Optimal mixing enhancement by local perturbation

Gary Froyland[1], Cecilia González-Tokman[2], and Thomas M. Watson[1]

[1]School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia
[2]School of Mathematics and Physics, The University of Queensland, St Lucia QLD 4072, Australia

October 14, 2015

Abstract

We consider the problem of how to apply local perturbations to optimally enhance the mixing of a (possibly time-dependent) dynamical system. We develop a flexible modelling approach based on the transfer operator of the dynamical system, and pose the problem in the language of convex optimisation. The optimal local perturbations can then be efficiently computed, at discrete time instants, by standard convex optimisation techniques. The local perturbations satisfy physical constraints, such as preservation of the invariant measure of the dynamics (for example, for incompressible fluid flow, the perturbations preserve volume), and a variety of other physical constraints can also be easily enforced. We show that one can achieve surprisingly large speedups in mixing via optimising the diffusion, as compared to fixed diffusion protocols. Finally, we indicate how one might alternatively try to use local perturbation to push a mass density toward a particular region of the domain.

1 Introduction

Mixing in fluids is a question of fundamental interest in engineering and natural sciences, with applications ranging from industrial and chemical mixing on small and large scales, to preventing the spreading of pollutants in geophysical flows. Depending on the situation, one may be interested in maximising the speed of mixing, slowing it down or even in directing a passive tracer towards a desired target distribution. The goal of this work is to place these problems within a flexible numerical framework, and to develop a solution strategy based on existing optimisation tools. We refer the reader to the book of Sturman *et al.* [SOW06] for mathematical foundations of advective (kinematic) mixing, and to the recent review [A$^+$14].

A powerful theoretical and numerical approach to studying mixing arises from the spectral theory of transfer operators, which capture the full advective or advective-diffusive evolution of the flow. Their spectrum lies inside the unit circle in the complex plane, and for steady flows the eigenfunction corresponding to the largest positive eigenvalue (namely unity) is the steady-state mass distribution. The magnitude of the second largest eigenvalue controls the exponential rate at which initial distributions approach the equilibrium distribution. An important global feature of mixing, *strange eigenmodes* [Pie94], arises directly from transfer operator constructions. Strange eigenmodes are eigenfunctions of the transfer operator that correspond to eigenvalues of large magnitude (close to magnitude 1), in the situation where there is a large spectral gap between these large magnitude eigenvalues and lower magnitude eigenvalues. This large spectral gap means that the strange eigenmodes decay much more slowly, and survive for much longer timescales, than lower eigenmodes. In the context of time-varying dynamics, the strange eigenmodes are themselves time-varying [FLS10, Fro13]. Strange eigenmodes have also been investigated in the context of open flows [GDT11, GDTR09].

Following ideas of Dellnitz and Junge [DJ99], Froyland and co-workers have worked on identifying metastable [FPET07] and coherent [FSM10, FHR$^+$12] structures for geophysical flows using Ulam-based approximations of the transfer operator. Ulam's method is a popular method of numerically approximating the transfer operator, and is based on a projection of the operator onto characteristic functions supported on a fine grid in phase space [Ula60]. A practical advantage of Ulam's method is that if the governing flow is steady, the associated matrix needs to be computed only once, and the approximate evolution of future distributions can be computed by matrix/vector multiplication.

Related to Ulam's method is the *mapping method* due to [SW51], which has been exploited in recent years to analyse, design and optimise mixing [SKMA08, GSA12]. The diffusive mapping method [GGA12b] and modified mapping method [SCU$^+$13] have been developed to incorporate diffusion to purely advective transport models. This approach is relevant for the study of a more general class of flows, in particular for those with relatively low Péclet number (including advective-diffusive transport in *microflows*), as well as for the treatment of numerical schemes, which unavoidably introduce diffusion. This numerical diffusion, quantified by the size of the underlying discretisation, has been recently exploited to simulate advection-diffusion transport problems at different Péclet numbers [GGA12a].

Quantifying mixing has been a topic of interest for many years. In one dimension, how well a bijection of the unit interval mixes has been measured by the rate at which an initial interval is chopped into disjoint subintervals. The rationale for this criterion is that in the presence of small diffusion, the more disconnected the collection of subintervals, the greater the effect of the diffusion. Prior work considering deterministic mixing only includes Krotter *et al.* [KCOL12], who carry out numerical experiments of interval exchange transformations (IET) of the interval, called *cutting and shuffling*. They focus on two finite-time measures of mixing – the number of interfaces introduced by the IET, and the percent unmixed – and propose design mechanisms for optimising mixing in these systems. Byott *et al.* [BHZ13] consider maps of the form $x \mapsto mx \pmod 1$, $m \in \mathbb{Z}$, composed with IETs. They conclude

that the mixing rate is either unaffected or slowed by the additional permutations. The rate of mixing is quantified by the magnitude of second largest eigenvalue of the Perron-Frobenius operator. Subsequently, [BLZ14] study a similar setup, now allowing for the orientation of some pieces to be reversed. They establish conditions on the interval exchange which make the full dynamics non-mixing. They also identify a map, the *zig-zag* map, which is associated with the worst *asymptotic mixing rate* within the family, and explicitly compute the rate of mixing for that case.

The combined effect of advective and diffusive dynamics in one dimension has been studied by several authors. Ashwin *et al.* [ANK02] combine interval exchanges of 3 to 5 equal length intervals with diffusion from a Gaussian heat kernel. Mixing rates were computed as the magnitude of the second eigenvalue of the Perron-Frobenius operator (acting on L^2), using a Galerkin scheme on Fourier modes. The authors also investigate the time it takes to achieve 95% mixing. The mixing rates observed were faster when advection and diffusion were combined than when only diffusion is present. Sturman [Stu12, Section 5.5] repeats the experiment of [ANK02] with a much tighter heat kernel diffusion and compares the result with a weak-mixing interval exchange. The mixing rates are reported in a *mix-norm* (introduced below), and again, the combination of advection and diffusion produces faster mixing than either advection or diffusion alone.

Early measures of mixing in the advection-diffusion setting have been based on quantities such as dispersion statistics and spatial variance of passive scalar concentrations; see e.g. [Pro99, HY00, LH04, Thi08]. Thiffeault, Doering and Gibbon [TDG04] use the long term average of the variance of a stirring protocol (source) as a way of measuring *mixing efficiency* in an advection-diffusion setting described by the equation $\partial_t\theta + \mathbf{u}\cdot\nabla\theta = \kappa\Delta\theta + s$, where κ represents molecular diffusivity, and $\mathbf{u}$ is an (incompressible) stirring velocity field. They derive bounds for mixing efficiency in terms of the variance at the steady state, κ, $\mathbf{u}$, and the scale on which u and s act.

Mathew, Mezić and Petzold [MMP05] introduced a multiscale measure for mixing on tori. This multiscale norm, called the mix-norm, is related to weak convergence in L^2, and is equivalent to a Sobolev norm of index $-1/2$. This and related norms have been receiving considerable attention in the mixing literature; we refer the reader to Thiffeault [Thi12] for a detailed review on multiscale norms, including an extensive list of references.

With established quantitative ways of studying mixing, the problem of *optimising* mixing has become a subject of investigation in recent years. The work [TP08] identifies optimal source distributions which are best mixed for given stirring field and diffusivity. The work [CAG08] propose *short time horizon* optimisation procedures, which minimise the mix-norm of the solution of an advection-diffusion equation over a range of different mixing protocols. In a recent review, Balasuriya [Bal15] surveys dynamical systems techniques to enhance mixing in the context of microfluids.

Our focus in this paper is to determine an optimal application of spatially localised diffusion, which, in combination with the underlying dynamics, minimises the L^2 distance from equilibrium. We investigate diffusion protocols for enhanced mixing in 1D and 2D case studies, ranging from the purely diffusive regime to advection-diffusion models. The

flexible numerical optimisation scheme we develop allows for the introduction of practical constraints, such as placing limits on the shape of the perturbation, limiting the perturbation budget, and limiting the number of perturbation sites. Our results show that optimising the diffusion protocol can significantly outperform standard protocols, such as uniform diffusion. Our technique is insensitive to the Péclet number.

The paper is arranged as follows. In Section 2 we introduce basic dynamical concepts and construct the relevant convex optimisation problem. Section 3 discusses discretisations of the dynamics and the discretised optimisation problem, in preparation for the numerics. Two case studies in one and two dimensions are carried out in Sections 4 and 5, respectively. We analyse several natural situations, including limits on the range and form of the diffusion, limits on the overall diffusion budget, and targetting nonequilibrium distributions.

2 Model setup and a convex optimisation problem

2.1 Dynamics and densities

Our domain is a bounded subset X of $\mathbb{R}^d$. The dynamics could be governed by several setups.

1. An ordinary differential equation $\frac{dx}{dt} = F(x,t)$ representing purely advective dynamics, where $F : X \to \mathbb{R}^d$ is sufficiently smooth to guarantee unique solutions.

2. A partial differential equation $\frac{\partial \rho}{\partial t} = -\nabla\cdot(F(x,t)\rho(x,t)) + \nabla\cdot(D(x,t)\nabla\rho(x,t))$, consisting of advective dynamics given by the vector field F and diffusive dynamics controlled by $D(x,t)$.

3. A non-singular[1] measurable map $T : X \to X$ representing purely advective dynamics.

4. A discrete-time advection-diffusion process, for example, $\mathcal{L}_\eta\rho(x) = \int \rho(x)k_\eta(x,Ty)\,d\ell(y)$, where $k : X \times X \to \mathbb{R}^+$ is a stochastic kernel and T is a measurable non-singular map.

While the above four examples seem somewhat disparate, they can be linked by considering the dynamical action on a density $\rho : X \to \mathbb{R}^+$. In the following sections, we will apply our local perturbations at discrete time intervals, so in the continuous-time examples 1. and 2. above, we choose a flow duration τ, and consider time-τ dynamics, yielding a map $T : X \to X$. Define a linear operator $\mathcal{L}_T\rho(x) = \rho(T^{-1}x)/|det(DT(T^{-1}x))|$, where DT is the spatial derivative of T. The operator $\mathcal{L}_T$ is known as the *Perron-Frobenius operator* or *transfer operator* of T and describes the discrete-time evolution of the density ρ under the action of the time-τ map T. In case 3., we can implicitly define a Perron-Frobenius operator for T by $\mathcal{L}_T : L^1(X) \circlearrowleft$ by $\int_A \mathcal{L}_T\rho\; d\ell = \int_{T^{-1}A} \rho\; d\ell$ for all measurable $A \subset X$. In the situation where T is smooth and invertible, this definition coincides with the one given above. Similarly, in case 4. we have the linear operator $\mathcal{L}_\eta$, which is the transfer operator for this dynamics. Thus, for each of these dynamical systems, we have a linear operator that describes the discrete-time evolution of the density ρ.

[1] A map $T : X \to X$ is non-singular if $\ell(A) = 0 \Rightarrow \ell(T^{-1}A) = 0$ (positive volume cannot be created out of zero volume by pullback with T).

2.2 Local perturbations and mixing

Let us first consider the problem of enhancing mixing; in this case trying to push ρ toward an equilibrium distribution. For simplicity, in each of the model dynamical systems 1.-4. we assume that there is a unique equilibrium density ρ^* satisfying $\mathcal{L}\rho^* = \rho^*$. In the pure advection setting the choice of Banach space $\mathcal{B}$ for the operators $\mathcal{L} : \mathcal{B} \to \mathcal{B}$ is typically very important, if one wishes to have a "spectral gap" between the eigenvalue 1 corresponding to ρ^* and the rest of the spectrum. The ideal case is when the eigenvalue 1 is the only eigenvalue on the unit circle in the complex plane and all other spectral points have magnitude strictly less than 1. This guarantees that any other initial density in the Banach space converges to ρ^* exponentially quickly (at a geometric rate determined by the magnitude of the spectral point closest to, but not on, the unit circle). Because our perturbations are of diffusion-type, these issues do not concern us and we consider our perturbed $\mathcal{L}$ as acting on L^2 functions. The space L^2 is simple to work with and for our diffusive perturbations, we obtain a spectral gap without having to employ more exotic Banach spaces.

We denote our target density ρ^T and our initial density ρ^0; in many situations $\rho^T = \rho^*$. At each step of the optimisation procedure, we seek a perturbation (a stochastic kernel) $k : X \times X \to \mathbb{R}^+$ so that by applying the perturbation, and then evolving the dynamics, we push ρ^0 as close as possible to ρ^T in the L^2 norm. Then, the optimisation proceeds iteratively, taking the outcome of the current step as the initial density for the next optimisation step. The reason that we apply the perturbation first and then evolve the dynamics is that we want the perturbation to take advantage of the knowledge of the (one step) future dynamics. Such a short-term prediction of the dynamics at the level of densities can be constructed directly whenever a numerical model of the system is available. Our basic version of the full optimization problem is:

$$\min_{k(x,y)\in L^2(X\times X)} \quad \left\| \mathcal{L}\left(\int_X \rho^0(x) k(x,y)\, d\ell(x) \right) - \rho^T \right\|_2 \tag{1}$$

$$\text{subject to} \quad \int_X k(x,y)\, d\ell(y) = 1 \quad \forall x \in X \tag{2}$$

$$\int_X \rho^*(x) k(x,y)\, d\ell(x) = \rho^*(y) \quad \forall x \in X \tag{3}$$

$$k(x,y) = 0 \quad \text{if } |x-y| > \epsilon \tag{4}$$

$$0 \le k(x,y) \le U \quad \forall x, y \in X. \tag{5}$$

The constraint (2) and the lower bound of (5) simply say that k indeed represents a stochastic kernel. The upper bound of (5), which may depend on ϵ, guarantees that the operator $Kf(y) := \int_X f(x)k(x,y)\, d\ell(x)$ is compact as an operator on $L^2(X, \ell)$. The constraints (3) insist that the kernel k preserve the equilibrium density of the dynamics. Even if the target density ρ^T is not the equilibrium density ρ^*, it is a reasonable condition that the diffusion does not alter the dynamics too much; for example in incompressible volume-preserving fluid flow, one should insist that the diffusion does not violate volume-preservation. The

constraints (4) restrict the spatial range of the perturbations to a maximum distance of ϵ, ensuring the perturbations are local or "small". Note that the volume of the support of $k(x,\cdot)$ is at least $1/U$. The least-radius perturbation corresponding to this volume is the radius of a unit ball with volume $1/U$, which is $r_U = \Gamma(d/2+1)^{1/d}/(\sqrt{\pi}U^{1/d})$, where Γ is the standard gamma function. Thus, for a given U, one should choose ϵ large enough so that $r_U \leq \epsilon$; a "most local" or "minimal" perturbation would be

$$k(x,y) = \mathbf{1}_{B_{r_U}(x)}(y). \tag{6}$$

It is relatively easy to introduce further constraints that are linear in k.
Bounding the kernel: One can easily place lower bounds on the kernel:

$$0 \leq L \leq k(x,y) \leq U < \infty \text{ for all } x,y \in X \text{ such that } |x-y| \leq \epsilon. \tag{7}$$

Bounding the variance of k: For each fixed x, one can limit the "spread" induced by $k(x,y)$ in the form of a variance constraint:

$$\int_X |x-y|^2 k(x,y)\ d\ell(y) \leq V \text{ for all } x \in X, \tag{8}$$

where $V \geq 0$ is an upper bound on the variance of the perturbation.
Perturbation budget: When the kernel is bounded above by U, one can talk about the cost of perturbations beyond the "minimal" perturbation (6). To include a total perturbation budget $0 \leq B \leq 1$ where $B = 1$ represents complete freedom, we can use the constraint

$$(1/\ell(X)) \int_X \int_{X \setminus B_{r_U}(x)} k(x,y)\ d\ell(y)\ d\ell(x) \leq B. \tag{9}$$

Thus, all perturbations lying outside the support of the minimal perturbation incur a cost. One could also add a distance term $|x-y|$ as a coefficient to $k(x,y)$ to charge according to distance perturbed, updating the value of B accordingly.
Uniform diffusion: When k is bounded above by U, to insist on a uniform diffusion (of variable intensity) beyond the support of the minimal perturbation, one can include the constraints

$$k(x,y) = k(x,z) \text{ for all } y,z \in B_\epsilon(x) \setminus B_{r_U}(x), \tag{10}$$

(uniform on an ϵ-ball centred at x outside the r_U-ball centred at x).

We note that the objective (1) is a convex quadratic function of k, and the constraints are all linear functions of k, placing the optimisation problem in a convenient class. Nevertheless the problem is infinite-dimensional, so to solve it numerically, one could either "optimize then discretize", namely write down a (e.g. gradient-based) scheme to solve the infinite-dimensional problem and then discretize that scheme, or "discretize then optimize", namely, discretize the problem and then use finite-dimensional solution methods. Because there are natural ways to discretise the action of $\mathcal{L}_T$, we take the latter approach and solve a finite-dimensional version of (1)–(5).

3 Discretising the dynamics and the optimisation problem

3.1 Discretised dynamics

We wish to have a finite-dimensional approximation of the evolution operator $\mathcal{L}$. A common approach is to perform a Galerkin projection on $\mathcal{L}$ with respect to some finite-dimensional approximating basis for $L^2(X)$. When the dynamics is governed by a purely advective map T, a common choice of basis is the set of characteristic functions $\{\mathbf{1}_{A_1}, \ldots, \mathbf{1}_{A_n}\}$, where $\{A_1, \ldots, A_n\}$ partition X. Typically the sets A_j, $j = 1, \ldots, n$ are formed via a fine mesh on X, so that they are often boxes of small diameter. The associated projection $\pi_n : L^2 \to \text{span}\{\mathbf{1}_{A_1}, \ldots, \mathbf{1}_{A_n}\}$ is defined by taking expectations on boxes: $\pi_n f = \sum_{i=1}^n (1/\ell(A_i)) \left(\int_{A_i} f \, d\ell \right) \mathbf{1}_{A_i}$. The matrix representation of $\pi_n \mathcal{L}$ restricted to $\text{span}\{\mathbf{1}_{A_1}, \ldots, \mathbf{1}_{A_n}\}$ is given by $Q_{ij} = \ell(A_i \cap T^{-1}A_j)/\ell(A_j)$. In the discrete setting, it is more convenient to evolve probability measures rather than density functions, so we use the matrix $P_{ij} = Q_{ij}\ell(A_j)/\ell(A_i)$. We remark that the matrix P is sparse, as the only nonzero elements in the i^{th} row are those indices j for which $T(A_i)$ intersects A_j. This approach is known as Ulam's method [Ula60]. The volumes of intersection $\ell(A_i \cap T^{-1}A_j)$ can be easily estimated numerically by sampling a large number of test points uniformly on A_i, applying T, and counting how many land in A_j. For example, let $z_{i,1}, \ldots, z_{i,r}$, $r = 1, \ldots, R$ denote uniformly distributed test points in A_i, $i = 1, \ldots, n$. We estimate $P_{ij} \approx \#\{r : z_{i,r} \in A_i \text{ and } Tz_{i,r} \in A_j\}/R$. Because $\|\pi_n f - f\|_1 \to 0$ as $n \to \infty$ for each $f \in L^1(X)$, one has that $\pi_n \mathcal{L} f \to \mathcal{L} f$ strongly in $L^1(X)$. In the case of an advection-diffusion equation, one can compute sample trajectories using Euler-Maruyama integration or similar techniques [KP92].

3.2 A finite-dimensional optimisation problem

We now describe a discrete version of (1)–(5) using the finite-dimensional basis $\text{span}\{\mathbf{1}_{A_1}, \ldots, \mathbf{1}_{A_n}\}$. In the formulation below, ϱ_i^*, $i = 1, \ldots, n$ is an estimate of $\int_{A_i} \rho^* \, d\ell$; because ρ^* is a density function, ϱ^* is a probability vector. We similarly define the vectors ϱ^0, ϱ^T, approximating continuous densities ρ^0, ρ^T. The array κ_{ij} estimates $(1/\ell(A_i)) \int_{A_i \times A_j} k(x, y) \, d(\ell \times \ell)$, $i, j = 1, \ldots, n$; by (2) one has $\sum_{j=1}^n \kappa_{ij} = 1$ for all $i = 1, \ldots, n$. Define $\mathcal{B}_i^\epsilon = \{i \in \{1, \ldots, n\} : |c_i - c_j| < \epsilon \text{ where } c_i \text{ is the centroid of } A_i\}$; this is the set of boxes reachable by an ϵ perturbation from A_i.

Thus, the finite-dimensional version of (1)–(5) becomes:

$$\min_{\kappa_{ij}} \quad \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \left(\sum_{m=1}^{n} \varrho_m^0 \kappa_{mi} \right) P_{ij} - \varrho_j^T \right)^2 \tag{11}$$

$$\text{subject to} \quad \sum_{j=1}^{n} \kappa_{ij} = 1 \quad \forall i = 1, \ldots, n \tag{12}$$

$$\sum_{i=1}^{n} \varrho_i^* \kappa_{ij} = \varrho_j^* \quad \forall j = 1, \ldots, n \tag{13}$$

$$\kappa_{ij} = 0 \quad i = 1, \ldots, n, j \notin \mathcal{B}_i^\epsilon \tag{14}$$

$$\kappa_{ij} \geq 0 \quad \forall i, j = 1, \ldots, n. \tag{15}$$

The objective (11) and constraints (12)–(15) are completely analogous to (1)–(5), respectively. The additional constraints (7)–(10) discussed in Section 2 have the following discrete forms.

Bounding κ: One can include the upper and/or lower bounds

$$0 \leq \ell(A_j)L \leq \kappa_{ij} \leq \ell(A_j)U < \infty \text{ for all } i, j = 1, \ldots, n. \tag{16}$$

By (12) and (15) we see that $\kappa_{ij} \leq 1$ for all $i, j = 1, \ldots, n$, thus one already has an effective upper bound of $U = 1/\ell(A_j)$ for κ_{ij}. This corresponds to the perturbation κ placing all mass in the set A_j, so that the density on A_j is $1/\ell(A_j)$.

Bounding the variance of κ: For each fixed i, one can limit the "spread" induced by κ_{ij} in the form of a variance constraint:

$$\sum_{j=1}^{n} |c_i - c_j|^2 \kappa_{ij} \leq V \text{ for all } 1 \leq i \leq n, \tag{17}$$

where $V \geq 0$ is an upper bound on the variance of the perturbation.

Perturbation budget: To include a total perturbation budget $0 \leq B \leq 1$ where $B = 1$ represents complete freedom, we can use the constraint

$$(1/\ell(X)) \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \ell(A_i)\kappa_{ij} \leq B. \tag{18}$$

One could also weight κ_{ij} by some function of $|c_i - c_j|$ to reflect higher cost for larger perturbations, adjusting B accordingly.

Uniform diffusion: For each box A_i, to insist on uniform diffusion (of varying intensity) centred at c_i and of radius approximately ϵ, one can include the constraints

$$\kappa_{ij}/\ell(A_j) = \kappa_{im}/\ell(A_m) \text{ for all } j, m \in \mathcal{B}_i^\epsilon \setminus \{i\}. \tag{19}$$

Restricted perturbation sites: For physical reasons it may be that one can only apply perturbations at specific spatial sites. Let $\mathcal{S} \subset \{1, \dots, n\}$ be the index set of allowed sites. To restrict to these sites, we add the constraints

$$\kappa_{ij} = 0, \text{ for all } i \in \{1, \dots, n\} \setminus \mathcal{S}, j \in \mathcal{B}_i^\epsilon. \tag{20}$$

If one can only use a maximum of $W < |\mathcal{S}|$ of the sites in $\mathcal{S}$, we additionally introduce binary variables $w_i \in \{0, 1\}$, $i \in \mathcal{S}$ with $w_i = 1$ representing usage of site i and non-use otherwise. We also require

$$\kappa_{ij} \leq w_i, \text{ for all } i \in \mathcal{S}, j \in \mathcal{B}_i^\epsilon. \tag{21}$$

If we have a fine grid, then n is large, however by constraint (14) the number of variables is approximately nM_ϵ, where M_ϵ is the average number of boxes within distance epsilon from a given box. The number of constraints is of order n. However, our optimisation problem is a finite-dimensional convex quadratic program with linear constraints, which can be solved efficiently with e.g. barrier methods. In the numerical experiments reported in the following sections, we used FICO Xpress Optimizer (version 7.5) with default settings on a desktop with Intel Core i7 (3.4 GHz) processor and 16 GB of RAM. The sparse matrix P was passed to Xpress in sparse format for memory efficiency.

4 Mixing in one dimension

We first consider the situation of deterministic invertible dynamics targetting the equilibrium distribution on the unit interval or the circle with unit circumference. We focus on three case studies: the identity transformation, a periodic interval exchange, and a weak mixing interval exchange. Each of these transformations is length-preserving and by a straightforward calculation[2], the L^2 norm of the difference between a non-equilibrium density and the equilibrium density is invariant under the dynamics: $\|\mathcal{L}_T f - \rho^*\|_{L^2(\ell)} = \|f - \rho^*\|_{L^2(\ell)}$. The addition of diffusion enables L^2-mixing and the rate of this mixing heavily depends on the deterministic dynamics.

We investigate the two interval exchanges studied in [Stu12], and compare the Gaussian diffusion used by [ANK02, Stu12] with uniform diffusion on an ϵ-ball, and with our L^2-optimised diffusion. Our results are quantified in both the L^2 norm and a mix-norm. We demonstrate that significant mixing speedups can be obtained by optimising the diffusion process.

4.1 Mixing by diffusion only

In order to better understand what our optimal L^2 perturbations are doing, we set the advective part of the dynamics to the identity transformation and aim to reduce the L^2

[2] $\|\mathcal{L}_T f - \rho^*\|^2_{L^2(\ell)} = \int_X |\mathcal{L}_T(f - \rho^*)|^2 \, d\ell = \int_X \left|f \circ T^{-1} - \rho^* \circ T^{-1}\right|^2 \, d\ell = \int_X |f - \rho^*|^2 \, d\ell = \|f - \rho^*\|^2_{L^2(\ell)}$, since ℓ is preserved by T.

error between an initial density

$$\rho^0(x) = \begin{cases} 2, & 0 \le x < 1/2 \\ 0, & 1/2 \le x < 1 \end{cases} \tag{22}$$

and the equilibrium density $\rho^* \equiv 1$, by applying diffusion alone.

We select a value of ϵ based on experiments of Sturman [Stu12], which used Gaussian diffusion with mean zero and variance $V = 2 \times 10^{-4}$. This is matched by the variance of a uniform noise kernel supported on $[-\epsilon, \epsilon]$, with $\epsilon = \sqrt{6} \cdot 10^{-2}$. We compare the evolution of ρ^0 under three different diffusion operators. Firstly, uniform diffusion: $\mathcal{D}^u_\epsilon f(y) := \frac{1}{2\epsilon} \int_X f(x) \mathbf{1}_{[-\epsilon,\epsilon]}(y-x)\, d\ell(x)$. Secondly, Gaussian diffusion: $\mathcal{D}^g_\epsilon f(y) := \frac{1}{\sqrt{2\pi V}} \int_X f(x) \exp(-(x-y)^2/2V)\, d\ell(x)$. Thirdly, iterative application of the diffusion operator generated by solutions of (11)–(15).

We partition the circle of unit circumference $X = S^1$ into 4096 equal subintervals. The uniform diffusion operator $\mathcal{D}^u_\epsilon$ is approximated by the matrix

$$D^u_{ij} = \begin{cases} \frac{1}{\#\mathcal{B}^\epsilon_i} = \frac{1}{200} & \text{if } j \in B^\epsilon_i \\ 0 & \text{otherwise} \end{cases}, \tag{23}$$

and the Gaussian diffusion operator $\mathcal{D}^g_\epsilon$ by the matrix

$$D^g_{ij} = \frac{1}{\sqrt{2\pi V}} \int_{|c_j - c_i| - r_j}^{|c_j - c_i| + r_j} e^{-x^2/2V}\, dx, \tag{24}$$

where c_j is the centre of the j-th subinterval, r_j its radius and $|c_j - c_i|$ denotes distance along the circle. The above integral is approximated in Matlab using the `normcdf` subroutine. We compute vectors $\varrho^{t+1,u} = \varrho^{t,u} D^u_\epsilon$, $\varrho^{t+1,g} = \varrho^{t,g} D^g_\epsilon$, $t = 0, \ldots, 29$, representing the evolution of the initial density $\varrho^0 = \varrho^{0,u} = \varrho^{0,g}$ under the uniform and Gaussian diffusion protocols. The optimised operator is simply multiplication by the optimal 4096×4096 perturbation matrix κ^*. We obtain $\varrho^1 = \varrho^0 \kappa^*$, where κ^* is the optimum of (11)–(15), and similarly compute $\varrho^t, t = 2, \ldots, 30$. Figure 1 shows that over 30 time steps, the uniform and Gaussian diffusion protocols approach equilibrium at an essentially identical rate, and that by optimising the diffusion locally, we can achieve a faster rate of approach to equilibrium. To illustrate the optimised protocol in more detail, we focus on the first application of κ^* at time $t = 0$. Figure 2 (upper) is an intensity plot of the 4096×4096 perturbation matrix κ^* resulting from an application of (11)–(15) with ρ^0 as in (22). The width of the diagonal band is about 200 boxes, as our choice of ϵ corresponds to 100.33 boxes of width 1/4096; that is, $\mathcal{B}^\epsilon_i = \{i - 100, \ldots, i - 1, \ldots, i + 1, \ldots, i + 100\}$ (modulo 4096). The inset boxes labelled (a) and (b) show zooms of the intensity of κ^* in the vicinity of the discontinuities of ρ^0. The optimal perturbation strategy in the neighbourhood of the discontinuity is approximately deterministic. At (b), the perturbation on a set $A_i \subset [1/2 - \epsilon, 1/2]$ is roughly to place 1/2 of its mass in the set A_{i+100}; there is a matching reverse perturbation with the set A_{i+100} placing half of its mass in the set A_i. After such perturbations, the value of $\varrho^0 \kappa^*$ is approximately 1 (the target value of the equilibrium density ϱ^*) in the interval $[1/2 - \epsilon, 1/2 + \epsilon]$. At (a)

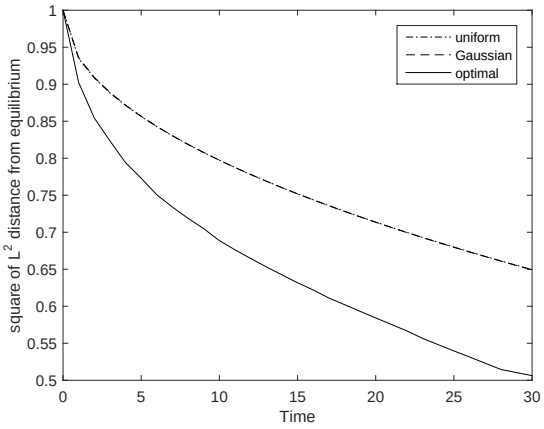


Figure 1: Plot of $\|\varrho^{t,u} - \varrho^*\|_{L^2}^2$, $\|\varrho^{t,g} - \varrho^*\|_{L^2}^2$, $\|\varrho^t - \varrho^*\|_{L^2}^2$, vs. t, for advective dynamics given by the identity transformation. The curves for $\|\varrho^{t,u} - \varrho^*\|_{L^2}^2$ and $\|\varrho^{t,g} - \varrho^*\|_{L^2}^2$ are almost identical.

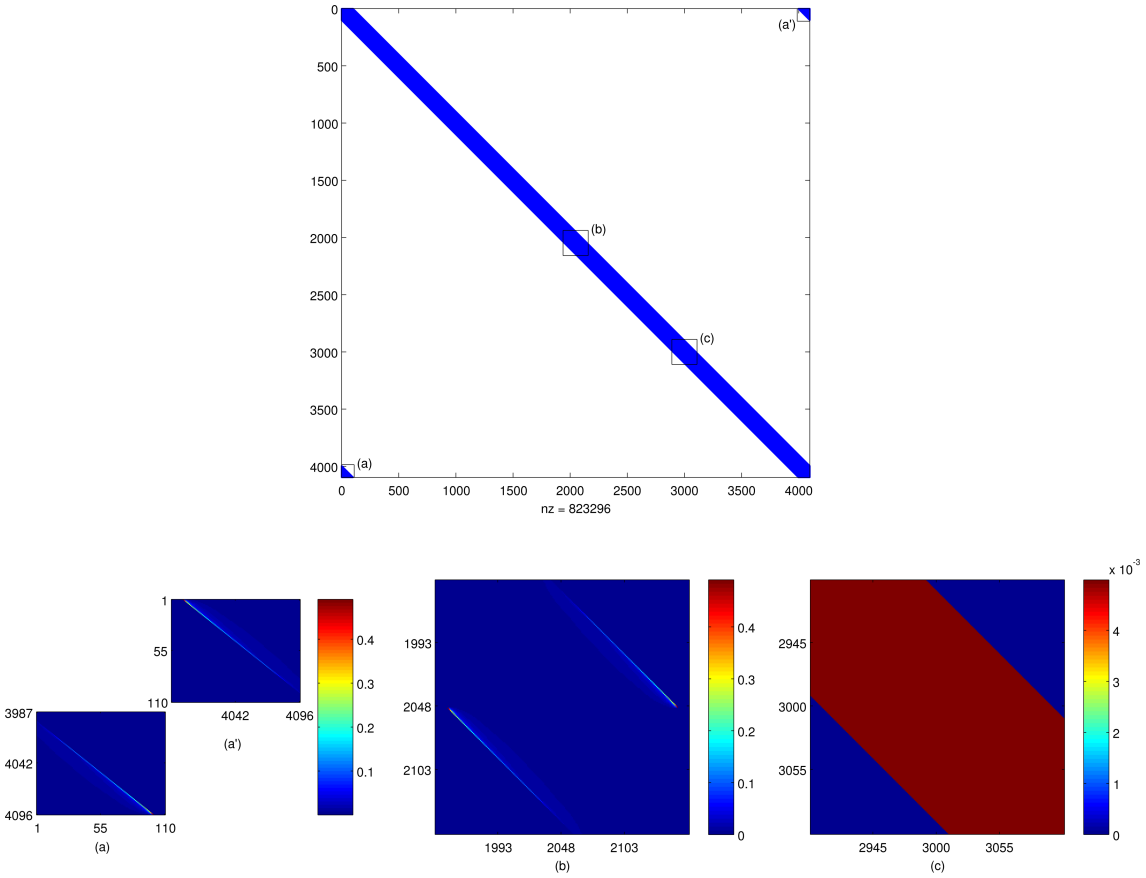
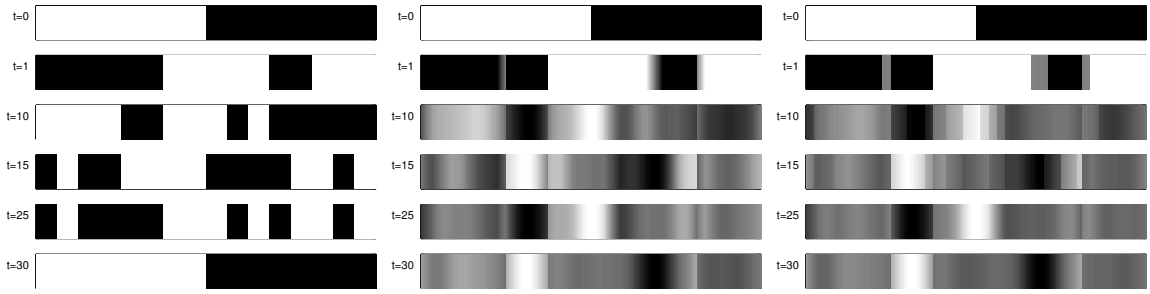


Figure 2: Upper: Support of the optimal perturbation matrix κ^* at time $t = 0$. There are 823296 nonzero entries in the 4096×4096 matrix (4.91% sparsity). Lower: (a),(b),(c) show numerical values of κ^* and are zooms of the corresponding insets in the upper figure.

similar perturbations occur in the interval $[-\epsilon, \epsilon]$. Away from the ϵ-neighbourhoods of the two discontinuities in ρ^0 at $x = 0, 1/2$, perturbations have no effect because ρ^0 is constant. The inset (c) shows a uniform perturbation, however, this choice has no effect on the objective function.

4.2 Periodic interval exchange dynamics

We now introduce advective dynamics, and firstly consider the periodic interval exchange based on intervals $[0, 3/16), [3/16, 10/16), [10/16, 12/16), [12/16, 1)$, which when labelled with the integers 1,2,3,4, respectively, undergo the permutation $(1, 2, 3, 4) \mapsto (4, 3, 2, 1)$. An interval exchange is a length preserving map, which simply interchanges the prescribed intervals under the given permutation, but preserves order and distance within each of them. This interval exchange is periodic, returning to the identity map after 30 iterations; see Figure 3(a), which shows the effect of the periodic interval exchange on the initial density ρ^0, defined on equation (22).



(a) Periodic interval exchange map

(b) Periodic interval exchange map with uniform diffusion

(c) Periodic interval exchange map with optimised diffusion

Figure 3: Evolution of the initial density $\rho^0 = 2\mathbf{1}_{[1/2,1]}$ at $t = 0$, under different diffusion protocols at times $t = 0, 1, 10, 15, 25, 30$. A density of value 2 is indicated in black and a density of 0 is indicated in white.

We now introduce diffusion and consider the evolution of ρ^0 under repeated application of $\mathcal{L}_T\mathcal{D}_\epsilon^u$ and $\mathcal{L}_T\mathcal{D}_\epsilon^G$, where $\mathcal{L}_T$ is the Perron-Frobenius operator for the interval exchange. These results are compared to optimised diffusion; see Figure 4. We use the same numerical and parameter setup as the previous section; the only change is the introduction of $\mathcal{L}_T$. In comparison to Figure 1, clearly the combination of advection and diffusion increases mixing rates when compared to diffusion alone, confirming results by [ANK02, Stu12]. Additionally we see that the optimised diffusion significantly improves the mixing rate over fixed diffusion protocols such as uniform diffusion or Gaussian diffusion.

Figure 4 shows the distance from equilibrium in a Sobolev norm, which is often used to quantify mixing for purely advective dynamics. We use the negative index Sobolev space H^{-1}, which is dual to the Sobolev space H^1 via the standard L^2 inner product. The latter consists of functions which are square integrable and have square integrable first derivatives.

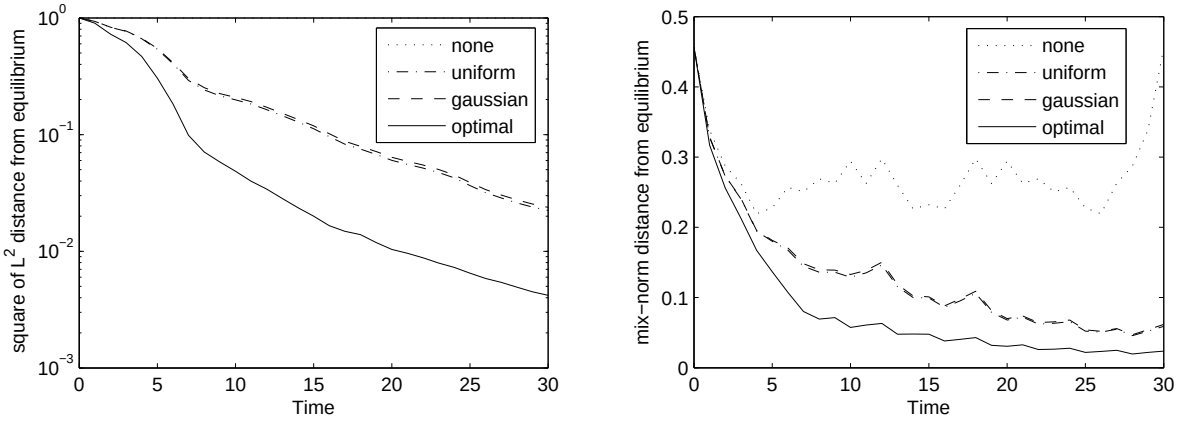


Figure 4: Plot of distance from equilibrium versus time for a periodic interval exchange with uniform, Gaussian, and optimised diffusion protocols with respect to the $\|\cdot\|_{L^2}^2$ norm (left) and $\|\cdot\|_{H^{-1}}^2$ (right).

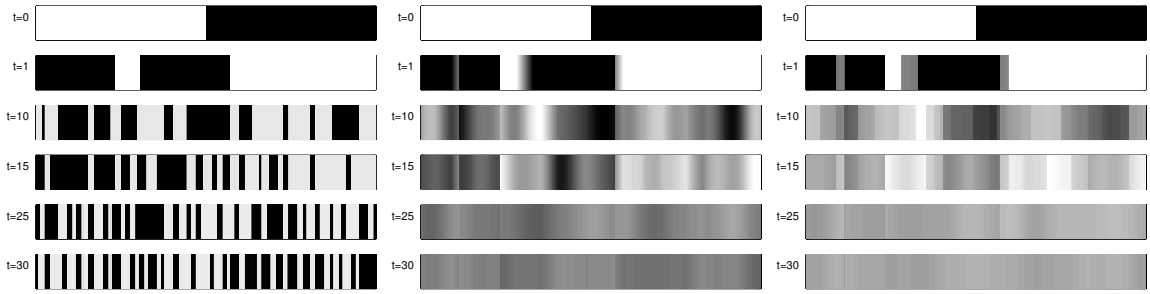
Thus H^{-1} contains L^2 and also contains functions which are much more irregular, enabling one to make sense of limits of infinite-time pushforwards of integrable densities ρ. The norm on H^{-1}, represented by Fourier coefficients[3] of ρ, is $\|\rho\|_{H^{-1}}^2 = \sum_{k=-\infty}^{\infty}(1+k^2)^{-1}|\hat{\rho}_k|^2$. The conclusions regarding mixing efficiency hold just as strongly when using the above mix-norm.

4.3 Weak mixing interval exchange dynamics

In this third one-dimensional case study, we consider a *weak mixing* interval exchange transformation. A length preserving map T is called weak mixing if $\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}|\ell(A\cap T^{-k}B)-\ell(A)\ell(B)|=0$. In the context of IETs, the weakly mixing property is typical [AF07]. However, for example if the lengths of all intervals are rational (as in the case of the periodic interval exchange in section 4.2), the resulting map is never weak mixing.

We set T to be the weak mixing interval exchange on 4 intervals identified in [SU05, Appendix A.1.3], where the lengths of the intervals permuted by T are based on roots of a quartic. To four decimal places, the interval lengths are $0.4276, 0.3382, 0.1196, 0.1144$, and the IET dynamics is defined by the permutation $(1,2,3,4)\mapsto(4,3,2,1)$. Figure 5 shows the effect of this weak mixing interval exchange on the initial density ρ^0. We repeat the experiments for the periodic interval exchange for this weak mixing interval exchange, and the rate of approach to equilibrium in the L^2 norm and mix-norm are shown in Figure 6.

[3]In our numerical experiments, ρ is represented by its values on 4096 equispaced grid points on the unit circle, and we use the fast Fourier transform to compute the values of $\hat{\rho}_k$, $k=0,\ldots,4096$. We adjust these values to obtain $\hat{\rho'}_k = \hat{\rho}_k\cdot(\sin(\pi kN)/(\pi kN))^2$, so as to yield the exact (continuous transform) Fourier coefficients of the unique function ρ' which linear interpolates the 4096 values of ρ. The values $\hat{\rho}_k$ are obtained via periodicity of the discrete Fourier transform: $\hat{\rho}_{-k}=\hat{\rho}_{N-k}$. We thus in practice compute $\|\rho'\|_{H^{-1}} = \sum_{k=-N+1}^{N-1}(1+k^2)^{-1}|\hat{\rho'}_k|^2$.



(a) Weak mixing interval exchange map only

(b) Weak mixing interval exchange map with uniform diffusion

(c) Weak mixing interval exchange map with optimised diffusion

Figure 5: Evolution of the initial density $\rho^0 = 2\mathbf{1}_{[1/2,1]}$ at $t = 0$, under different diffusion protocols at times $t = 0, 1, 10, 15, 25, 30$. A density of value 2 is indicated in black and a density of 0 is indicated in white.

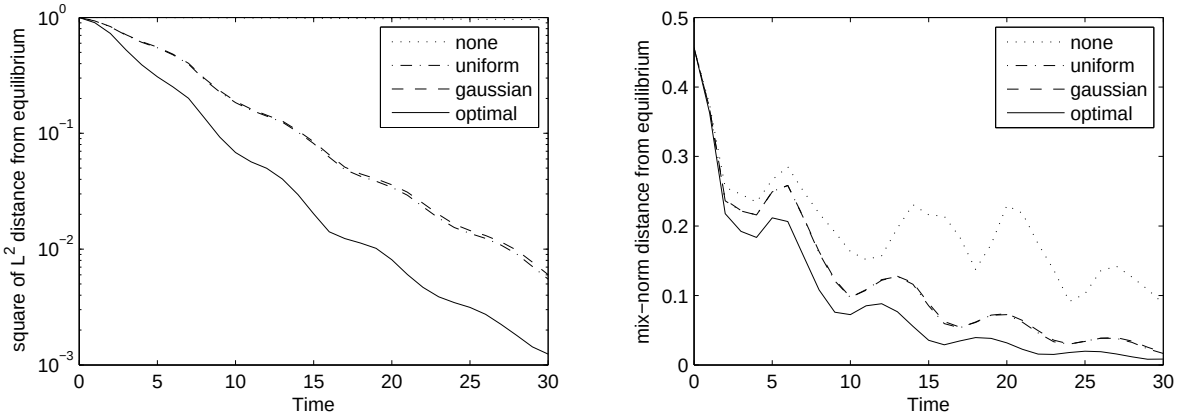


Figure 6: Plot of distance from equilibrium versus time for a weak mixing interval exchange with uniform, Gaussian, and optimised diffusion protocols with respect the $\|\cdot\|_{L^2}^2$ norm (left) and $\|\cdot\|_{H^{-1}}^2$ (right). Note that in the left figure, the "none" curve takes the constant value 1.

5 Mixing in two dimensions

On the rectangular domain $X = [0,1] \times [0,2]$ we consider the time-dependent system of differential equations [SLM05]

$$\begin{aligned} \dot{x} &= -\pi A \sin(\pi f(x,t)) \cos(\pi y) \\ \dot{y} &= \pi A \cos(\pi f(x,t)) \sin(\pi y) \frac{df}{dx}(x,t), \end{aligned} \tag{25}$$

where $f(x,t) = \beta \sin(\omega t) x^2 + (1 - 2\beta \sin(\omega t))x$. We fix parameter values $A = 0.25$, $\beta = 0.25$ and $\omega = 2\pi$ and obtain a flow of period $\tau = 1$. For $\beta = 0$, (25) models two rotating gyres of incompressible fluid, centred at the equilibria $(0.5, 0.5)$ and $(1.5, 0.5)$ respectively. For small $\beta > 0$ the vertical separatrix between these gyres oscillates in the x-direction with a period of $2\pi/\omega$. The stable and unstable manifolds of hyperbolic periodic orbits on the upper and lower boundaries intersect transversally, forming so-called "lobes" [RKLW90, RKW90, Wig92] and giving rise to chaotic dynamics. In addition to chaotic motion there are also regions of regular

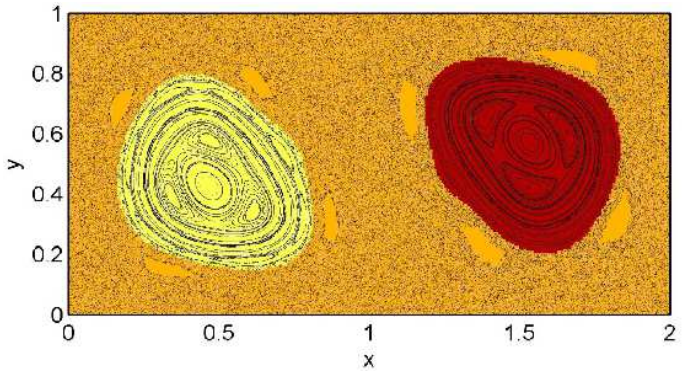


Figure 7: Several long orbits of the time-1 map of the double-gyre flow ©2009 Froyland & Padberg [FP09].

motion. Figure 7 shows several long orbits of the time-1 flow map (one period) of the time-dependent ODE (25). The oscillating separatrix and lobe dynamics allows slow mixing between the left and right halves of the domain, while the regular regions preclude complete phase space mixing. This obstruction to fast mixing is intimately related to the structures appearing in the dominant eigenmode of numerical approximations of the transfer operator for the time-1 flow map, shown in Figure 8.

5.1 Enhancing mixing

For the computations we approximate $\mathcal{L}$ for the time-τ flow map of (25) using Ulam's method based on a partition of the domain into $n = 2^{15}$ equally-sized squares. We form the matrix P by numerically integrating 400 uniformly distributed test points in each box from time $t = 0$ to time $t = 1$ using a fourth-order Runge-Kutta integrator. In all of the following experiments

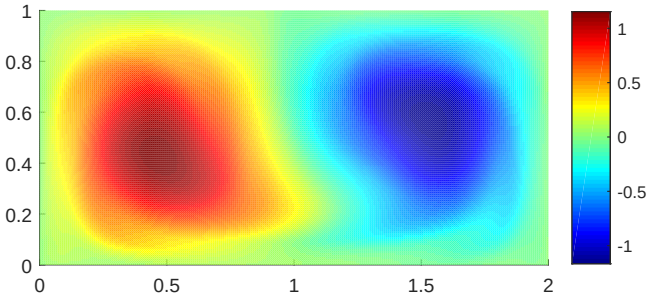


Figure 8: Dominant nontrivial eigenmode of the time-1 flow map of (25), computed as the left eigenvector of Ulam matrix P corresponding to the second largest (in magnitude) numerical eigenvalue.

we set $\epsilon = 0.05$. The associated perturbation sets $\mathcal{B}_i^\epsilon$, $i = 1, \ldots, n$, each contain 137 boxes, except near the boundary of X, where they contain fewer boxes to avoid perturbations that leave the domain. Each set $\mathcal{B}_i^\epsilon$ is about 0.4% of the area of the domain and therefore represents a small perturbation. As in the one-dimensional case, we use a uniform diffusion operator $\mathcal{D}_\epsilon^u f(y) = \int_X f(x)k(x,y)\, d\ell(x)$, where $k(x,y) = \mathbf{1}_{B_\epsilon(y)}(x)$, to compare the uniform diffusion protocol with our optimised diffusion protocol. Numerically, we form a discretised version of $k(x,y)$:

$$\kappa_{ij}^u = \begin{cases} 1/\#\mathcal{B}_i^\epsilon, & \text{for } j \in \mathcal{B}_i^\epsilon; \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

We set our initial density ϱ_0 to a uniform density on a disk of radius $1/4$, centred at $(1/2, 1/4)$. We chose this initial density to partly intersect both regions of regular dynamics and chaotic dynamics; see Figure 7. The dashed curve in Figure 9 shows the evolution of the squared L^2 distance between the equilibrium density $\varrho^* \equiv 1/2$ and $\varrho^{t,u} = \varrho^0(\kappa^u P)^t$, $t = 0, \ldots, 30$. We

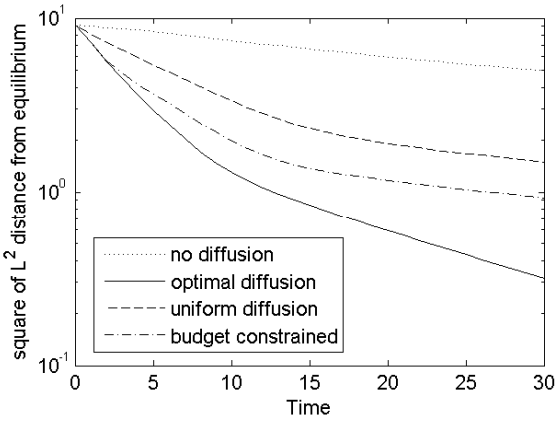


Figure 9: Plot of $\|\varrho^{t,u} - \varrho^*\|_{L^2}^2$ and $\|\varrho^t - \varrho^*\|_{L^2}^2$, vs. t for the double-gyre system (25).

remark that the transfer operator $\mathcal{L}$ for the flow map is invertible, but P is not an invertible

matrix. The Ulam scheme thus produces small numerical diffusion, shown by the dotted curve in Figure 9.

We now run our optimisation procedure for 30 iterations. The solid curve in Figure 9 shows the evolution of the squared L^2 distance between ϱ^* and ϱ^t, where $\varrho^{t+1} = \varrho^t(\kappa^t)^* P$, and $(\kappa^t)^*$ is the solution of (11)–(15), replacing ϱ^0 with ϱ^t, for $t = 0, \ldots, 29$. Our optimisation procedure has a significant impact on the rate of mixing over a uniform diffusion. After 30 steps the L^2 distance from equilibrium using our optimised protocol is five times smaller than the uniform diffusion protocol, despite the fact that both diffusion processes have the same diffusion radius.

In both the solid and dashed curves of Figure 9 we see an initially steep decay of the L^2 distance, followed by a shallower decay. This transition is complete by $t = 11$ for the optimised mixing protocol and at a later time (around $t = 15$ or 16) under the uniform diffusion protocol. This two-phase mixing is likely due to rapid initial spreading of ϱ^t out of the yellow invariant set in Figure 7 to cover $[0,1] \times [0,1]$, followed by a slower exchange between $[0,1] \times [0,1]$ and $[1,2] \times [0,1]$ via diffusion-enhanced lobe dynamics. Figure 10 shows ϱ^{11}; one sees good spreading on the left half of the domain, but relatively little penetration into the right half.

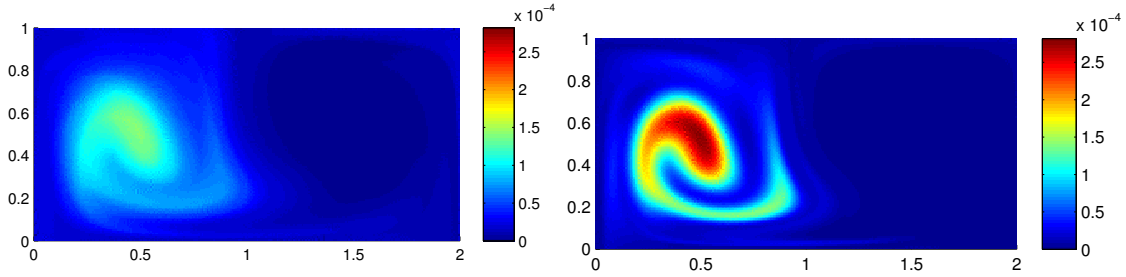


Figure 10: Density ϱ^{11} at time $t = 11$ after 11 iterations of our optimised protocol (left). Density $\varrho^{11,u}$ at time $t = 11$ after 11 iterations of the uniform diffusion protocol (right).

We now interpret the strategy that the optimisation takes at the first iteration. In Figure 11 (upper left) we have plotted arrows that point from the centroid of a box A_i to the mean spatial location of the perturbation described by κ^*_{ij}. In Figure 11 (upper left) the main translational (as opposed to dispersive) effects of the perturbations are concentrated at the boundary of the support of the initial density ϱ^0, and parallel to the gradient of ϱ^0. Mass contained in the support of ϱ^0 and within ϵ of the boundary of the support is pushed out of the support, while "zero" mass in boxes outside the support of ϱ^0 but within ϵ of the boundary is pushed into the support. Both effects are necessary in order to conserve Lebesgue measure and the combined effect is to average the density ϱ^0 in an 2ϵ-annulus containing the boundary of the support of ρ^0. Figure 11 (lower left) shows the result ϱ^1 after the perturbation and advection with P. The same principle is applied in Figure 11 (upper right), where the overall translational motion of the local perturbations is parallel to the gradient of the current density, and in both directions, in order to preserve Lebesgue measure. Note that the distribution of mass in Figure 11 (lower right) reflects the dominant

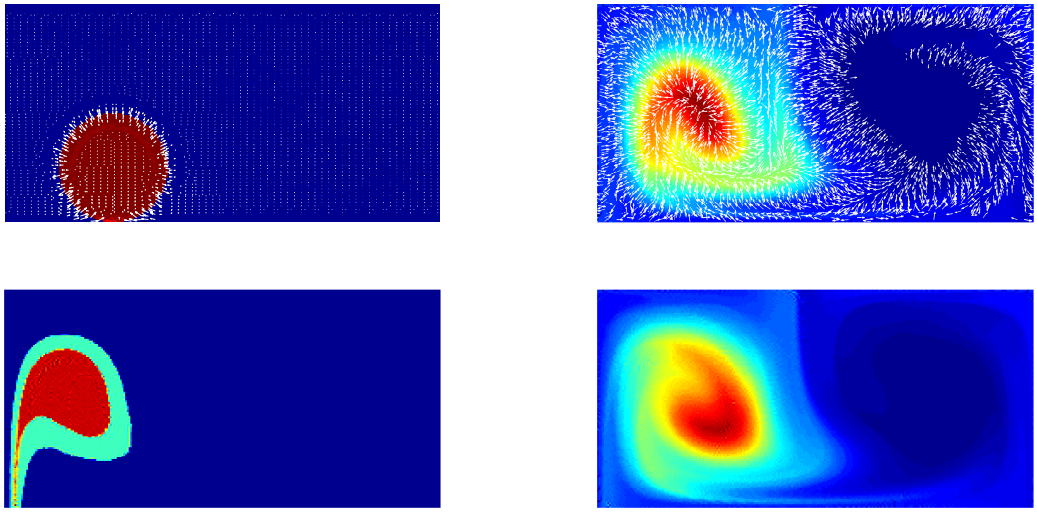


Figure 11: Quiver plot of optimised perturbations (white arrows) overlaid on the initial density ϱ^0 at $t = 0$ (upper left) and the optimised density ϱ^{11} at $t = 11$ (upper right). Optimised densities at the following iteration at times $t = 1$ (lower left) and $t = 12$ (lower right). Note that the color scale has been rescaled for each figure to show extra detail.

obstacles to mixing shown in the signed distribution of the strange eigenmode in Figure 8.

5.2 Enhancing mixing with a budget constraint

Our restriction to local perturbations of maximum diameter ϵ reflects a limited influence that one may be subject to in applications. Until now, we have not imposed an overall budget on the effort expended in the perturbations, as described in (18), and we now explore such a restriction. In the following experiments we set B in (18) to be 0.1. Firstly, Figure 9 shows the resulting approach to equilibrium in the L^2 norm (dash-dot curve). By imposing this restriction on our optimised diffusion protocol, we slow down the rate at which we approach equilibrium, but Figure 9 shows that even with only 10% of the total perturbation effort of the uniform diffusion protocol, we still obtain a faster approach to equilibrium than uniform diffusion provides.

One can measure whether the optimised perturbations are predominantly translational or dispersive by computing quantities such as the mean displacement of the perturbation on A_i: $s_i = |c_i - \bar{c}_i|$, where $\bar{c}_i = \sum_{j=1} \kappa^*_{ij} c_j$, and the entropy of the perturbation on A_i: $H_i = -\sum_{j=1}^{n} \kappa^*_{ij} \log \kappa^*_{ij}$; see Figure 12. The application of the global budget constraint forces the optimiser to be parsimonious about its perturbations. Figure 12 shows that right at the boundary of the support of ϱ^0, there is a peak of translational perturbations (the red ring in the annulus shown in Figure 12 (middle left), exactly corresponding to the blue ring in the same location in Figure 12 (lower left)). The perturbations on this ring have high displacement but low entropy, meaning they are approximately deterministic perturbations.

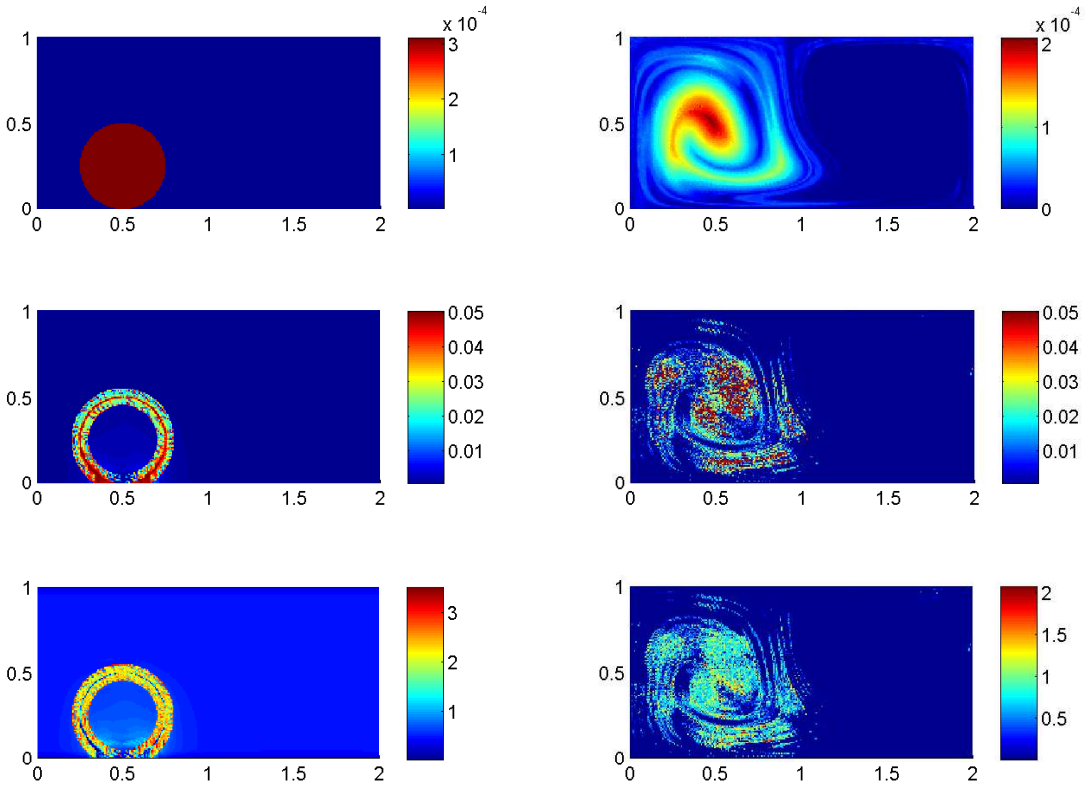


Figure 12: Densities at $t = 0$ (upper left) and $t = 11$ (upper right) when iterated using the optimised diffusion with the budget constraint $B = 0.1$. Mean displacement fields $s_i, i = 1, \ldots, 2^{15}$ at times $t = 0$ (middle left) and $t = 11$ (middle right). Entropy fields $H_i, i = 1, \ldots, 2^{15}$ at times $t = 0$ (lower left) and $t = 11$ (lower right). Note that the color scale has been rescaled for each figure to show extra detail.

The region away from the boundary but still within ϵ of the boundary show a mixture of determinism and stochasticity as the optimiser attempts to mix, but is constrained by area preservation. In Figure 12 (middle right, lower right) we see that the budget constraint focuses the perturbations on areas of high gradient in the density ϱ^{11} (upper right).

5.3 Containing a passive tracer

In this final section, we discuss how one might try to use the previous optimisation methodology to direct (or contain) a passive tracer in a bounded region $S \subset X$. This is more difficult than enhancing mixing toward an equilibrium density because the containment region S may be far from invariant under the advective dynamics, and so we are attempting to fight against advection with a small amount of diffusion. For this reason, building some future information into the optimisation will be beneficial. We do this by penalising mass in the complement of the containment region S and additionally in the preimages of the comple-

ment of S. The preimage $T^{-k}S^c$ can be estimated numerically by advecting backwards in time $\mathbf{1}_{S^c}$, the indicator function on S^c. If S is a union of boxes, this backward advection can be estimated using the transpose of P (a discrete approximation of the Koopman operator). We approximate S^c by the support of $S_k^c := \mathbf{1}_{S^c}(P^\top)^k$, where here we abuse notation, thinking of $\mathbf{1}_{S^c}$ is the indicator vector on box indices corresponding to S^c. We replace the objective (11) with an objective which penalises mass both in S^c and in the preimages of S^c, as mass in these latter sets will shortly enter S^c itself. We have found experimentally that penalising mass in $S^c, T^{-1}S^c, \ldots, T^{-K}S^c$ works well for modest K, as does a decreasing penalty with increasing $0 \leq k \leq K$ (the further into the future mass sits outside S, the lower the penalty). We denote the penalty coefficients by $\alpha_k, k = 0, \ldots, K$.

$$\min_{\kappa_{ij}} \sum_{k=1}^{K} \alpha_k \left(\sum_{j \in S_k^c} \sum_{i=1}^{n} \left(\sum_{m=1}^{n} \rho_m^0 \kappa_{mi} \right) P_{ij} \right). \tag{27}$$

In the experiment we report, we set $S = [1,2] \times [0,1]$, $K = 3$, and $\alpha_k = 10 \times 2^{-k}$. Figure 13 shows the value of (27) for $\kappa_{ij} = \delta_{ij}$ (no perturbation). We initialise with the density

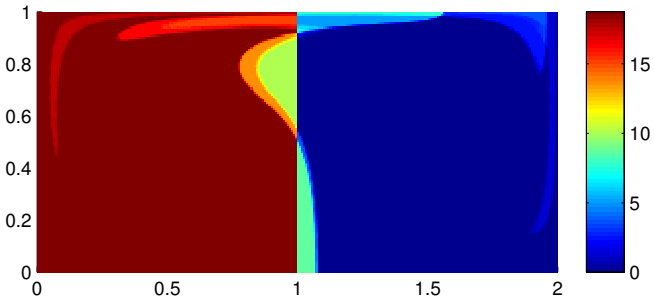


Figure 13: Value of (27) for $\kappa_{ij} = \delta_{ij}$. Note that the highest penalty is applied to points that do not enter S in three steps, while the lowest penalty is applied to points that remain in S for three steps.

shown in Figure 12 (upper left) and iteratively minimise the objective (27), subject to the constraints (2)–(5) for 30 time steps. Figure 14 shows the proportion of mass in S over the 30 time steps, as compared to advection only, the uniform diffusion of Section 5.1, and the optimal mixing protocol of Section 5.1. While the protocol optimised for rapid mixing (solid line) also naturally moves mass into S, we can achieve a faster transfer of mass into S via our containment objective (27) (upper dashed line). Figure 15 shows the results of our containment protocol versus advection after 30 steps.

6 Acknowledgements

GF is supported by an ARC Future Fellowship (FT120100025). CGT's research was supported by an ARC Discovery Project (DP110100068) at UNSW. TMW has been supported

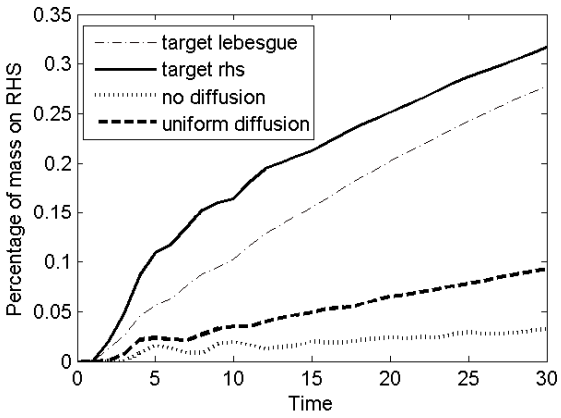


Figure 14: The proportion of mass in the target set S vs. time for various diffusion protocols.

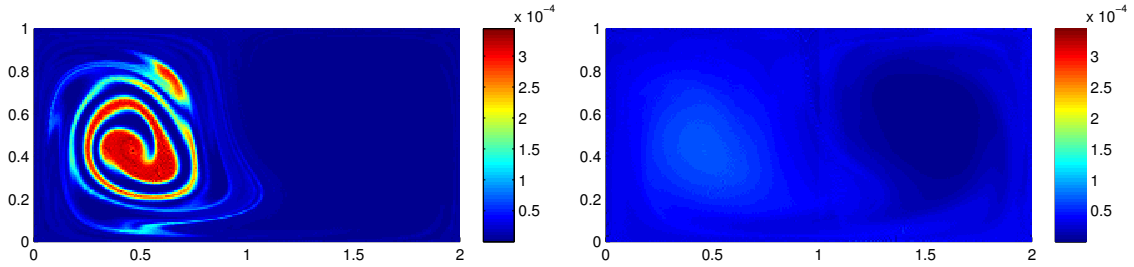


Figure 15: Left: Density after 30 time steps of advection only. Right: Density after 30 iterative applications of the optimised containment objective (27).

by the ARC Centre of Excellence for the Mathematics and Statistics of Complex Systems (MASCOS), an Australian Postgraduate Award, and the UNSW School of Mathematics and Statistics. The authors acknowledge Guoyin Li for useful conversations.

References

[A+14] H. Aref et al. Frontiers of chaotic advection. arXiv:1403.2953, 2014.

[AF07] A. Avila and G. Forni. Weak mixing for interval exchange transformations and translation flows. *Ann. of Math. (2)*, 165(2):637–664, 2007.

[ANK02] P. Ashwin, M. Nicol, and N. Kirkby. Acceleration of one-dimensional mixing by discontinuous mappings. *Phys. A*, 310(3-4):347–363, 2002.

[Bal15] S. Balasuriya. Dynamical systems techniques for enhancing microfluidic mixing. Preprint, 2015.

[BHZ13] N. P. Byott, M. Holland, and Y. Zhang. On the mixing properties of piecewise expanding maps under composition with permutations. *Discrete Contin. Dyn. Syst.*, 33(8):3365–3390, 2013.

[BLZ14] N. P. Byott, C. Lin, and Y. Zhang. On the mixing properties of piecewise expanding maps under composition with permutations, II: Maps of non-constant orientation. arXiv:1405.0481, 2014.

[CAG08] L. Cortelezzi, A. Adrover, and M. Giona. Feasibility, efficiency and transportability of short-horizon optimal mixing protocols. *Journal of Fluid Mechanics*, 597:199–231, 2 2008.

[DJ99] M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. *SIAM J. Numer. Anal.*, 36(2):491–515, 1999.

[FHR+12] G. Froyland, C. Horenkamp, V. Rossi, N. Santitissadeekorn, and A. Sen Gupta. Three-dimensional characterization and tracking of an Agulhas Ring. *Ocean Modelling*, 52–53(0):69 – 75, 2012.

[FLS10] G. Froyland, S. Lloyd, and N. Santitissadeekorn. Coherent sets for nonautonomous dynamical systems. *Physica D*, 239(16):1527–1541, 2010.

[FP09] G. Froyland and K. Padberg. Almost-invariant sets and invariant manifolds—connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D*, 238(16):1507–1523, 2009.

[FPET07] G. Froyland, K. Padberg, M.H. England, and A.M. Treguier. Detection of coherent oceanic structures via transfer operators. *Physical review letters*, 98(22):224503, 2007.

[Fro13] G. Froyland. An analytic framework for identifying finite-time coherent sets in time-dependent dynamical systems. *Physica D*, 250:1–19, 2013.

[FSM10] G. Froyland, N. Santitissadeekorn, and A. Monahan. Transport in time-dependent dynamical systems: Finite-time coherent sets. *Chaos*, 20(4):043116, 2010.

[GDT11] E. Gouillart, O. Dauchot, and J.-L. Thiffeault. Measures of mixing quality in open flows with chaotic advection. *Physics of Fluids*, 23(1):013604, 2011.

[GDTR09] E. Gouillart, O. Dauchot, J.-L. Thiffeault, and S. Roux. Open-flow mixing: Experimental evidence for strange eigenmodes. *Physics of Fluids*, 21(2):023603, 2009.

[GGA12a] O. Gorodetskyi, M. Giona, and P. D. Anderson. Exploiting numerical diffusion to study transport and chaotic mixing for extremely large Péclet values. *EPL (Europhysics Letters)*, 97(1):14002, 2012.

[GGA12b] O. Gorodetskyi, M. Giona, and P. D. Anderson. Spectral analysis of mixing in chaotic flows via the mapping matrix formalism: Inclusion of molecular diffusion and quantitative eigenvalue estimate in the purely convective limit. *Physics of Fluids*, 24(7):073603, 2012.

[GSA12] O. Gorodetskyi, M. F. M. Speetjens, and P. D. Anderson. An efficient approach for eigenmode analysis of transient distributive mixing by the mapping method. *Physics of Fluids*, 24(5):053602, 2012.

[HY00] G. Haller and G. Yuan. Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D: Nonlinear Phenomena*, 147(3–4):352 – 370, 2000.

[KCOL12] M. K. Krotter, I. C. Christov, J. M. Ottino, and Richard M. Lueptow. Cutting and shuffling a line segment: mixing by interval exchange transformations. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 22(12):1230041, 18, 2012.

[KP92] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer Science & Business Media, 1992.

[LH04] W. Liu and G. Haller. Strange eigenmodes and decay of variance in the mixing of diffusive tracers. *Physica D: Nonlinear Phenomena*, 188(1–2):1 – 39, 2004.

[MMP05] G. Mathew, I. Mezić, and L. Petzold. A multiscale measure for mixing. *Phys. D*, 211(1-2):23–46, 2005.

[Pie94] R.T. Pierrehumbert. Tracer microstructure in the large-eddy dominated regime. *Chaos, Solitons & Fractals*, 4(6):1091 – 1110, 1994. Special Issue: Chaos Applied to Fluid Mixing.

[Pro99] A. Provenzale. Transport by coherent barotropic vortices. *Annual Review of Fluid Mechanics*, 31(1):55–93, 1999.

[RKLW90] V. Rom-Kedar, A. Leonard, and S. Wiggins. An analytical study of transport, mixing and chaos in an unsteady vortical flow. *Journal of Fluid Mechanics*, 214:347–394, 1990.

[RKW90] V. Rom-Kedar and S. Wiggins. Transport in two-dimensional maps. *Archive for Rational Mechanics and Analysis*, 109:239–298, 1990.

[SCU+13] C. P. Schlick, I. C. Christov, P. B. Umbanhowar, J. M. Ottino, and R. M. Lueptow. A mapping method for distributive mixing with diffusion: Interplay between chaos and diffusion in time-periodic sine flow. *Physics of Fluids*, 25(5):052102, 2013.

[SKMA08] M. K. Singh, T. G. Kang, H. E. H. Meijer, and P. D. Anderson. The mapping method as a toolbox to analyze, design, and optimize micromixers. *Microfluidics and Nanofluidics*, 5(3):313–325, 2008.

[SLM05] S.C. Shadden, F. Lekien, and J.E. Marsden. Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Phys. D*, 212(3-4):271–304, 2005.

[SOW06] R. Sturman, J. M. Ottino, and S. Wiggins. *The mathematical foundations of mixing*, volume 22 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2006.

[Stu12] R. Sturman. The role of discontinuities in mixing. In Herman J.H. Clercx and Michel F.M. Speetjens, editors, *Advances in Applied Mechanics*, volume 45, pages 51–90. Academic Press, 2012.

[SU05] Y. G. Sinai and C Ulcigrai. Weak mixing in interval exchange transformations of periodic type. *Letters in Mathematical Physics*, 74(2):111–133, 2005.

[SW51] R. S. Spencer and R. M. Wiley. The mixing of very viscous liquids. *Journal of Colloid Science*, 6(2):133 – 145, 1951.

[TDG04] J.-L. Thiffeault, C. R. Doering, and J. D. Gibbon. A bound on mixing efficiency for the advection-diffusion equation. *Fluid Mechanics*, 521:105–114, 2004.

[Thi08] J.-L. Thiffeault. Scalar decay in chaotic mixing. In JeffreyB. Weiss and Antonello Provenzale, editors, *Transport and Mixing in Geophysical Flows*, volume 744 of *Lecture Notes in Physics*, pages 3–36. Springer Berlin Heidelberg, 2008.

[Thi12] J.-L. Thiffeault. Using multiscale norms to quantify mixing and transport. *Nonlinearity*, 25(2):R1, 2012.

[TP08] J.-L. Thiffeault and G. A. Pavliotis. Optimizing the source distribution in fluid mixing. *Physica D: Nonlinear Phenomena*, 237(7):918 – 929, 2008.

[Ula60] S. M. Ulam. *A collection of mathematical problems*. Interscience Tracts in Pure and Applied Mathematics, no. 8. Interscience Publishers, New York-London, 1960.

[Wig92] S. Wiggins. *Chaotic Transport in Dynamical Systems*. Springer-Verlag, New York, NY, 1992.