# Optimal mixing in two-dimensional plane Poiseuille flow at finite Peclet number

D. P. G. Foures, C. P. Caulfield, Peter J. Schmid

*J. Fluid Mech.* (2014), *vol.* 748, *pp.* 241–277.  
doi:10.1017/jfm.2014.182

# Optimal mixing in two-dimensional plane Poiseuille flow at finite Péclet number

**D. P. G. Foures[1], C. P. Caulfield[2,1,†] and P. J. Schmid[3]**

[1]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

[2]BP Institute, University of Cambridge, Madingley Rise, Madingley Road, Cambridge, CB3 0EZ, UK

[3]Laboratoire d'Hydrodynamique (LadHyX), CNRS-École Polytechnique, 91128 Palaiseau, France

(Received 8 November 2013; revised 8 November 2013; accepted 31 March 2014; first published online 28 April 2014)

We consider the nonlinear optimisation of the mixing of a passive scalar, initially arranged in two layers, in a two-dimensional plane Poiseuille flow at finite Reynolds and Péclet numbers, below the linear instability threshold. We use a nonlinear-adjoint-looping approach to identify optimal perturbations leading to maximum time-averaged energy as well as maximum mixing in a freely evolving flow, measured through the minimisation of either the passive scalar variance or the so-called mix-norm, as defined by Mathew, Mezić & Petzold (*Physica D*, vol. 211, 2005, pp. 23–46). We show that energy optimisation appears to lead to very weak mixing of the scalar field whereas the optimal mixing initial perturbations, despite being less energetic, are able to homogenise the scalar field very effectively. For sufficiently long time horizons, minimising the mix-norm identifies optimal initial perturbations which are very similar to those which minimise scalar variance, demonstrating that minimisation of the mix-norm is an excellent proxy for effective mixing in this finite-Péclet-number bounded flow. By analysing the time evolution from initial perturbations of several optimal mixing solutions, we demonstrate that our optimisation method can identify the dominant underlying mixing mechanism, which appears to be classical Taylor dispersion, i.e. shear-augmented diffusion. The optimal mixing proceeds in three stages. First, the optimal mixing perturbation, energised through transient amplitude growth, transports the scalar field across the channel width. In a second stage, the mean flow shear acts to disperse the scalar distribution leading to enhanced diffusion. In a final third stage, linear relaxation diffusion is observed. We also demonstrate the usefulness of the developed variational framework in a more realistic control case: mixing optimisation by prescribed streamwise velocity boundary conditions.

**Key words:** flow control, mixing enhancement, variational methods

## 1. Introduction

The study of mixing mechanisms in a flow is a long-standing problem in fluid mechanics due both to its inherent complexity and its wide-ranging applicability. Indeed, understanding and modelling the ocean's circulation (Wunsch & Ferrari

† Email address for correspondence: c.p.caulfield@bpi.cam.ac.uk

2004), the prediction of the spreading of a pollutant in the atmosphere, the ocean, a lake or a river (Lekien *et al.* 2005), the natural ventilation in a room or a building (Linden 1999), the study of mixing of air and fuel in an engine (Annaswamy & Ghoniem 1995) or the mixing in microfluidic devices (Hessel, Löwe & Schönfeld 2005; Nguyen & Wu 2005) are all problems where mixing holds a central role. Mixing quantification and in some cases control (inhibition or enhancement) are very important for many geophysical, environmental or industrial applications.

The mixing of a passive scalar was first identified as requiring two ordered processes by Eckart (1948). First, relatively large-scale 'stirring' creates filaments of passive scalar and in a second stage molecular diffusion acts to smooth out the filamented concentration field. Mixing is therefore strongly linked to the concurrent action of advection and diffusion. Taylor (1953) studied the action of a shear flow on a passive scalar field, and was able to show that diffusion is enhanced in zones of large shear (see also Rhines & Young 1983) by a mechanism now commonly known as Taylor dispersion. Later, with the development of the dynamical systems description of fluid flows, the possibility of mixing by chaotic advection was investigated and described by Aref (1984) with his famous blinking vortex system. It was shown that stretching and folding mechanisms (typical of chaotic flows) are very effective at mixing (for a review of mixing processes, see Ottino 1990). Despite a large body of literature on the subject, no unified theory for mixing has yet emerged, principally because of a lack of a way to quantify the phenomenon.

In fact, one of the major challenges when considering the mixing of a passive scalar in a flow consists of defining a way to measure the mixing properly and quantitatively. There is no universally accepted measure of mixing, and although ergodic theory defines a 'mixing flow' rigorously (see Thiffeault 2012), this definition is far from being practical in real fluid problems. Therefore, other measures have been applied to assess the mixedness of a flow. Variance has historically been used in order to quantify mixing (Danckwerts 1952) but this approach fails when there is no molecular diffusion, because it is unable to quantify pure stirring effects. D'Alessandro, Dahleh & Mezic (1999) proposed quantifying mixing using an entropy definition, but this approach is unfortunately not adapted to a continuous passive scalar field (see Mathew, Mezić & Petzold 2005 for an overview of mixing quantification methods). Mathew *et al.* (2005) introduced a new measure to quantify mixing in spatially periodic domains: the mix-norm. Based on Sobolev norms of negative index (as defined more carefully in (2.5)) which de-emphasise the role of small scales, this class of measures is sensitive to both stirring and diffusion, and is derived from the theoretical definition of mixing given by the ergodic theory.

Naturally, the need to quantify mixing arises when one wants to control this process, either by enhancing or preventing it. As noted above, D'Alessandro *et al.* (1999) studied the problem of optimal mixing via an entropy maximisation approach. However, the flow was not governed by the Navier–Stokes equations, but was prescribed as two orthogonal shear flows. An optimal steady Stokes flow ($Re \to 0$) obtained by superposition of two flows (array of vortices) was obtained by Mathew *et al.* (2007) by optimising the mix-norm. Once again, there was no incorporated dynamics for the velocity flow and the model does not embed the molecular diffusion of the passive scalar. A mixed cost functional based on variance minimisation was also used by Liu (2008). This model embedded diffusion for the passive scalar but no flow dynamics: the velocity field was once again prescribed, and was not a freely evolving solution of the Navier–Stokes equations. Yet another mix-norm optimisation was performed by Lin, Thiffeault & Doering (2011). The mixing flow was optimally

found via an adjoint-based optimisation but was still not required to be a solution of the Navier–Stokes equations. (A very clear and complete review on mixing measures and optimisation is given in Thiffeault 2012.)

In parallel, because of the lack of an obvious way to measure mixing, optimal destabilisation based on energy optimisation has often been considered as a proxy for mixing optimisation. This originates from the hypothesis that an energetic perturbation being associated with a strong advection field will induce stirring and hence ultimately effective mixing. Indeed mixing via destabilisation of the flow was thoroughly studied by Aamo and co-workers (Aamo & Krstić 2003; Aamo, Krstić & Bewley 2003; Aamo & Krstić 2004; Balogh, Aamo & Krstić 2005). The mixing of discrete particles was considered and the quality of the mixing assessed via a box-counting algorithm. In this case, the flow was a solution of the Navier–Stokes equations, and control was performed through wall blowing and suction. Despite the control not being designed for mixing optimisation, effective mixing was indeed observed. We may however wonder whether the same mixing quality can be obtained with less effort with a direct optimisation of 'mixedness'.

Furthermore, because mixing enhancement can only be performed via stirring control which is an inherently nonlinear phenomenon, we have to consider the full nonlinear Navier–Stokes operator in any variational formulation of the optimisation problem. A variational formulation based on a Lagrange multipliers approach to enforce the Navier–Stokes equation was recently applied to nonlinear flows (Cherubini *et al.* 2010; Pringle & Kerswell 2010; Juniper 2011; Monokrousos *et al.* 2011; Rabin, Caulfield & Kerswell 2012, 2014). We will therefore follow and generalise the formulations presented in those studies. However, previous nonlinear optimisation studies were carried out to investigate nonlinear energetic optimal perturbations, typically to identify an 'optimal' route to turbulence from a laminar state in various shear flows. Therefore, those studies focused on appropriate measures to distinguish turbulent and laminar flow, either the finite-time maximisation of the energy defined as the $L^2$-norm of the velocity vector squared or the time-averaged dissipation (Monokrousos *et al.* 2011).

However, in this paper, we are interested in a finite-time optimal value of 'mixedness'. This concept is defined using various norms of the passive scalar field (as we briefly outlined earlier) which we wish to homogenise. By exploring different mixing measures we are able to characterise their respective usefulness. In particular, we are interested in whether measures based on the mix-norm are appropriate proxies for mixing in bounded flows with finite diffusivities, thus demonstrating a generalisation of their applicability from their original theoretical development. We apply the nonlinear variational framework to find nonlinear optimal initial perturbations maximising the final time value of mixedness. Also, in order to investigate the question of effective mixing via an energy amplification mechanism, we find optimal initial perturbations maximising the time-average value of the energy. We find that energy-optimal perturbations are not able to homogenise the passive scalar, in contrast to both (variance and mix-norm) mixing-optimal perturbations.

Furthermore, generalising Eckart's two-stage concept, we identify a generic three-stage mechanistic picture for long-time optimal mixing. The first stage is an energisation/transport stage during which the perturbation velocity displaces the concentration interface from the centre of the domain towards the walls as it is energised by transient growth mechanisms and grows in amplitude. Then, mixing occurs due to both perturbation advection (stirring) and background shear (Taylor-dispersion-enhanced diffusion) during a second dispersion/diffusion stage,

before a third final stage during which the flow 'relaxes' and purely diffuses. We conjecture that the relative strength of the two phenomena of stirring and dispersion is closely related to the initial amount of energy injected into the flow. Furthermore, optimal boundary conditions can also be optimised using the same variational framework (Fursikov, Gunzburger & Hou 1998; Mao, Blackburn & Sherwin 2012). Therefore, we demonstrate that a forcing strategy for the wall velocity boundary condition can be identified which leads to effective mixing and relies on the same three-stage mechanistic picture as identified for the optimal initial perturbation case.

The remainder of the paper is organised as follows. In § 2 we define several measures to quantify mixing and derive their time-evolution equations in order to gain some insight into their evolution subject to both advection and diffusion. We also present the classical two-stage mixing process due to both advection and diffusion. In order to study the behaviour of the several defined mixing norms, we study a spatially periodic toy model (chequerboard flow)which consists of an analytical mixing flow embedding both advection and molecular diffusion. In § 3.1 we present the governing equations and the flow configuration. Armed with the insight gained in § 3.1, we present in § 3.2 a variational frameworkdesigned to optimise either finite-time mixing or time-averaged energy in order to investigate the possibility of energy optimisation as a proxy for mixing optimisation. We present the results in § 4 for different control strategies (either by prescribing a finite-size initial perturbation or by imposing a streamwise velocity boundary condition). We eventually draw our conclusions in § 5.

## 2. Mathematical definitions of mixing

### 2.1. *Measures of mixing*

'Mixing' is a concept which is commonly used to characterise the properties of the spatial distribution of a species in a given domain. We focus our attention on the case where this distribution can entirely be described by a scalar field representing the species concentration. Even though a rigorous mathematical definition of a 'mixing flow' exists in the sense of ergodic theory (see Mathew *et al.* 2005; Thiffeault 2012), we believe it is fair to say that there is no universally accepted measure of mixing. A mixing measure should be an indicator of how uniformly a scalar field is distributed across the domain; loosely, it should be a measure of homogeneity or 'mixedness'. Let us consider a case where the scalar field is equal to its maximum value in a convex subspace of the domain $\Omega$ and vanishes everywhere else. This should naturally correspond to a minimal mixing measure (given the subspace is not too elongated). On the other hand, if the scalar field is equal to its spatial average everywhere in the domain, this homogeneous state should yield a maximal value of the mixing measure. Considering these two extreme situations, one is tempted to adopt the variance of the field as an appropriate measure for mixing, the first case of minimal mixing corresponding to a maximum variance value and the second case of maximal mixing corresponding to the minimum variance.

Without any loss of generality, we can consider a zero-mean scalar field $\theta(\boldsymbol{x}, t)$, such that the variance can simply be expressed by the $L^2$-norm of the scalar field:

$$\text{Var}\ \theta = \frac{1}{V_\Omega}\|\theta\|_2{}^2 = \frac{1}{V_\Omega}\int_\Omega \theta(\boldsymbol{x}, t)^2\ \mathrm{d}\Omega, \tag{2.1}$$

where $V_\Omega$ is the volume of the domain $\Omega$. This definition of mixing is the most natural one, because it measures the fluctuations of the scalar field from its mean. We consider the case where $\theta$ is governed by a (non-dimensional) advection–diffusion equation such as

$$\partial_t \theta + \boldsymbol{U} \cdot \boldsymbol{\nabla} \theta - Pe^{-1} \nabla^2 \theta = 0, \tag{2.2}$$

where $Pe = U_0 L_0 / \kappa$ is the Péclet number with $U_0$ and $L_0$ being characteristic velocity and length scales of the flow, and $\kappa$ is the diffusion coefficient of the fluid. Here $\boldsymbol{U}$ is the non-dimensional velocity field transporting the quantity $\theta$, either prescribed or a dynamically varying solution of evolution equations. We further require $\boldsymbol{\nabla} \cdot \boldsymbol{U} = 0$. With such an evolution equation for $\theta$, we can derive a differential equation describing the time evolution of its variance:

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\theta\|_2{}^2 = -Pe^{-1} \|\boldsymbol{\nabla} \theta\|_2{}^2. \tag{2.3}$$

The theoretical challenge arises when $Pe \to +\infty$ ($\kappa = 0$, no molecular diffusion), as this equation leads to $\|\theta(\boldsymbol{x}, t)\|_2 = \|\theta_0(\boldsymbol{x})\|_2$ for any time $t$. With the variance independent of time, mixing (in the variance sense) is not possible for any total velocity field $\boldsymbol{U}$.

However, we expect a mixing measure to be sensitive to the effect of pure advection of the scalar field, a process often referred to as 'stirring'. For a finite Péclet number $Pe$ ($\kappa \neq 0$), we see that the rate of decrease of the square of the variance is proportional to the norm of the concentration gradient squared. We can also derive an equation for the time evolution of the gradient norm in which the velocity field explicitly appears. We obtain

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\boldsymbol{\nabla} \theta\|_2{}^2 = -\int_\Omega \boldsymbol{\nabla} \theta \cdot \boldsymbol{\nabla} \boldsymbol{U} \cdot \boldsymbol{\nabla} \theta \, \mathrm{d}\Omega - Pe^{-1} \left\| \nabla^2 \theta \right\|_2^2. \tag{2.4}$$

By combining (2.3) and (2.4), we see that in a diffusive case the time evolution of the variance does depend on the velocity field $\boldsymbol{U}$. Therefore, it seems straightforward that increasing the concentration gradients via the velocity field $\boldsymbol{U}$ will promote a more effective diffusion process. In other words, the velocity gradient $\boldsymbol{\nabla U}$ acts as a 'catalyst' for diffusion.

Specifically, diffusion can thus be enhanced by shear, a phenomenon known as Taylor dispersion (see Taylor 1953). Shear-enhanced diffusion can be exploited by creating a velocity field (or by taking advantage of the background velocity field) which transports the concentration field in such a way as to increase the length of large-gradient zones separating regions of different concentration. The concentration field hence becomes very filamented. Qualitatively, a filament is an elongated zone of given concentration value, embedded in a fluid with a very different concentration value. In practice, the longer the filament, or the more filaments there are, the more effective will be the mixing of the flow. Mathematically, such a filamented concentration field is associated with a large value of the gradient-norm. When the typical size of a filament approaches the diffusion length scale, the overall diffusion is enhanced and the sharp gradients are rapidly smoothed, leading to a very mixed concentration field. Therefore, we conclude that with diffusion present in the system, variance appears to be an acceptable mixing measure as it is able to account for mixing due to both diffusion and advection (stirring). Despite this, it is unable to quantify mixing in the limit where $Pe \to +\infty$.

To overcome this issue, Mathew *et al.* (2005) proposed making use of Sobolev norms of negative index, referred to as mix-norms. Such measures have strong links to the ergodic-theory definition of mixing. A common choice is to define the mix-norm as

$$\text{Mix } \theta = \frac{1}{V_\Omega}\left\|\nabla^{-1}\theta\right\|_2^2 = \frac{1}{V_\Omega}\int_\Omega \left(\nabla^{-1}\theta(\boldsymbol{x},t)\right)^2 \, \mathrm{d}\Omega. \tag{2.5}$$

In this expression, the operator $\nabla^{-1}$ can be defined using the inverse of the Laplacian operator $\nabla^{-1} = \nabla\nabla^{-2}$. In practice, to compute the quantity $\nabla^{-1}\theta$, we solve a Poisson equation $\nabla^2\Theta = \theta$, with appropriate boundary conditions (see appendix A), and identify $\nabla^{-1}\theta$ with the gradient $\nabla\Theta$. In the simplest case, when the domain is doubly periodic, the operator $\nabla^{-1}$ can be defined by its action on the Fourier transform $\hat{\theta}$:

$$\nabla^{-1}\hat{\theta} = \frac{\boldsymbol{k}}{k^2}\hat{\theta}, \tag{2.6}$$

where $\boldsymbol{k}$ is the wavenumber with magnitude $k$. Considering the magnitude of this quantity (as in (2.5)), we find $|\nabla^{-1}\hat{\theta}|^2 = k^{-2}|\hat{\theta}|^2$. The pre-factor $k^{-2}$ shows that this operator emphasises the role of large scales (small $k$); in contrast, the small scales (large $k$) will contribute less to the norm. The generalisation of this quantity to non-periodic domains requires the introduction of a boundary condition for the quantity $\Theta$. This condition will naturally arise when expressing the optimisation problem as a variational formulation. We conjecture that the practical usefulness of the mix-norm to quantify mixing extends from periodic to bounded domains.

By downplaying the role of small scales, this norm is able to measure the degree of 'mixing' of a flow even in the absence of diffusion. Indeed, the mix-norm will be large in a completely unmixed situation and will decrease as advection transports the scalar to distribute it equally across the domain, by for example creating filaments of large concentration surrounded by low-concentration areas. In the limit where $Pe \to +\infty$, an optimal 'mixing' flow (i.e. one with minimised mix-norm) creates increasingly smaller scales for the scalar field (Lin *et al.* 2011), reminiscent of fractals.

To summarize, an effective mixing flow corresponds to a velocity field which creates a very filamented concentration field, by increasing gradients of concentration and generating progressively smaller length scales in the passive scalar field. Then, when the concentration gradients are sufficiently large, if $Pe$ is finite, diffusion is greatly enhanced and quickly renders the concentration field uniform. For $Pe \to +\infty$, mixing can never be accomplished through the action of diffusion; no matter how fine the filaments, the concentration field will never be uniform, but will rather tend towards uniformity if the stirring is very effective, such as when stretching-and-folding mechanisms are present. The mix-norm (2.5), by being able to measure how filamented a flow is, can measure the efficacy of stirring, even in the absence of diffusion. In the following section, we will introduce and study a spatially periodic toy model for a mixing flow at different values of the Péclet number $Pe$. One part of the mixing is due to stirring (mixing due to advection only) which creates a filamented flow corresponding to large gradients (or more specifically a long contact surface between areas of very different concentration). To reach a perfectly mixed state eventually, we also need to have some irreversible smoothing process, i.e. molecular diffusion.

### 2.2. *Mixing toy model: the 'chequerboard flow'*

We propose an analytical toy model for the mixing of a passive scalar due to both advection and molecular diffusion. As presented in the previous section, the mixing of

a passive scalar can be accomplished through two independent mechanisms: stirring and diffusion. One can act only on stirring (through the advecting velocity field) since diffusion is entirely governed by the fluid properties and described by the only parameter $Pe$. A velocity field is able to mix a passive scalar effectively if it is able to increase the gradient of concentration in the $L^2$-sense (more precisely, increase the total length of the interface between areas of different concentration), which will allow, in a subsequent stage, the variance of the concentration to decrease quickly when the diffusion length scale is reached by the advection process.

We propose a flow field which monotonically mixes a passive scalar field $\theta$ by the action of a prescribed time-dependent velocity field which continuously creates chequerboard patterns. As time evolves, the size of the chequerboard squares decreases, yielding finer and finer length scales. This flow is considered in a doubly periodic unit square. We consider a finite-Péclet-number version of the toy model flow presented in Lunasin *et al.* (2012). This model will give insight into how different mixing measures behave in the presence of diffusion, which is essential for effective mixing. In fact, in real flows, no matter how weak the molecular diffusion, it cannot be neglected by assuming $\kappa = 0$ because of the singular behaviour discussed in § 2.1. Moreover, despite the absence of dynamics – the prescribed flow is not a solution of the Navier–Stokes equations – we will be able to draw conclusions about the temporal evolution of the norms in a mixing situation.

We define a one-dimensional rectangular diffusive 'packet' of width $w$, centred at $x = c$, as:

$$r_{c,w}(x,t) = \frac{1}{2}\sum_{k=-\infty}^{+\infty}\left[\operatorname{erf}\left(\frac{x+k-c+w/2}{2\sqrt{Pe^{-1}t}}\right) - \operatorname{erf}\left(\frac{x+k-c-w/2}{2\sqrt{Pe^{-1}t}}\right)\right], \tag{2.7}$$

where $Pe$ is the inverse of the non-dimensional diffusivity coefficient. We use the error function because it is an exact solution of the diffusion equation. The summation over $k$ allows us to consider periodic boundary conditions, but can in practice be truncated to a few terms. Due to the integral definition of the error function, we notice the following property for a rectangular diffusive packet:

$$r_{c_3,w_3} = r_{c_1,w_1} + r_{c_2,w_2}, \tag{2.8}$$

with $c_3 = (c_1 + c_2)/2$ and $w_3 = w_1 + w_2$. The equivalent two-dimensional rectangular diffusive packet of width $\boldsymbol{w} = (w_x, w_y)^\top$ and centred at $\boldsymbol{c} = (c_x, c_y)^\top$, can be defined for all $\boldsymbol{x} = (x, y)^\top$ and $t$ as:

$$R_{\boldsymbol{c},\boldsymbol{w}}(\boldsymbol{x},t) = r_{c_x,w_x}(x,t)\, r_{c_y,w_y}(y,t). \tag{2.9}$$

The concentration field $\theta$ is defined as a sum of such rectangular functions:

$$\theta(\boldsymbol{x},t) = \sum_{i=1}^{N(t)} R_{\boldsymbol{c}_i,\boldsymbol{w}_i}(\boldsymbol{x} - \boldsymbol{u}_i(t), t) - \frac{1}{2}, \tag{2.10}$$

where $\boldsymbol{u}_i(t) = (u_i(t),\ v_i(t))^\top$ are the horizontal and vertical velocities of the $i$th diffusive packet, and $N(t)$ their number. We subtract $1/2$ in order to have a zero-mean concentration field. At time $t = 0$, the upper half of the unit square has a concentration equal to $\theta(y > 0.5, 0) = -0.5$ and the bottom half has the opposite concentration value

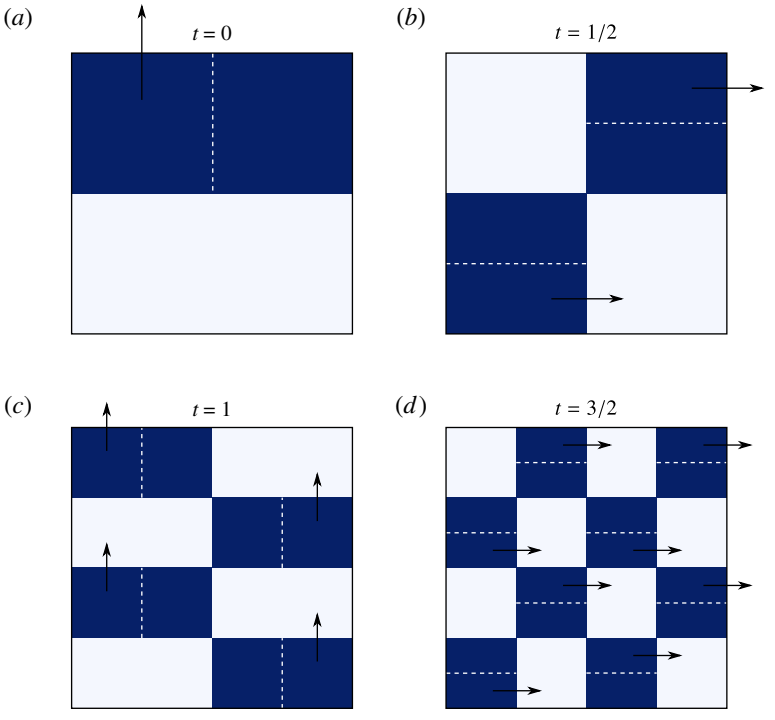


FIGURE 1. (Colour online) Schematic of the chequerboard flow from $t = 0$ to $t = 3/2$. ($a$) Initial concentration pattern ($t = 0$) consisting of two bands of opposite concentration. ($b$) $2 \times 2$ chequerboard pattern after a time $t = 1/2$. ($c$) Concentration field at $t = 1$, composed of four squares which are in the same configuration as the initial condition. ($d$) Chequerboard pattern of size $4 \times 4$ at $t = 3/2$. The evolution between these phases is a continuous process.

$\theta(y < 0.5, 0) = 0.5$, as sketched in figure 1($a$). This field corresponds to the expression

$$\theta_d(\boldsymbol{x}, t) = R(\boldsymbol{x}, t)_{\boldsymbol{c}_0, \boldsymbol{w}_0} - \tfrac{1}{2}, \tag{2.11}$$

for $t = 0$, $\boldsymbol{c}_0 = (L/2, L/4)^\top$ and $\boldsymbol{w}_0 = (L, L/2)^\top$.

This simple model has the ability to create increasingly smaller scales due to the self-similar properties of the chequerboard flow and also to account for diffusion of the scalar field since the functions used are exact solutions of the two-dimensional diffusion equation. Therefore, this toy model, even though no dynamics is included, contains many of the ingredients required to describe the main properties of mixing when only advection and diffusion are at play. At $t = 0$ (this is the first period, so $n = 1$) we split the single rectangle into two squares (by virtue of formula (2.8)). The left square is then advected upward over a distance $l_u(1) = 1/2$ with a velocity $u(1) = 1$, such that at $t = 1/2$, the concentration field has the pattern of a $2 \times 2$ chequerboard, as shown in figure 1($b$). We can then split each of these squares (along the $y$ direction) into $N = 4$ identical rectangles. This manipulation allows us to define a horizontal velocity field (still with a magnitude of $u(1) = 1$) such that the bottom and top rectangles are advected to the right over a distance $l_u(1)$ while the two middle rectangles remain at the same location, leading after time $t = 1$ to the pattern shown in figure 1($c$). We recognise that this field is composed of four squares which are

in the same position as the initial condition at $t = 0$ or in its symmetric counterpart with respect to the horizontal axis. The same process can thus be applied to create an $8 \times 8$ chequerboard at $t = 5/2$. By repeating this process iteratively $n$ times, the flow continuously shapes a concentration field into the form of a chequerboard of size $2^n \times 2^n$ at times $t = (n - 1/2)$. The advection length during the $n$th step is naturally given by $l_u(n) = 1/2^n$. Therefore, the time duration of the $n$th period is $2u/l_u$. In order to keep the time period constant and equal to one, we have to choose the velocity magnitude during the $n$th stage to be $u(n) = 1/2^{n-1}$. We here stress that we depart from the fixed-power (constant velocity $u(n)$) approach adopted in Lunasin *et al.* (2012). The steps corresponding to $n < t < n + 1/2$ can be seen as a folding stage whereas a stretching motion is observed when $n + 1/2 < t < n + 1$. This flow can therefore be considered as a classical example of stretching and folding flows (Ottino 1990).

The only parameter of this flow is the Péclet number $Pe$. We study three different values for this parameter, corresponding to strong diffusion ($Pe = 10^3$), medium diffusion ($Pe = 10^5$) and weak diffusion ($Pe = 10^7$). Finally we discuss the no-diffusion limiting case $Pe \to +\infty$. We present in figure 2 snapshots of the concentration fields at different times for the case $Pe = 10^3$. The velocity field acts as desired, continually creating an increasingly fine chequerboard pattern for times corresponding to half-integer numbers of the fundamental period. In addition, the diffusion acts to smooth the initially discontinuous distribution of concentration. We notice that the variance (that we can infer qualitatively by looking at the intensity of the colours) decreases dramatically around $t = 2$. This can be explained because the length scales created by advection only (the inviscid chequerboard flow) have decreased sufficiently for diffusion to become effective. Indeed, the gradient-norm first increases, after which, when mixing by diffusion becomes relevant, it starts to decrease again (around $t \simeq 1.5$ for $Pe \simeq 10^3$ and $t \simeq 4$ for $Pe = 10^5$). The mixing time can be evaluated by predicting when the flow length scale $l_f(t) = 1/2^{t+1}$ (a continuous extension of the advection length scale $l_u(n)$) and the diffusion length scale $l_d(t) = (Pe^{-1}t)^{1/2}$ are equal. Equating these two length scales gives a mixing time of $t_m = 3.15$ for $Pe = 10^3$, $t_m = 6.01$ for $Pe = 10^5$ and $t_m = 9.04$ for $Pe = 10^7$.

We now perform several mixing simulations with the various norms introduced in § 2.1. In effect, we measure, for each specified case, the variance of the concentration, the norm of its gradient and the mix-norm as functions of time. We normalise each of these quantities with the corresponding quantities, evaluated for the concentration field $\theta_d$ (defined in (2.11)) corresponding to the field obtained when no advection is imposed. Normalising with respect to this pure-diffusion case allows us to compare the evolution of the different norms without focusing on simple diffusion processes but rather stressing the impact of the imposed advection field. All these quantities are plotted against time in figure 3. First, the predicted mixing times $t_m$ correspond remarkably well to the time when the variance curve can no longer be distinguished from the abscissa. This observation strongly supports the validity of the assumption made about the leading mixing mechanism: when the diffusion and advection length scales become of the same order, mixing (due to diffusion) becomes very effective.

Several interesting trends can be deduced from observing these three quantities, all expressing in a different manner the passive scalar distribution. Both the variance and the mix-norm curves are monotonic functions of time while the gradient norm (for all considered cases) is initially larger than for the purely diffusive case ($\|\nabla\theta\|_2 > \|\nabla\theta_d\|_2$). This period lasts longer for smaller diffusion coefficients (larger $Pe$). The increase of the gradient norm reflects the creation of small length scales

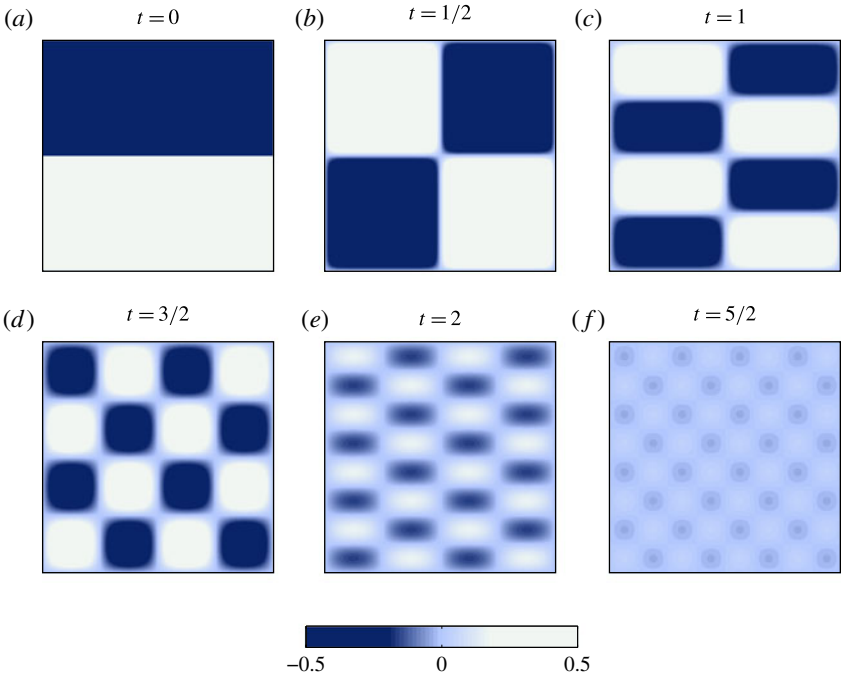


FIGURE 2. (Colour online) Chequerboard flow for $Pe = 10^3$. We plot the concentration field for several times. We notice that for times corresponding to a half-integer numbers of periods, the flow has the shape of a chequerboard, more and more refined as time increases. Around $t = 2$, the variance decreases dramatically due to the fine length scales created by the chequerboard flow, enabling diffusion to act more effectively.

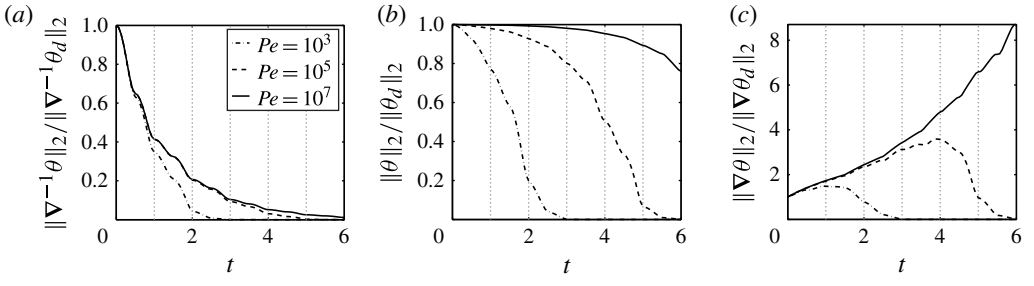


FIGURE 3. Time evolution of various quantities characterising the passive scalar $\theta$ spatial distribution. (*a*) Normalised mix-norm $\|\nabla^{-1}\theta\|_2/\|\nabla^{-1}\theta_d\|_2$. (*b*) Normalised variance $\|\theta\|_2/\|\theta_d\|_2$. (*c*) Normalised gradient norm $\|\nabla\theta\|_2/\|\nabla\theta_d\|_2$. The normalising field $\theta_d$ corresponds to a field obtained with no advection, i.e. by the pure diffusion of the initial profile at $t = 0$, see (2.11). Both variance and mix-norm can be considered as good mixing measures because of their monotonic decay in time.

(thus increasing the interface length between zones of low and high concentration) by the chequerboard flow, ultimately leading to very effective diffusion, which explains why all norms decrease suddenly after the gradient norm has reached its maximum. For the case $Pe = 10^7$, the large-gradient stage lasts longer (as predicted by the large mixing time $t_m = 9.04$); it is therefore not completely shown in our time window

which is limited by computational memory constraints. We nevertheless expect the same behaviour as observed for larger diffusivities, but for longer times. The fact that the evolution of the gradient-norm is not monotonic, even though our toy model has been designed to mimic monotonic mixing in time, thus shows that this norm is not a good measure of mixing. Based on this criterion, both variance and the mix-norm can be considered as good mixing measures.

However, in the no-diffusion case ($Pe \to +\infty$), we would have a constant variance, which eliminates this quantity as an appropriate mixing measure in this limit. Therefore only the mix-norm (amongst these three quantities) should be used as a robust mixing measure. In the case where $Pe$ is finite, even if both the mix-norm and the variance can characterise mixing correctly, the variance seems to be far more sensitive to a change in diffusivity coefficient whereas the mix-norm evolution remains very similar when the Péclet number varies by four orders of magnitude. For very large values of $Pe$, the variance decreases very slowly for short times and consequently its variations are harder to measure. The mix-norm, however, clearly decreases monotonically from $t = 0$, showing as argued by Mathew *et al.* (2005), that the mix-norm can capture 'stirring' processes as precursors to real 'mixing'.

We thus conclude that for cases where diffusion is very small, the mix-norm is a more robust mixing measure than the variance, even though, for finite $Pe$ the variance, due to its monotonic decay, is as plausible a measure as the mix-norm.

## 3. Mixing optimisation in two-dimensional plane Poiseuille flow

### 3.1. *Governing equations and flow configuration*

In order to demonstrate the technique of mixing optimisation, we choose to study a two-dimensional flow test case: the mixing of a passive scalar by an incompressible flow in a two-dimensional channel of dimension $L = 4\pi$, $H = 2$ (non-dimensionalised by the channel half-width $h$). The flow is driven by a constant pressure gradient. The flow variables (velocity and pressure) are governed by the two-dimensional Navier–Stokes equations, characterised by the momentum diffusivity $\nu$ (i.e. the kinematic viscosity), while the scalar field satisfies an advection–diffusion equation, where the transport is imposed by the velocity field and the diffusion is governed by the molecular diffusivity coefficient $\kappa$. By denoting the centreline velocity $u_{max}$, we can derive two non-dimensional numbers: the Reynolds number $Re$ and the Péclet number $Pe$, measuring the relative strength of advection over diffusion (momentum diffusion for the Reynolds number and molecular diffusion for the Péclet number). These numbers can be expressed for this particular flow as

$$Re = \frac{u_{max} h}{\nu}, \quad Pe = \frac{u_{max} h}{\kappa}. \tag{3.1a,b}$$

We consider the simplest case $Re = Pe$, corresponding to a Schmidt number $Sc = \nu/\kappa = 1$.

With this non-dimensionalisation, the solution to the steady Navier–Stokes equations in the channel can be expressed as

$$\overline{\boldsymbol{u}} = (1 - y^2)\, \boldsymbol{e_x}, \tag{3.2}$$

where $y$ is the cross-stream variable of the spatial vector $\boldsymbol{x} = (x, y)^\top$, and $\boldsymbol{e_x}$ denotes the unit vector in the streamwise direction. We consider a finite-size perturbation $q(\boldsymbol{x}, t) = (\boldsymbol{u}(\boldsymbol{x}, t), p(\boldsymbol{x}, t))^\top$. The scalar field is denoted by $\theta(\boldsymbol{x}, t)$ and is advected by both

the background and disturbance velocities, $\overline{\boldsymbol{u}}$ and $\boldsymbol{u}$, respectively. For clarity, we will define the total velocity field as $\boldsymbol{U} = \overline{\boldsymbol{u}} + \boldsymbol{u}$. The equations governing these quantities are the nonlinear disturbance equations given as

$$\partial_t \boldsymbol{u} + \boldsymbol{U} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \overline{\boldsymbol{u}} + \boldsymbol{\nabla} p - Re^{-1} \nabla^2 \boldsymbol{u} = 0, \tag{3.3a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0, \tag{3.3b}$$

$$\partial_t \theta + \boldsymbol{U} \cdot \boldsymbol{\nabla} \theta - Pe^{-1} \nabla^2 \theta = 0. \tag{3.3c}$$

We impose the boundary conditions

$$\boldsymbol{u}(\boldsymbol{x}_\pm, t) = u_\pm \boldsymbol{e_y}, \tag{3.4a}$$

$$\partial_y \theta(\boldsymbol{x}_\pm, t) = 0, \tag{3.4b}$$

$$\partial_y p(\boldsymbol{x}_\pm, t) = 0, \tag{3.4c}$$

with $\boldsymbol{x}_\pm$ representing the coordinates of the walls ($\boldsymbol{x}_\pm = (x, \pm 1)$) and $\boldsymbol{e_y}$ denoting the cross-stream unit vector. We keep homogeneous Dirichlet boundary conditions for the cross-stream velocity to avoid injecting or pumping fluid into the channel (which would imply that the scalar field average is not conserved). Periodic boundary conditions are enforced in the $x$-direction. This choice enables the study of mixing for relatively long times without the requirement of a very large computational domain. However, this type of boundary condition might introduce some artificial mixing due to the overlap of the perturbation with itself after a certain time. We limit the unwanted consequence of the periodic boundary condition by choosing a sufficiently long channel length of $L = 4\pi$.

Naturally, this system of equations requires a set of initial conditions. These can be expressed as

$$\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{u}_0(\boldsymbol{x}), \tag{3.5a}$$

$$\theta(\boldsymbol{x}, 0) = \theta_0(\boldsymbol{x}). \tag{3.5b}$$

The initial condition $\boldsymbol{u}_0(\boldsymbol{x})$ as well as the wall boundary condition $u_\pm$ will be considered as control variables, and will thus be used in our optimisation scheme to achieve optimal mixing. The initial concentration $\theta_0(\boldsymbol{x})$ will be fixed for all simulations performed in this paper: we start from a configuration of concentration which corresponds to a non-mixed state consisting of two stripes of opposite concentration with a common interface in the middle of the domain, at $y = 0$. The interface between these two layers is considered to be a diffusive interface which, in the absence of advection, can be described by the scalar field $\theta_0(\boldsymbol{x})$:

$$\theta_0(\boldsymbol{x}) = \frac{1}{2} \mathrm{erf}\left(\frac{y}{l_0}\right). \tag{3.6}$$

As a discontinuity might be expected to lead to numerical instabilities, we choose $l_0 = 5\Delta y$ (with $\Delta y$ as the grid spacing in the $y$ direction) to obtain a smooth profile on our computational grid. The initial concentration field then consists of two horizontal stripes of opposite concentration, of magnitude $\pm 0.5$. The interface between these two stripes is centred at $y = 0$.

### 3.2. *Variational formulation of optimal mixing*

We have defined in (2.1) and (2.5) two measures (variance and mix-norm) which are able to determine the level of mixing in a given flow, based on the spatial distribution

of a passive scalar $\theta$. We have explored in § 2.2 how these measures behave in the case of a mixing flow, in the presence of molecular diffusion for the passive scalar. Here, we will study a more realistic flow: pressure-driven two-dimensional plane Poiseuille flow. Contrary to previous mixing studies, scalar diffusion is embedded in the model. Also, instead of a prescribed velocity field, we will consider a dynamical flow, governed by the incompressible Navier–Stokes equations. Above all, we would like to investigate the sensitivity of optimal control solutions (initial conditions and boundary conditions) with respect to the chosen norm. In previous studies, such as in Aamo & Krstić (2003), mixing was thought to be strongly enhanced as the energy of the flow was maximised. To verify whether this assumption holds, we want to compare final-time mixing optimisation against time-integrated energy optimisation. More specifically, we choose a time-averaged energy optimisation since mixing – a continuous process – takes place continuously over the time interval $[0, T]$. Moreover, the stirring part of the mixing is entirely governed by the nonlinear terms in the Navier–Stokes equations, whose strength at time $t$ can be measured by evaluating the disturbance energy

$$E(t) = \tfrac{1}{2}\|\boldsymbol{u}(\boldsymbol{x}, t)\|_2{}^2. \tag{3.7}$$

Therefore, it seems sensible to consider time-averaged rather than finite-time energy optimisation. In fact, we study time-averaged energy, final-time mix-norm and variance optimisation. We investigate whether the most energetic perturbation is the one leading to maximum mixing and assess the appropriateness of the variance and the mix-norm for quantifying mixing processes.

We pose a general cost functional $\mathscr{J}$ which allows for several measures based on either the kinetic energy or directly on the passive scalar $\theta$. The cost functional can be written as

$$\mathscr{J}\ (\boldsymbol{u}, \theta) = \frac{1-\alpha}{2} \int_0^T \|\boldsymbol{u}(\boldsymbol{x}, t)\|_2{}^2\ \mathrm{d}t + \frac{\alpha}{2}\left\|\boldsymbol{\nabla}^{-\beta}\theta(\boldsymbol{x}, T)\right\|_2{}^2, \tag{3.8}$$

with the $L^2$-norm $\|\cdot\|_2$ classically defined as in (2.1). The parameter $\alpha$ (we only consider the values $\alpha = 0$ and $\alpha = 1$) determines to what degree energy or mixing is optimised. When $\alpha = 0$, the time-integrated energy is maximised. As already noted, mixing is a continuous process which might be enhanced by nonlinear advection (or stirring). Therefore if one wants to reach a finite-time optimal mixing, it is sensible to look for the perturbation which maximises the total amount of energy within this time-interval. This justifies the choice of the time-averaged energy rather than its finite-time evaluation. For a discussion on the relation between the optimisation of the finite-time and the time-averaged energy, see Duguet *et al.* (2013). For $\alpha = 1$, mixing is the optimised quantity. In this case, the norm is entirely determined by $\beta$. As presented in § 2.1, $\beta = 1$ corresponds to a Sobolev norm of negative index ($L^2$-norm of $\boldsymbol{\nabla}^{-1}\theta$, also referred to as the mix-norm), $\beta = 0$ corresponds to the variance ($L^2$-norm of $\theta$). The case $\beta = -1$ corresponding to the gradient $L^2$-norm will not be considered for optimisation, as it was found to yield an inappropriate measure of mixing owing to its non-monotonic temporal evolution.

After defining the cost functional $\mathscr{J}$ we need to decide what kind of 'control' we will apply to the flow in order to reach an extremal value of $\mathscr{J}$. In this paper, we will consider two different types of optimal control: optimal initial conditions (3.5) and optimal boundary conditions (3.4). The term 'optimal control' may be misleading, in as far as we will only optimally identify the initial (or wall boundary) condition, and then let the flow evolve freely, solely governed by the Navier–Stokes equations.

The cost functional $\mathscr{J}(\boldsymbol{u}, \theta)$ is a function of the velocity and concentration fields, which are themselves implicit functions of the initial condition $\boldsymbol{u}_0$ and boundary condition $u_\pm$. Since this dependence cannot be expressed explicitly, we have to embed the dependence of the variables $\boldsymbol{u}$ and $\theta$ on the control variables in an augmented cost functional where the evolution constraints equation (3.3), and both initial (3.5) and boundary conditions (3.4) appear explicitly. This augmented functional $\mathscr{L}$ is given by

$$\begin{aligned}\mathscr{L}(\boldsymbol{u}, p, \theta, \boldsymbol{u}_0, u_\pm, \boldsymbol{u}^\dagger, p^\dagger, \theta^\dagger, \boldsymbol{u}_0^\dagger, u_\pm^\dagger)\\ = \mathscr{J}(\boldsymbol{u}, \theta) - \left\langle \boldsymbol{u}^\dagger, \partial_t \boldsymbol{u} + \boldsymbol{U}\cdot\nabla\boldsymbol{u} + \boldsymbol{u}\cdot\nabla\overline{\boldsymbol{u}} + \nabla p - Re^{-1}\nabla^2\boldsymbol{u}\right\rangle\\ - \left\langle p^\dagger, \nabla\cdot\boldsymbol{u}\right\rangle - \left\langle \theta^\dagger, \partial_t\theta + \boldsymbol{U}\cdot\nabla\theta - Pe^{-1}\nabla^2\theta\right\rangle\\ - \left(\boldsymbol{u}_0^\dagger, \boldsymbol{u}(\boldsymbol{x}, 0) - \boldsymbol{u}_0\right)\\ - \left[u_\pm^\dagger, u(\boldsymbol{x}_\pm, t) - u_\pm\right],\end{aligned} \tag{3.9}$$

where we have used the scalar products $\langle\cdot,\cdot\rangle$, $(\cdot,\cdot)$ and $[\cdot,\cdot]$ defined as

$$\left.\begin{aligned}\left\langle a(\boldsymbol{x}, t), b(\boldsymbol{x}, t)\right\rangle &= \int_0^T \int_\Omega a(\boldsymbol{x}, t)^\top b(\boldsymbol{x}, t)\ \mathrm{d}\Omega\ \mathrm{d}t,\\ \left(c(\boldsymbol{x}), d(\boldsymbol{x})\right) &= \int_\Omega c(\boldsymbol{x})^\top d(\boldsymbol{x})\ \mathrm{d}\Omega,\\ \left[e(x, t), f(x, t)\right] &= \int_0^T \int_0^L e(x, t)^\top f(x, t)\ \mathrm{d}x\ \mathrm{d}t,\end{aligned}\right\} \tag{3.10}$$

for arbitrary (possibly vectorial) functions $a$, $b$, $c$, $d$, $e$ and $f$ of the indicated variables.Moreover, we recall that $\boldsymbol{U} = \overline{\boldsymbol{u}} + \boldsymbol{u}$. We here stress that the nonlinear dynamics of the Navier–Stokes operator has been included in the variational formulation.

The adjoint variables (indicated with a superscript †) are Lagrange multipliers enforcing all constraints of the system: $\boldsymbol{u}^\dagger$ enforces the Navier–Stokes equations (3.3*a*), $p^\dagger$ is associated with the incompressibility condition (3.3*b*), and $\theta^\dagger$ is the adjoint concentration field imposing the passive scalar advection–diffusion equation (3.3*c*). Finally, $\boldsymbol{u}_0^\dagger$ and $u_\pm^\dagger$ impose the initial and boundary condition constraints (3.5) and (3.4), respectively.

To ensure the optimality of the solution, we need to enforce $\delta\mathscr{L} = 0$. As detailed in appendix A, $\delta\mathscr{L}$ depends on several variables, all of which are independent of each other. We thus require that all partial derivatives of the functional $\mathscr{L}$ vanish. It is easily verified that the variations with respect to the adjoint variables result in the enforcement of the various dynamical constraints.

First variations with respect to the 'direct' variables $\boldsymbol{u}$, $p$ and $\theta$ lead to the definition of a set of partial differential equations describing the evolution of the adjoint variables $\boldsymbol{u}^\dagger$, $p^\dagger$ and $\theta^\dagger$:

$$\partial_t\boldsymbol{u}^\dagger + (\boldsymbol{U}\cdot\nabla\boldsymbol{u}^\dagger - \boldsymbol{u}^\dagger\cdot\nabla\boldsymbol{U})^\top + \nabla p^\dagger + Re^{-1}\nabla^2\boldsymbol{u}^\dagger = \theta^\dagger\nabla\theta - (1-\alpha)\,\boldsymbol{u}, \tag{3.11a}$$

$$\nabla\cdot\boldsymbol{u}^\dagger = 0, \tag{3.11b}$$

$$\partial_t\theta^\dagger + \boldsymbol{U}\cdot\nabla\theta^\dagger + Pe^{-1}\nabla^2\theta^\dagger = 0, \tag{3.11c}$$

where $[\boldsymbol{u}^\dagger\cdot\nabla\boldsymbol{U}^\top]_i = u_j^\dagger\partial_i U_j$. As is conventional, these equations are anti-diffusion equations and therefore have to be solved backwards in time, from $t = T$ to

$t = 0$. Moreover, because the direct equations (3.3) are nonlinear and because the time-averaged energy is considered in the objective functional, the direct disturbance velocity field $\boldsymbol{u}$ explicitly appears in the adjoint equations. This requires saving the velocity field during the direct time integration. Due to memory constraints, we adopt a classical check-pointing approach in order to avoid saving the direct field at every iteration. Some boundary integrals (temporal and spatial) arise when calculating the partial derivatives of $\mathscr{L}$. By setting each of the independent expressions to zero, we find the so-called natural adjoint boundary conditions:

$$\boldsymbol{u}^\dagger(\boldsymbol{x}_\pm, t) = 0, \quad \partial_y \theta^\dagger(\boldsymbol{x}_\pm, t) = 0. \tag{3.12a,b}$$

No natural boundary condition for the pressure is found, and so we adopt, as in the direct case,

$$\partial_y p^\dagger(\boldsymbol{x}_\pm, t) = 0. \tag{3.13}$$

The adjoint terminal conditions are found to be

$$\boldsymbol{u}^\dagger(\boldsymbol{x}, T) = 0, \quad \theta^\dagger(\boldsymbol{x}, T) = (-1)^\beta \alpha \nabla^{-2\beta} \theta(T). \tag{3.14a,b}$$

The fact that the adjoint velocity field at the final time vanishes may seem unusual, but it stems from the fact that we do not evaluate the disturbance velocity magnitude (energy) at the final time $t = T$. Instead, we consider a time-averaged measure which leads to a forcing term in the adjoint Navier–Stokes equation (3.11*a*). For $t < T$ and $\alpha = 0$ (i.e. when energy is optimised), the adjoint velocity field is no longer zero because it is forced by the production term $-(1-\alpha)\boldsymbol{u}$. In this case, the adjoint scalar field $\theta^\dagger$ is identically zero, because its terminal condition vanishes and there is no external forcing in (3.11*c*), while the extra forcing term $\theta^\dagger \boldsymbol{\nabla}\theta$ in (3.11*a*) vanishes as well. However, when $\alpha = 1$, the adjoint scalar field $\theta^\dagger$ is no longer zero, and the momentum equation (3.11*a*) is forced by the non-zero term $\theta^\dagger \boldsymbol{\nabla}\theta$. For all choices of $\alpha$, with the above set of boundary and terminal conditions, we also find the following equations:

$$\boldsymbol{u}_0^\dagger = \boldsymbol{u}(\boldsymbol{x}, 0), \tag{3.15a}$$

$$u_\pm^\dagger = \mp Re^{-1} \partial_y \boldsymbol{u}^\dagger(\boldsymbol{x}_\pm, t). \tag{3.15b}$$

These expressions are directly related to the gradient of the cost functional with respect to the control variables:

$$\boldsymbol{\nabla}_{\boldsymbol{u}_0} \mathscr{J} = \boldsymbol{u}_0^\dagger, \tag{3.16a}$$

$$\boldsymbol{\nabla}_{u_\pm} \mathscr{J} = u_\pm^\dagger. \tag{3.16b}$$

A more detailed derivation of these equations is given in appendix A. By employing a gradient descent algorithm, we are able to reach the optimal control strategies $\boldsymbol{u}_0$ or $u_\pm$ such that $\delta\mathscr{L}$ vanishes.

Furthermore, we need to constrain the size of the control variables $\boldsymbol{u}_0$ and $u_\pm$. This is accomplished by normalising the general $L^p$-norms (defined on their respective spaces):

$$\|\boldsymbol{u}_0\|_p^p = \int_\Omega |\boldsymbol{u}_0|^p \, \mathrm{d}\Omega = 2E_0, \tag{3.17a}$$

$$\|u_\pm\|_p^p = \int_0^T \int_0^L |u_\pm|^p \, \mathrm{d}x \, \mathrm{d}t = 2E_w, \tag{3.17b}$$

where $E_0$ and $E_w$, respectively, represent their initial $L^p$-norms. It is important to appreciate that only for $p = 2$ do these two quantities represent actual energies. In what follows we will consider the optimisation of the initial condition or the wall

boundary condition separately, which requires the normalisation of only one of these variables. A convenient way to enforce the normalisation of such a quantity in the optimisation process is detailed in the appendix of Foures, Caulfield & Schmid (2013) for the case $p=2$, and is here generalised for any $p$. Therefore, our variational formulation can be applied to time-averaged energy or mixing optimisation, either by prescribing an initial perturbation or a wall disturbance.

### 3.3. *Numerical implementation*

The numerical code used is a modified version of the Navier–Stokes solver presented in Foures *et al.* (2013). Spatial discretisation uses Fourier modes in $x$ and second-order finite differences in $y$. Advection and diffusion are treated using a mixed fourth-order Runge–Kutta Crank–Nicolson method. The incompressibility condition is enforced using a standard incremental pressure-correction scheme (Guermond, Minev & Shen 2006). We use 200 grid points in the $y$ direction and 400 Fourier modes in the $x$ direction. The time step for the time integration scheme is chosen according to a classical CFL condition. In practice, we have $\mathrm{d}t \approx 10^{-3}$. The results were not sensitive to an increase of either spatial or temporal resolution, therefore establishing the convergence of the numerical scheme.

## 4. Results

### 4.1. *Optimal initial perturbation*

In this section, we study optimal perturbations for various measures in the Poiseuille flow defined in § 3.1 at $Re = 500$. It has been conjectured that mixing is closely related to nonlinear energy amplification (Aamo & Krstić 2003). Thus, in order to investigate the mixing properties of an energetically optimal flow, we will first focus on the optimal initial perturbations yielding an extremal value of the time-averaged energy. This problem is rather classical (see, for example, the linear study by Guégan, Schmid & Huerre 2006). By choosing $\alpha = 0$ in the cost functional (3.8) and optimising with respect to $\boldsymbol{u}_0$ only (homogeneous Dirichlet boundary conditions are prescribed at the wall), we perform a series of nonlinear optimisations for several time horizons $T$. The initial energy amplitude is constrained using the $L^2$-norm ($p = 2$ in (3.17*b*)) and such that $E_0 = 10^{-2}$, which corresponds to an initial velocity value one order of magnitude lower than the base-flow velocity amplitude. This ensures that the nonlinear terms in (3.3*a*) and, more importantly, (3.3*c*) are not negligible, and therefore stirring by advection is possible. For small values of $E_0$, no mixing is possible by advection because the nonlinear terms can be neglected. However, for increasing $E_0$ values, we expect the mixing due to stirring to be more and more effective because a larger amount of energy is provided to the system. We restrict ourselves to the intermediate case $E_0 = 10^{-2}$, as this choice is associated with the nonlinear advection having the same order of magnitude as the linear base-flow advection.

We present in figure 4 the initial vorticity field $\boldsymbol{\nabla} \times \boldsymbol{u}_0$ and terminal concentration field $\theta(T)$ corresponding to the time-averaged energy optimisation ($\alpha = 0$) over the time interval $[0, T]$ for several optimisation times $T$. The optimised initial perturbation velocity field consists of arrays of elongated vortices tilted against the base-flow shear which allows for energy transfer from the base flow to the perturbation via the Orr mechanism (see Orr 1907; Lindzen 1988). The final concentration field consists of large-scale sinuous structures which are characteristic of the optimal advection field. We notice that the scalar field is not well-mixed, even after a rather long integration time. Indeed, the two initial layers of opposite concentration are still clearly defined, showing that the large vortices generated during the time evolution of

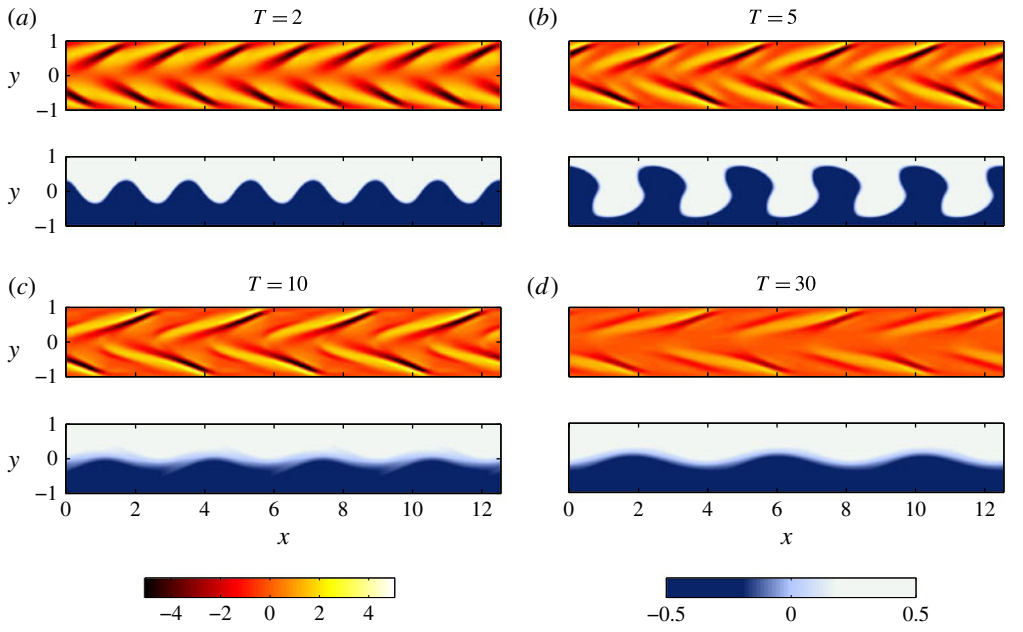


FIGURE 4. (Colour online) Initial vorticity $\nabla \times \boldsymbol{u}_0$ (top plot of each pair) and terminal concentration $\theta(T)$ (bottom plot of each pair) distributions corresponding to the time-averaged energy optimisation over the time interval $[0, T]$ for four optimisation times $T \in [2; 5; 10; 30]$. Colourbars for the vorticity (left) and the concentration field (right) are displayed at the bottom of the figure.

the time-averaged energy-optimal perturbations are unable to mix effectively the two regions (top and bottom) of the flow. Based on this observation, mixing enhancement via energy optimisation seems to be a highly suboptimal method.

We define the time-averaged gain as

$$G(t) = \frac{\dfrac{1}{t}\displaystyle\int_0^t \|\boldsymbol{u}\|_2{}^2(\tau)\ \mathrm{d}\tau}{\|\boldsymbol{u}_0\|_2{}^2}. \tag{4.1}$$

This gain is simply a rescaled time-integrated energy which is the objective functional for $\alpha = 0$. We plot in figure 7(*a*) the time-averaged gain as a function of time for the energy-optimal solutions. First, we note that the time evolution of the optimal solutions defines an optimal time-averaged gain envelope $G_{opt}(T)$ (plotted with a dashed line), establishing an upper limit of energy amplification that a perturbation (satisfying $E_0 = 10^{-2}$) can experience. Any curve $G(t)$ has to lie below this optimal envelope and be tangent to it at precisely the relevant target time $t = T$.

We also define normalised quantities in order to quantify mixing. Let us define $M(t)$ and $V(t)$, the normalised mix-norm and variance, as

$$M(t) = \frac{\left\|\nabla^{-1}\theta\right\|_2{}^2}{\left\|\nabla^{-1}\theta_d\right\|_2{}^2}, \tag{4.2a}$$

$$V(t) = \frac{\|\theta\|_2{}^2}{\|\theta_d\|_2{}^2}, \tag{4.2b}$$

where the scalar field $\theta_d$ is the solution of the pure diffusion equation (3.3$c$) with $\boldsymbol{U} = \boldsymbol{0}$ and initial condition $\theta(\boldsymbol{x}, 0) = \theta_0(\boldsymbol{x})$. With such a time-dependent normalisation, we only measure the additional mixing that can be attributed to the advection field (as in § 2.2). Under the scaling (4.2), a purely diffusive process yields constant values for both $M(t)$ and $V(t)$.

The curves of both $M(t)$ and $V(t)$ corresponding to the time-averaged energy optimisations are plotted in figures 7($b$) and 7($c$). At this stage, we note that short-time energy-optimal solutions are more likely to lead eventually to low values of either the mix-norm or the variance. Indeed, smaller advection length scales observed for small values of $T$ mix the scalar field more effectively compared to the very large length scales which arise when the time integration interval is longer. As a general remark, we can already conclude that energy-optimal perturbations fail to induce an effective finite-time mixing mechanism, especially for long time horizons.

To confirm, both qualitatively and quantitatively, that mixing optimisation cannot be achieved by maximising energy growth, we perform a series of mixing optimisations, for the same time intervals as used for the energy optimisation case. We thus choose $\alpha = 1$ to define the cost functional and investigate the cases $\beta = 1$ and $\beta = 0$, respectively corresponding to mix-norm and variance optimisation. Contrary to the time-averaged energy maximisation, now the quantity $\mathscr{J}$ has to be minimised.

Let us first focus on the most sophisticated mixing definition: the mix-norm as defined in (2.5) (with $\alpha = 1$, $\beta = 1$). As before, we present in figure 5 several optimal initial perturbations and final scalar fields for typical horizon times $T = 2$, 5, 10 and 30. Even if their general structures and typical length scales are different from the optimal energy case, the optimal mix-norm perturbations still consist of elongated vortices tilted against the mean shear for all considered horizon times $T$. This means that the strategy for optimal mixing (in the mix-norm sense) relies to some extent on energy harvesting from the base flow. For short times, we notice that large blobs of the passive scalar are periodically protruding into the region of opposite concentration value, quite similarly to the short-time energy-optimal solutions. Since the mix-norm downplays the role of small length scales, large-scale structures promoting the exchange of the passive scalar between the top and bottom part of the channel are preferred. For later times, the optimal initial perturbation does not vary significantly; nonetheless, the scalar field becomes progressively well mixed. It seems to be clear that final-time mix-norm optimisation leads to a better mixed scalar field at the end of the optimisation interval, compared to the case of energy optimisation. Indeed, the scalar fields at time $t = T$ presented in figure 5 are more homogeneous than their energy-optimal counterparts. We notice, however, that the global stratification from negative to positive concentration has been conserved, due to the no-slip condition at the wall.

We plot in figure 7($d$) the energy evolution $G(t)$ corresponding to the mix-norm optimisation initial perturbation. This plot, displaying a transient growth of energy for the mix-norm-optimal perturbation, confirms that the mix-norm-optimal mixing strategy does depend on energy amplification, but does not maximise it. Similarly to the energy optimisation case, we plot in figure 7($e$) the optimal curves $M(t)$ for the optimal mix-norm initial perturbations, while in figure 7($f$) we plot the normalised variance time evolution $V(t)$ corresponding to the mix-norm optimal perturbation. The curves in figure 7($e$) define the optimal variance envelope $M_{opt}(T)$. We note that the individual normalised mix-norm time evolution $M(t)$ of each perturbation is not monotonic. Instead, there is a generic behaviour for the curves $M(t)$: after an initial strong decrease, the normalised mix-norm value displays an oscillatory behaviour.

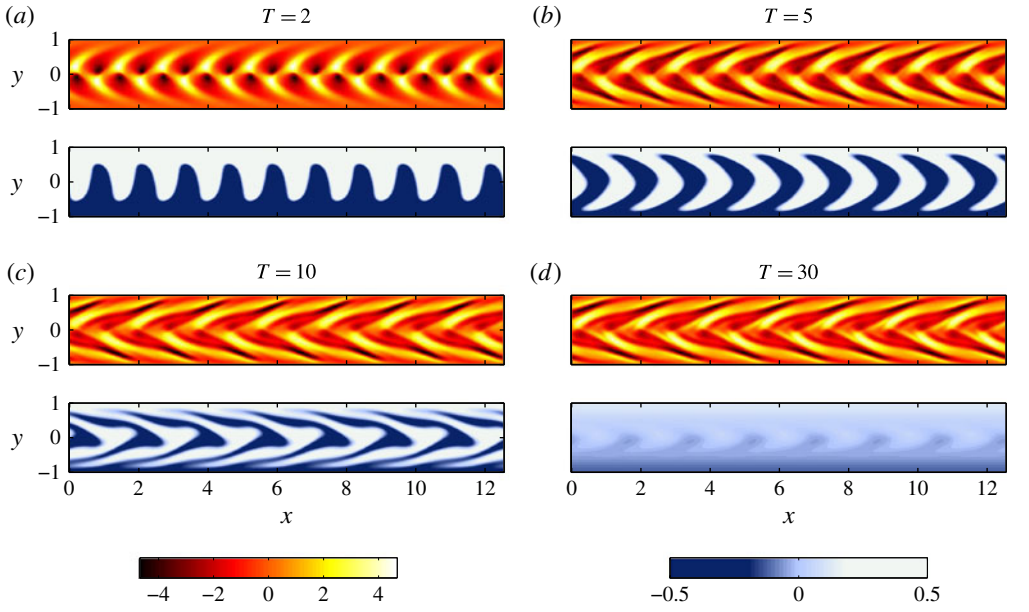


FIGURE 5. (Colour online) Initial vorticity $\boldsymbol{\nabla} \times \boldsymbol{u}_0$ (top) and terminal concentration $\theta(T)$ (bottom) profiles corresponding to the mix-norm optimisation over the time interval $[0, T]$ for four optimisation times $T \in [2; 5; 10; 30]$. Colourbars for the vorticity (left) and the concentration field (right) are displayed at the bottom of the figure.

We note, on each of these curves, a temporal stage during which the normalised mix-norm starts to increase anew. In terms of mixing, this would correspond to a 'demixing' or ordering phenomenon. This effect is however very small compared to the initial decrease of the mix-norm for earlier times, demonstrating that strong mixing is taking place in the flow.

Finally, in figure 6 we display the optimal initial vorticity and the final passive scalar fields corresponding to variance minimisation (with $\alpha = 1$, $\beta = 0$). For short times, the optimal solutions (initial perturbations) consist of small-scale vortices localised in the middle of the domain which are able to mix effectively the two layers of the passive scalar in the close vicinity of their common interface. For such time horizons, variance optimisation leads to optimal initial perturbations with a far smaller typical wavelength than both the energy- and mix-norm-optimal solutions. For longer times, the typical length scale of the perturbation increases (but still remains smaller than for the energy optimisation), and the localisation of the initial disturbance at the concentration interface persists even though the vorticity field spreads in the cross-stream direction as $T$ increases. Besides, it is noteworthy that the $T = 30$ variance optimal initial perturbation is almost identical to the mix-norm-optimal initial perturbation for the same optimisation time. This means that, in contrast to short-time strategies, long-time mixing strategies are very similar in both norms.

It is clear from the analysis of both the mix-norm- and variance-optimal solutions that for short times, mix-norm optimisation leads to a more globally mixed flow, without taking advantage of molecular diffusion at small scales, while the variance optimisation yields an optimal perturbation focusing on the interface and entirely relying on the effects of small-scale diffusion, rather than on energy amplification. This is confirmed in figure 7(*g*) where we observe thatfor short times,

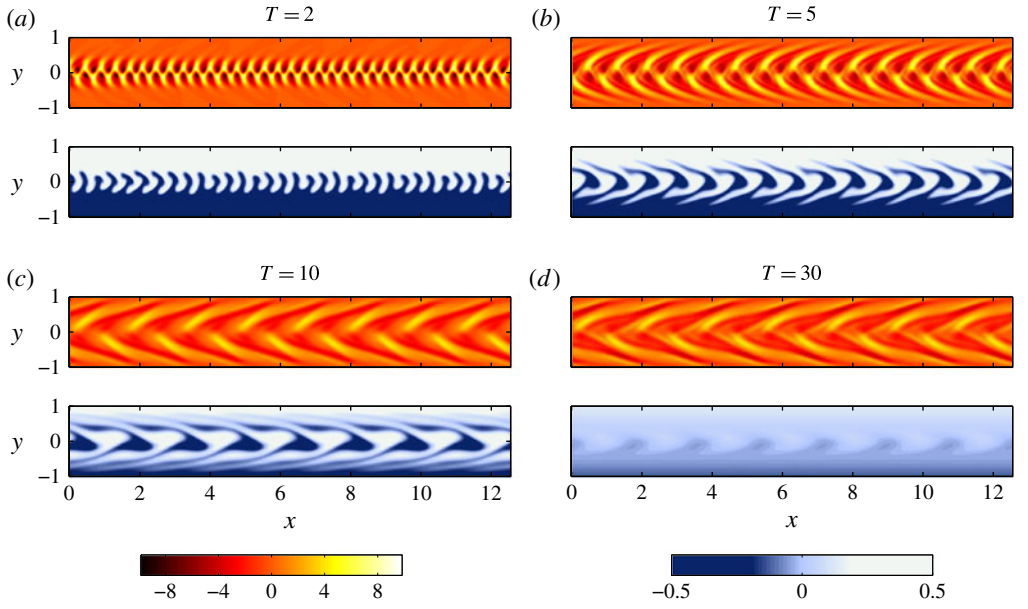


FIGURE 6. (Colour online) Initial vorticity $\boldsymbol{\nabla} \times \boldsymbol{u}_0$ (top) and terminal concentration $\theta(T)$ (bottom) profiles corresponding to the variance optimisation over the time interval $[0, T]$ for four optimisation times $T \in [2; 5; 10; 30]$. Colourbars for the vorticity (left) and the concentration field (right) are displayed at the bottom of the figure.

the variance-optimal solutions do not amplify energy at all, instead preferring a different strategy based on localising the perturbation at the interface in order to increase its length and thus to maximise diffusion. For longer time though, both types of optimisation (mix-norm and variance) seem to lead to very similar results, as is apparent through comparison of figures 7(*e*) and 7(*h*), where we plot the normalised mix-norm time evolution $M(t)$ for the mix-norm optimal perturbation and the variance optimal perturbation respectively.

In figure 7(*i*), we plot the optimal curves $V(t)$ defining the optimal envelope $V_{opt}(T)$. We notice, as a general trend, that, after the optimal time, the time evolution of the normalised variance $V(t)$ reaches a plateau, implying that at this stage the mixing is almost entirely due to a pure diffusion process. The solution for $T = 2$ even displays a positive slope for $V(t)$ after $t \simeq 6$. After this time, the flow is diffusing the scalar field less effectively than the pure diffusion equation, suggesting that there is a late drawback for optimising variance for very early times. This is the price to pay when the mixing strategy is not based on energy amplification but rather on diffusion optimisation close to the interface. The resulting mechanism is effective for short times, but the perturbation amplitude subsequently decays too quickly to have any further effect on the scalar field.

Overall, even if the energy-optimal perturbation is a suboptimal candidate for enhancing mixing in a flow, we note that nearly all optimal mixing (variance or mix-norm) initial perturbations are composed of several vortices tilted against the mean shear, taking advantage of the two-dimensional Orr mechanism in order to extract energy from the base flow. However, the optimal wavelengths of these structures depend on the type of optimisation chosen. We present in table 1 the dominant observed wavelength $m$ as a function of the time integration length $T$ and

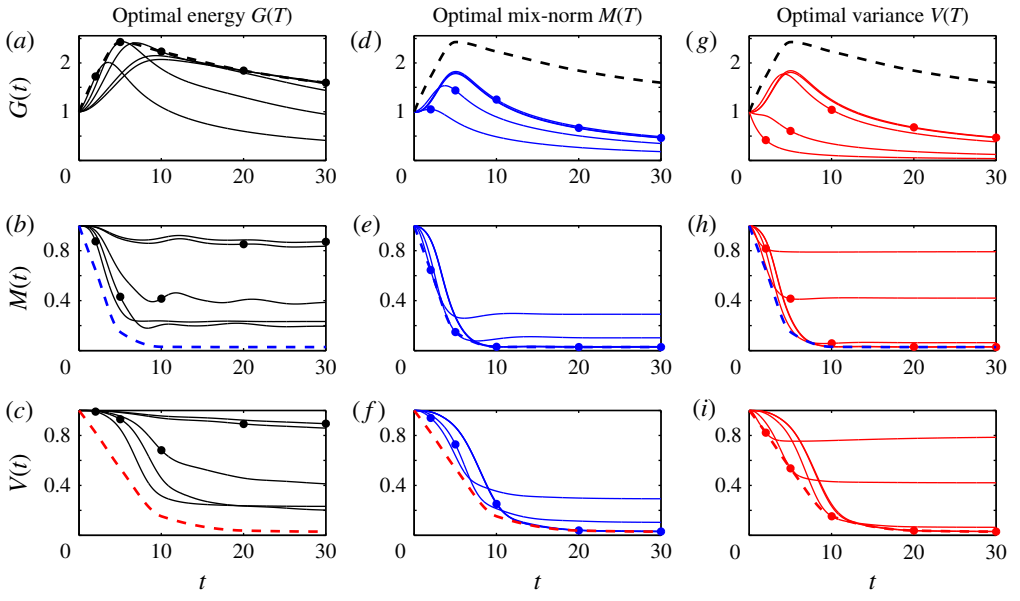


FIGURE 7. (Colour online) Time evolution of the considered measures $G(t)$ (top row), $M(t)$ (middle row) and $V(t)$ (bottom row) corresponding to the various identified optima (as defined in (4.1), (4.2*a*) and (4.2*b*)). The first, second and third columns correspond respectively to the time evolution of the optimal initial perturbations of the time-averaged energy $G(T)$, the mix-norm $M(T)$ and the variance $V(T)$. For each series of optimisations, five different horizon times were considered: $T \in [2; 5; 10; 20; 30]$. The dots on each curve indicate the optimisation times. The figures located on the diagonal correspond to the time evolution of the actual optimised quantities, from which we can define the optimal envelopes $G_{opt}(T)$, $M_{opt}(T)$ and $V_{opt}(T)$ (obtained by cubic interpolation of the five computed optima). These optimal envelopes are plotted with dashed lines on the panels of each row.

| | | Horizon time $T$ | | | | |
|---|---|---|---|---|---|---|
| | | 2 | 5 | 10 | 20 | 30 |
| Optimised measure | $G(T)$ | 7 | 5 | 4 | 3 | 3 |
| | $M(T)$ | 10 | 9 | 7 | 7 | 7 |
| | $V(T)$ | 33 | 15 | 8 | 7 | 7 |

TABLE 1. Dominant wavenumber $m$ for each of the performed optimisations.

the type of optimisation performed. Since these perturbations are not monochromatic, $m$ is defined as the most energetic wavenumber contained in $\boldsymbol{u}_0$ (for example, $m = 0$ corresponds to a constant and $m = 1$ to a sinusoidal signal of period $L = 4\pi$). As a general observation, no matter what the nature of the objective functional is, short-time optimisation leads to high wavenumbers (small-scale structures) while long-time optimisation is associated with low wavenumbers (large-scale structures). Also, for a fixed optimisation time, we observe a hierarchy from low to high wavenumbers when passing from energy to mix-norm and, finally, to variance optimisation.

Finally, because of the nonlinear nature of the optimisation problem and the complex objective functional, optimality is unsurprisingly difficult to reach. To characterise the convergence properties of our solutions, we define the gradient residual $r$ as

$$r = \frac{\left\| \nabla_{\boldsymbol{u}_0} \mathscr{J}^{\perp} \right\|_2^2}{\left\| \nabla_{\boldsymbol{u}_0} \mathscr{J} \right\|_2^2}. \tag{4.3}$$

This residual measures the component of the gradient (3.16) along the normalisation constraint (3.17), and is further discussed in the appendix of Foures *et al.* (2013). We observed for all performed optimisations a decrease of the residual down to $r \sim O(10^{-4})$. We believe that the size of this residual shows that each of the identified solutions lies in the close vicinity of an at least locally optimal solution.

### 4.2. *Time evolution analysis and physical interpretation*

Our conclusion from the previous section is that the energy-optimal initial perturbation is less effective at mixing a flow than the optimal solution based on an appropriate mixing measure (variance or mix-norm). In order to understand the principal difference between the energy-optimal perturbations and mixing-optimal perturbations, we will analyse in more detail the time evolution of two typical optimal solutions. We plot in figure 8 the instantaneous vorticity and scalar fields corresponding to the variance-optimal initial disturbance ($\alpha = 1$ and $\beta = 0$ and time horizon $T = 30$) for several time instants $t \in [0; 30]$. The vorticity colourmap is deliberately saturated at 50 % of the maximum and minimum to visualise the evolution of the perturbation amplitude. Even though energy cannot directly be deduced from these plots, the more vivid the colours, the larger the size of the disturbance. Inspection of both the vorticity and scalar field snapshots shows that the initial vortices are responsible for a periodic vertical transport of scalar from the bottom to the top, and vice versa, creating at $t \simeq 5$ a pattern consisting of a succession of vertical stripes of minimum and maximum value.Then, for $t > 5$, the perturbation loses a significant amount of energy, indicating that the transport is then almost completely due to the parabolic background flow. By advecting the scalar field as close to the walls as possible, the optimal perturbation exploits Taylor dispersion due to the background flow. Indeed, close to the walls, where shear is important, diffusion can be enhanced via the production term in (2.4). It is thus very beneficial to advect the scalar field towards the walls.

For comparison, in figure 9 we display the time evolution of the vorticity and scalar fields for the energy-optimal perturbation for a horizon time of $T = 5$. We choose this particular solution, since amongst the energy optima, it is the solution which is most able to mix the scalar field (see figures 7*b* and 7*c*). Initially, the flow organises itself in a series of vortices tilted against the mean shear, which subsequently get amplified owing to the non-normality of the Navier–Stokes equations. These vortices create a spatially periodic disturbance of the scalar field which is advected away from its side of origin in the channel (i.e. the dark fluid lifted to the top of the channel and the light fluid brought to the bottom of the channel). Even though the shape of the perturbation is different for the energy and variance optima, the early stage ($0 < t < 10$) is rather similar for both solutions. However, we notice that the energy-optimal perturbation does not transport the scalar field as close to the walls as the variance-optimal perturbation. After this initial stage of re-organisation of the scalar field into an array of alternating positive- and negative-concentration patches, the energy-optimal

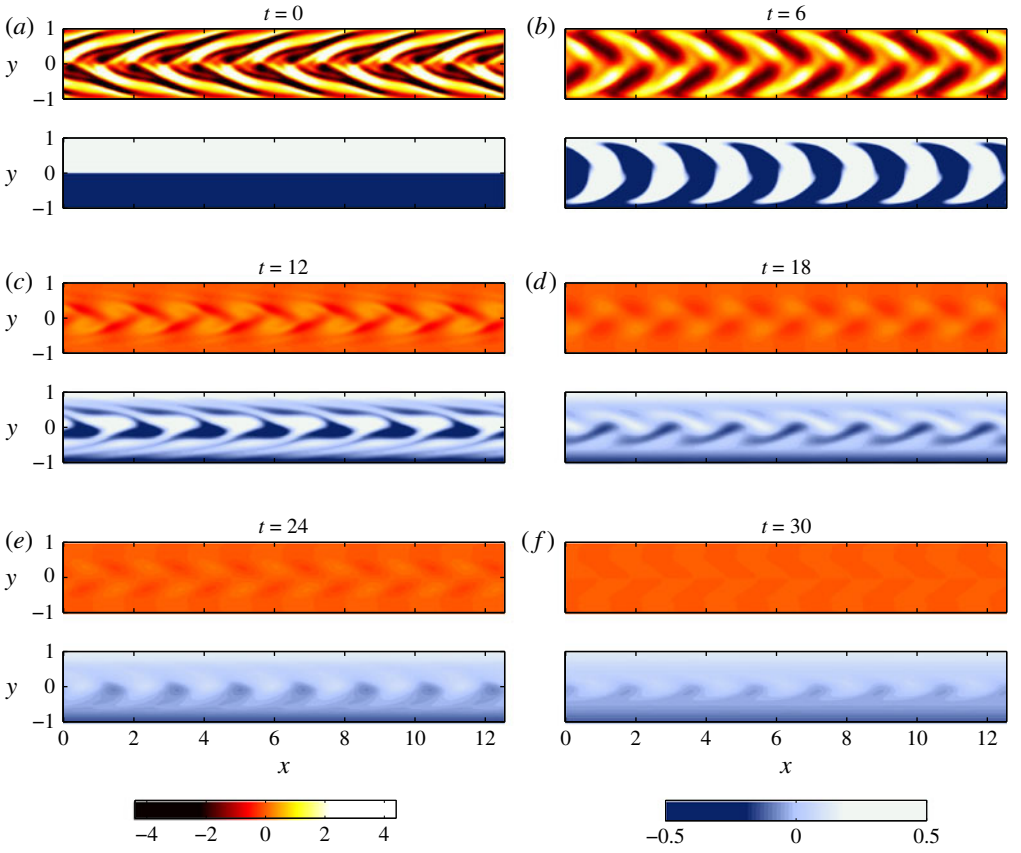


FIGURE 8. (Colour online) Time evolution of the variance-optimal perturbation ($T = 30$) for six different times from $t = 0$ to $t = 30$. Top figures show the perturbation vorticity field $\boldsymbol{\nabla} \times \boldsymbol{u}$ (left colourbar) and bottom figures show the scalar field $\theta$ (right colourbar).

perturbation does not decay as quickly as the variance-optimal perturbation, as we can deduce from the vorticity colour scales in figures 8 and 9. For $t > 10$, the advection then has a negative effect on mixing as the vortices continue to transport the scalar field in a roughly circular manner, which eventually leads, once again, to a segregation of the negative and positive layers of the passive scalar.

We observed in § 2.2 that the passive scalar gradient norm does not show a monotonic time evolution during a typical mixing process. This prevented us from using it as a mixing measure, but amongst the norms considered, it is the one exhibiting the richest dynamics. Therefore, in an effort to characterise the different mixing stages, we will focus on its rate of production given in (2.4). We define $x$-averaged quantities corresponding to the several production terms observed in this equation as

$$P_{\overline{\boldsymbol{u}}}(y, t) = -\frac{1}{L}\int_0^L \boldsymbol{\nabla}\theta \cdot \boldsymbol{\nabla}\overline{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\theta \ \mathrm{d}x, \tag{4.4a}$$

$$P_{\boldsymbol{u}}(y, t) = -\frac{1}{L}\int_0^L \boldsymbol{\nabla}\theta \cdot \boldsymbol{\nabla}\boldsymbol{u} \cdot \boldsymbol{\nabla}\theta \ \mathrm{d}x, \tag{4.4b}$$

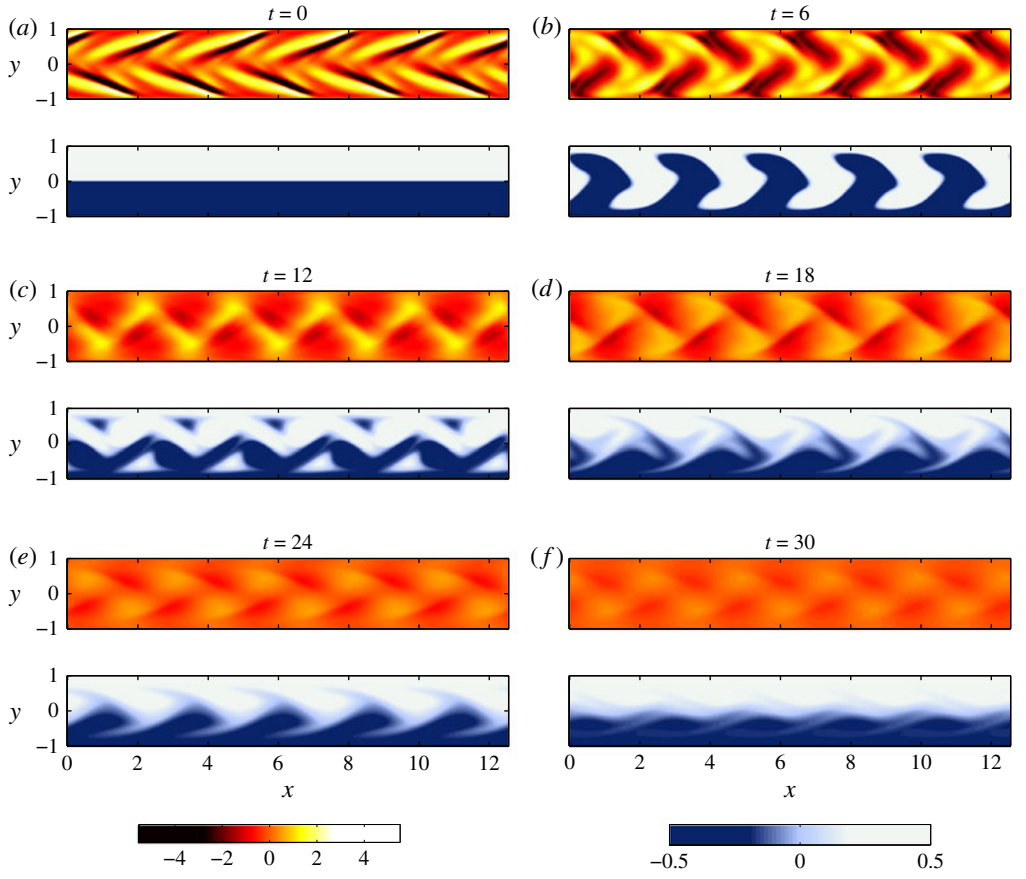


FIGURE 9. (Colour online) Time evolution of the energy-optimal perturbation ($T = 5$) for different times from $t = 0$ to $t = 30$. Top figures show the perturbation vorticity field $\boldsymbol{\nabla} \times \boldsymbol{u}$ (left colourbar) and bottom figures show the scalar field $\theta$ (right colourbar).

$$P_d(y, t) = -\frac{1}{Pe\, L} \int_0^L |\boldsymbol{\nabla}\theta|^2 \, \mathrm{d}x, \tag{4.4c}$$

where $P_{\overline{\boldsymbol{u}}}$ and $P_{\boldsymbol{u}}$ represent respectively the production term due to the base flow $\overline{\boldsymbol{u}}$ and the production term due to the perturbation velocity $\boldsymbol{u}$; $P_d$ represents the negative-definite production due to molecular diffusion. Moreover, the term $P_{\overline{\boldsymbol{u}}}$ can be simplified using the expression for the base flow given in (3.2). We find

$$P_{\overline{\boldsymbol{u}}}(y, t) = \frac{y}{L} \int_0^L \partial_x\theta \partial_y\theta \, \mathrm{d}x. \tag{4.5}$$

We also define the $x$-integrated total rate of change of the scalar gradients $P$ as

$$P = P_{\overline{\boldsymbol{u}}} + P_{\boldsymbol{u}} + P_d. \tag{4.6}$$

Here we focus on the optimal mixing case (corresponding to the $T = 30$ variance-optimal perturbation) represented in figure 10. There is an initial strong

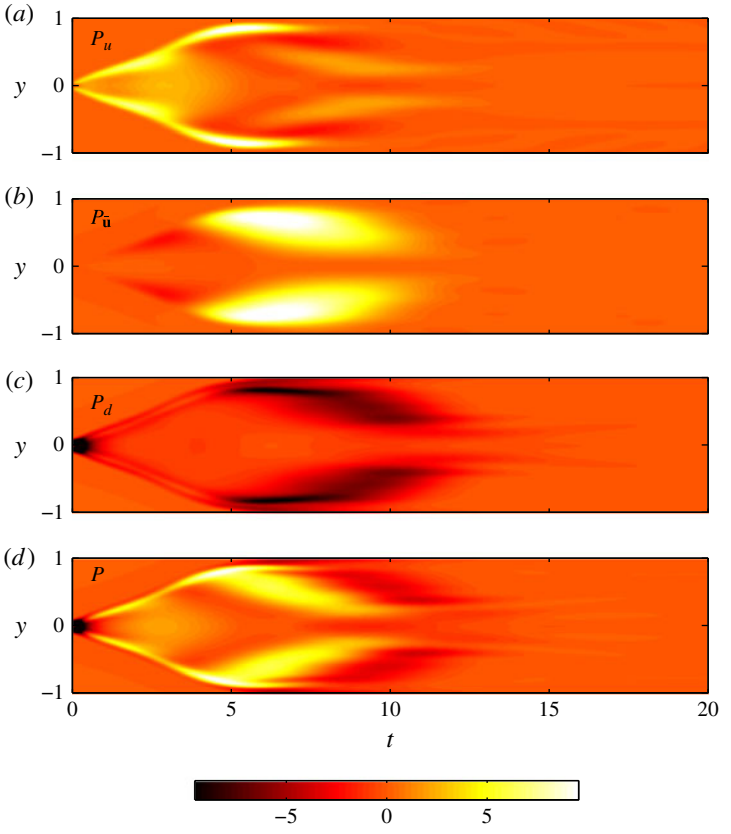


FIGURE 10. (Colour online) Variation with time of: the $x$-integrated production of scalar gradients (as defined in (4.4)) due to ($a$) the velocity perturbation ($P_{\boldsymbol{u}}$), ($b$) the base-flow velocity ($P_{\bar{\boldsymbol{u}}}$), ($c$) the molecular diffusion ($P_d$); ($d$) the total $x$-integrated scalar gradient production term $P$ as defined in (4.6). These plots correspond to the $T = 30$ variance-optimal initial perturbations.

diffusion process occurring at the initially sharp interface ($y \simeq 0$) for very short times. This occurs in every simulation and is independent of the optimal mixing process we are about to describe. Moreover, the contribution of this process to the overall mixing is very limited. The production of gradients (and therefore the mixing process) then occurs in three overlapping stages. First, an energisation/transport stage occurs for $0 < t \lesssim 4.5$. During this stage, the interface is advected towards the top and bottom walls by the perturbation velocity $\boldsymbol{u}$ as we can deduce from the two rays departing from the centreline to the walls in the $P_{\boldsymbol{u}}$ plot (figure 10$a$). We furthermore deduce from the plot of $P_d$ (figure 10$c$) that the transport of the interface towards the walls is not associated with enhanced diffusion. This transport stage is correlated in time with the energetic transient growth of the perturbation velocity $\boldsymbol{u}$ which reaches its maximal energy for $t = 5$.

A second stage then starts during which the passive scalar is effectively mixed by several co-operating mechanisms. First, the velocity perturbation (which reaches its maximum energy) further advects the passive scalar field from $4 \lesssim t \lesssim 7$. This advection is no longer a displacement of the interface, but rather a stretching and stirring process, mostly located in the vicinity of the wall. The signature of this

phenomenon can be seen by the strong production of gradients close to the wall in the $P_{\boldsymbol{u}}$ plot, for $4 \lesssim t \lesssim 7$. Secondly, because of the previous transport stage, the concentration field is in a configuration which is able to take advantage of the constant shear offered by the background flow. Indeed, the concentration vertical interfaces (at $t \simeq 4$) are progressively deformed into parabolae due to the base-flow advection, therefore increasing their total length. The background shear action occurs for $4 \lesssim t \lesssim 10$ and is the leading passive scalar gradient production mechanism as can be observed in the $P_{\bar{u}}$ plot (figure 10*b*). These two gradient production mechanisms (perturbation advection stirring and interface elongation due to the background shear) are both eventually followed by a slightly delayed strong molecular diffusion as we can see in the $P_d$ plot of figure 10(*c*).Finally, the third and last stage starts around $t \simeq 10$ and is a relaxation stage where the perturbation is too small to have a significant impact, and the scalar field has almost become $x$-independent close to the wall, preventing the background shear from further mixing (the term in (4.5) vanishes when $\partial_x \theta = 0$). The concentration field is therefore mostly controlled by a linear diffusion process which continues to homogenise it. A second-order effect also responsible for the quality of mixing is a weak contribution of the term $P_{\boldsymbol{u}}$ in the interior of the domain for $6 \lesssim t \lesssim 13$. This corresponds to a weak stirring.

Therefore, in this case, the identified leading mechanism for mixing is the production of concentration gradients by the background shear, followed up by a strong molecular diffusion. This mechanism is commonly referred to as shear-enhanced diffusion or Taylor dispersion. However, the exploitation of the base-flow shear is only possible because of the initial energisation and transport stage performed by the velocity perturbation $\boldsymbol{u}$.

In this particular case of channel flow at $Re = 500$, the strategy to mix effectively relies on an initial nonlinear advective transport (with coincident transient perturbation energy growth) of the scalar field towards the walls of the channel where, in a second stage, a linear mechanism driven by the parabolic background flow takes over to stretch and disperse the scalar field therefore enhancing molecular diffusion processes. Eventually pure diffusion (i.e. with no advective enhancement) prevails and further homogenises the concentration field in a final relaxation stage. This three-stage mixing strategy takes advantage of the shear-augmented diffusion (i.e. Taylor dispersion) which the background flow provides close to the walls.

### 4.3. *Mixing through wall forcing*

The variational method presented in § 3.2 is formulated to accommodate both optimal perturbations and optimal streamwise velocity boundary conditions. We study the latter problem in this section. We focus on the problem of wall actuation because of its great interest in engineering applications, where prescribing an initial perturbation, or acting directly in the bulk of the flow is not feasible or at least difficult to realise. Therefore we act on the fluid through its boundaries with the exterior domain, i.e. the walls. The norm chosen for the normalisation of the initial wall forcing is no longer the $L^2$-norm, but instead we choose an $L^p$-norm with a relatively large value of $p$ ($p = 20$) to approximate to the $L^\infty$-norm and preserve the continuity of the associated unit sphere. We choose an $L^p$-norm constraint to constrain the maximum value of the streamwise velocity at the wall. If we had instead opted for an $L^2$-norm constraint, the energy could be strongly localised in the $(\boldsymbol{x}, t)$ space and therefore the streamwise velocity could locally be substantially stronger than the base flow, which is a situation we wish to avoid. We set $E_w = 1/2$ (as defined in (3.17*b*)). Therefore, we ensure

that the boundary velocity is never larger than the base-flow advection, i.e. $|u_\pm| \lesssim 1$, without constraining the total amount of energy $\|u_\pm\|_2{}^2/2$ injected into the flow, which quantity we measure *a posteriori*. Also, the length of the channel is reduced to $L = 2\pi$ to reduce the size of optimisation space.

Owing to the incompressibility condition, we can demonstrate that any gradient in the $x$-axis direction of the boundary velocity $u_\pm$ will induce a vertical velocity in the flow, without violating the homogeneous Dirichlet condition for $v$. We recall that we keep the homogeneous Dirichlet boundary condition for the vertical velocity in order to avoid any pumping or injection of fluid in the channel, therefore conserving the total concentration of passive scalar. The vertical velocity close to the wall can thus be deduced to be

$$v_\pm \propto \mp \partial_x u_\pm, \tag{4.7}$$

where the proportionality factor is the inverse of a typical length scale in the $y$-direction. We will not compute the exact magnitude of the cross-stream velocity, but thanks to this scaling we can derive its shape. As an example, we choose $\alpha = 1$, $\beta = 0$ and $T = 20$. We thus optimise the mixing at $T = 20$ of the variance of the scalar field.

We have found that the nonlinear nature of the optimisation problem poses great challenges to effective numerical algorithms converging to a satisfactory solution. Multiple solutions, local rather than global maxima of the cost functional and algorithmic difficulties in finding a proper descent direction and step length all contribute to problems which are non-trivial to address and overcome. In the previous section, both the cost functional and its gradient have been reduced to acceptable levels such that an optimal solution could be claimed. For the wall forcing problem, such a favourable outcome eluded us despite considerable efforts. Nonetheless, a substantial reduction in the cost functional could be achieved before numerical difficulties prevented any further reduction, and we present the solution obtained at this point. Even though optimality cannot be claimed due to a non-vanishing gradient, we were able to identify a wall forcing strategy which resulted in a noticeable mixing efficacy. The variational problem thus suggests an improved, albeit suboptimal, solution to achieve a better mixed state by wall forcing. The effectiveness of the incomplete optimisation thus gives valuable results and physical insight into mixing processes. We present in figure 11 a typical result of a time-dependent boundary condition designed to 'optimise' the finite-time mixing in the flow.

First of all, we remark that the top and bottom wall boundary conditions are very similar even though they are completely independent variables in the optimisation process. They are actually the same field, shifted by a phase of $\phi = \pi$, therefore promoting sinuous modes. The boundary condition identified is composed of streaks of gradients of streamwise velocity, and therefore we observe that the vertical velocity close to the wall is composed of jets of very small streamwise extent (associated with sharp streamwise gradients in the $u_\pm$ plots in figure 11), evolving along different lines of almost constant velocity. The velocity of these jets is, for the vast majority of the time, positive (going in the same direction as the base-flow velocity). We give some reference lines for the velocity of the jet $u_{jet}$ (velocity of the moving streaks of gradient along the walls) in figure 11(*a*). It appears that the velocity of the jets is limited by the base-flow advection, i.e. $u_{jet} \lesssim 1$.

We plot in figure 12 the time evolution of both the vorticity field and the concentration field from $t = 0$ to $t = 40$. Despite the fact that the optimisation has not converged, we observe very good mixing of the scalar field in the flow. This suggests that for practical applications, finding the optimal solution is not always

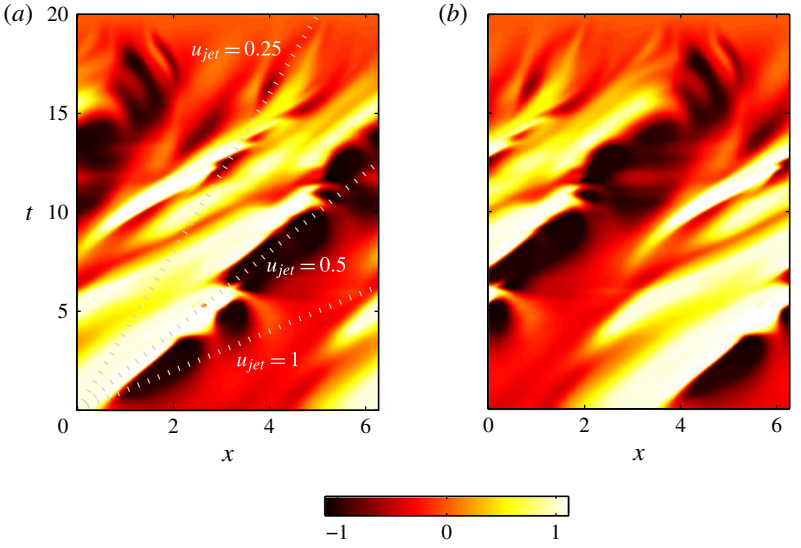


FIGURE 11. (Colour online) The boundary condition aiming to minimise variance at time $T = 20$: ($a$) bottom boundary condition $u_-$, ($b$) top boundary condition $u_+$.

needed and that an approximate solution (which is cheaper to obtain) can indeed lead to a very satisfactory level of mixing.

In order to compare how much energy we have injected into the flow, we compute the value of the $L^2$-norm of the wall velocity. We find $\|u_\pm\|_2{}^2/2 \simeq 3.8 \times 10^{-1}$. Therefore, we have injected roughly 40 times more energy than in the initial perturbation study described in § 4.1. Naturally, forcing the flow from the boundary is much less effective at mixing than applying a perturbation directly within it. Secondly, we apply a streamwise velocity perturbation, even though the cross-stream direction is the one in which we wish to mix. Finally, we recall that the wall boundary condition is not optimal, since we were unable to obtain convergence.

Similarly to how we defined the normalised variance and mix-norms, we define the normalised gradient norm as

$$D(t) = \frac{\|\nabla\theta\|_2{}^2}{\|\nabla\theta_d\|_2{}^2}. \tag{4.8}$$

We plot in figure 13 the time evolution of several normalised measures: gradient norm $D$, variance $V$, mix-norm $M$ and energy of the perturbation $E$ (defined respectively in (4.8), (4.2$b$), (4.2$a$) and (3.7)). We confirm the qualitative observation made in figure 12 that the mixing is indeed very effective because the three scalar field measures have decreased by a factor of $10^4$ at the end of the time integration. Furthermore, by further analysing these time evolution plots, we notice that the mixing process once again occurs in three separate stages. First, for $0 \lesssim t \lesssim 5$, the perturbation energy increases from zero to approximately $\|\boldsymbol{u}\|_2{}^2 = 5 \times 10^{-2}$. This increase of energy in the bulk of the flow is associated with a purely diffusive evolution of the passive scalar norms (horizontal line) but a substantial 'transport' of the scalar field (as shown in figure 12$b$,$c$). From $t \simeq 5$ to $t \simeq 11$, the gradient norm increases by approximately an order of magnitude, and in the meantime, the

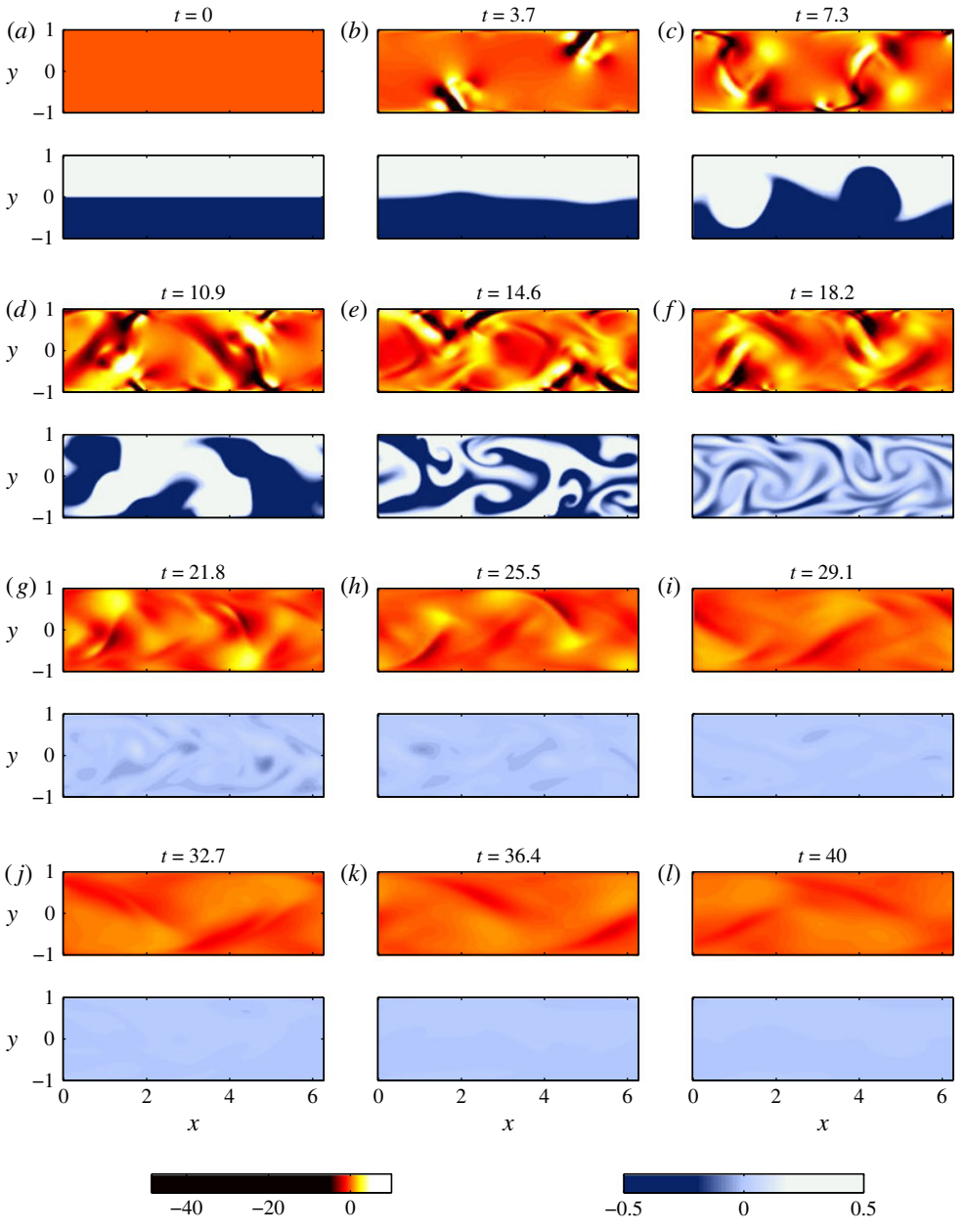


FIGURE 12. (Colour online) Time evolution for the boundary condition aiming to minimise variance (optimised for $T = 20$) for different times from $t = 0$ to $t = 40$. Top figures show the perturbation vorticity field $\boldsymbol{\nabla} \times \boldsymbol{u}$ and bottom figures show the scalar field $\theta$. The vorticity colourmap (bottom left) is adjusted to vary over the range $[-5; 5]$ so as to highlight variations in the bulk vorticity. The bottom right colourmap represents the concentration intensity.

mix-norm starts to decrease. This is the signature of a second dispersion/diffusion stage during which the length of the 'interfaces' is increased. Filaments of dye are created in the flow but the typical size of a filament has not yet reached the diffusion

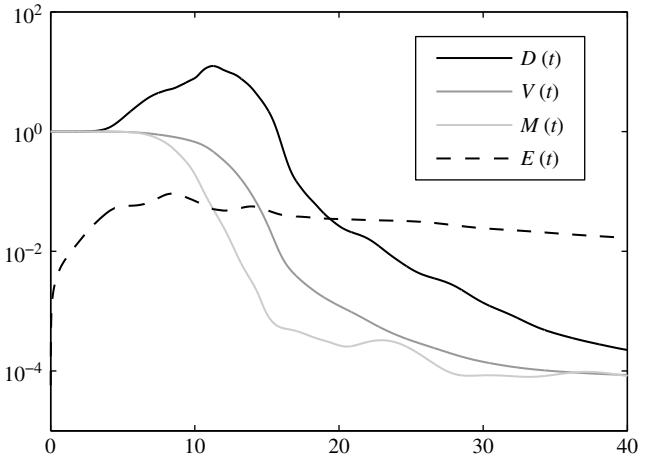


FIGURE 13. Time evolution of various measures for the wall boundary condition case: gradient norm $D(t)$, variance $V(t)$ and mix-norm $M(t)$. The figure also displays the energy of the perturbation $E(t)$. These various measures are respectively defined in (4.8), (4.2*a*), (4.2*b*) and (3.7). After the initial increase of the flow energy $E(t)$, the gradient norm $D(t)$ increases during the passive scalar filamentation stage. For longer times, effective mixing occurs as shown by the strong decrease of all passive scalar norms.

length scale (figure 12*d*,*e*). This occurs around $t \simeq 11$ where we see that the gradient norm starts to decrease anew, this being associated with a stronger decrease of the variance of the passive scalar (figure 12*f*,*g*). Finally, a relaxation/diffusion stage is observed where both the velocity and passive scalar field are governed by essentially linear diffusion processes (figure 12*h*–*l*). We therefore notice a hierarchy of the norms in terms of sensitivity to mixing. Indeed, the mix-norm is the first to experience a decrease, before the variance and eventually the gradient norm decrease as well. We conjecture that this scenario is generic for a real (i.e. finite $Pe$) mixing process.

Similarly to our analysis approach for the mixing-optimal initial perturbation presented in § 4.2, we plot in figure 14 the several contributions, as defined in (4.4), governing the production of concentration gradients. The very beginning of the first stage, as observed in figure 13, corresponds to energy amplification and transport from the wall where it is injected into the interior of the domain. Therefore, for $t \lesssim 5$, because the concentration interface is not yet significantly perturbed on small scales, no gradient production mechanism is observed. From $t \simeq 5$, the interface is advected and 'transported' towards the walls of the domain (e.g. figure 12*c*), very similarly to what has been observed in the optimal initial perturbation case. Then, for $10 \lesssim t \lesssim 20$, we observe the second dispersion/diffusion stage, with a very strong contribution to the gradients production from the velocity perturbation $\boldsymbol{u}$, both at the walls and throughout the bulk of the domain. Concentration gradient production is associated with a slightly delayed strong diffusion as we can notice in the $P_d$ plot in figure 14(*c*), once again the signature of the classical filamentation-then-diffusion process. The Taylor dispersion phenomenon is also observed (term $P_{\overline{u}}$, figure 14*b*), but is not the dominant concentration gradient production mechanism, in contrast to what has been observed in the initial perturbation case. This can be explained by the high levels of perturbation energy we provide in the boundary forcing situation. Indeed, with more energy available, not only is the transport of the scalar towards

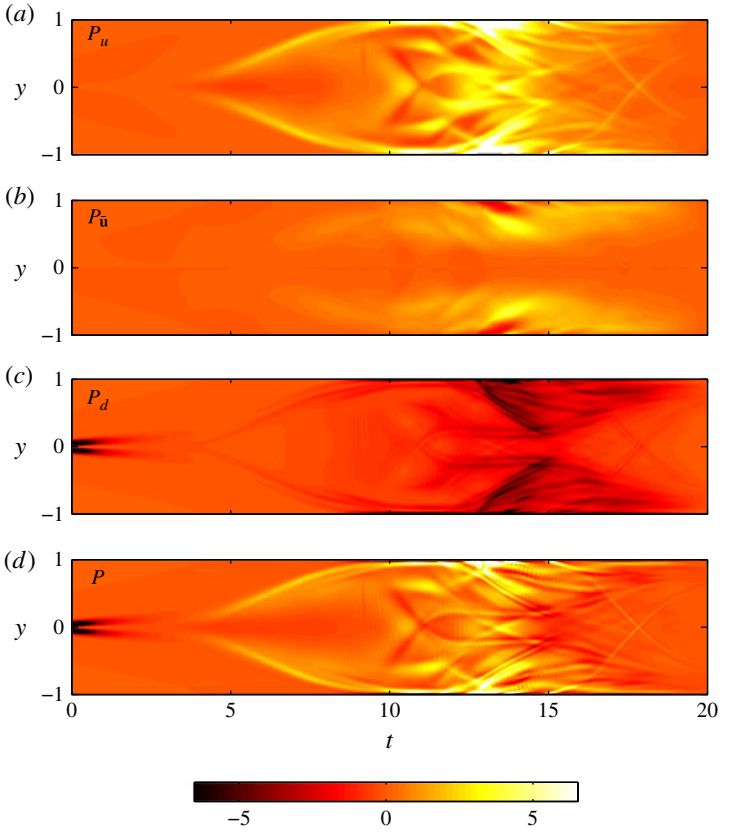


FIGURE 14. (Colour online) Variation with time of: the $x$-integrated production of scalar gradients (as defined in (4.4)) due to ($a$) the velocity perturbation ($P_{\boldsymbol{u}}$), ($b$) the base-flow velocity ($P_{\bar{\boldsymbol{u}}}$) ($c$) the molecular diffusion ($P_d$); ($d$) the total $x$-integrated scalar gradient production term $P$ as defined in (4.6). These plots correspond to the flow obtained with the boundary condition aiming to minimise variance at $T = 20$.

the walls to take advantage of the Taylor dispersion possible, but a powerful interior stirring driven by the perturbation velocity $\boldsymbol{u}$ is present as well (see figure 12*e*,*f* in particular). This can be observed on both the concentration snapshots in figure 12 and in the $P_{\boldsymbol{u}}$ plot of figure 14(*a*). The final late relaxation stage for $t \gtrsim 20$ is also clear.

## 5. Conclusion and perspectives

Here, we have focused on the problem of the mixing of a passive scalar in a two-dimensional plane Poiseuille flow, governed by the Navier–Stokes equations for the velocity field and a classical advection–diffusion equation for the concentration field with finite values of the Reynolds and Péclet numbers. We have presented a toy model for mixing in the presence of both advection and diffusion in order to characterise the behaviour of several scalar field measures during the time evolution of an idealised mixing flow. We conclude that in the presence of diffusion, both the variance and the mix-norm as defined in (2.5) can be used successfully in order to assess the degree of mixing in the flow. When diffusion is not present however, amongst the considered norms, only the mix-norm is sensitive to 'mixing', which is then reduced to stirring

only. We also conclude that measuring the passive scalar gradient norm cannot be linked to the degree of mixing in a flow because of its non-monotonic time evolution during a pure mixing process.

We have formulated a nonlinear variational framework in order to identify control strategies (in terms of initial perturbations or boundary conditions) optimising either the time-averaged energy of the velocity perturbation or a measure of the mixing performed on the scalar field: the variance or the mix-norm. We have investigated the possibility of mixing through energy amplification as was proposed in Aamo & Krstić (2003). Our results show that the energy-optimal initial perturbation is not able to produce effective mixing in the channel at this subcritical Reynolds number. One of the main differences between our study and previous works by Aamo and co-workers is the stability of the system. Indeed, we have considered a stable flow whereas in Aamo *et al.* (2003) for example, the channel flow is considered at a supercritical Reynolds number. We conclude that for stable flows, mixing cannot be systematically produced by optimal control strategies promoting energy amplification. We conjecture that this result extends to unstable laminar flows, but this of course has yet to be shown.

However, the identified solutions which are associated with maximum mixing do have some characteristics of the energy-optimal solution, and in particular rely (for early stages) on the well-known Orr energy amplification mechanism. By comparing initial perturbations obtained by either optimising the variance or the mix-norm we conclude that for short times, the variance-optimal perturbation is associated with disturbances of very small length scales which are able to enhance mixing through diffusion before the optimal target time. On the other hand, the mix-norm-optimal perturbation (since it downplays the role of small scales) is associated with larger length scales and the leading mechanism is coherent transport (stirring) by the perturbation. For longer times however, the optimisation of both norms leads to the identification of the same type of optimal solutions.

Through the analysis of the different terms in the scalar gradient time-evolution equation, we identify a generic long-time mixing scenario associated with this long-time mixing-optimal perturbation. The mixing process in this particular case of a stable two-dimensional Poiseuille flow consists of three stages. First, the perturbation velocity is ‘energised’ by undergoing transient energy growth and advects or ‘transports’ the concentration interface from the centreline of the channel towards the walls where the background shear is large, in order to take advantage of the well-known Taylor dispersion mechanism (or shear-augmented diffusion). After this first transport and energisation stage, the concentration field is ‘dispersed’ and stretched by the background shear, therefore enhancing diffusion and leading to effective mixing. This is essentially classical Taylor dispersion. During this dispersion/diffusion stage, some stirring is also performed by the velocity perturbation, but this mechanism is weaker than the Taylor-enhanced diffusion process for the flow associated with an optimal initial perturbation, although this nonlinear stirring due to the perturbation velocity would be likely to be stronger and more effective at mixing with a larger initial perturbation energy. Finally, a third relaxation/diffusion stage is observed where energy and mixing measures decay due to linear diffusion processes.

We also have identified a wall boundary condition able to produce very effective finite-time mixing. In this case, because the initial energy is zero, the transport and transient energy growth stage must be preceded by a forced energisation where energy is transferred from the wall to the bulk of the flow. Then transport of the passive scalar interface from the middle of the domain towards the walls is once again

observed, still associated with a transient energy growth of the velocity perturbation. After this transport/energisation stage, because the total amount of injected kinetic energy is larger than in the flow associated with an optimal initial perturbation, the perturbation velocity continues to act on the passive scalar field by performing a strong stirring, breaking up large scales into smaller scales where diffusion is more effective. Taylor-dispersion-enhanced diffusion is also present but is no longer the dominant mixing mechanism in this intermediate dispersion/diffusion stage, which is once again followed by a final relaxation stage.

The mechanistic picture described in this paper only holds for the two-dimensional case considered. In three dimensions, more complex fluid mechanisms are present such as vortex stretching, and effective energy amplification strategies do not rely on the Orr mechanism alone anymore. Therefore, different mixing strategies might be observed. We conjecture that the two main results of this paper – namely the ineffective mixing through energy amplification and the exploitation of Taylor dispersion due to the background flow – extend to the three-dimensional case. However, the exact nature of the nonlinear mixing due to the perturbation velocity field is likely to be very different. Furthermore, the picture is even more uncertain if the flow is susceptible to primary instabilities.

We consider this study to be a first attempt to optimise the mixing of a passive scalar in a flow governed by the nonlinear Navier–Stokes equations. However, in order to apply such methods in real engineering situations, many improvements must still be made. First of all, this optimal control study was performed in a full-state-information context where we assumed that we know the state vector everywhere and at any time exactly. Of course, in applications, this is not possible and the information can only be collected by sensors which force us to develop a partial-state-information framework. Also, even in the wall forcing flow, we considered continuous functions. However, in reality, the control of the flow can only be done at discrete locations, for example by blowing and suction. The type of actuator needs to be embedded in the model in order to be able to mimic realistic situations. Also, the presence of noise coming from the measurements needs to be taken into account.

Finally, the ultimate goal of such control problems is to be able to control the flow in real time. Therefore, the optimisation algorithm needs to be very effective in order to identify the control strategy as the flow evolves. Developing effective optimisation algorithms is therefore crucial. Variational methods for optimisation have been intensively developed and applied recently in the fluid mechanics literature. Applying these theoretical studies by performing control experiments or designing optimal mixing devices is probably the next task to carry out in order to push forward real-world flow control possibilities.

**Acknowledgement**

The research activities of D.P.G.F. and C.P.C. are supported by EPSRC Research Grant No. EP/H050310/1 'AIM (Advanced Instability Methods) for Industry'.

## Appendix A. Derivation of adjoint equations and gradients

In this appendix we derive the gradient of the cost functional $\mathscr{J}$ defined in (3.8), as well as the adjoint equations together with the associated terminal and boundary conditions. We rewrite the cost functional (3.8) using the Einstein summation convention:

$$\mathscr{J}(u_i,\theta) = \frac{1-\alpha}{2}\int_0^T\int_\Omega u_i(\boldsymbol{x},t)u_i(\boldsymbol{x},t)\,\mathrm{d}t\,\mathrm{d}\Omega + \frac{\alpha}{2}\int_\Omega \partial_j^{-\beta}\theta(\boldsymbol{x},T)\partial_j^{-\beta}\theta(\boldsymbol{x},T)\,\mathrm{d}\Omega. \tag{A 1}$$

In subsequent steps, we will use for simplicity the following notation:

$$\left.\begin{aligned}(a(\boldsymbol{x},t_0),b(\boldsymbol{x},t_0)) &= (a,b)_{t=t_0},\\ [c(\boldsymbol{x}_0,t),d(\boldsymbol{x}_0,t)] &= [c,d]_{\boldsymbol{x}=\boldsymbol{x}_0}.\end{aligned}\right\} \tag{A 2}$$

The augmented Lagrangian functional with embedded constraints is

$$\begin{aligned}\mathscr{L}(u_i,u_i^\dagger,p,p^\dagger,\theta,\theta^\dagger,u_\pm,u_\pm^\dagger) = {}& \mathscr{J}(u_i,\theta) - \left\langle u_i^\dagger,\partial_t u_i + U_j\partial_j u_i + u_j\partial_j\overline{u}_i\right.\\ &\left. + \partial_i p - Re^{-1}\partial_{jj}u_i\right\rangle - \left\langle p^\dagger,\partial_j u_j\right\rangle\\ &- \left\langle\theta^\dagger,\partial_t\theta + U_j\partial_j\theta - Pe^{-1}\partial_{jj}\theta\right\rangle - \left(u_{0i}^\dagger,u_{0i}-u_i\right)_{t=0}\\ &- \left[u_{\pm i}^\dagger,u_{\pm i}-u_i\right]_{\boldsymbol{x}=\boldsymbol{x}_\pm}.\end{aligned} \tag{A 3}$$

In order to ensure the optimality of the solution, we need to enforce $\delta\mathscr{L}=0$ with

$$\begin{aligned}\delta\mathscr{L} = {}& \left\langle\frac{\delta\mathscr{L}}{\delta\boldsymbol{u}},\delta\boldsymbol{u}\right\rangle + \left\langle\frac{\delta\mathscr{L}}{\delta p},\delta p\right\rangle + \left\langle\frac{\delta\mathscr{L}}{\delta\theta},\delta\theta\right\rangle + \left(\frac{\delta\mathscr{L}}{\delta\boldsymbol{u}_0},\delta\boldsymbol{u}_0\right) + \left[\frac{\delta\mathscr{L}}{\delta u_\pm},\delta u_\pm\right]\\ &+ \left\langle\frac{\delta\mathscr{L}}{\delta\boldsymbol{u}^\dagger},\delta\boldsymbol{u}^\dagger\right\rangle + \left\langle\frac{\delta\mathscr{L}}{\delta p^\dagger},\delta p^\dagger\right\rangle + \left\langle\frac{\delta\mathscr{L}}{\delta\theta^\dagger},\delta\theta^\dagger\right\rangle + \left(\frac{\delta\mathscr{L}}{\delta\boldsymbol{u}_0^\dagger},\delta\boldsymbol{u}_0^\dagger\right) + \left[\frac{\delta\mathscr{L}}{\delta u_\pm^\dagger},\delta u_\pm^\dagger\right].\end{aligned} \tag{A 4}$$

With all terms in (A 4) independent of each other, we require all partial derivatives of the functional $\mathscr{L}$ to vanish. The second line of (A 4) corresponds to the derivative of $\mathscr{L}$ with respect to the adjoint variables, enforcing the direct equations and constraints. Let us compute the differential with respect to the direct variables term by term. We obtain

$$\begin{aligned}\left\langle\frac{\delta\mathscr{L}}{\delta u_i},\delta u_i\right\rangle = {}& \left\langle\frac{\delta\mathscr{J}}{\delta u_i},\delta u_i\right\rangle - \left\langle u_i^\dagger,\partial_t\delta u_i + \delta u_j\partial_j U_i + U_j\partial_j\delta u_i\right.\\ &\left. - Re^{-1}\partial_{jj}\delta u_i\right\rangle - \left\langle p^\dagger,\partial_j\delta u_j\right\rangle\\ &- \left\langle\theta^\dagger,\delta u_j\partial_j\theta\right\rangle - \left(u_{0i}^\dagger,\delta u_i\right)_{t=0} - \left[u_{\pm i}^\dagger,\delta u_i\right]_{\boldsymbol{x}=\boldsymbol{x}_\pm},\end{aligned} \tag{A 5}$$

which can be rewritten after some rearranging and integration by parts, and by explicitly computing the term $\delta\mathscr{J}/\delta u_i$:

$$\begin{aligned}\left\langle\frac{\delta\mathscr{L}}{\delta u_i},\delta u_i\right\rangle = {}& \left\langle(1-\alpha)\,u_i + \partial_t u_i^\dagger - u_j^\dagger\partial_i U_j + U_j\partial_j u_i^\dagger + \partial_i p^\dagger + Re^{-1}\partial_{jj}u_i^\dagger - \theta^\dagger\partial_i\theta,\delta u_i\right\rangle\\ &- \left(u_i^\dagger,\delta u_i\right)_{t=T} + \left(u_i^\dagger - u_{0i},\delta u_i\right)_{t=0}\\ &- \left[u_i^\dagger U_j n_j + p^\dagger n_i + Re^{-1}\partial_j u_i^\dagger n_j + u_{\pm i}^\dagger,\delta u_i\right]_{\boldsymbol{x}=\boldsymbol{x}_\pm}\\ &+ \left[Re^{-1}u_i^\dagger n_j,\partial_j\delta u_i\right]_{\boldsymbol{x}=\boldsymbol{x}_\pm},\end{aligned} \tag{A 6}$$

where $n_j$ is the $j$th component of the outward unit vector normal to $\partial\Omega$. The differentiation with respect to the pressure $p$ gives, after integration by parts,

$$\left\langle \frac{\delta\mathscr{L}}{\delta p}, \delta p \right\rangle = -\left\langle \partial_i u_i^\dagger, \delta p \right\rangle + \left[ u_i^\dagger n_i, \delta p \right]_{\boldsymbol{x}=\boldsymbol{x}_\pm}. \tag{A 7}$$

The differentiation with respect to the scalar concentration $\theta$ yields, after integration by parts,

$$\begin{aligned}\left\langle \frac{\delta\mathscr{L}}{\delta\theta}, \delta\theta \right\rangle = {} & \left\langle \partial_t\theta^\dagger + U_j\partial_j\theta^\dagger + Pe^{-1}\partial_{jj}\theta^\dagger, \delta\theta \right\rangle \\ & - \left(\theta^\dagger - (-1)^\beta\alpha\partial_{jj}^{-\beta}\theta, \delta\theta\right)_{t=T} + \left(\theta^\dagger, \delta\theta\right)_{t=0} \\ & - \left[\theta^\dagger U_j n_j + Pe^{-1}\partial_j\theta^\dagger n_j, \delta\theta\right]_{\boldsymbol{x}=\boldsymbol{x}_\pm} + \left[Pe^{-1}\theta^\dagger n_j, \partial_j\delta\theta\right]_{\boldsymbol{x}=\boldsymbol{x}_\pm}. \end{aligned} \tag{A 8}$$

During the derivation of the previous expression, we used the direct boundary condition on the concentration field (3.4*c*) when $\beta = -1$. When $\beta = 1$ however, by requiring the boundary contributions (arising from integration by parts) to be zero, we find the boundary condition

$$\partial_y^{-1}\theta(\boldsymbol{x}_\pm, t) = \partial_y\Theta(\boldsymbol{x}_\pm, t) = 0, \tag{A 9}$$

where $\Theta$ is the solution to the Poisson equation $\nabla^2\Theta = \theta$. This boundary condition is needed to define the Poisson operator used to compute $\boldsymbol{\nabla}^{-1}\theta$.

From the space and time integrals, we recover the field adjoint equations:

$$\left.\begin{gathered} \partial_t u_i^\dagger + U_j\partial_j u_i^\dagger - u_j^\dagger\partial_i U_j + \partial_i p^\dagger + Re^{-1}\partial_{jj}u_i^\dagger = \theta^\dagger\partial_i\theta - (1-\alpha)\,u_i, \\ \partial_j u_j^\dagger = 0, \\ \partial_t\theta^\dagger + U_j\partial_j\theta^\dagger + Pe^{-1}\partial_{jj}\theta^\dagger = 0. \end{gathered}\right\} \tag{A 10}$$

Many simplifications can be made in the boundary terms thanks to the direct boundary and initial conditions. After these simplifications, we find the following boundary conditions:

$$\boldsymbol{u}^\dagger(\boldsymbol{x}_\pm, t) = 0, \quad \partial_y\theta^\dagger(\boldsymbol{x}_\pm, t) = 0. \tag{A.11$a$,$b$}$$

No boundary condition is found for the pressure, and we thus decide to prescribe a homogeneous Neumann condition identical to the direct system:

$$\partial_y p^\dagger(\boldsymbol{x}_\pm, t) = 0. \tag{A 12}$$

We also find the following equalities for the adjoint variables:

$$\left.\begin{gathered} u_\pm^\dagger = \mp Re^{-1}\partial_y u^\dagger(\boldsymbol{x}_\pm, t) \\ \boldsymbol{u}_0^\dagger = \boldsymbol{u}^\dagger(\boldsymbol{x}, 0). \end{gathered}\right\} \tag{A 13}$$

The adjoint terminal condition is found to be

$$\left.\begin{gathered} \boldsymbol{u}^\dagger(\boldsymbol{x}, T) = 0, \\ \theta^\dagger(\boldsymbol{x}, T) = (-1)^\beta\alpha\nabla^{-2\beta}\theta(\boldsymbol{x}, T). \end{gathered}\right\} \tag{A 14}$$

We stress here that these formulae do not hold for any value of $\beta$, but only for $\beta \in [\![-1, 0, 1]\!]$.

Finally, we need to enforce that the terms $(\delta\mathscr{L}/\delta\boldsymbol{u}_0, \delta\boldsymbol{u}_0)$ and $[\delta\mathscr{L}/\delta u_\pm, \delta u_\pm]$ are equal to zero. We easily observe that

$$\frac{\delta\mathscr{L}}{\delta\boldsymbol{u}_0} = \boldsymbol{u}_0^\dagger \quad \text{and} \quad \frac{\delta\mathscr{L}}{\delta u_\pm} = u_\pm^\dagger. \tag{A.15a,b}$$

There is no reason why $\boldsymbol{u}_0^\dagger$ and $u_\pm^\dagger$ should be equal to zero. However, by noticing that

$$\boldsymbol{\nabla}_{\boldsymbol{u}_0}\mathscr{J} = \frac{\delta\mathscr{L}}{\delta\boldsymbol{u}_0} \quad \text{and} \quad \boldsymbol{\nabla}_{u_\pm}\mathscr{J} = \frac{\delta\mathscr{L}}{\delta u_\pm}, \tag{A.16a,b}$$

we can formulate an expression for the gradient of the cost functional with respect to each of the control variables $\boldsymbol{u}_0$ and $u_\pm$ as

$$\boldsymbol{\nabla}_{\boldsymbol{u}_0}\mathscr{J} = \boldsymbol{u}_0^\dagger, \tag{A 17a}$$

$$\boldsymbol{\nabla}_{u_\pm}\mathscr{J} = u_\pm^\dagger. \tag{A 17b}$$

By employing a gradient descent algorithm, we thus are able to reach the optimal control strategies $\boldsymbol{u}_0$ or $u_\pm$ such that the norm of the gradient projected onto the normalisation constraint surface vanishes (see Foures *et al.* 2013). In practice, we do not reach machine precision, but we do observe a large decrease of the residual as defined in (4.3).

## REFERENCES


AAMO, O. M. & KRSTIĆ, M. 2003 *Flow Control by Feedback: Stabilization and Mixing*. Springer.

AAMO, O. M. & KRSTIĆ, M. 2004 Feedback control of particle dispersion in bluff body wakes. *Intl J. Control* **77**, 1001–1018.

AAMO, O. M., KRSTIĆ, M. & BEWLEY, T. R. 2003 Control of mixing by boundary feedback in 2d channel flow. *Automatica* **39**, 1597–1606.

ANNASWAMY, A. M. & GHONIEM, A. F. 1995 Active control in combustion systems. *IEEE Control Syst.* **15**, 49–63.

AREF, H. 1984 Stirring by chaotic advection. *J. Fluid Mech.* **143**, 1–21.

BALOGH, A., AAMO, O. M. & KRSTIĆ, M. 2005 Optimal mixing enhancement in 3-D pipe flow. *IEEE Trans. Control Syst. Technol.* **13**, 27–41.

CHERUBINI, S., DE PALMA, P., ROBINET, J.-C. & BOTTARO, A. 2010 Rapid path to transition via nonlinear localised optimal perturbations in a boundary-layer flow. *Phys. Rev.* E **82**, 066302.

D'ALESSANDRO, D., DAHLEH, M. & MEZIC, I. 1999 Control of mixing in fluid flow: a maximum entropy approach. *IEEE Trans. Autom. Control* **44**, 1852–1863.

DANCKWERTS, P. V. 1952 The definition and measurement of some characteristics of mixtures. *Appl. Sci. Res.* A **3**, 279–296.

DUGUET, Y., MONOKROUSOS, A., BRANDT, L. & HENNINGSON, D. S. 2013 Minimal transition thresholds in plane Couette flow. *Phys. Fluids* **25**, 084103.

ECKART, C. 1948 An analysis of the stirring and mixing processes in incompressible fluids. *J. Mar. Res.* **7**, 265–275.

FOURES, D. P. G., CAULFIELD, C. P. & SCHMID, P. J. 2013 Localization of flow structures using $\infty$-norm optimization. *J. Fluid Mech.* **729**, 672–701.

FURSIKOV, A. V., GUNZBURGER, M. D. & HOU, L. 1998 Boundary value problems and optimal boundary control for the Navier–Stokes system: the two-dimensional case. *SIAM J. Control Optim.* **36** (3), 852–894.

GUÉGAN, A., SCHMID, P. J. & HUERRE, P. 2006 Optimal energy growth and optimal control in swept Hiemenz flow. *J. Fluid Mech.* **566**, 11–45.

GUERMOND, J. L., MINEV, P. & SHEN, J. 2006 An overview of projection methods for incompressible flows. *Comput. Meth. Appl. Mech. Engng* **195**, 6011–6045.

HESSEL, V., LÖWE, H. & SCHÖNFELD, F. 2005 Micromixers - a review on passive and active mixing principles. *Chem. Engng Sci.* **60**, 2479–2501.

JUNIPER, M. P. 2011 Triggering in the horizontal Rijke tube: non-normality, transient growth and bypass transition. *J. Fluid Mech.* **667**, 272–308.

LEKIEN, F., COULLIETTE, C., MARIANO, A. J., RYAN, E. H., SHAY, L. K., HALLER, G. & MARSDEN, J. 2005 Pollution release tied to invariant manifolds: a case study for the coast of Florida. *Physica* D **210**, 1–20.

LIN, Z., THIFFEAULT, J.-L. & DOERING, C. R. 2011 Optimal stirring strategies for passive scalar mixing. *J. Fluid Mech.* **675**, 465–476.

LINDEN, P. F. 1999 The fluid mechanics of natural ventilation. *Annu. Rev. Fluid Mech.* **31**, 201–238.

LINDZEN, R. S. 1988 Instability of plane parallel shear flow (toward a mechanistic picture of how it works). *Pure Appl. Geophys.* **126**, 103–121.

LIU, W. 2008 Mixing enhancement by optimal flow advection. *SIAM J. Control Optim.* **47**, 624–638.

LUNASIN, E., LIN, Z., NOVIKOV, A., MAZZUCATO, A. & DOERING, C. R. 2012 Optimal mixing and optimal stirring for fixed energy, fixed power, or fixed palenstrophy flows. *J. Math. Phys.* **53**, 115611.

MAO, X., BLACKBURN, H. M. & SHERWIN, S. J. 2012 Optimal inflow boundary condition perturbations in steady stenotic flow. *J. Fluid Mech.* **705**, 306–321.

MATHEW, G., MEZIC, I., GRIVOPOULOS, S., VAIDYA, U. & PETZOLD, L. 2007 Optimal control of mixing in Stokes fluid flows. *J. Fluid Mech.* **580**, 261–281.

MATHEW, G., MEZIĆ, I. & PETZOLD, L. 2005 A multiscale measure for mixing. *Physica* D **211**, 23–46.

MONOKROUSOS, A., BOTTARO, A., BRANDT, L., DI VITA, A. & HENNINGSON, D. S. 2011 Nonequilibrium thermodynamics and the optimal path to turbulence in shear flows. *Phys. Rev. Lett.* **106**, 134502.

NGUYEN, N.-T. & WU, Z. 2005 Micromixers: a review. *J. Micromech. Microengng* **15**, R1.

ORR, W. M. F. 1907 The stability or instability of the steady motions of a perfect liquid and of a viscous liquid. Part I: a perfect liquid, part II: a viscous liquid. *Proc. R. Irish Acad.* A **27**, 9–138.

OTTINO, J. M. 1990 Mixing, chaotic advection, and turbulence. *Annu. Rev. Fluid Mech.* **22**, 207–254.

PRINGLE, C. C. T. & KERSWELL, R. R. 2010 Using nonlinear transient growth to construct the minimal seed for shear flow turbulence. *Phys. Rev. Lett.* **105**, 154502.

RABIN, S. M. E., CAULFIELD, C. P. & KERSWELL, R. R. 2012 Triggering turbulence efficiently in plane Couette flow. *J. Fluid Mech.* **712**, 244–272.

RABIN, S. M. E., CAULFIELD, C. P. & KERSWELL, R. R. 2014 Designing a more nonlinearly stable laminar flow via boundary manipulation. *J. Fluid Mech.* **738**, R1.

RHINES, P. B. & YOUNG, W. R. 1983 How rapidly is a passive scalar mixed within closed streamlines? *J. Fluid Mech.* **133**, 133–145.

TAYLOR, G. I. 1953 Dispersion of soluble matter in solvent flowing slowly through a tube. *Proc. R. Soc. Lond.* A **219**, 186–203.

THIFFEAULT, J.-L. 2012 Using multiscale norms to quantify mixing and transport. *Nonlinearity* **25**, R1.

WUNSCH, C. & FERRARI, R. 2004 Vertical mixing, energy, and the general circulation of the oceans. *Annu. Rev. Fluid Mech.* **36**, 281–314.