error $\epsilon = u_1^k - v_1^k$ is zero, and its standard deviation σ_ϵ equals

$$\sigma_\epsilon = \sqrt{\sum_{i=1}^{P} \left(\frac{\exp(-k^2 \nu t_i)}{\displaystyle\sum_{j=1}^{P} \exp(-2k^2 \nu t_j)} \sigma_v \right)^2} = \frac{\sigma_v}{\sqrt{\displaystyle\sum_{i=1}^{P} \exp(-2k^2 \nu t_i)}}. \tag{3.16}$$
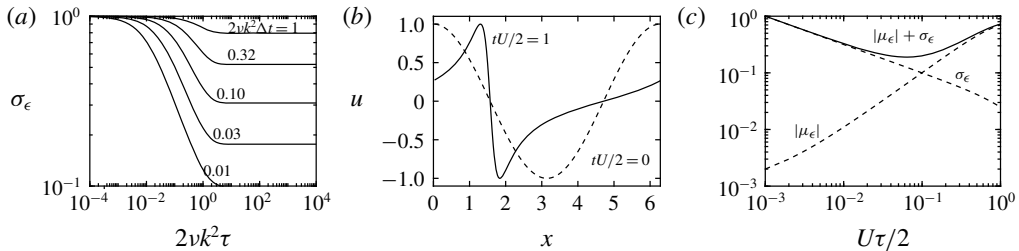
FIGURE 1. (*a*) Theoretical standard deviation of the reconstruction error σ_ϵ (3.17) for a diffusing Fourier mode, relative to the standard deviation of the measurement error of that mode σ_v as a function of the dimensionless assimilation time $2\nu k^2\tau$ for various dimensionless sampling times $2\nu k^2\Delta t$. (*b*) The evolution of a scalar profile, with $u(t = 0) = \cos(x)$, that is advected by an inhomogeneous velocity field $U\cos(x)$. (*c*) Theoretical standard deviation σ_ϵ and absolute value of mean $|\mu_\epsilon|$ of the reconstruction error (3.13) for the $k = 1$ mode in non-uniform advection (3.18) based on a measurement filter, that only passes the $k = 1$ mode (3.7).

By recalling that $t_i = \Delta t(i - 1)$, where $\Delta t = \tau/(P - 1)$ is the sampling time, and $\tau = t_P - t_1$ is the assimilation time, and by using that $\sum_{i=0}^{P-1} x^i = (x^P - 1)/(x - 1)$, we find

$$\sigma_\epsilon = \sigma_v \sqrt{\frac{\exp(-\Delta t') - 1}{\exp(-\tau' - \Delta t') - 1}}, \qquad (3.17)$$

where $\tau' = 2\tau\nu k^2$ and $\Delta t' = 2\nu k^2 \Delta t$ are the assimilation time and the sampling time, non-dimensionalised with the correlation time $\mathscr{T} = (2k^2\nu)^{-1}$. For $\Delta t' \ll \tau' \ll 1$, equation (3.17) predicts the expected behaviour, that the standard deviation of the noise error σ_ϵ scales inversely with the square root of the number of samples P, i.e. $\sigma_\epsilon = \sigma_v\sqrt{\Delta t'/\tau'} = \sigma_v/\sqrt{P}$. For $\tau' \gg 1 \gg \Delta t'$ on the other hand, the reconstruction is affected only by a fraction of the total number of samples P, whose time distances are smaller than the correlation time. The number of contributing samples is then equal to the correlation time divided by the sampling time, which in non-dimensional units reads: $1/\Delta t' \ll P$. In this case σ_ϵ scales inversely with the square root of this number, i.e. it grows with the square root of the sampling time $\sigma_\epsilon = \sigma_v\sqrt{\Delta t'}$.

Equation (3.17) is plotted in figure 1(*a*) as a function of τ' for various $\Delta t'$ showing the transition from inverse square root to constant. The graph illustrates that the reconstruction does not improve when information is added beyond the correlation time of the system $\mathscr{T} = (2k^2\nu)^{-1}$, i.e. measurements at time $t_i \gg \mathscr{T}$ do not influence the reconstruction at time $t = 0$.

3.3. *Advection*

Next we consider the advection of a scalar u in an inhomogeneous velocity field $U\cos(x)$, with an amplitude U and a unit wavenumber on a domain of size 2π. The evolution of a scalar profile $u(x, t)$ with $u(x, t = 0) = \cos(x)$ is sketched in figure 1(*b*). The Fourier representation of the corresponding operator $S(u) = -U\cos(x)\partial_x u$ is given by

$$M^{kl} = -\frac{\mathrm{i}lU}{2}(\delta^{l,k-1} + \delta^{l,k+1}). \qquad (3.18)$$

We compute the error (3.13) in the small time limit $\tau \|M\| \ll 1$, such that G_i (3.8b) is approximated by

$$G_i \approx \delta + i\Delta t M, \tag{3.19}$$

which to leading order in $\tau \|M\|$ gives

$$u_1 - v_1 = \left[F - \delta + \frac{\tau}{2}(F \cdot M - M \cdot F)\right] \cdot v_1 + P^{-1} \sum_{i=1}^{P} F \cdot v_i', \tag{3.20}$$

where we have used that $\tau = \Delta t P$. Using (3.7), (3.18), we see that the noise error (second term on the right-hand side of (3.20)) equals

$$P^{-1} \sum_{i=1}^{P} F^{kl} v_i'^l = \begin{cases} P^{-1} \sum_{i=1}^{P} v_i'^k, & \text{if } k \leqslant k_{filter}, \\ 0, & \text{if } k > k_{filter}, \end{cases} \tag{3.21}$$

and the leading order of the filter error (first term on the right-hand side of (3.20)) equals

$$\left[F^{kl} - \delta^{kl} + \frac{\tau}{2}(F^{km}M^{ml} - M^{km}F^{ml})\right] v_1^l = \begin{cases} 0, & \text{if } k < k_{filter}, \\ -\frac{1}{4}\mathrm{i}U\tau(k+1)v^{k+1}, & \text{if } k = k_{filter}, \\ v_1^k, & \text{if } k > k_{filter}. \end{cases} \tag{3.22}$$

Equation (3.22) shows a 0 % and a 100 % error for $k < k_{filter}$ and $k > k_{filter}$, and an intermediate error for $k = k_{filter}$, which is explained as follows. When acting on the k_{filter} mode at $t = 0$, the advection operator (3.18) creates $k_{filter} + 1$ and $k_{filter} - 1$ modes at $t = \tau$. Since the $k_{filter} + 1$ mode is cut off by the filter (3.7), the resulting filtered signal at $t = \tau$ contains insufficient information, to fully reconstruct the k_{filter} mode at $t = 0$, which explains the intermediate filter error for $k = k_{filter}$ in (3.22). Combining (3.21), (3.22), we write for the total reconstruction error $\epsilon = u_1^{k_{filter}} - v_1^{k_{filter}}$ for the $k = k_{filter}$ mode

$$u_1^{k_{filter}} - v_1^{k_{filter}} = -\frac{\mathrm{i}U\tau(k_{filter} + 1)}{4} v_1^{k_{filter}+1} + P^{-1} \sum_{i=1}^{P} v_i'^{k_{filter}}. \tag{3.23}$$

The mean of the error μ_ϵ is due to the filtering

$$\mu_\epsilon = -\frac{\mathrm{i}U\tau(k_{filter} + 1)}{4} v_1^{k_{filter}+1}, \tag{3.24a}$$

and the standard deviation of the error σ_ϵ is due to the noise

$$\sigma_\epsilon = \sigma_v \sqrt{\frac{\Delta t}{\tau}}, \tag{3.24b}$$

where we have used that $P = \tau/\Delta t$, and that $v_i'^{k_{filter}}$ are uncorrelated random variables with a standard deviation σ_v. The (absolute value of the) filter error (3.24a) increases with the assimilation time τ, since the advection operator (3.18) re-distributes the

Fourier modes from the resolved part of the spectrum $k \leqslant k_{filter}$ to the unresolved part $k > k_{filter}$. The noise error (3.24b), on the other hand, decreases with τ, i.e. it decreases as the inverse square root of the number of samples $P \sim \tau$, similar as in the diffusion problem (3.17).

It is recalled, that μ_ϵ and σ_ϵ in (3.24) are obtained by linearly expanding $\boldsymbol{G}_i$ in $i\Delta t$ (3.19). In figure 1(c) we plot μ_ϵ and σ_ϵ, based on the full nonlinear operator (3.8b), and by numerically evaluating (3.7), (3.13), (3.18), where we have used $k_{filter} = 1$. In agreement with (3.24), the figure shows that the (absolute value of the) filter error $|\mu_\epsilon|$ increases, and the noise error σ_ϵ decreases, both as functions of τ.

4. Two-dimensional turbulence

4.1. *Set-up*

In this section we study the performance of the STIM (2.2), to improve the accuracy of synthetic PIV data for a 2-D incompressible, decaying turbulent velocity field $\boldsymbol{v}$, on a square bi-periodic domain of size $L = 2\pi$. The flow starts at $t = 0$ from a random velocity field, with a velocity norm of $\mathscr{U} = \|\boldsymbol{v}\| = 1$.

It is emphasised, that the present test case is a purely 2-D flow, which may be realised in the laboratory using large aspect ratio flow cells (e.g. Shats *et al.* 2007) or soap films (e.g. Gillissen *et al.* 2018). Since the reconstruction principles are equivalent in two and three dimensions, these principles are studied most effectively in two dimensions. Purely 2-D flow is not to be confused with planar PIV measurements of a 3-D flow. In this context, it is noted, that measuring turbulence in a single plane is insufficient to reconstruct volumetric turbulence, which instead requires three-dimensional PIV, involving for instance the use of four cameras, and tomographic reconstruction algorithms (e.g. Scarano 2012). It is further noted that the characteristics of the measurement errors in tomographic PIV are different than in standard PIV, and the ability of STIM to reduce these errors deserves further study.

We non-dimensionalise our variables using the velocity scale $\mathscr{U}$, and the length scale $L/(2\pi)$, which corresponds to a time scale $L/(2\pi\mathscr{U})$. Starting from random conditions at $t = 0$, the flow takes a few time units to develop physical structures. We apply the reconstruction thereafter, i.e. within the period $2 < t < 2 + T$, where the reconstruction interval is chosen to be $T = 10$. We define the following PIV errors for the kinetic energy $\tilde{\epsilon}_K$:

$$\tilde{\epsilon}_K = T^{-1} \int_2^{2+T} \frac{\displaystyle\int (\tilde{\boldsymbol{v}} - \boldsymbol{v})^2 \, \mathrm{d}V}{\displaystyle\int \boldsymbol{v}^2 \, \mathrm{d}V} \, \mathrm{d}t, \tag{4.1}$$

for the pressure $\tilde{\epsilon}_p$

$$\tilde{\epsilon}_p = T^{-1} \int_2^{2+T} \frac{\displaystyle\int p^2(\tilde{\boldsymbol{v}} - \boldsymbol{v}) \, \mathrm{d}V}{\displaystyle\int p^2(\boldsymbol{v}) \, \mathrm{d}V} \, \mathrm{d}t, \tag{4.2}$$

and for the energy dissipation $\tilde{\epsilon}_{\mathscr{D}}$

$$\tilde{\epsilon}_{\mathscr{D}} = T^{-1} \int_2^{2+T} \frac{\displaystyle\int \mathscr{D}(\tilde{\boldsymbol{v}} - \boldsymbol{v}) \, \mathrm{d}V}{\displaystyle\int \mathscr{D}(\boldsymbol{v}) \, \mathrm{d}V} \, \mathrm{d}t, \tag{4.3}$$

where it is recalled that v is the true velocity field and $\tilde{v}$ is the synthetic PIV velocity field. The dissipation field is given by

$$\mathscr{D}(u) = \nu \nabla u : \nabla u, \tag{4.4}$$

and the pressure field is computed from the equation of state

$$\nabla^2 p(u) = -\nabla\nabla : uu, \tag{4.5}$$

and reads in Fourier space

$$p^k(u) = -\frac{kk : u^k u^k}{k^2}, \tag{4.6}$$

where the superscript indices refer to the Fourier wavevectors. Similarly we define the following reconstruction errors for the kinetic energy ϵ_K

$$\epsilon_K = T^{-1} \int_2^{2+T} \frac{\int (u - v)^2 \, dV}{\int v^2 \, dV} \, dt, \tag{4.7}$$

for the pressure ϵ_p

$$\epsilon_p = T^{-1} \int_2^{2+T} \frac{\int p^2(u - v) \, dV}{\int p^2(v) \, dV} \, dt, \tag{4.8}$$

and for the energy dissipation $\epsilon_{\mathscr{D}}$

$$\epsilon_{\mathscr{D}} = T^{-1} \int_2^{2+T} \frac{\int \mathscr{D}(u - v) \, dV}{\int \mathscr{D}(v) \, dV} \, dt, \tag{4.9}$$

where it is recalled that u is the reconstructed velocity field.

The reconstruction time interval $2 < t < 2 + T$ is split into segments of size τ, referred to as the assimilation time. In each of these segments, a reconstruction problem is solved, which finds the velocity field at the start of that segment u_1. While the initial guess for the initial conditions of the first segment is zero: $u_1 = 0$, a considerable improvement of the reconstruction is achieved by using for the initial guess for the initial conditions of the following segments, the final conditions of each previous segment (Gillissen *et al.* 2018). Figure 3(*a*) illustrates this improvement, by plotting the time dependence of the instantaneous reconstruction error in the kinetic energy ϵ'_K

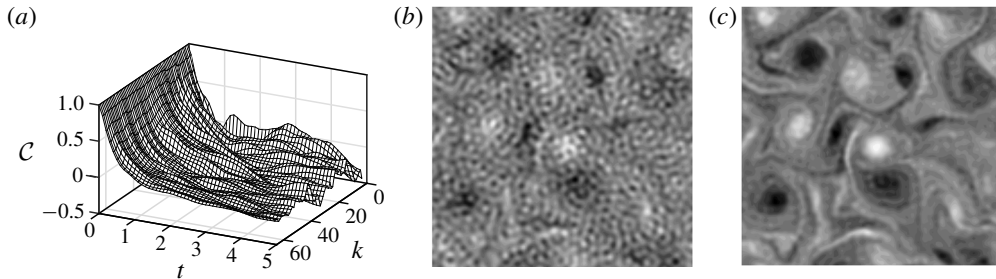$$\epsilon'_K = \frac{\int (u - v)^2 \, dV}{\int v^2 \, dV}. \tag{4.10}$$

FIGURE 2. (*a*) Auto-correlation $\mathcal{C}$ of the various modes (with absolute wavevector k), of the true vorticity field $\nabla \times \boldsymbol{v}$, as a function of the time distance t between the pair of values, obtained by analysing data in the time interval $2 < t < 12$. (*b*) Vorticity field at time $t = 2$ of the synthetic PIV data $\nabla \times \tilde{\boldsymbol{v}}$ using a noise level of $\sigma_v = 0.3$, and a reciprocal filter width of $k_{filter} = 32$. (*c*) Vorticity field at time $t = 2$ of the corresponding STIM reconstruction $\nabla \times \boldsymbol{u}$, using an assimilation time of $\tau = 0.16$ and a sampling time of $\Delta t = 5 \times 10^{-3}$.

In this figure each line segment corresponds to one assimilation window with a time interval of τ.

The synthetic PIV data $\tilde{\boldsymbol{v}}$ are generated by adding noise to the true velocity field $\boldsymbol{v}$. The true velocity field is generated by numerically simulating the Navier–Stokes equations (2.1), which in the absence of forcing mechanisms, and starting from random initial conditions, results in freely decaying turbulence. The numerical method uses Fourier functions to compute spatial derivatives (e.g. Canuto *et al.* 1988). Time integration is performed using the second-order explicit Adams–Bashforth scheme for the advection terms and the second-order implicit Crank–Nicolson scheme for the diffusion terms. The number of grid points is $N^2 = 128^2$, which is sufficient to resolve all length scales, and the numerical integration time step is 1×10^{-3}, which is sufficiently small, such that the numerical solution does not change upon reducing the time step. The initial random velocity field is constructed by assigning random numbers to the Fourier modes, whose absolute wavevectors $|\boldsymbol{k}| \leqslant 8$, while the remaining Fourier modes are assumed zero. The initial velocity field is normalised, such that $\mathcal{U} = \|\boldsymbol{v}\| = 1$ at $t = 0$. The viscosity is $\nu = 7.3 \times 10^{-4}$, which corresponds to a Reynolds number, based on the velocity scale $\mathcal{U} = 1$ and on the domain size $L = 2\pi$, of $Re = \mathcal{U}L/\nu = 8.6 \times 10^3$. The correlation time of the resulting turbulent flow is around $\mathcal{T} \sim 1$, see figure 2(*a*).

To generate synthetic PIV data $\tilde{\boldsymbol{v}}$, we add noise to the simulated, true velocity field $\boldsymbol{v}$. In this work we consider two types of noise. First, we consider in § 4.2 white PIV noise, which is an idealisation compared to realistic PIV noise. This simplification, however, allows for a systematic variation of the noise properties and a clear interpretation of the method performance, through a comparison with the analytical results in § 3. In addition, we consider in § 4.3, realistic PIV noise, that are generated by computationally mimicking the PIV measurement process. To this end we employ a simplistic PIV algorithm, without employing state of the art PIV noise reduction techniques. The resulting synthetic PIV error is highly correlated, which provides an extremely stringent test case for STIM. This is opposite to the white noise error in § 4.2, which is tackled effectively by STIM. The extreme cases of white noise and simplistic PIV noise mark the upper and lower boundaries for the effectiveness of the STIM.
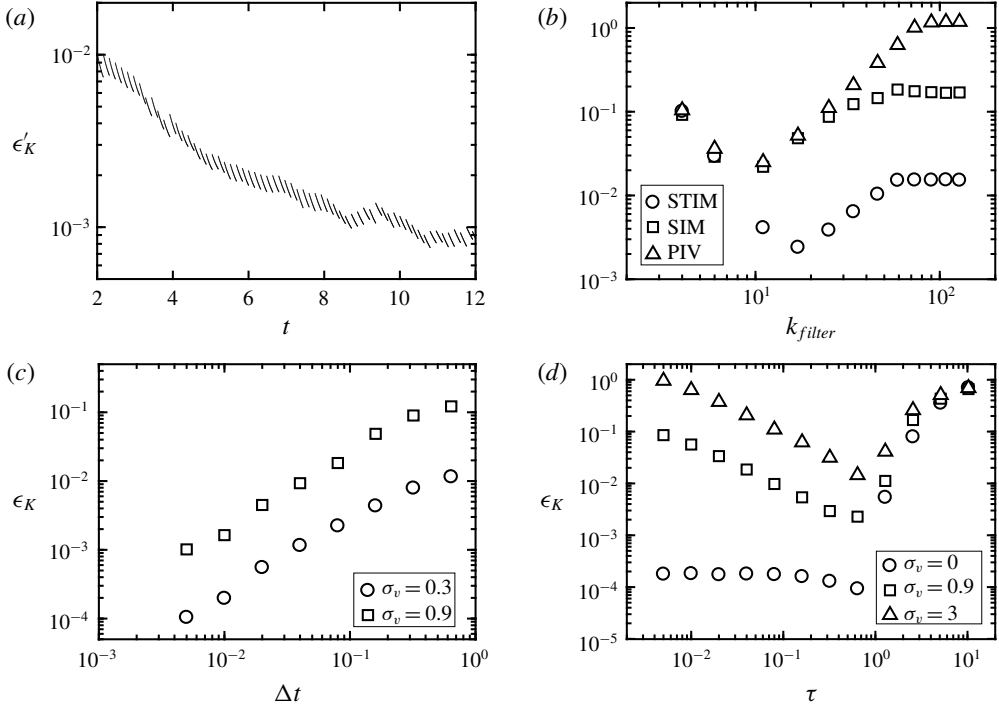
FIGURE 3. STIM performance for idealised synthetic PIV data with white noise. (*a*) Instantaneous STIM reconstruction error in the kinetic energy ϵ'_K (4.10) as a function of time t, using a reciprocal filter width of $k_{filter} = 16$, a noise level of $\sigma_v = 0.3$, an assimilation time of $\tau = 0.16$ and a sampling time of $\Delta t = 5 \times 10^{-3}$. (*b*) PIV error in the kinetic energy $\tilde{\epsilon}_K$ (4.1) (triangles) and reconstruction error in the kinetic energy ϵ_K (4.7) using STIM (circles) and SIM (squares), as functions of k_{filter}, using $\sigma_v = 0.9$, $\tau = 0.16$ and $\Delta t = 5 \times 10^{-3}$. (*c*) STIM reconstruction error ϵ_K as a function of Δt, using $k_{filter} = 32$, $\tau = 0.64$ and varying σ_v. (*d*) STIM reconstruction error ϵ_K as a function of τ, using $k_{filter} = 32$, $\Delta t = 5 \times 10^{-3}$ and varying σ_v. In (*b*–*d*), ϵ_K and $\tilde{\epsilon}_K$ are averaged over five reconstructions, using different true velocity fields v with different (random) initial conditions.

4.2. *White PIV noise*

First we consider idealised PIV data $\tilde{v}$, which are constructed by adding white noise to the true velocity field v, and subsequently applying a filter (3.5). The standard deviation of the noise σ_v is referred to as the noise level, and the reciprocal filter width is denoted $k_{filter} = 2\pi/\ell$.

Figure 2(*b*) visualises the resulting synthetic PIV vorticity field $\nabla \times \tilde{v}$ at $t = 2$, using a noise level of $\sigma_v = 0.3$ and a reciprocal filter width of $k_{filter} = 32$. The corresponding reconstructed vorticity $\nabla \times u$, that is produced by the STIM, is shown in figure 2(*c*), where we have used an assimilation time of $\tau = 0.16$ and a sampling time of $\Delta t = 5 \times 10^{-3}$. The noise in the reconstruction is significantly reduced, as compared to the synthetic PIV data in figure 2(*b*).

Figure 3(*b*) shows the PIV error in the kinetic energy $\tilde{\epsilon}_K$ (4.1) as a function of the reciprocal filter width k_{filter} for a noise level of $\sigma_v = 0.9$. It is seen that $\tilde{\epsilon}_K$ depends non-monotonically on k_{filter}, i.e. an initial decrease is followed by an increase. This

non-monotonicity reflects, that low wavenumber modes contain more flow structure than noise, while large wavenumber modes contain more noise than structure. Low wavenumber modes therefore improve the measurement, while high wavenumber modes deteriorate the measurement.

Figure 3(*b*) also shows the corresponding STIM reconstruction error in the kinetic energy ϵ_K (4.7) as a function of k_{filter}, where we have used an assimilation time of $\tau = 0.16$ and a sampling time of $\Delta t = 5 \times 10^{-3}$. It is seen that ϵ_K follows the non-monotonic trend of the PIV error $\tilde{\epsilon}_K$. For $k_{filter} \leqslant 4$, the STIM does not improve the PIV data, i.e. $\tilde{\epsilon}_K \approx \epsilon_K$, while for $k_{filter} \geqslant 4$, the STIM improves the accuracy of the data, with a maximum error reduction of two decades in the absence of spatial filtering, i.e. for $k_{filter} = N = 128$.

Next we compare the performance of the STIM (2.2) to that of the SIM (2.9). Whereas STIM reduces measurement noise by fitting a solution to the Navier–Stokes equation to time-dependent measurement data, SIM achieves this, by applying a hyper diffusion process to the measurement data. The hyper diffusivity is referred to as the regularisation strength κ. This parameter must be chosen sufficiently large, as to dampen the small-scale noise, and sufficiently small, as to not affect the large-scale, energy-containing eddies. A parametric study (not shown) revealed that $\kappa = 10^{-6}$ $\mathcal{U}[L/(2\pi)]^3$ is a suitable value, for the test case at hand. Figure 3(*b*) shows the resulting SIM reconstruction error in the kinetic energy ϵ_K (4.7) as a function of k_{filter}. As expected, SIM is less effective in reducing the measurement error than STIM. In the absence of spatial filtering ($k_{filter} = N = 128$) SIM achieves a tenfold reduction in measurement error, whereas STIM achieves a hundredfold reduction under identical conditions.

As STIM fits a solution to the Navier–Stokes equation to time-dependent measurement data, the performance of STIM, in the case of white measurement noise, increases with the sampling frequency, i.e. the STIM reconstruction error ϵ_K (4.7) increases with the sampling time Δt. This feature is illustrated in figure 3(*c*), where we plot ϵ_K as a function of Δt, and where we have used a reciprocal filter width of $k_{filter} = 32$, an assimilation time of $\tau = 0.64$ and various noise levels σ_v. The data in figure 3(*c*) show that the error in the kinetic energy ϵ_K depends linearly on Δt which is consistent with the square root dependence of the scalar error on Δt for the linear diffusion problem (3.17).

Figure 3(*d*) shows the STIM reconstruction error ϵ_K (4.7) as a function of the assimilation time τ for a reciprocal filter width $k_{filter} = 32$, a sampling time of $\Delta t = 5 \times 10^{-3}$ and for various noise levels σ_v. It is seen that ϵ_K depends non-monotonically on τ, where for small τ the error decreases, while for large τ the error increases, in qualitative agreement with linear theory (3.17), (3.24). This error increase reflects that information de-correlates over long times, making the inverse problem of minimising (2.2) ill-posed. In the linear advection problem, a similar increase in the reconstruction error ϵ_K with τ was observed ((3.24*b*); figure 1*c*), which is related to the loss of information from the resolved modes into the unresolved (filtered) modes. This loss of information therefore bears an analogy with the de-correlation of information in the 2-D turbulent flow, which corresponds to a dissipation of kinetic energy, via the energy cascade, into heat.

4.3. *Realistic PIV noise*

Next we generate synthetic velocimetry data with realistic PIV noise by computing trajectories of passive point particles. The total number of particles $N_p(L/\ell)^2$ is

chosen such, that there are on average N_p particles inside the PIV interrogation window of size ℓ. The interrogation window corresponds to 32^2 image pixels. The image size of the total domain is therefore $(32L/\ell)^2$ pixels. Each particle generates a light intensity image, which is a Gaussian function with a standard deviation of 0.7 pixels. A PIV image is constructed by superimposing the Gaussian images of all the particles. To compute the PIV velocity field at $t = t_i$, we generate two particle images at $t = t_i - \frac{1}{2}\Delta t_{PIV}$ and $t = t_i + \frac{1}{2}\Delta t_{PIV}$. The image time difference equals $\Delta t_{PIV} = 0.02\ell/\mathscr{U}$. To find the PIV velocity vector at position $\boldsymbol{x}$ we apply a mask of width ℓ and centre $\boldsymbol{x}$ to these images, and maximise the correlation between the masked images at $t = t_i - \frac{1}{2}\Delta t_{PIV}$ and $t = t_i + \frac{1}{2}\Delta t_{PIV}$, as given in (1.1). The PIV velocity field is determined using a sliding interrogation window on $N^2 = 128^2$ spatial points, which is the same spatial grid, as is used in the fluid flow simulations. The location of the correlation maximum is determined with sub-pixel accuracy by fitting a paraboloid to the correlation function in a region of 3×3 pixels around the pixel with the largest intensity.

We apply the PIV algorithm described above, to the simulated two-dimensional turbulent flow, which is described in §4.1. The resulting synthetic PIV error $\tilde{\boldsymbol{v}} - \boldsymbol{v}$ is correlated with the true velocity $\boldsymbol{v}$, as shown by the joint probability density function of $\tilde{\boldsymbol{v}} - \boldsymbol{v}$ and $\boldsymbol{v}$ in figure 4(a). Figures 4(b), 4(c) and 4(d) compare the corresponding true velocity field $\boldsymbol{v}$ the synthetic PIV velocity field $\tilde{\boldsymbol{v}}$ and the reconstructed velocity field $\boldsymbol{u}$ using STIM, respectively. The comparison shows, that STIM significantly improves the accuracy of the PIV data. The data in figure 4 are taken at time $t = 4$, using a reciprocal filter width of $k_{filter} = 2\pi/\ell = 32$, a PIV particle density of $N_p = 2$, an assimilation time of $\tau = 0.64$ and a sampling time of $\Delta t = 5 \times 10^{-3}$.

In figure 5(a), we study the PIV error in the kinetic energy $\tilde{\epsilon}_K$ (4.1) as a function of N_p, using $k_{filter} = 32$, $\tau = 0.64$ and $\Delta t = 5 \times 10^{-3}$. As expected $N_p < 1$ is ineffective, resulting in a 100 % error $\tilde{\epsilon}_K \approx 1$. When N_p exceeds unity, $\tilde{\epsilon}_K$ drops sharply, and for large N_P the error saturates at $\tilde{\epsilon}_K \sim 10^{-2}$.

In figure 5(b) we study the PIV error in the kinetic energy $\tilde{\epsilon}_K$ (4.1) in the pressure $\tilde{\epsilon}_p$ (4.2) and in the energy dissipation $\tilde{\epsilon}_{\mathscr{D}}$ (4.3) as functions of k_{filter}, using $N_p = 2$. The PIV error in the pressure $\tilde{\epsilon}_p$ is similar to that in the kinetic energy $\tilde{\epsilon}_K$, and these errors increase with decreasing k_{filter}. This reflects, that a larger interrogation window ℓ cuts off more energy, resulting in larger errors. The PIV error for the dissipation $\tilde{\epsilon}_{\mathscr{D}}$ is two orders of magnitude larger than that for the energy and the pressure, and $\tilde{\epsilon}_{\mathscr{D}}$ is independent of k_{filter}. This independence reflects, that the dissipative scales are poorly captured by the PIV image correlation algorithm, irrespective of k_{filter}, in the range considered.

In addition to the PIV errors, we also plot in figure 5(b), the corresponding STIM reconstruction errors for the kinetic energy ϵ_K (4.7), the pressure ϵ_p (4.8) and the dissipation $\epsilon_{\mathscr{D}}$ (4.9), where we used $\tau = 0.64$ and $\Delta t = 5 \times 10^{-3}$. It is seen, that STIM reduces the PIV error by a factor of around three for both the energy and the pressure, while it reduces the error for the dissipation by a factor of up to 100. This result illustrates, that STIM significantly improves the structure of the dissipating eddies, which are generally poorly represented by PIV. This conclusion is also supported by the snapshots in figure 4.

STIM (2.2) fits a solution of the Navier–Stokes equations to time-dependent measurement data. We have seen in figure 3(d), that, in the case of white measurement noise, the accuracy of the fit can be made arbitrarily small, by increasing the data density, i.e. by reducing the sampling time Δt. For realistic PIV noise, the situation is different, however. This is illustrated in figure 5(c), showing, that for synthetic
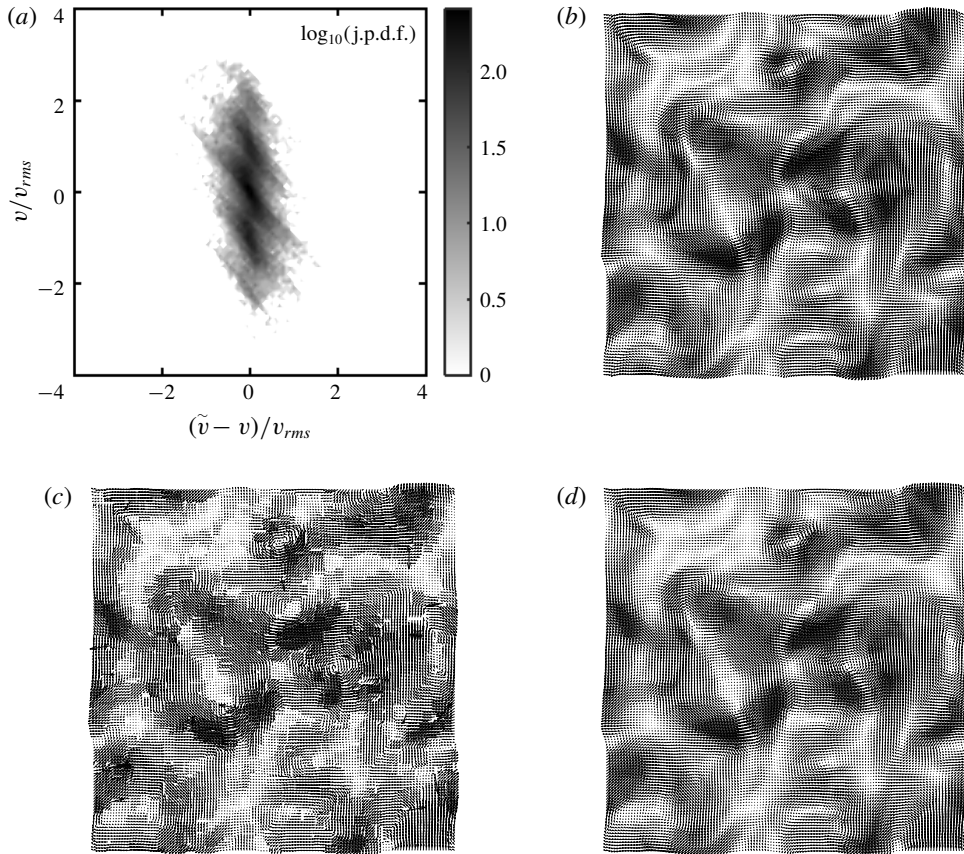
FIGURE 4. (*a*) Joint probability density function (j.p.d.f.) of $\tilde{v} - v$ and v, where $\tilde{v}$ is the synthetic PIV velocity field and v is the true velocity field. (*b*) Snapshot of the true velocity field v. (*c*) Snapshot of the synthetic PIV velocity field $\tilde{v}$. (*d*) Snapshot of the STIM reconstructed velocity field u. The data in (*a–d*) are taken at time $t = 4$, using a reciprocal filter width of $k_{filter} = 2\pi/\ell = 32$, an assimilation window of $\tau = 0.64$ and a sampling time of $\Delta t = 5 \times 10^{-3}$.

velocimetry data with realistic PIV noise, the STIM reconstruction error in the kinetic energy ϵ_K does not depend on Δt, provided that Δt is smaller than the flow correlation time $\mathscr{T} \sim 1$ (figure 2*a*). This means, that the maximum sampling time, that still produces a reasonable reconstruction, is of the order of the flow correlation time, and adding more intermediate samples does not necessarily improve the reconstruction.

5. Conclusions

We have derived a space–time integral method to reduce the PIV measurement noise. The performance of the method is studied using synthetic PIV data in a 2-D turbulent flow. Under the assumption of white measurement noise, there is a non-monotonic relationship between the reconstruction error ϵ and the assimilation time τ, where ϵ decreases with τ, when τ is smaller than the turbulent correlation time $\mathscr{T}$, while ϵ increases with τ, for $\tau \gtrsim \mathscr{T}$. This supports the notion that predicting the flow at time
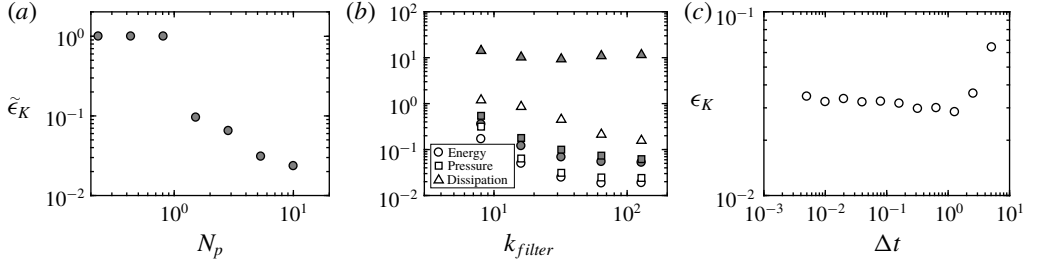
FIGURE 5. STIM performance for synthetic velocimetry data with realistic PIV noise. (*a*) The PIV error in the kinetic energy $\tilde{\epsilon}_K$ (4.1) as a function of the PIV particle density N_p, using a reciprocal filter width of $k_{filter} = 2\pi/\ell = 32$, an assimilation time of $\tau = 0.64$ and a sampling time of $\Delta t = 5 \times 10^{-3}$. (*b*) The PIV error in the kinetic energy $\tilde{\epsilon}_K$ (grey circles; (4.1)), in the pressure $\tilde{\epsilon}_p$ (grey squares; (4.2)) and in the energy dissipation $\tilde{\epsilon}_\mathcal{D}$ (grey triangles; (4.3)) and the STIM reconstruction error in the kinetic energy ϵ_K (white circles; (4.7)), in the pressure ϵ_p (white squares; (4.8)) and in the energy dissipation $\epsilon_\mathcal{D}$ (white triangles; (4.9)) as functions of k_{filter}, using $N_p = 2$, $\tau = 0.64$ and $\Delta t = 5 \times 10^{-3}$. (*c*) The STIM reconstruction error in the kinetic energy ϵ_K (4.7) as a function of Δt, using $k_{filter} = 32$, $N_p = 2$ and $\tau = \Delta t$. In (*a*–*c*), the presented errors are averaged over five simulations, using different true velocity fields $\boldsymbol{v}$ with different (random) initial conditions.

t, based on flow measurements at time $t + \tau$, is an ill-posed problem, when τ exceeds the flow correlation time.

To interpret these numerical results, we compare them to linear problems, for which we have derived analytical expressions for ϵ, which show a similar non-monotonic dependence on τ. This non-monotonicity is explained by a decomposition of ϵ into a noise contribution, which decreases with τ, due to an increase in the number of samples, and a filter contribution, which increases with τ, due to a loss of information from resolved to unresolved modes.

We also apply the method to synthetic velocimetry data with realistic PIV noise in a 2-D turbulent flow. The method is shown to reduce the realistic PIV noise and to improve the small scale structures, although to a lesser extent as for the idealised PIV data, with white measurement noise.

Acknowledgements

REFERENCES

ADRIAN, R. J. & WESTERWEEL, J. 2011 *Particle Image Velocimetry*. Cambridge University Press.

CANUTO, C., HUSSAINI, M. Y., QUARTERONI, A. M. & THOMAS, A. JR 1988 *Spectral Methods in Fluid Dynamics*. Springer.

FOURES, D. P. G., DOVETTA, N., SIPP, D. & SCHMID, P. J. 2014 A data-assimilation method for Reynolds-averaged Navier–Stokes-driven mean flow reconstruction. *J. Fluid Mech.* **759**, 404–431.

GILLISSEN, J. J. J., VILQUIN, A., KELLAY, H., BOUFFANAIS, R. & YUE, D. K. P. 2018 A space–time integral minimisation method for the reconstruction of velocity fields from measured scalar fields. *J. Fluid Mech.* **854**, 348–366.

GRONSKIS, A., HEITZ, D. & MÉMIN, E. 2013 Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. *J. Comput. Phys.* **242**, 480–497.

GUNZBURGER, M. D. 2003 *Perspectives in Flow Control and Optimization*. SIAM.

HUMPHREY, J. A. C. 1993 Interactive computational-experimental methodologies (ICEME) for thermofluids research: application to the optimized packaging of heated electronic components. In *Computers and Computing in Heat Transfer Science and Engineering* (ed. K. T. Yang & W. Nakayama), pp. 293–317. CRC and Begell House.

MONS, V., CHASSAING, J.-C., GOMEZ, T. & SAGAUT, P. 2016 Reconstruction of unsteady viscous flows using data assimilation schemes. *J. Comput. Phys.* **316**, 255–280.

RUHNAU, P., KOHLBERGER, T., SCHNÖRR, C. & NOBACH, H. 2005 Variational optical flow estimation for particle image velocimetry. *Exp. Fluids* **38** (1), 21–32.

RUHNAU, P., STAHL, A. & SCHNÖRR, C. 2007 Variational estimation of experimental fluid flows with physics-based spatio-temporal regularization. *Meas. Sci. Technol.* **18** (3), 755–763.

SCARANO, F. 2012 Tomographic PIV: principles and practice. *Meas. Sci. Technol.* **24** (1), 012001.

SHATS, M. G., XIA, H., PUNZMANN, H. & FALKOVICH, G. 2007 Suppression of turbulence by self-generated and imposed mean flows. *Phys. Rev. Lett.* **99** (16), 164502.

SUZUKI, T. 2012 Reduced-order Kalman-filtered hybrid simulation combining particle tracking velocimetry and direct numerical simulation. *J. Fluid Mech.* **709**, 249–288.

SYMON, S., DOVETTA, N., MCKEON, B. J., SIPP, D. & SCHMID, P. J. 2017 Data assimilation of mean velocity from 2D PIV measurements of flow over an idealized airfoil. *Exp. Fluids* **58** (5), 61.

TALAGRAND, O. & COURTIER, P. 1987 Variational assimilation of meteorological observations with the adjoint vorticity equation. I. Theory. *Q. J. R. Meteorol. Soc.* **113** (478), 1311–1328.

TIKHONOV, A. N. & ARSENIN, V. Y. 1977 *Solutions of Ill-Posed Problems*. Winston.