Designing a Finite-Time Mixer: Optimizing Stirring For Two-Dimensional Maps

by

R. A. Mitchell

B.A., Colorado College, 2012

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Applied Mathematics

2017

This thesis entitled:
Designing a Finite-Time Mixer: Optimizing Stirring For Two-Dimensional Maps
written by R. A. Mitchell
has been approved for the Department of Applied Mathematics

_____

Prof. James D. Meiss

_____

Prof. Keith Julien

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Mitchell, R. A. (Ph.D., Applied Mathematics)

Designing a Finite-Time Mixer: Optimizing Stirring For Two-Dimensional Maps

Thesis directed by Prof. James D. Meiss

Mixing of a passive scalar in a fluid flow results from a two part process in which large gradients are first created by advection and then smoothed by diffusion. We investigate methods of designing efficient stirrers to optimize mixing of a passive scalar in a two-dimensional nonautonomous, incompressible flow over a finite time interval. The flow is modeled by a sequence of area-preserving maps whose parameters change in time, defining a mixing protocol. Stirring efficiency is measured by the mix norm, a negative Sobolev seminorm; its decrease implies creation of fine-scale structure. A Perron-Frobenius operator is used to numerically advect the scalar for three examples: compositions of Chirikov standard maps, of Harper maps, and of blinking vortex maps. In the case of the standard maps, we find that a protocol corresponding to a single vertical shear composed with horizontal shearing at all other steps is nearly optimal. For the Harper maps, we devise a predictive, one-step scheme to choose appropriate fixed point stabilities and to control the Fourier spectrum evolution to obtain a near optimal protocol. For the blinking vortex model, we devise two schemes: A one-step predictive scheme to determine a vortex location, which has modest success in producing an efficient stirring protocol, and a scheme that finds the true optimal choice of vortex positions and directions of rotation given four possible fixed vortex locations. The results from the numerical experiments suggest that an effective stirring protocol must include not only steps devoted to decreasing the mix norm, but also steps devoted to preparing the density profile for future steps of mixing.

Dedication

To Mom and Dad

Acknowledgements

Thank you for the support from NSF grant DMS-1211350. Thank you to my advisor, Prof. Jim Meiss for your guidance, support, and company for the past three years. It's been a pleasure working with you. Thank you to my other committee members, Professors Keith Julien, Juan Restrepo, John Crimaldi, and Roseanna Neupauer. I greatly appreciate your time and thoughtfulness in your consideration of my research. Thank you to Prof. Jim Curry for his genuine interest and his insightful and helpful comments. Thank you to Emily O'Connor for her help and patience, and to all the Applied Math staff for all the hard work they do. Thank you Mom and Dad for your unwavering support at confidence in me. Thank you to all my roommates, friends, and peers who have been with me through this whole process.

Contents

Chapter

1 Introduction **1**

1.1 Finite-Time Mixers 1

1.2 A Measure of Mixing 5

1.3 Two-Dimensional Maps 6

2 Mathematical Formulation **9**

2.1 The Mix Norm 9

2.2 Advection by Two-Dimensional Maps 14

2.3 Compositions of Shears 15

2.3.1 Computing the Perron-Frobenius Operator and the Mix Norm 17

2.4 Blinking Vortices 19

2.4.1 Computing the Perron-Frobenius Operator and the Mix Norm 24

3 Results of Chirikov's Standard Map and the Harper Map **29**

3.1 Chirikov Map 29

3.1.1 Single Step Stirring 29

3.1.2 Constant Amplitude Stirring 31

3.1.3 Varying Amplitudes: Random Optimization Method 32

3.1.4 Sparse Stirring 33

3.1.5 Other Initial Conditions 40

3.2 Harper Map 41
3.2.1 Fixed Points 42
3.2.2 Constant Shear Amplitude 43
3.2.3 Random Optimal Protocol 46
3.2.4 Stepwise Minimization 50
3.2.5 Horizontal or Vertical Shearing 53

4 Results of Blinking Vortex Map **61**
4.1 Roaming Vortex Problem 62
4.2 Four Vortex Problem 70

5 Conclusions **78**

Bibliography **82**

Tables

Table

3.1 Mix norms for the Harper map with initial condition (3.1), energy $E_H = 0.5$, and total steps $T = 12$ for the optimization schemes discussed in §3.2.2-3.2.5. The initial condition has $\|\rho^0\|_{-1} = 0.2505$. The last column gives the percentage of protocols from the random trials of Fig. 3.22 that perform better, i.e., have a smaller mix-norm. 49

3.2 Optimal, mean and standard deviations of the mix-norms at time $T = 12$ for the Harper map from the distributions of Fig. 3.32(a). The initial conditions are normalized such that $\|\rho^0\|_{-1} = 1$. . . . 60

Figures

Figure

1.1 Examples of finite-time mixers in which discrete mechanisms are used to mix fluid flowing through three-dimensional pipe. Plot (a) shows the Kenics Static Mixer [22]. Plot (b) shows the rotated-arc mixer (RAM) [42]. 2

2.1 Decay of H^q norms for $q = -\frac{1}{2}, -1, -\frac{3}{2}, -2, -\frac{5}{2}, -3$ for a particular stirring protocol applied for $t = 1, \ldots, 12$. 13

2.2 Depiction of the shearing maps (2.5) for the Chirikov standard map (2.8) with $a = 0.75$. 16

2.3 Depiction of the shearing maps (2.5) for the Harper map (2.9) with $a = b = 0.75$. 16

2.4 Diagram of (2.25) with $v = 0.4 + i0.5$, $s = 1$, and $t = 4$. The black curve indicates the invariant boundary at the unit circle $|z| = 1$. The blue points mark the vortex and its image. The red curves show the rotation direction of the vortices. The green curves show resulting trajectories for different initial particle positions. The green points mark the initial positions of the particles and the green arrowheads show the final positions at time $t = 4$. 21

2.5 Grids of polar coordinates for $N_r = N_\theta = 32, 64, 128, 256$. 27

2.6 Plot (a) is ρ^0 as described by (2.39). Plot (b) is $\tilde{\rho}^0$, which is (2.33) applied to the coefficients produced by (2.38). Plot (c) is the radial error $\rho_e = \rho^0 - \tilde{\rho}^0$. 28

3.1 The H^{-1} seminorm (3.3) after one step of the standard map for initial state $\rho^0(x, y) = \exp(2\pi i(kx + ly))$ (left panel) and one that has a Cauchy or Gaussian Fourier spectrum (right panel) as a function of the shear strength a_1. 30

3.2 Phase portrait of the autonomous Chirikov standard map with $a_t = a_0 = 0.05$. 31

3.3 Evolution of the density $\rho^t(x, y)$ for $t = 0, 1, 2, 3$, applying the standard map (2.8) with constant amplitude $a_t = 0.05$ with initial condition (3.1) using $N = 301$ grid points. For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/ChirikovConst.mp4`. 32

3.4 Decay of $\|\rho^t\|_{-1}$ for the standard map (2.8) with constant amplitude $(a_0, a_0, \ldots, a_0)$. The different curves are computed with grid sizes N, as shown in the legend. 32

3.5 (a) Histogram of $\|\rho^T\|_{-1}$ for $\sim 2.2 \times 10^6$ trials of the RO algorithm applied to the standard map (2.8) with $E_C = T + 0.03$, $T = 12$, and initial condition (3.1). The green curve indicates the best fit normal distribution with mean 0.1429 and standard deviation 0.03704. The red point indicates the norm achieved for the constant case $(a_0, \ldots, a_0)$. (b) Comparison between the mix-norms as a function of time for the constant, RO, and RW protocols. 34

3.6 Amplitude vectors for the best twenty protocols found from the RO algorithm described in Fig. 3.5. The black curve indicates the RO protocol and a movie of the spatial evolution of this case is at `http:/amath.colorado.edu/faculty/jdm/movies/ChirikovMC.mp4` 34

3.7 Amplitude vectors of the best ($d = 1$) and the best twenty ($d > 1$) protocols found from the sparse RO method. The open circles/black line indicates the optimal protocol for each case. 35

3.8 Comparison of $\|\rho^t\|_{-1}$ for the random optimal protocol with the sparse optimal protocols for $d = 1, 2, 3, 4$. 35

3.9 Phase portraits of the map (3.4) for $T = 2$, 3, and 4 with $a_2 = -0.1732$. For $T \geq 6$ (not shown), the dynamics appears to be nearly uniformly chaotic, except for small period-six island chains along lines of slope -1. 36

3.10 Amplitude vectors for the worst ($d = 1$) and the worst twenty ($d > 1$) stirring protocols up to $d = 4$. The open circles/black line indicates the worst protocol for each case. 37

3.11 Evolution of $\rho^t(x, y)$ after one step of shear ($a = 0$). Top row shows evolution in real space. Bottom row shows the evolution of magnitudes of Fourier coefficients of degree up to 2. One step of shear results in a vertical shift of the Fourier coefficients, except for those corresponding to x wavenumber $= 0$, which remain constant. 38

3.12 Evolution of $\rho^t(x,y)$ for the $d=1$ sparse RO protocol of Fig. 3.7. Top row shows the evolution in physical space. Bottom row shows the evolution of magnitudes of Fourier coefficients with mode numbers up to 2; the value of the invariant $(0,0)$ mode, $\hat{\rho}_{0,0}$ set to zero. The numbered pairs indicate the wave numbers and the greyscale indicates the coefficient magnitude, black being the maximum and white being zero. For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/ChirikovSparse.mp4`. 39

3.13 Evolution of $\rho^t(x,y)$ for the $d=1$ sparse RW protocol. The eight panels are otherwise equivalent to those in Fig. 3.12. For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/ChirikovSparseBad.mp4`. 40

3.14 Density profiles of the form (3.7), with the randomly chosen parameters $(u,v) \in [0,1)^2$, $\varepsilon \in [0.1,5]$, $K,L \in \{1,5\}$, $n_l \in \{-10,10\}^2$, $\phi_l \in [0,2\pi]$, C_l less that 20% of the maximum bump amplitude. Each density profile is scaled such that $\|\rho^0\|_{-1} = 1$. 41

3.15 (a) Histogram of time at which largest $|a_t|$ occurs for $d=2$ RO protocols. (b) The fraction of the RO protocols that satisfy $a_{t_{max}}^2 > rE$ as a function of r. We perform 10^4 trials to find the $d=2$ RO protocols for 1200 different initial conditions of the form (3.7) on a grid of size $N=128$ with (3.2). 42

3.16 Level sets of the Hamiltonian (3.8) with various relative values of (a,b). 43

3.17 Phase portraits of the Harper map (2.9) with constant amplitudes, (a) $a_t = b_t = 0.1443$ and (b) $(a_t,b_t) = (-0.211881, 0.15726)$. With the relatively small amplitudes in (a) the evolution resembles the Hamiltonian flow shown in Fig. 3.16, with a thin chaotic separatrix layer. For the larger amplitudes in (b) there is more chaos. 44

3.18 Evolution of $\rho^t(x,y)$ for the Harper map (2.9) with initial condition (3.1) for $t=0,1,2,3$ (top row) and $t=6,8,10,12$ (bottom row) with constant strengths $a_t = b_t = 0.1443$. Contours correspond to the Hamiltonian (3.8). For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/HarperConst.mp4`. 45

3.19 Evolution of $\rho^t(x,y)$ for the Harper map (2.9) with initial condition (3.1) for $t = 0, 1, 2, 3$ (top row) and $t = 6, 8, 10, 12$ (bottom row) with constant strengths $a_t = -b_t = 0.1443$. Contours correspond to the Hamiltonian (3.8). For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/HarperConstNeg.mp4`. 45

3.20 Decay of $\|\rho^t\|_{-1}$ for the Harper map (2.9) with initial condition (3.1) for twelve iterates with constant strength $a_t = b_t = 0.1443$ (left) and $a_t = -b_t = 0.1443$ (right). The values in the legend denote the maximum mode number, M, with grid size $N = 2M + 1$. 46

3.21 Level sets of constant mix-norm for $(a_t, b_t) = (a, b)$ with $T = 12$ and initial condition (3.1) for a grid of amplitudes in the square $-0.2 \le a, b \le 0.2$. The color bar shows the value of the mix-norm: dark colors correspond to less mixing, and the optimal case is white. The blue points indicate the optimal protocol as a function of energy $E_H = T(a^2 + b^2)$. 47

3.22 Distribution of $\|\rho^T\|_{-1}$ after $\sim 1.4 \times 10^6$ trials of the random optimization method applied to the Harper map (2.9) with $T = 12$ and initial condition (3.1). The green curve indicates the best fit normal distribution with mean 0.1486 and standard deviation 0.05636. The red and orange dots indicate $\|\rho^T\|_{-1}$ for the constant amplitude elliptic, Fig. 3.18, and hyperbolic, Fig. 3.19, cases respectively. 47

3.23 Evolution of $\rho^t(x,y)$ for the Harper map (2.9) with initial condition (3.1) for $t = 0, 1, 2, 3, 6, 8, 10, 12$ with best protocol of $\sim 1.7 \times 10^7$ random trials. Overlaid on each frame, except the last, are the contours of the Hamiltonian (3.8) for the *upcoming* map step. For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/HarperMC.mp4`. 48

3.24 Energy profile of the RO solution for initial density (3.1), $E_H = 0.5$, and $T = 12$. 49

3.25 (a) Mix-norm as a function of time for the RO protocol for the Harper map (2.9) with $E_H = 0.5$, solid (black) curve. Also shown are the elliptic constant, $a_t = b_t = \gamma_o$, (dashed red curve) and the hyperbolic constant, $a_t = -b_t = \gamma_o$ (dotted blue curve) protocols of §3.2.2. (b) Pareto frontier, showing the attainable mix-norm as a function of energy for the Harper map, solid (black) curve. The dashed (red) curve shows the worst trial, RW, and the mix-norm achieved by the hyperbolic constant protocol is shown as the dotted (blue) curve. 50

3.26 Comparison of RO and SW protocols of the Harper map for $t = 1, 2, 3$ using the initial condition (3.1). The first column shows $\rho^t(x, y)$ for first three steps of the RO protocol, and the second gives that of the stepwise protocol applied to the RO state ρ^{t-1}. The color scale shows the mix-norm as a function of (a, b) for step t. The dotted circle is (3.9), from which the parameters for the SW protocol are selected. The green point on this circle corresponds to the RO and the blue point to the SW parameters. 52

3.27 (a) Progression of $\|\rho^t\|_{-1}$ for the RO and SW protocols described in Figure 3.26. For a movie of the SW protocol steps, see `http:/amath.colorado.edu/faculty/jdm/movies/HarperSW.mp4`. (b) Progression of $\|\rho^t\|_{-1}$ for the RO, CS SW, and CS RO protocols. The latter used $\sim 4 \times 10^5$ trials. The two constant shear protocols are identical until $t = 3$. 53

3.28 Density plot of the number of times choosing $a \neq 0$ results in better one step mixing than choosing $b \neq 0$ with respect to mode numbers (n_1, n_2) for initial condition (3.11) with 200 random choices of phase shift ϕ. Here the shear amplitudes, $|a|$ or $|b|$, are fixed to $\gamma_0 = 0.2041$, but varying γ_0 within reasonable bounds does not meaningfully change the results. 55

3.29 The intensity of the color at a given point (u, v) is proportional to the percentages of sign choices that yielded the best mixing over different ε ranging from 0.1 to 5 using the initial density (3.6). . 57

3.30 Evolution of $\rho^t(x, y)$ for the stirring protocol produced by CS Full and CS FS: alternating steps of $(0, \gamma_o)$ and $(-\gamma_o)$, $\gamma_o =\approx 0.204$, with initial condition (3.1) for $t = 1, 2, 3, 4$ (top row) and $t = 6, 8, 10, 12$ (bottom row) on a 128×128 grid. 59

3.31 Decay of $\|\rho^t\|_{-1}$ for the RO protocol, the SW minimizing protocol, and the CS FS protocol shown in Fig. 3.30. 59

3.32 (a) Distribution of $\|\rho^T\|_{-1}$ using the three shearing techniques, for random choices of (3.7) where $(u, v) \in [0, 1)^2$, $\varepsilon \in [0.1, 5.0]$, and $n \in \{-10, 10\}^2$. We use energy $E_H = 0.5$ and $T = 12$. (b) Means and (c) standard deviations of $\|\rho^t\|_{-1}$ for the CS SW, CS Full, and CS FS solvers. 60

4.1 Evolution of ρ^t for $t = 0, \ldots, 3$ with initial density (4.1), $N_r = N_\theta = 128$, and blinking vortex protocol vectors (4.2). 62

4.2 Initial Gaussian density profile (4.1) with $N_r = N_\theta = 128$ subjected to vortices at different distances: (a) $v = 0.05$, $\|\rho\|_{-1}$ defined by (2.36) decreases by 0.2%, (b) $v = 0.35$, $\|\rho\|_{-1}$ decreases by 5.4%, (c) $v = 0.65$, $\|\rho\|_{-1}$ decreases by 3.6%. 63

4.3 Density profile (4.3) with $(a, b) = (0.5, 0.125)$ with fitted covariance ellipse $\mathcal{E}_{z_C,\varrho,\alpha}$ with $z_C = 0$, $\varrho = 0.5$. The black point marks z_C and the black curve marks $\mathcal{E}_{z_C,\varrho,\alpha}$. 65

4.4 Optimizing one-step mixing of the RV Problem for an elliptic unmixed region with initial condition (4.3) and $(a, b) = (0.5, 0.125)$ depicted by Fig. 4.3. Plot (a) shows the variance (4.4). Plot (b) shows the mix norm ρ^1 with respect to v. Plot (c) shows the density profile ρ^1 after applying the variance optimizing vortex. Plot (d) shows ρ^1 after applying the mix norm optimizing vortex. The purple point marks the variance optimizing vortex $v = -0.2381 - 0.1429i$. The green point marks the mix norm optimizing vortex $v = -0.3333 - 0.2381i$. 67

4.5 Evolution of ρ^t for $t = 0, 1, 2, 4, 6, 8, 10, 12$ with initial condition (4.1) and $N_r = N_\theta = 128$. Parameters of the RV algorithm are $N_z = 1005$, $\varrho = 0.2$, $\alpha = 0.75$, and $N_v = 88$. 68

4.6 Evolution of ρ^t for $t = 0, 1, 2, 4, 6, 8, 10, 12$ with initial condition (4.1) and $N_r = N_\theta = 128$. Parameters of the RV algorithm are $N_z = 1005$, $N_v = 88$ and ϱ and α vary according to (4.5). 69

4.7 Decay of $\|\rho^t\|_{-1}$ for the RV stirring protocols with constant parameters (Fig. 4.5) and nonconstant parameters (Fig. 4.6). 70

4.8 One-step evolution of initial density (4.1) with default first vortex choice $\omega_1 = +r$ and $N_r = N_\theta = 2^8$. Plot $t = 0$ shows ρ^0 with a square marking the vortex $v_1 = \frac{1}{2}$. Plot $t = 1$ shows ρ^1 after vortex at v_1 was applied counterclockwise. 72

4.9 Evolution of ρ^t for the optimal stirring protocols of the FV problem for $N = 2, \ldots, 8$ with initial condition (4.1) and $N_r = N_\theta = 2^7$. Squares mark where a vortex will be activated counterclockwise, diamonds mark where a vortex will be activated clockwise. 73

4.10 ρ^4 after applying the same v_t as the optimal 4-step FV scheme (Third row of Fig. 4.9), with all possible other choices of rotation direction s_t. 74

4.11 Mix norms of the FV optimal stirring protocols for $N = 1, \dots 8$ shown in Figures 4.8 and 4.9. Plot (a) shows the decay of $\|\rho^t\|_{-1}$ with respect to t . Plot (b) shows the decay of $\|\rho^N\|_{-1}$ with respect to N. 75

4.12 Results of the Random Optimal algorithm applied with 10^4 trials to find the proportion vector $(\gamma_{\frac{1}{2}}, \gamma_{\frac{i}{2}}, \gamma_{-\frac{1}{2}}, \gamma_{-\frac{i}{2}})$ which satisfies $\gamma_v \in [0, 1]$ and (4.9) and the sign vector $(s_{\frac{1}{2}}, s_{\frac{i}{2}}, s_{-\frac{1}{2}}, s_{-\frac{i}{2}})$ which minimize $\|\rho^{100}\|_{-1}$ with $T = 1$ and $\tau = \frac{1}{100}$. The vectors produced by the algorithm are $\gamma = (0.47, 0.4, 0.01, 0.11)$ and $s = (1, -1, 1, -1)$. Plot (a) shows the density profile after the scheme is applied. Plot (b) shows the contours of the Hamiltonian (4.8). 77

Chapter 1

Introduction

The design of an efficient mixer for a passive scalar in an incompressible flow requires effective mechanical "stirring" of the fluid that eventually will lead to homogenization through diffusion. When stirring gives rise to fine-scale filaments, even a small diffusivity can be highly effective in homogenization. As was noted by Spencer and Wiley [43]:

> When the materials to be mixed are very viscous liquids... neither diffusion nor turbulence can assist very much in mixing. Rather, mixing must be effected by some complex, continuous deformation which serves to disperse the components to the desired degree.

Thus, as a first step in designing a mixer one can ignore diffusion and study only advection by an incompressible flow. The results are expected to be "transportable" to the diffusive case. Cortelezzi conducted experiments to support to this claim by devising stirring schemes first by taking diffusion into account and then by ignoring diffusion [10]. He found that, provided the Péclet number, or the ratio of the rate of advection to diffusion, is sufficiently large, the stirring schemes produced by the pure advection case closely resembled those of the full advection-diffusion problem.

1.1 Finite-Time Mixers

In this thesis we are interested in the finite-time case: How can one design the best mixer for a given initial state and a given finite time? There have been previous investigations of finite-time mixers in which fluid is injected at one end of a cylindrical pipe, flows past a finite number of mixing elements, and is emitted at the other end. For example, the Kenics Static Mixer discussed

in Galaktionov et al. consists of a cylindrical pipe with twisted blades placed at fixed locations along the pipe [22]. The twist directions and angles of the blades are the parameters that can be modified to improve mixing efficiency. Examples of the mixer with different blade angles and orientations are shown in Fig. 1.1(a). Another example is the rotated-arc mixer (RAM) discussed in Speetjens et al., which consists of two cocentric cylinders: a stationary inner cylinder and a rotating outer cylinder [42]. A window is cut into the inner cylinder, so that the fluid is mixed by drag created by exposure to the outer cylinder. A picture of the RAM mixer is shown in Fig. 1.1(b). Other examples of this type of mixing include "making up rubber formulations on compounding rolls, using an extruder as a mixer, kneading bread dough, and pulling taffy" [43]. Our goal is to develop methods that can be used to optimize the geometry of the elements to maximize mixing.

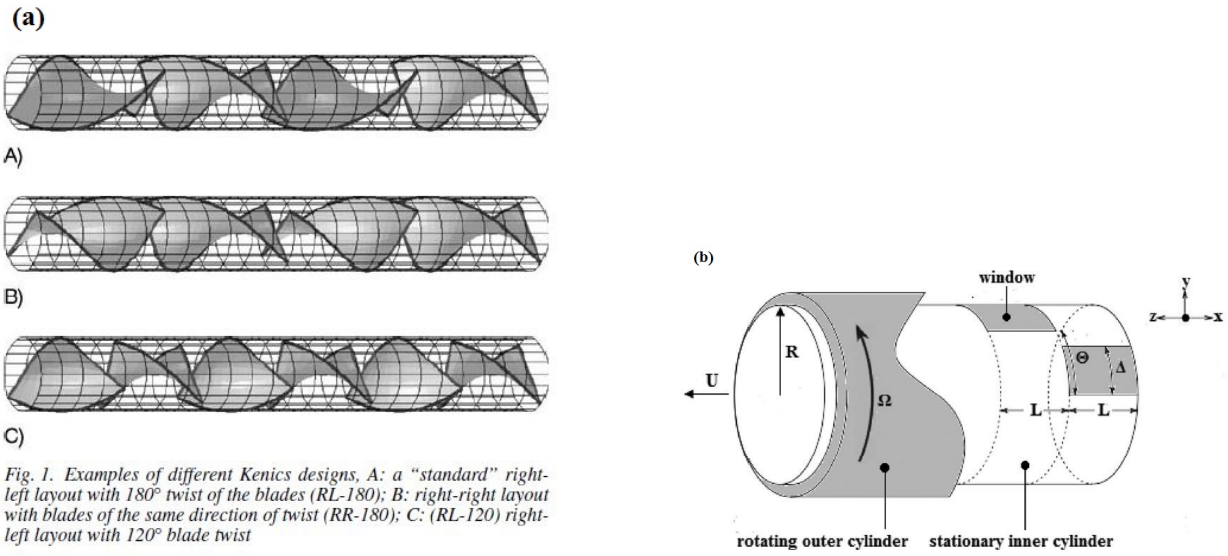

Figure 1.1: Examples of finite-time mixers in which discrete mechanisms are used to mix fluid flowing through three-dimensional pipe. Plot (a) shows the Kenics Static Mixer [22]. Plot (b) shows the rotated-arc mixer (RAM) [42].

We view the stirring process as a finite sequence of steps; each corresponding to one element of the mixer. The design corresponds mathematically to the selection of a finite sequence $f_t : D \to D$, $t = 1, 2, \ldots, T$, of maps on a domain D. Each map can be thought of as the result of a flow for

a time τ_t of an incompressible fluid so that each f_t is a volume-preserving diffeomorphism. Such a blinking flow could be a reasonable model of a low Reynolds number fluid in a cavity where the side walls are alternatively moved, or for flows with magnetohydrodynamic forcing [33, 10].

We will study compositions of time dependent transport maps, see §2.3 and §2.4. The mixer acts on a passive scalar, represented by a density $\rho : D \to \mathbb{R}$. Our goal is to find a sequence $\{f_t\}$ that takes a fixed "unmixed" initial state, ρ^0, to one that is "optimally mixed" after flowing through the full system, i.e., under the composition

$$F_T = f_T \circ f_{T-1} \circ \cdots \circ f_2 \circ f_1.$$

Ignoring diffusion, sources, and reactions, the evolution of the density is determined by the Perron-Frobenius operator $\mathcal{P}$,

$$\rho^t(z) = \mathcal{P}_t[\rho^{t-1}](z) = \rho^{t-1}(f_t^{-1}(z)), \quad t = 1, 2, \ldots, T. \tag{1.1}$$

It is common to study mixing using autonomous dynamics, where the maps f_t are either identical or repeated in a periodic sequence. Mixing for this case can be measured using concepts appropriate for the infinite-time limit, i.e., Lyapunov exponents or measure-theoretic entropy. Lyapunov exponents indicate the presence of chaos, so they can be used as an indicator of effective mixing. D'Alessandro, Dahleh, and Mezić consider a two-dimensional piecewise flow and use a control theory approach to maximize the entropy of the system [12]. They show that the entropy can be found by computing the positive Lyapunov exponents of the flow. Entropy is an intuitive measure of mixing because it describes the randomness of a system. They conclude that alternating maps of vertical and horizontal shears maximizes the entropy.

Aref noted that effective stirring can arise from chaotic dynamics [2]. He used a piecewise constant blinking vortex model to stir particles in a two-dimensional circular domain. He investigated the emergence of chaos relative to the period of switching between activation of the two vortices. He discovered that as the period approaches zero, the flow approaches an integrable system, and therefore the regions of chaotic orbits shrink to zero, resulting in very poor mixing. He

concluded that a large enough period was required so that the dynamical instability created by the vortex switching would increase the size of the chaotic regions. In §2.4, we introduce a stirring model very similar to Aref's blinking vortices, and we present results for this model in Chapter 4.

The effect of regular islands in the phase space on mixing is discussed in [16] and [17], where it is observed that the emergence of islands in "a sea of chaos" greatly detracts from the mixing efficiency. These works use symmetries of the stirring maps to identitiy elliptic points that will indicate the emergence of an island, then use this knowledge to recursively rotate and manipulate the symmetries to destroy the islands. The first work applies this method to a cavity flow, in which two walls of a cavity containing a fluid are moved in opposite directions. The second work applies the method to egg beater flows, which use piecewise alternating vertical and horizontal shearing similar to the model discussed in D'Alessandro, Dahleh, and Mezić [12].

There are a number of works that investigate the relationship between chaos and mixing with applications to the optimal design of a periodic sequence of obstacles in a channel or pipe. One is the Kenics Static Mixer discussed above [22]. Others include Micro Total Analysis Systems, which use twisting microchannels to mix miniscule quantities of fluid [29]. The Anisotropic Unstructured Meshes Mapping Method attempts to mix microfluids by specifying a spatial pattern of bends in a pipe to induce chaos [23]. A longitudinal vortex generators approach can be used to design a T-shaped channel micromixer [28].

It is less clear what role chaos plays in the finite-time, nonautonomous case. The study of aperiodic protocols was initiated by [33], using as a model, the "sine flow", equivalent to the Harper map in §2.3. The flow is modeled by a succession of horizontal or vertical shears with fixed amplitudes. The mixing is optimized for a discrete set of possibilities: a choice to switch—or not—from one shear to another with a fixed time step. This work is extended in [10], where short-time horizons are used to determine an optimal scheme for switching, in which the algorithm only looks a few steps in the future to reduce the number of protocols to choose from. These studies found that even though a periodic switching protocol can have a large Lyapunov exponent (when repeated for infinite time), it is typically not the optimal choice for mixing. Later studies allowed

continuous shifts in the phase of the shears [26] and its shape [25]. We will extend their results below by considering more general protocols in which the shear amplitudes, or equivalently their actuation times, can vary.

An alternative approach to finite-time optimization, based on control theory, was used in [19] for the advection-diffusion problem; in this work the advective part is fixed and diffusion is allowed to locally vary to enhance mixing. The authors describe a target equilibrium density profile that is invariant under advection, and they seek to to minimize the L^2 distance from the target density to the computed density after some finite time.

Another approach is to fix the diffusivity and use control theory to optimize a time-dependent linear combination of two vortical flows [11]. The problem becomes to minimize a function that consists of the sum of the measure of mixedness with the cost of the energy of the flows. The process is approximated by discrete operations, and then subjected to a control feedback loop to determine the optimal advective scheme. The approach of optimizing mixing with limits on the allowed energy of the advective action will be introduced in §2.3, and used for the results in Chapter 3.

More recently, Sturman and Thiffeault have investigated the bounding of Lyapunov exponents for random products of two shear matrices [46]. These random products can be thought of as non autonomous composition of shearing maps. The Lyapunov exponents can be a good indication of mixing, so computing bounds for them is a way of predicting the decay of mixing.

1.2 A Measure of Mixing

In order to select an "optimal" mixing protocol, we must define the stirring effectiveness in some way. Mixing, in a measure theoretic sense, is defined through an infinite time limit [31]. A transformation $f : D \to D$ that preserves a measure μ is *mixing* if for any two measurable sets A and B, the measure of the overlap $\mu(A \cap f^{-t}(B)) \to \mu(A)\mu(B)/\mu(D)$ as $t \to \infty$. Equivalently, f is mixing if for any initial density $\rho^0 \in L^1(D)$, $\rho^t(x)$, determined by (1.1), converges *weakly* as $t \to \infty$

to its mean

$$\langle \rho \rangle = \frac{1}{|D|} \int_D \rho \, d\mu, \tag{1.2}$$

where $|D| = \mu(D)$ is the volume of D.

For our purposes, measure-theoretic mixing is not an appropriate yardstick for two reasons. First, mixing implies that every L^2 initial state, not just those of physical interest, approaches a uniform state. Moreover, even if the entropy is zero, a physical state may still develop fine-scale structure [36]. Second, we are interested in a given state becoming "mixed-up" after a *finite* number of **different** transformations, not in the limit $t \to \infty$ for a single map.

As was noted by Danckwerts, in the presence of diffusion, the decrease of variance

$$\mathrm{Var}(\rho) = \left\langle \rho^2 \right\rangle - \langle \rho \rangle^2 , \tag{1.3}$$

is one measure of mixing [13]. Indeed, diffusion will cause $\mathrm{Var}(\rho)$ to decrease at a rate proportional to $\|\nabla \rho\|^2$, the L^2 norm of the gradient (See §2.1) [48]. Thus to be effective, a stirrer should increase gradients so that diffusion is activated. This "intensity of segregation" has been used to optimize mixing in diffusive flows [30, 22, 23]. However, in the absence of diffusivity and sources, the variance, and each L^p-norm, of the density is constant.

One class of mixing measures include the mix-norm of Mathew, Mezić, and Petzold [36] and equivalent Sobolev seminorms [48]. For functions that have zero mean, the squared H^q-seminorm is defined by

$$\|\rho\|_q^2 = \|(-\Delta)^{\frac{q}{2}} \rho\|^2.$$

When q is not a positive integer, the Laplacian Δ is to be interpreted by its Fourier symbol. A more extensive discussion of the mix norm is given in §2.1.

1.3 Two-Dimensional Maps

Our primary interest is the design of efficient, finite-time mixers in a three-dimensional domain. However, in this thesis, we treat a simpler problem: we assume that the axial velocity is

constant and that f corresponds to the map from a two-dimensional cross-section, through a mixing element, to another such cross-section. Since the geometry of the mixing elements will vary, subsequent maps will differ. Each of the f_t will be chosen from a given family that satisfies some geometric constraints. For example the RAM consists of an inner pipe surrounded by a spinning outer pipe, and the mixing elements correspond to windows cut into the inner pipe [22]. Each map represents the effect of the angularly sheared flow induced by the boundary drag. The maps, f_t, will depend on parameters that represent the window width, length, and rotation rate.

In this thesis, we study two families of shearing maps in a square domain, Chirikov and Harper maps, see §2.3 and Chapter 3, and a family of blinking vortex maps, see §2.4 and Chapter 4, inspired by Aref's blinking vortex model [2]. Each map depends on one or more parameters, say a_t. Optimization then corresponds to selecting the best sequence from a given family, that is, to select a vector $a = (a_1, a_2, \ldots, a_T)$ of parameters for maps in the family. Each family of maps will have certain constraints on the parameter choices. In the case of the standard and Harper maps, we assume that the total energy expended in the mixing process is bounded. The simple version of this is to assume that some norm, $\|a\|$, is fixed. When the norm is small, any fixed map in one of our families is only weakly chaotic on a small subset of phase space. However, varying the parameters from step-to-step can nevertheless result in effective stirring. In the case of the blinking vortex maps, there are more parameter choices, such as vortex location, strength, and rotation direction. We consider different schemes in which we allow some parameters to vary and others to be held fixed.

For the shear maps, the simplest method to find an optimal protocol is to exhaustively search all possible cases. This corresponds to selecting an optimal parameter vector a that satisfies an energy constraint. Thus, we can simulate the optimization process by a random optimization method: simply select the best result from a large number of trials that are equidistributed on the sphere $\|a\|^2 = E$, corresponding to a fixed energy. For the blinking vortex maps, there are two many parameters for a random optimal approach to be effective. One method used is to devise a one-step predictive scheme. Another is to limit the parameter choices, so that the possible parameter vectors

become finite and can be explored through an exhaustive search.

Our investigation seeks to understand mixing by focusing only on advection created by a sequence of stirring events, accounting for diffusion in the mixing measure itself. Through numerical simulations using simple area-preserving maps, and through analytical considerations based on Fourier analysis and key characteristics of the maps, we gain some understanding of how to design efficient stirrers.

In Chapter 2 we describe the action of advection in terms of compositions of maps. We show how the shearing and blinking vortex maps can be derived from flows, and we explain how we carry out the application of the maps numerically. In Chapter 3 we present results from numerical experiments with the Chirikov standard and Harper maps. In Chapter 4 we present results from numerical experiments with the blinking vortex map. In Chapter 5 we summarize the results and suggest courses for future study.

Chapter 2

Mathematical Formulation

In this section, we justify and explain the implementation of the mathematical and numerical methods used in this thesis. In §2.1 we justify the use of the negative Sobolev seminorm (2.1) in the place of tracking diffusion. In §2.2 we show how a continuous flow described by the pure advection equation can be represented by the composition of discrete maps. In §2.3 and §2.4 we explain the mathematics and numerics associated with applying the two mixing models: the shear maps and the blinking vortex map.

2.1 The Mix Norm

In this section, we justify the use of negative Sobolev-seminorms as a measure of mixing for a passive scalar subjected to an incompressible flow with small diffusion. For an understanding of what is meant by "small diffusion", see [10]. For simplicity, we assume ρ has zero mean, such that $\langle\rho\rangle = 0$, since if $\langle\rho\rangle = 0$ we could redefine the function $\rho = \rho - \langle\rho\rangle$. The mean of the density profile is irrelevant when measuring mixing, so we need not consider it when computing the mix norm. Recall the general definition of the H^q Sobolev norm:

$$\|\rho\|_q^2 = \|(-\Delta)^{\frac{q}{2}}\|^2 = \frac{1}{|D|}\int_D |(-L^2\Delta)^{q/2}\rho|^2 \, d\mu, \tag{2.1}$$

where $\rho(x, y, t)$ is a passive scalar defined over two-dimensional domain D at time t, and L is a normalizing length scale. Recall from §1.2 that $|D| = \mu(D)$. For the case $q < 0$, the operator Δ is interpreted by its Fourier symbol.

Let's consider the incompressible velocity field $u(x, y, t)$ which represents the stirring mechanism, and diffusion coefficient κ. The evolution of ρ through time can be described by the advection-diffusion equation with the incompressiblity constraint:

$$\frac{\partial}{\partial t}\rho + u \cdot \nabla, \rho = \kappa \Delta \rho. \tag{2.2}$$

$$\nabla \cdot u = 0 \tag{2.3}$$

We apply a no-flux boundary condition, where if ∂D describes the boundary of D, we require that

$$\nabla \rho \cdot \hat{n} = 0 \quad \text{on } \partial D$$

where $\hat{n}$ is the outward pointing unit normal of the surface ∂D. In addition, we assume we always either have a Dirichlet boundary condition

$$\rho = 0 \quad \text{on } \partial D,$$

or a Neumann boundary condition

$$\nabla \rho = 0 \quad \text{on } \partial D.$$

The variance (1.3) seems to be an intuitive way of measuring mixing because it measures deviations from the mean, so small variance indicates that the concentration is mostly uniform throughout the domain. To understand how the variance $\mathrm{Var}(\rho)$ decays in time, we recall the definitions (1.2) and (1.3) from Chapter 1. First, we can show that the mean of the concentration remains constant, by substituting (2.2) into (1.2) to get

$$\frac{d}{dt}\langle \rho \rangle = \frac{1}{|D|}\int_D \frac{\partial \rho}{\partial t} d\mu \overset{(2.2)}{=} \frac{1}{|D|}\int_D (\kappa \Delta \rho - u \cdot \nabla \rho) d\mu = \frac{1}{|D|}\left(\kappa \int_D \overset{(1)}{\Delta \rho} d\mu - \int_D \overset{(2)}{u \cdot \nabla \rho} d\mu \right).$$

The integral (1) becomes zero by applying the divergence theorem and the boundary conditions:

$$(1) = \int_D \nabla \cdot \nabla \rho d\mu \overset{\text{Dv Thm}}{=} \int_{\partial D} \nabla \rho \cdot \hat{n} dS \overset{\text{Bnd Cnd}}{=} 0.$$

The integral (2) becomes zero by applying product rule, the incompressibility condition (2.3), the divergence theorem, and the boundary conditions:

$$(2) \overset{\text{Prd Rl}}{=} \int_D (\nabla \cdot (\rho u) - \rho(\nabla \cdot u) d\mu \overset{(2.3)}{=} \int_D \nabla \cdot (\rho u) d\mu \overset{\text{Dv Thm}}{=} \int_{\partial D} \rho u \cdot \hat{n} dS \overset{\text{Bnd Cnd}}{=} 0.$$

Therefore,

$$\langle \rho \rangle = 0.$$

A similar calculation, which uses Green's Theorem, will show that $\frac{d}{dt}\|\rho\|^2 = -2\kappa\|\nabla\rho\|^2$, since

$$\begin{aligned}
\frac{d}{dt}\|\rho\|^2 = \frac{d}{dt}\frac{1}{|D|}\int_D \rho^2 d\mu &= 2\frac{1}{|D|}\int_D (\kappa\rho\Delta\rho - \rho(u\cdot\nabla\rho))d\mu \\
&\overset{\substack{\text{Prd Rl}\\(2.3)}}{=} 2\frac{1}{|D|}\int_D (\kappa\rho\Delta\rho - \nabla\cdot(\rho u))d\mu \\
&\overset{\substack{\text{Dv Thm}\\\text{Bnd Cnd}}}{=} 2\kappa\frac{1}{|D|}\int_D \rho\Delta\rho d\mu \\
&\overset{\text{Grn Thm}}{=} 2\kappa\frac{1}{|D|}\int_{\partial D}\rho(\nabla\rho\cdot\hat{n})dS - 2\kappa\frac{1}{|D|}\int_D \nabla\rho\cdot\nabla\rho d\mu \\
&\overset{\text{Bnd Cnd}}{=} -2\kappa\|\nabla\rho\|^2.
\end{aligned}$$

Combining the results gives the decay rate of the variance:

$$\frac{d}{dt}\mathrm{Var}(\rho) = \frac{d}{dt}\langle\rho\rangle^2 + \frac{d}{dt}\|\rho\|^2 = -2\kappa\|\nabla\rho\|^2.$$

For κ near zero, as in the types of problems we are concerned with, the calculation of the rate of change of the variance will become more difficult, as it will depending on the computation of gradients on small scales. Moreover, if $\kappa = 0$, then the variance is conserved, so we require a different measure of mixing. To understand why the H^{-1} norm in particular accomplishes this, we consider the time evolution of the H^1 norm, defined as

$$\|\rho\|_{H^1} = \|(-\Delta)^{\frac{1}{2}}\rho\| = \|\nabla\rho\|,$$

so that the time derivative of the square norm is

$$\begin{aligned}
\frac{d}{dt}\|\rho\|_{H^1}^2 = \frac{1}{|D|}\int_D \frac{d}{dt}|\nabla\rho|^2 d\mu = 2\frac{1}{|D|}\int_D \nabla\rho\cdot\frac{d}{dt}(\nabla\rho)\,d\mu &= 2\frac{1}{|D|}\int_D \nabla\rho\cdot\nabla\left(\frac{\partial\rho}{\partial t}\right)d\mu \\
&\overset{\text{Prd Rl}}{=} 2\frac{1}{|D|}\int_D \left(\nabla\left(\frac{\partial\rho}{\partial t}\nabla\rho\right) - \Delta\rho\frac{\partial\rho}{\partial t}\right)d\mu \\
&\overset{\substack{\text{Dv Thm}\\\text{Bnd Cnd}}}{=} -2\frac{1}{|D|}\int_D \Delta\rho\frac{\partial\rho}{\partial t}d\mu \\
&= -2\left\langle\Delta\rho\frac{\partial\rho}{\partial t}\right\rangle.
\end{aligned}$$

Substituting (2.2) gives

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|\rho\|_{H_1}^2 &= -\left\langle \Delta\rho \frac{\partial\rho}{\partial t} \right\rangle = -\frac{1}{|D|}\int_D \Delta\rho(\kappa\Delta\rho - u\cdot\nabla\rho)d\mu \\
&= -\kappa\|\Delta\rho\|^2 + \frac{1}{|D|}\int_D \Delta\rho(u\cdot\nabla\rho)d\mu \\
&\overset{\text{Prd Rl}}{=} -\kappa\|\Delta\rho\|^2 + \frac{1}{|D|}\int_D \left[\nabla\cdot((u\cdot\nabla\rho)\nabla\rho) - \nabla\rho\cdot\nabla(u\cdot\nabla\rho)\right]d\mu \\
&\overset{\substack{\text{Dv Thm}\\ \text{Bnd Cnd}}}{=} -\kappa\|\Delta\rho\|^2 - \frac{1}{|D|}\int_D \nabla\rho\cdot\nabla(u\cdot\nabla\rho)d\mu \\
&= -\left\langle \nabla\rho\cdot\nabla(u\cdot\nabla\rho)\right\rangle - \kappa\|\Delta\rho\|^2.
\end{aligned}$$

Therefore, as gradients on fine scales get large, as would be the case with effective mixing, the H^1 norm will diverge. Applying the Cauchy-Schwartz inequality gives the following relationship between the H^1 and H^{-1} norms:

$$\|\rho\|^2 = \left\langle \nabla\rho\cdot\nabla^{-1}\rho\right\rangle \le \|\rho\|_{H^1}\|\rho\|_{H^{-1}}.$$

If we consider the pure advective case with $\kappa = 0$, the quantity $\|\rho\|^2$ will remain constant. We know that good mixing requires the H^1 norm to diverge. If the norm H^{-1} were to converge to zero, the H^1 would diverge to satisfy the inequality above. Therefore the convergence of the H^{-1} norm to zero implies good mixing [48].

For the purposes of this thesis, we use the H^{-1} norm, abbreviated as $\|\cdot\|_{-1}$, though any negative Sobolev seminorm can be used as a measure of mixing. The following theorem, presented and proved in [32], states that convergence of any negative Sobolev norm to zero implies weak convergence to the mean.

Theorem 1 A time-dependent function $\rho(z,t)$, where $\rho(\cdot,t) \in L^2(D)$ has mean zero and is bounded in L^2 uniformly in time, is weakly convergent to zero if and only if

$$\lim_{t\to\infty}\|\rho(\cdot,t)\|_{H^q} = 0, \quad \text{for any } q < 0.$$

There are numerical implications for different choices of negative q values. Define ρ' to be the density profile ρ with length scales half the size. Thiffeault notes in [48] that

$$\|\rho'\|_{H^q} = 2^q\|\rho\|_{H^q}.$$

Halving the length scales will cause the norm to decay at a rate of 2^q. For q closer to zero, this will require the use of finer scales to get an accurate measure of mixing. For large negative q, the norm decays dramatically with decreasing length scales. Figure 2.1 shows the time evolution of H^q norms for different values of q for a particular stirring protocol. As q becomes a larger negative number, the high order mode terms of (2.1) decay more rapidly, and stirring applied to density with fine scale structure will have no significant impact. While the norms for $q = -\frac{1}{2}, -1$ continue to decay for steps $t = 8, \ldots, 12$, the norms for $q = -\frac{5}{2}, -3$ stay almost the same. Using a norm with q as too large a negative number is not advisable, since this simulates high diffusion, which is not within the scope of the problems we are interested in.

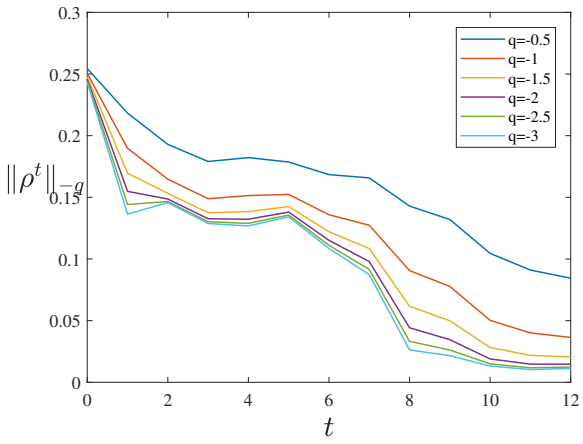

Figure 2.1: Decay of H^q norms for $q = -\frac{1}{2}, -1, -\frac{3}{2}, -2, -\frac{5}{2}, -3$ for a particular stirring protocol applied for $t = 1, \ldots, 12$.

Mathew, Mezić and Petzold define their mix norm to be the $H^{-\frac{1}{2}}$ norm, by defining the following functions on an n-dimensional torus $\mathbb{T}^n$:

$$d(\rho, z, \varrho) = \frac{\int_{B_\varrho(z)} \rho d\mu}{|B_\varrho|}, \quad \pi(\rho, \varrho) = \left(\int_{\mathbb{T}^n} d^2(\rho, z, \varrho) d\mu \right)^{\frac{1}{2}}, \quad \Pi(\rho) = \left(\int_0^1 \pi^2(\rho, \varrho) d\mu \right)^{\frac{1}{2}},$$

where $B_\varrho(z)$ is the closed ball of radius ϱ centered around the point $z \in \mathbb{T}^n$ [36]. The function $d(\rho, z, \varrho)$ is the mean value of ρ in the ball $B_\varrho(z)$, $\pi(\rho, \varrho)$ is the L^2 norm of $d(\rho, z, \varrho)$ for a fixed ball size ϱ, and $\Pi(\rho)$, the mix norm, is the L^2 norm of $\pi(\rho, \varrho)$. Mathew et al. explain, "The basic idea behind the mix norm is to parametrize all sub-intervals of $[\mathbb{T}^n]$ and to take the root mean square

of the average values of $[\rho]$ over these sub-intervals" [36]. In other words, the norm computes local averages throughout the domain for different length scales, then measures the total impact of all these averages. They refer to the application of Φ on the function $\rho - \langle\rho\rangle$ as the mix variance. The mix norm and the mix variance are equivalent for the case $\langle\rho\rangle = 0$. Since $\rho - \langle\rho\rangle$ has zero mean, the quantity $\Pi(\rho - \langle\rho\rangle)$ will give a measure of the total difference between local means and the global mean $\langle\rho\rangle$. By rewriting the above equations with inner products, and providing a spectral form of a mix-norm operator, Mathew et al. show that $\Pi(\rho)$ is equivalent to the $H^{-\frac{1}{2}}$ norm [36].

2.2 Advection by Two-Dimensional Maps

Consider the pure advection equation defined by

$$\frac{\partial}{\partial t}\rho + u \cdot \nabla\rho = 0, \tag{2.4}$$

over a two-dimensional domain D, where $\rho(x, y, t) = \rho^t(x, y)$ is the concentration of a passive scalar, and $u(x, y, t) = (u_x(x, y), u_y(x, y))$ is an incompressible velocity field. We now ignore diffusion when evolving ρ in time, recall §2.1. Given the initial condition $\rho(x, y, 0) = \rho^0(x, y)$, the velocity field u produces the flow φ such that $(x_t, y_t) = \varphi_t(x_0, y_0)$. The general solution of (2.4) is

$$\rho^t(x, y) = \rho^0(\varphi_{-t}(x, y)).$$

We define a map $f : D \to D$ to carry out the discrete action of the flow for a unit time interval:

$$(x_t, y_t) = f(x_{t-1}, y_{t-1}) = \varphi_1(x_{t-1}, y_{t-1}).$$

We allow the velocity field u to change at each time step, so we consider the sequence of velocity fields $(u_1, u_2, \ldots, u_T)$ and the corresponding sequence of maps $(f_1, f_2, \ldots, f_T)$. To find $\rho^t(x, y)$ we apply the Perron-Frobenius operator $\mathcal{P}$ which takes compositions of the sequence of

inverse maps in reverse order:

$$\begin{aligned}\rho^t(x,y) &= \mathcal{P}_t[\rho^{t-1}(x,y)] \\ &= \rho^{t-1}(f_t^{-1}(x,y)) \\ &= \rho^0(f_1^{-1}\circ\cdots\circ f_t^{-1}(x,y)).\end{aligned}$$

2.3 Compositions of Shears

As a simple model, we study a family of two-dimensional maps given by the composition of a pair of shears on a periodic domain. Taking the domain to be the torus with unit period, $D = \mathbb{T}^2 = [0,1)^2$, and $z = (x,y) \in \mathbb{T}^2$, the vertical and horizontal shears have the forms

$$\begin{aligned}s_y(x,y) &= (x\,, y + X(x)), \\ s_x(x,y) &= (x + Y(y)\,, y),\end{aligned} \tag{2.5}$$

where $X, Y : \mathbb{S} \to \mathbb{S}$ are circle maps. Composing the shears (2.5) gives the one-element map

$$f(x,y) = s_x \circ s_y(x,y) = (x + Y(y + X(x))\,,\, y + X(x)). \tag{2.6}$$

When the shears are smooth, (2.6) is an area-preserving diffeomorphism with the inverse

$$f^{-1}(x,y) = (x - Y(y)\,, y - X(x - Y(y))).$$

The map (2.6) is the solution of the advection problem

$$(\dot{x},\dot{y}) = v(x,y,t) = \begin{cases} (0, g(x)), & 0 \le t \le \tau_y \\ (h(y), 0), & \tau_y < t \le \tau_y + \tau_x \end{cases}, \tag{2.7}$$

where $X(x) = \tau_y g(x)$ and $Y(y) = \tau_x h(y)$. Thus, the amplitudes of the functions X and Y represent both the elapsed times, τ_y and τ_x, and the strengths of the components of the velocity field v.

The family (2.6) includes many well-studied cases, including Arnold's cat map [3] with $X(x) = x$ and $Y(y) = y$, Chirikov's standard map [9, 37], depicted in Fig. 2.2, for which

$$\begin{aligned}X(x) &= -a\sin(2\pi x), \\ Y(y) &= y,\end{aligned} \tag{2.8}$$

and the Sine-flow or Harper map [33, 4, 39, 6], depicted in Fig. 2.3, for which

$$\begin{aligned} X(x) &= -a\sin(2\pi x), \\ Y(y) &= b\sin(2\pi y). \end{aligned} \tag{2.9}$$

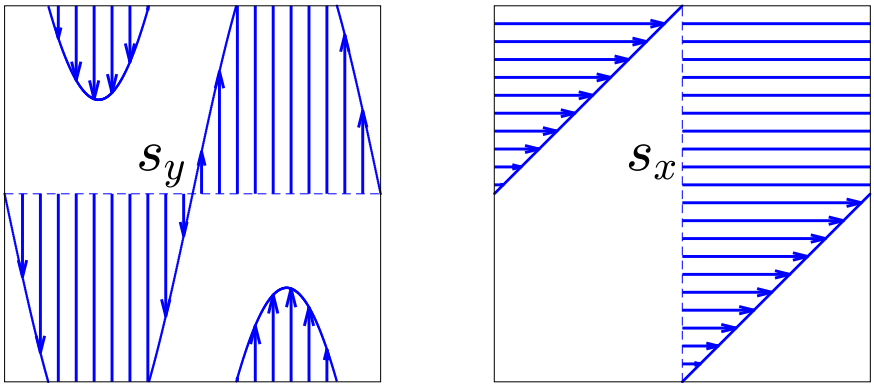

Figure 2.2: Depiction of the shearing maps (2.5) for the Chirikov standard map (2.8) with $a = 0.75$.

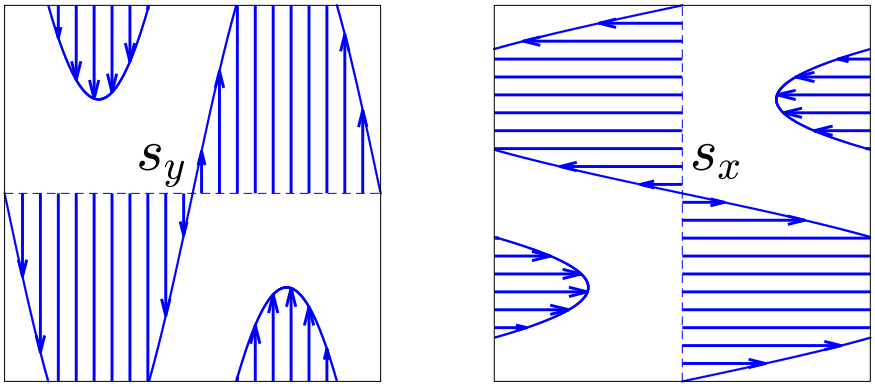

Figure 2.3: Depiction of the shearing maps (2.5) for the Harper map (2.9) with $a = b = 0.75$.

The design of a mixer consists of the choice of a mixing protocol, that is, to a sequence of T maps $f_T \circ f_{T-1} \circ \cdots \circ f_2 \circ f_1$. Formally, T may be infinite, but we will typically take $T = \mathcal{O}(10)$. The goal is to optimize the mixing for some maximal total "energy" of the flow. As a measure of this energy we take a squared norm of the shear:

$$E_s = \sup_D \|s(x,y) - (x,y)\|^2. \tag{2.10}$$

For example, for the Harper maps, $E_{s_x} = b^2$, $E_{s_y} = a^2$. Since s_x corresponds to the velocity (2.7), the energy is proportional to the kinetic energy of the fluid. The parameters (a_t, b_t) of the Harper maps are constrained by

$$\sum_{t=1}^{T}(a_t^2 + b_t^2) = E_H, \tag{2.11}$$

where E_H is some given maximum energy. The energy for a sequence of Chirikov maps includes that required to create the horizontal shear s_x, thus for (2.8)

$$\sum_{t=1}^{T} a_t^2 + T = E_C. \tag{2.12}$$

2.3.1 Computing the Perron-Frobenius Operator and the Mix Norm

To compute the image $\rho^t(x, y)$ and the seminorm (2.1) numerically, one must represent the density in some finite basis set. One possible technique is the "mapping" [43] or "Ulam" method [49, 18, 20, 15]: choose a grid of $N \times N$ points $(x_i, y_j) = (\frac{i}{N}, \frac{j}{N})$ on the torus, with $\rho_{ij} = \rho(x_i, y_j)$ and approximate the operator $\mathcal{P}$ (1.1) by an $N^2 \times N^2$ stochastic matrix P,

$$\rho_{i,j}^t = P_{ij,i'j'}\rho_{i'j'}^{t-1}.$$

Here $P_{ij,i'j'}$ is the fraction of the $i'j'$ cell that is mapped onto the ij cell by f. This can be estimated—for a fixed map f—by a one step iteration of a large number of trial points in each cell. Such a discretization introduces an effective diffusion that will decrease the variance (1.3) [30], and explicit diffusion can also be included in the process [24]. However, this method is not practical for our purposes, since the stochastic matrix would need to be recomputed for each new map in the sequence $f_1, \ldots, f_T$.

A second discrete method is to use a Fourier basis. This seems natural since the Sobolev norm that we seek to minimize is most easily evaluated in Fourier space. On the two-dimensional periodic domain $\mathbb{T}^2 = [0, 1)^2$, the H^{-1} norm which we use as the mix norm becomes

$$\|\rho\|_{-1} = \left(\sum_{\substack{m \in \mathbb{Z}^2 \\ |m| \neq 0}} \frac{|\hat{\rho}_m|^2}{|m|^2} \right)^{\frac{1}{2}}, \tag{2.13}$$

where $\hat{\rho}_m$ are the Fourier coefficients of ρ defined by

$$\hat{\rho}_m = \int_{\mathbb{T}^2} \rho(z) e^{-i2\pi mz} d^2z, \tag{2.14}$$

so that

$$\rho(z) = \sum_{m\in\mathbb{Z}^2} \hat{\rho}_m e^{i2\pi mz}, \tag{2.15}$$

and $m = (m_1, m_2)$ is a two-dimensional vector containing horizontal wave numbers m_1 and vertical wave number m_2. We exclude the $(0,0)$ term, since this represents the mean of ρ and does not impact the mixing measure. Each shear (2.5) has a relatively simple representation in Fourier space. Using (2.14), we obtain

$$\begin{aligned}(\hat{\mathcal{P}}_{s_y}[\rho])_m &= \int_{\mathbb{T}^2} \rho(x, y - X(x)) e^{-2\pi im\cdot z} d^2z = \sum_{n\in\mathbb{Z}} \hat{\rho}_{n,m_2} \mathcal{F}_{n-m_1,m_2}[X], \\ (\hat{\mathcal{P}}_{s_x}[\rho])_m &= \int_{\mathbb{T}^2} \rho(x - Y(y), y) e^{-2\pi im\cdot z} d^2z = \sum_{n\in\mathbb{Z}} \hat{\rho}_{m_1,n} \mathcal{F}_{n-m_2,m_1}[Y],\end{aligned} \tag{2.16}$$

where for each $(j,k) \in \mathbb{Z}^2$, we define a functional $\mathcal{F}_{j,k} : C^0(\mathbb{S}) \to \mathbb{C}$ that gives a Fourier coefficient of the exponential of a function:

$$\mathcal{F}_{j,k}[g] \equiv \int_0^1 e^{2\pi i(jx - kg(x))} dx.$$

For a linear shear $id(x) = x$ this gives

$$\mathcal{F}_{j,k}[id] = \delta_{j,k}, \tag{2.17}$$

the Kronecker-delta, while for a sinusoidal shear such as $X(x)$ in (2.8) and (2.9) we have

$$\mathcal{F}_{j,k}[-a\sin(2\pi x)] = J_j(-2\pi ka) = J_{-j}(2\pi ka), \tag{2.18}$$

using the standard identity $e^{iz\sin\theta} = \sum_{j\in\mathbb{Z}} J_j(z) e^{ij\theta}$ for the Bessel function. In general, $\mathcal{F}_{j,0}[g] = \delta_{j,0}$, which implies, in particular, that the $(0,0)$-Fourier mode is preserved—as expected for any volume-preserving map.

Composing the two results (2.16) gives the general formula for updating Fourier coefficients for the map (2.6):

$$\hat{\rho}_m^t = \hat{\mathcal{P}}_{s_x} \circ \hat{\mathcal{P}}_{s_y}[\rho^{t-1}] = \sum_{n\in\mathbb{Z}^2} \hat{\rho}_n^{t-1} \mathcal{F}_{n_1-m_1,n_2}[X] \mathcal{F}_{n_2-m_2,m_1}[Y], \tag{2.19}$$

since $f^{-1} = s_y^{-1} \circ s_x^{-1}$. For example, for the Chirikov map, (2.19) becomes, using (2.18),

$$\hat{\rho}_m^t = \sum_{n\in\mathbb{Z}^2} \hat{\rho}_n^{t-1} J_{m_1-n_1}(2\pi n_2 a)\delta_{n_2-m_2,m_1} = \sum_{k\in\mathbb{Z}} \hat{\rho}_{k,m_1+m_2}^{t-1} J_{m_1-k}\left(2\pi(m_1+m_2)a\right). \tag{2.20}$$

For the Harper map, (2.19) becomes

$$\hat{\rho}_m^T = \sum_{n\in\mathbb{Z}^2} \hat{\rho}_n^{t-1} J_{m_1-n_1}(2\pi n_2 a) J_{n_2-m_2}(2\pi m_1 b). \tag{2.21}$$

To implement this method numerically, one must truncate the Fourier basis, keeping say, the $N^2 = (2M+1)^2$ modes with $m_i \in [-M, M]$. Since the sums in (2.19) naturally generate modes that are not in the basis set, these must be discarded. Conversely, modes that are not in the truncated basis also may contribute to those in the truncation, and these give errors in the numerical scheme.

A final technique for implementing (1.1) is to simply evaluate an analytically defined function ρ^0 at a point $f^{-1}(x, y)$. Given a grid of points (x_i, y_j) we can define

$$\rho_{i,j}^t = \rho^t(x_i, y_j) = \rho^0(f_1^{-1} \circ f_2^{-1} \circ \cdots \circ f_t^{-1}(x_i, y_j)) \tag{2.22}$$

whenever the initial state ρ^0 is known everywhere. Numerically, this is efficient if one is interested only in the final mixed state, ρ^T, but less so if one would also like intermediate states $t < T$, since ρ_{ij}^t cannot be used to compute the next image. Nevertheless, to the extent that one can accurately compute the composition of the maps—subject to the growth of errors due to sensitive dependence on initial conditions—the resulting values of ρ^t are accurate when ρ^0 is Lipchitz, though of course we only know the values on the selected grid. This is the method used subsequently in this thesis.

We choose a regular grid of N^2 points, so that the mix norm (2.13) can be easily evaluated using a fast Fourier transform. Using a discrete lattice with N^2 points, we have $z_k = \frac{k}{N}$ for $k \in [0, N)^2$, and $\rho_k = \rho(z_k)$, with

$$\hat{\rho}_m = \sum_{k\in[0,N)^2} \rho_k e^{-2\pi i \frac{m\cdot k}{N}}, \quad \rho_k = \frac{1}{N^2} \sum_{m\in[0,N)^2} \hat{\rho}_m e^{2\pi i \frac{m\cdot k}{N}}. \tag{2.23}$$

2.4 Blinking Vortices

Another map considered is based on Aref's blinking vortex model [2]. For convenience we describe the map in the complex plane, since it is equivalent to the two-dimensional real plane. We

place a rotating point vortex at position $v \in \mathbb{C}$, defined in polar coordinates as

$$v = \nu e^{i\theta}, \quad \nu = |v|, \quad \theta = \arg(v).$$

Assuming a counterclockwise rotation, the motion of z can be described by the single vortex equation,

$$\dot{\bar{z}} = \frac{i}{v - z}, \tag{2.24}$$

where the time-derivative of z is denoted $\dot{z}$. The solution of this simple ODE is that the particle z traces the path of the circle

$$|v - z| = p,$$

where p is the distance between v and the initial point $z(0)$. This can be verified by noting that when (2.24) is satisfied,

$$\frac{d}{dt}|v - z|^2 = 0.$$

Therefore, z can be written in the form $z = v + pe^{i\phi}$ for some angle ϕ. Differentiating and substituting into (2.24) gives

$$-ipe^{-i\phi}\dot{\phi} = \frac{i}{v - (v + pe^{i\phi})}, \quad pe^{-i\phi}\dot{\phi} = \frac{e^{-i\phi}}{p}, \quad \dot{\phi} = \frac{1}{p^2}.$$

Integrating and applying the initial position angle $\phi(0)$ gives

$$\phi(t) = \phi(0) + \frac{t}{p^2}.$$

The angular speed $\dot{\phi}$ of particles varies inversely with the square distance from the vortex, and there is a singularity at the vortex itself. If $z = x + iy$ and $v = u + iv$, we can use the equations above to derive the following equations for the time evolution of the real and imaginary parts of z:

$$x(t) = u + (x - u)\cos(\frac{t}{p^2}) - (y - v)\sin(\frac{t}{p^2})$$
$$y(t) = v + (x - u)\sin(\frac{t}{p^2}) + (y - v)\cos(\frac{t}{p^2}).$$

The single vortex model (2.24) pertains to the whole complex plane. We wish to consider a vortex model confined to the unit disc centered at the origin. We place an imaging vortex at

the position $\frac{1}{\bar{v}}$, where $\bar{v}$ is the complex conjugate of v. The image vortex $\frac{1}{\bar{v}}$ rotates in the opposite direction of v. Activating v and its image simultaneously creates an invariant boundary at the unit circle. The new stirring mechanism is described by the equation

$$\dot{\bar{z}} = is\left(\frac{1}{z-v} - \frac{1}{z-\frac{1}{\bar{v}}}\right) = is\frac{v-\frac{1}{\bar{v}}}{(z-v)(z-\frac{1}{\bar{v}})}. \tag{2.25}$$

The value s indicates the direction of rotation of v such that

$$s = \begin{cases} 1 & \text{counterclockwise} \\ -1 & \text{clockwise} \end{cases}.$$

Figure 2.4 illustrates the action of a counterclockwise spinning vortex at $0.4 + i0.5$ and its image vortex activated for four time units, along with the trajectories of a few different initial particle positions. The speed of trajectories slows as the distance between the particle positions and vortex increases.

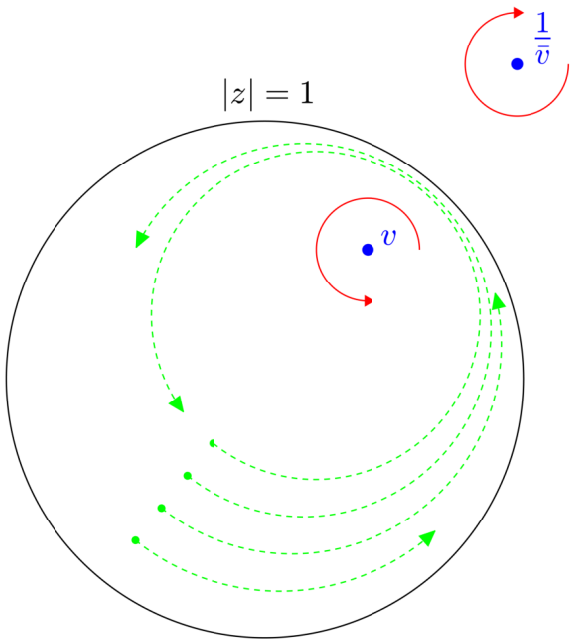

Figure 2.4: Diagram of (2.25) with $v = 0.4 + i0.5$, $s = 1$, and $t = 4$. The black curve indicates the invariant boundary at the unit circle $|z| = 1$. The blue points mark the vortex and its image. The red curves show the rotation direction of the vortices. The green curves show resulting trajectories for different initial particle positions. The green points mark the initial positions of the particles and the green arrowheads show the final positions at time $t = 4$.

The particle z traces the path of a circle of the form

$$\left|\frac{z-v}{z-\frac{1}{\bar{v}}}\right| = \lambda, \tag{2.26}$$

where λ is found by applying the initial position $z(0)$. Similar to the single vortex case, this is verified by noting that

$$\frac{d}{dt}\left[\left|\frac{z-v}{z-\frac{1}{\bar{v}}}\right|\right] = 0$$

is satisfied by (2.25). Recalling that $|v| = \nu$ and $\arg(v) = \theta$, the circle $\mathcal{C}$ described by (2.26) has radius

$$p = \frac{\lambda\left(\frac{1}{\nu} - \nu\right)}{1-\lambda^2},$$

and center z_c

$$z_c = \frac{v - \frac{\lambda^2}{\bar{v}}}{1-\lambda^2},$$

such that

$$z = z_c + pe^{i\phi}. \tag{2.27}$$

To see this, we use the identities $v - z_c = p\lambda e^{i\theta}$ and $\frac{1}{\bar{v}} - z_c = \frac{p}{\lambda}e^{i\theta}$ to show that

$$\left|\frac{z-v}{z-\frac{1}{\bar{v}}}\right|^2 = \left|\frac{z_c + pe^{i\phi} - v}{z_c + pe^{i\phi} - \frac{1}{\bar{v}}}\right|^2 = \lambda^2\left|\frac{e^{i\phi} - \lambda e^{i\theta}}{\lambda e^{i\phi} - e^{i\theta}}\right|^2 = \lambda^2\left|\frac{1 - 2\cos(\phi-\theta) + \lambda^2}{\lambda^2 - 2\cos(\phi-\theta) + 1}\right| = \lambda^2.$$

Substituting (2.27) into (2.25) gives

$$\begin{aligned}
\dot{\phi}pe^{-i\phi} &= -s\frac{v - \frac{1}{\bar{v}}}{(z-v)(z-\frac{1}{\bar{v}})} \\
\dot{\phi}e^{-i\phi}\frac{(e^{i\phi} - \lambda e^{i\theta})(e^{i\theta} - \frac{1}{\lambda}e^{i\theta})}{p(\nu - \frac{1}{\nu})e^{i\theta}} &= -\frac{s}{p^2} \\
-\dot{\phi}\frac{e^{i(\phi-\theta)} - (\frac{1}{\lambda} + \lambda) + e^{i(\theta-\phi)}}{\frac{1-\lambda^2}{\lambda}} &= -\frac{s}{p^2} \\
\dot{\phi}\left(1 - \frac{2\lambda}{1+\lambda^2}\cos(\phi-\theta)\right) &= -\frac{s}{p^2}\frac{1-\lambda^2}{1+\lambda^2}.
\end{aligned}$$

Integrating, and applying the initial position angle $\phi(0)$ gives the following implicit formula for updating ϕ:

$$\phi(t) - \frac{2\lambda}{1+\lambda^2}\sin(\phi(t)-\theta) = \phi(0) - \frac{2\lambda}{1+\lambda^2}\sin(\phi(0)-\theta) - t\frac{s}{p^2}\frac{1-\lambda^2}{1+\lambda^2}. \tag{2.28}$$

Taking the derivative of the term on the left with respect to $\phi(t)$ gives

$$1 - \frac{2\lambda}{1+\lambda^2}\cos(\phi(t) - \theta),$$

which is nonnegative, so the left hand side of (2.28) is monotonic and has a unique solution $\phi(t)$ for any $\phi(0) \in \mathbb{R}$ that can be found through an iterative solver, such as Newton's method. Therefore, there exists a one-to-one map of the form

$$\phi(t) = F_{v,t,s}(\phi(0)).$$

Considering that we will be updating an array of vortices in succession, we rewrite this as a map from ϕ_{t-1} to ϕ_t, where we activate vortex v_t at time step t for time τ_t and rotation direction s_t. The map then becomes:

$$\phi_t = F_{v_t,\tau_t,s_t}(\phi_{t-1}).$$

To recover the point z_t, we define the map $G_v : [0, 2\pi) \to \mathcal{C}_v$, such that

$$z = G_v(\phi) = z_c + pe^{i\phi}, \quad \phi = G_v^{-1}(z) = \arg(z - z_c),$$

and $\mathcal{C}_v$ is the corresponding circle in the complex plane. The inverse function G_v^{-1} can be computed by means of the atan2 function. We can combine the maps F, G, and G^{-1} to get the map $H_{v,\tau,s} : \mathcal{C}_v \to \mathcal{C}_v$, which carries out the action of (2.25):

$$z_t = H_{v_t,\tau_t,s_t}(z_{t-1}) = (G_{v_t} \circ F_{v_t,\tau_t,s_t} \circ G_{v_t}^{-1})(z_{t-1}). \tag{2.29}$$

The inverse of H is

$$z_{t-1} = H^{-1}_{v_t,\tau_t,s_t}(z_t) = ([G_{v_t}^{-1}]^{-1} \circ F^{-1}_{v_t,\tau_t,s_t} \circ G_{v_t}^{-1})(z_t) = (G_{v_t} \circ F^{-1}_{v_t,\tau_t,s_t} \circ G_{v_t}^{-1})(z_t),$$

where F^{-1} is defined by reversing the sign of the last term of the update formula (2.28) such that

$$\phi(t) - \frac{2\lambda}{1+\lambda^2}\sin(\phi(t) - \theta) = \phi(0) - \frac{2\lambda}{1+\lambda^2}\sin(\phi(0) - \theta) + t\frac{s}{p^2}\frac{1-\lambda^2}{1+\lambda^2},$$

or in other words, the inverse map is the forward map with the sign of s switched:

$$H^{-1}_{v_t,\tau_t,s_t} = H_{v_t,\tau_t,-s_t}.$$

2.4.1 Computing the Perron-Frobenius Operator and the Mix Norm

In the case of the blinking vortex problem, the domain is the unit disk centered at the origin, so our passive scalar ρ is defined in terms of polar coordinates (r, θ). To evaluate the mix norm (2.1) of $\rho(r, \theta)$ on the disk, we should consider solutions to the Helmoltz equation:

$$\Delta\rho + k^2\rho = 0,$$

because it is equivalent to diagonalizing the Laplacian operator into its complete set of eigenfunctions, which will make (2.1) a straightforward calculation. Expanding the Helmoltz equation in polar coordinates gives

$$\Delta\rho + k^2\rho = \rho_{rr} + \frac{1}{r}\rho_r + \frac{1}{r^2}\rho_{\theta\theta} + k^2\rho = 0. \tag{2.30}$$

We impose the periodicity constraint: $\rho(r, \theta) = \rho(r, \theta + 2\pi)$ to ensure continuity, and we assume the Dirichlet boundary condition

$$\rho(1, \theta) = 0.$$

Our only constraint is that the boundary $r = 1$ is invariant, so we choose the Dirichlet condition out of convenience. The results below are easily generalized to a Neumann boundary condition,

$$\rho_r(1, \theta) = 0.$$

Equation (2.30) can be solved by separation of variables by assuming $\rho(r, \theta) = R(r)\Theta(\theta)$ and substituting to get

$$R''\Theta + \frac{1}{r}R'\Theta + \frac{1}{r^2}R\Theta'' + k^2R\Theta = 0$$
$$\frac{1}{R}(r^2R'' + rR' + r^2k^2R) = -\frac{\Theta''}{\Theta} = n^2,$$

where n is a constant. The angular part becomes

$$\Theta'' = -n^2\Theta.$$

Applying the periodicity constraint, the solution is

$$\Theta_n(\theta) = A_n e^{in\theta}, \tag{2.31}$$

where $n \in \mathbb{Z}$ and $A_n \in \mathbb{C}$. The radial part becomes Bessel's equation

$$r^2 R'' + rR' + (r^2 k^2 - n^2)R = 0,$$

which has the (bounded) solution

$$R_n(r) = B_n J_n(kr), \tag{2.32}$$

where $B_n \in \mathbb{R}$ and J_n is the nth-order Bessel function of the first kind. Applying the Dirichlet boundary condition gives

$$0 = R_n(1) = B_n J_n(k).$$

This is satisfied for $k = u_{m,n}$, where $u_{m,n}$ is the mth (positive) root of J_n. Combining these results with (2.31) and (2.32), we get the general solution

$$\rho(r,\theta) = \sum_{m\in\mathbb{N},n\in\mathbb{Z}} \hat{\rho}_{m,n} J_n(u_{m,n} r) e^{in\theta}. \tag{2.33}$$

To compute the coefficients $\hat{\rho}_{m,n}$, we use the orthogonality relation [1]:

$$\int_0^1 J_n(u_{m,n}r) J_n(u_{k,n}r) r dr d\theta = \delta_{m,k} \frac{1}{2} [J_{n+1}(u_{m,n})]^2. \tag{2.34}$$

Applying (2.34) to (2.33) gives

$$\int_{-\pi}^{\pi}\int_0^1 \rho(r,\theta) J_n(u_{m,n}r) e^{-in\theta} r dr d\theta = \sum_{k\in\mathbb{N},l\in\mathbb{Z}} \hat{\rho}_{k,l} \int_{-\pi}^{\pi} e^{i\theta(l-n)} d\theta \int_0^a J_l(u_{k,l}r) J_n(u_{m,n}r) r dr$$

$$= \sum_{k\in\mathbb{N}} \hat{\rho}_{k,n} 2\pi \int_0^r J_n(u_{k,n}r) J_n(u_{m,n}r) r dr = \hat{\rho}_{m,n} \pi [J_{n+1}(u_{m,n})]^2.$$

By rearranging the above result, we obtain the following formula for finding the coefficients of (2.33):

$$\hat{\rho}_{m,n} = \frac{1}{\pi [J_{n+1}(u_{m,n})]^2} \int_{-\pi}^{\pi}\int_0^1 \rho(r,\theta) J_n(u_{m,n}r) e^{-in\theta} r dr d\theta. \tag{2.35}$$

For the unit disk domain, the squared H^q-seminorm (2.1) can be expanded as

$$\|\rho\|_{H^q}^2 = \int_{-\pi}^{\pi}\int_0^1 |(-1)^{\frac{q}{2}} L^q \Delta^{\frac{q}{2}} \rho|^2 r dr d\theta = (-1)^q L^{2q} \int_{-\pi}^{\pi}\int_0^1 |\Delta^{\frac{q}{2}} \rho|^2 r dr d\theta.$$

Recall that $\rho_{m,n}(r,\theta)$ satisfies $\Delta\rho_{m,n} = -(u_{m,n})^2\rho_{m,n}$. Defining $L = 1$ and substituting (2.33), (2.35), and the orthogonality (2.34) gives

$$\begin{aligned}
\|\rho\|_{H^q}^2 &= (-1)^q 1^{2q} \int_{-\pi}^{\pi}\int_0^1 \left| \sum_{m\in\mathbb{N},n\in\mathbb{Z}} (-1)^{\frac{q}{2}} (u_{m,n})^q \rho_{m,n}(r,\theta) \right|^2 r dr d\theta \\
&= \int_{-\pi}^{\pi}\int_0^1 \left(\sum_{m\in\mathbb{N},n\in\mathbb{Z}} u_{m,n}^q \hat{\rho}_{m,n} J_n(u_{m,n}r) e^{in\theta} \right) \left(\sum_{k\in\mathbb{N},l\in\mathbb{Z}} u_{k,l}^q \hat{\rho}_{k,l}^* J_l(u_{k,l}r) e^{-il\theta} \right) r dr d\theta \\
&= \sum_{\substack{m\in\mathbb{N},n\in\mathbb{Z} \\ k\in\mathbb{N},l\in\mathbb{Z}}} u_{m,n}^q u_{k,l}^q \hat{\rho}_{m,n}\hat{\rho}_{k,l}^* \int_{-\pi}^{\pi} e^{i\theta(n-l)} d\theta \int_0^1 J_n(u_{m,n}r) J_l(u_{k,l}r) r dr \\
&= 2\pi \sum_{\substack{m\in\mathbb{N},n\in\mathbb{Z} \\ k\in\mathbb{N}}} u_{m,n}^q u_{k,n}^q \hat{\rho}_{m,n}\hat{\rho}_{k,n}^* \int_0^1 J_n(u_{m,n}r) J_n(u_{k,n}r) r dr \\
&= \pi \sum_{m\in\mathbb{N},n\in\mathbb{Z}} u_{m,n}^{2q} [J_{n+1}(u_{m,n})]^2 |\hat{\rho}_{m,n}|^2.
\end{aligned}$$

Therefore, the H^{-1} norm defined on the radius-one disk is

$$\|\rho\|_{-1} = \left(\pi \sum_{m\in\mathbb{N},n\in\mathbb{Z}} \frac{1}{u_{m,n}^2} [J_{n+1}(u_{m,n})]^2 |\hat{\rho}_{m,n}|^2 \right)^{\frac{1}{2}}. \tag{2.36}$$

The norm (2.36) must be applied to a function with zero mean. In the case of a function with nonzero mean, we define the mix norm as $\|\rho - \langle\rho\rangle\|_{-1}$. To compute (2.36) numerically, we define the grid sizes N_r to be the number of radial points, and N_θ to be the number of angular points, so that the total number of grid points will be $N_r \times N_\theta$. We choose the radial points such that they are equally spaced. The use of polar coordinates means that the grid will be distorted, in that points will be closer together near the origin and farther apart near the edge of the disc. Figure 2.5 shows the polar grids for different values of N_r and N_θ.

To evaluate (2.35) numerically, first we arrange the integral so that we evaluate the angular part first and the radial part second:

$$\hat{\rho}_{m,n} = \frac{1}{\pi[J_{n+1}(u_{m,n})]^2} \int_0^1 \left(\int_{-\pi}^{\pi} \rho(r,\theta) e^{-in\theta} d\theta \right) J_n(u_{m,n}r) r dr.$$

The inner integral is simply a one-dimensional Fourier coefficent that can be evaluated by a fast

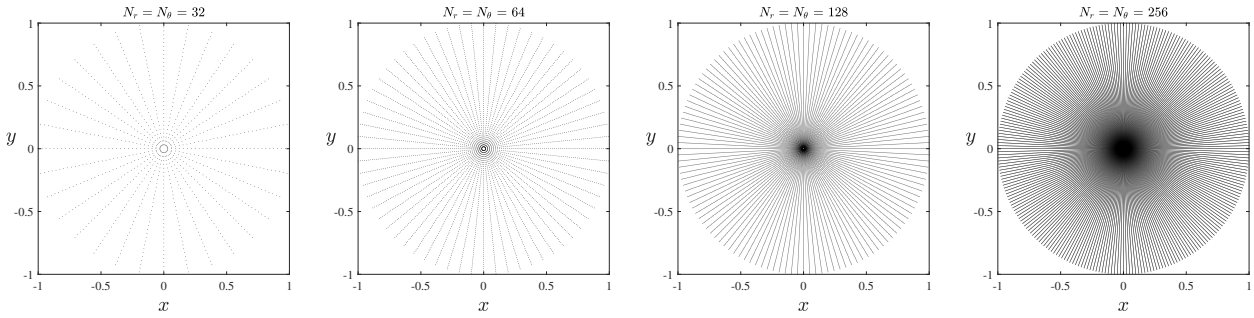

Figure 2.5: Grids of polar coordinates for $N_r = N_\theta = 32, 64, 128, 256$.

Fourier transform as in (2.23) to get

$$\begin{aligned}\hat{\rho}_{m,n} &= \frac{1}{\pi[J_{n+1}(u_{m,n})]^2}\int_0^1 \left(2\pi \sum_{k\in[0,N_\theta)} \rho(r, \frac{k}{N_\theta})e^{-i2\pi\frac{mk}{N_\theta}}\right) J_n(u_{m,n}r)rdr \\ &= \frac{2}{[J_{n+1}(u_{m,n})]^2}\sum_{k\in[0,N_\theta)} e^{-i2\pi\frac{k}{N_\theta}}\int_0^1 \rho(r, \frac{mk}{N_\theta})J_n(u_{m,n}r)rdr.\end{aligned}$$

The distortion of the grid created by using polar coordinates makes it necessary to use a higher order integration method to evaluate the radial integral. For the radial component, we apply a trapezoid method for improved accuracy in the approximated integral. This gives the following formula for numerically evaluating (2.35):

$$\hat{\rho}_{m,n} = \frac{2}{[J_{n+1}(u_{m,n})]^2}\sum_{k\in[0,N_\theta)} e^{-i2\pi\frac{mk}{N_\theta}}\frac{1}{2}\sum_{l\in[0,N_r)}\left(\rho(\frac{l}{N_r}, \frac{k}{N_\theta})J_n(u_{m,n}\frac{l}{N_r})\frac{l}{N_r} + \rho(\frac{l+1}{N_r}, \frac{k}{N_\theta})J_n(u_{m,n}\frac{l+1}{N_r})\frac{l+1}{N_r}\right) \tag{2.37}$$

$$= \frac{1}{[J_{n+1}(u_{m,n})]^2}\sum_{\substack{k\in[0,N_\theta)\\ l\in[0,N_r)}} e^{-i2\pi\frac{mk}{N_\theta}}\left(\rho(\frac{l}{N_r}, \frac{k}{N_\theta})J_n(u_{m,n}\frac{l}{N_r})\frac{l}{N_r} + \rho(\frac{l+1}{N_r}, \frac{k}{N_\theta})J_n(u_{m,n}\frac{l+1}{N_r})\frac{l+1}{N_r}\right). \tag{2.38}$$

Figure 2.6 displays the results of applying the Bessel Fourier transform to the Gaussian

$$\rho^0(z) = \exp\left(-\varepsilon\|z - z_o\|^2\right), \quad z_o = (0.5, 0.5), \quad \varepsilon = 5, \tag{2.39}$$

with grid size $N_r = N_\theta = 128$. Figure 2.6(a) shows the Gaussian profile ρ^0. Figure 2.6(b) shows $\tilde{\rho}^0$, which is found by applying (2.33) to the coefficents found by evaluating (2.38). Figure 2.6(c) shows the radial error $\rho_e = \rho^0(r, 0) - \tilde{\rho}^0(r, 0)$. Note that

$$\rho^0(r, \theta) - \tilde{\rho}^0(r, \theta) = \rho^0(r, 0) - \tilde{\rho}^0(r, 0) \quad \text{for all } \theta \in [0, 2\pi),$$

since both ρ^0 and $\tilde{\rho}^0$ are radially symmetric. The error is very small, except for at a singularity at the origin and a discontinuity at the boundary that results from the imposed Dirichlet boundary condition. However, this discontinuities only affect high order modes, and do not have a significant impact on the mix norm.

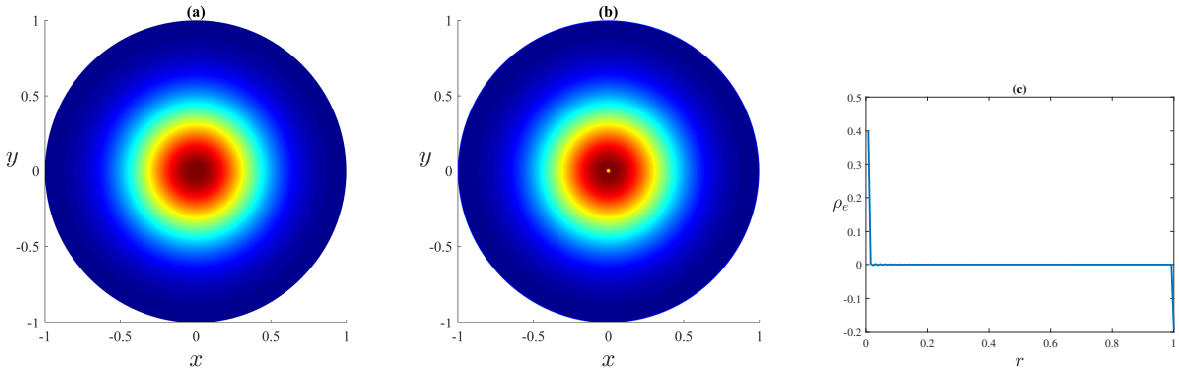

Figure 2.6: Plot (a) is ρ^0 as described by (2.39). Plot (b) is $\tilde{\rho}^0$, which is (2.33) applied to the coefficients produced by (2.38). Plot (c) is the radial error $\rho_e = \rho^0 - \tilde{\rho}^0$.

Chapter 3

Results of Chirikov's Standard Map and the Harper Map

In this chapter, we present results for the shear maps presented in §2.3. In particular, we investigate the Chirikov standard map (2.8) and the Harper map (2.9). The results presented below can be found in the author's and Meiss's paper [38].

3.1 Chirikov Map

In this section, we discuss various techniques to find optimal mixing protocols for a sequence of Chirikov maps (2.8). For most of the results in this section, we use an initial Gaussian profile,

$$\rho^0(z) = \exp\left(-\varepsilon\|z - z_o\|^2\right), \quad z_o = (0.5, 0.5), \quad \varepsilon = 10. \tag{3.1}$$

This distribution has standard deviation $\sigma = \sqrt{\frac{1}{2\varepsilon}} \approx 0.2236$. We will also assume that

$$T = 12, \quad \text{and} \quad \sum_{t=1}^{12} a_t^2 = 0.03, \tag{3.2}$$

so that the total energy (2.12) is $E_C = 12.03$. Unless otherwise specified, we use a 301×301 grid and choose the maximum Fourier mode $M = 150$ for the computations.

3.1.1 Single Step Stirring

If the initial state consists of a single, nonconstant Fourier mode, $\rho^0(x, y) = \hat{\rho}^0_{k,l} e^{2\pi i(kx+ly)}$, then, using (2.20), the squared seminorm (2.13) at $t = 1$ is

$$\|\rho^1\|^2_{-1} = |\hat{\rho}^0_{k,l}|^2 \sum_{m=-\infty}^{\infty} \frac{J^2_{m-k}(2\pi l a_1)}{m^2 + (m-l)^2}. \tag{3.3}$$

This results in a function that asymptotically decreases as the shear amplitude, a_1, grows; however, it does not decrease monotonically because of the oscillations of the Bessel functions, see Fig. 3.1(a). Indeed, if the total energy (2.10) is small enough, then the optimally mixed protocol, after one step, may be that with $a_1 = 0$.

When a full spectrum initial condition is used in (2.20), the oscillations seen in Fig. 3.1(a) are mostly averaged away. Figure 3.1(b) shows the seminorm for Cauchy (black curve) and Gaussian (red curve) initial spectra. Note that the resulting curves are not symmetric about $a_1 = 0$; for these real initial spectra, a positive shear is more effective than a negative shear. Of course, even for $a_1 = 0$, the density is still sheared since s_x is not the identity for (2.8). For large shear, the Sobolev norm decreases, but—though it is not seen in the figure—the decrease saturates. As $a_1 \to \infty$, $\|\rho^1\|_{-1} \to 0.758$, for the Gaussian initial spectrum, and $\to 0.809$ for the Cauchy distribution. It is not possible to "completely" mix after one step, even with arbitrarily large energy.

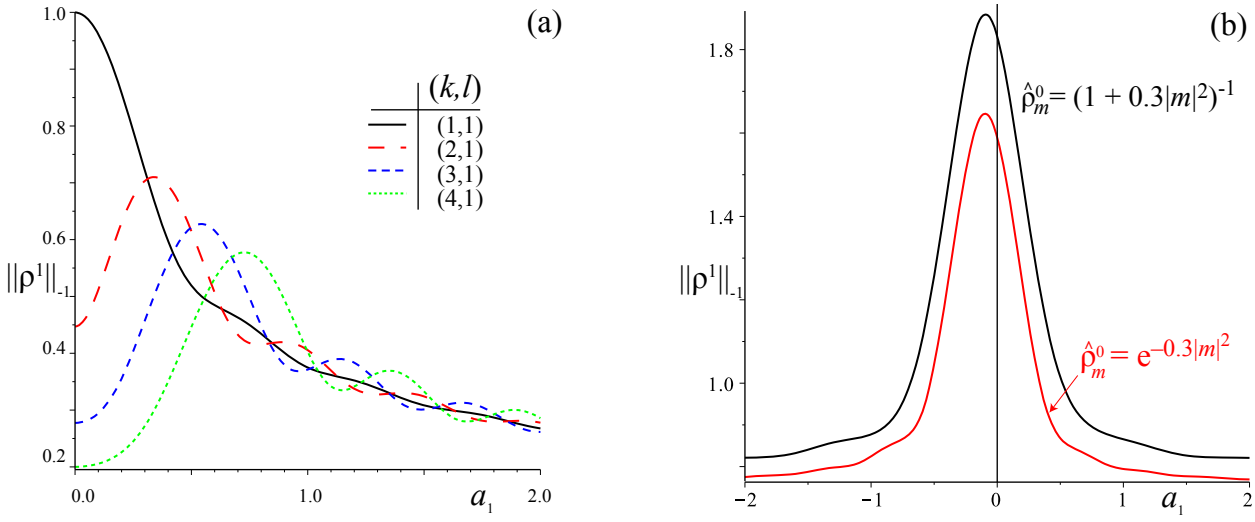

Figure 3.1: The H^{-1} seminorm (3.3) after one step of the standard map for initial state $\rho^0(x, y) = \exp(2\pi i(kx+ly))$ (left panel) and one that has a Cauchy or Gaussian Fourier spectrum (right panel) as a function of the shear strength a_1.

3.1.2 Constant Amplitude Stirring

When the amplitude is constant, $a = (a_o, a_o, \ldots, a_o)$, and satisfies the energy constraint (3.2), then $a_0 = 0.05$. If we think of this as the autonomous standard map, then its phase portrait is almost completely regular; its only large-scale features are elliptic islands around the fixed point at $(x, y) = (0, 0)$ and the period-two orbit at $(\frac{1}{2}, \frac{1}{2})$, see Fig. 3.2. The initial density (3.1) is centered over this period-two orbit, and evolves as shown in Fig. 3.3. The computations were done using the grid method (2.22) on an $N = 2M + 1 = 301$ squared grid. It is clear from the figure that the stirring is dominated by the horizontal shear s_x with $Y(y) = y$ from (2.8), though some stretching from the saddle fixed point at $(\frac{1}{2}, 0)$ can also be seen.

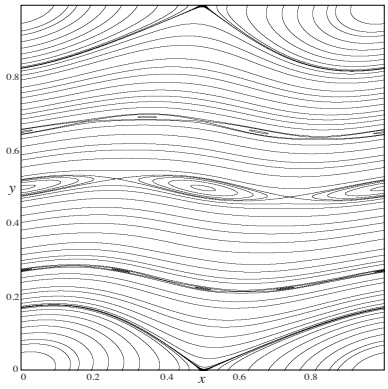

Figure 3.2: Phase portrait of the autonomous Chirikov standard map with $a_t = a_0 = 0.05$.

Figure 3.4 shows the decay of the Sobolev norm (2.13) with time. Here the computations were done with a varying number of grid points N, as shown in the figure. Note that as N is increased, the computations quickly converge: the value of $\|\rho^{12}\|_{-1}$ varies by less than 10^{-3} for $N \geq 51$. This indicates that our standard choice $N = 301$ gives sufficient accuracy for $T = 12$. The decrease in the mix-norm saturates with time, just as the single step case of §3.1.1 saturated with increasing amplitude. Repeatedly reapplying the same shear amplitudes may not be the best mixing strategy. This constant case will serve as the benchmark of comparison for our optimization methods.

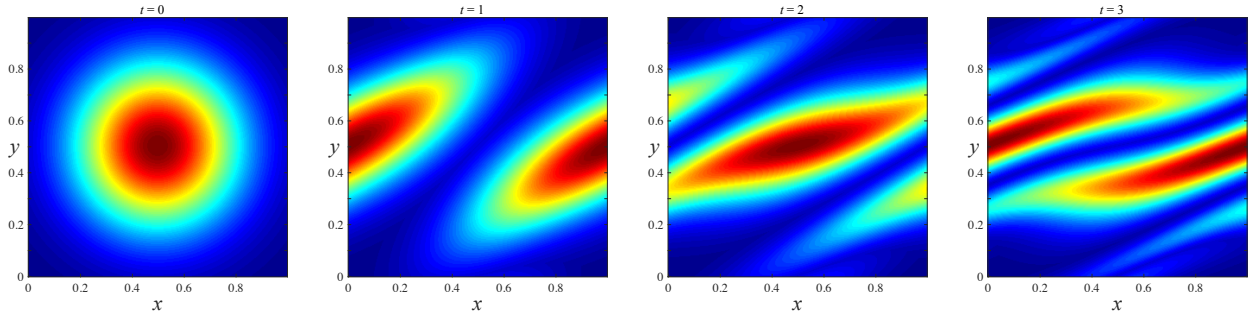

Figure 3.3: Evolution of the density $\rho^t(x, y)$ for $t = 0, 1, 2, 3$, applying the standard map (2.8) with constant amplitude $a_t = 0.05$ with initial condition (3.1) using $N = 301$ grid points. For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/ChirikovConst.mp4`.

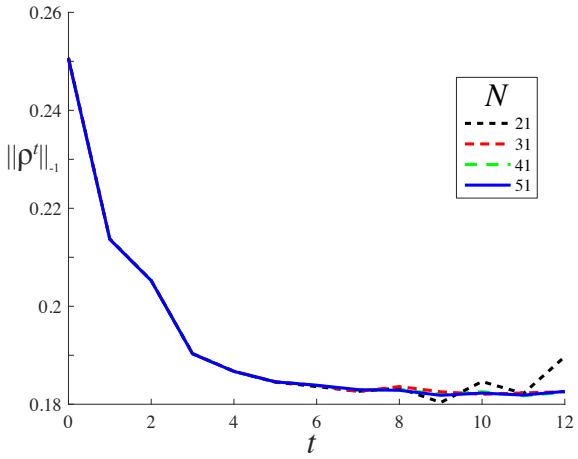

Figure 3.4: Decay of $\|\rho^t\|_{-1}$ for the standard map (2.8) with constant amplitude $(a_0, a_0, \ldots, a_0)$. The different curves are computed with grid sizes N, as shown in the legend.

3.1.3 Varying Amplitudes: Random Optimization Method

Choosing parameters at random from some distribution can serve as a crude optimization scheme: simply select the best protocol—the solution with smallest seminorm (2.13)—from the trials. For the standard map (2.8), there is one parameter, a_t, for each step of the map, so the parameter space will have $n = T$ dimensions. For convenience we think of the stirring parameters as a vector, letting

$$a = (a_1, a_2, \ldots, a_T)$$

denote a given stirring protocol. The energy (2.12) then becomes

$$T + \|a\|^2 = E_C.$$

To satisfy the energy constraint, we must sample amplitude vectors from the surface of a ball in $\mathbb{R}^n$. To ensure uniform sampling on $\mathbb{S}^{n-1}$, we randomly select n-dimensional vectors with components independently chosen from a normal distribution, centered at zero with standard deviation one. Scaling each realization to ensure it has the correct energy then gives a distribution that is uniform on the energy sphere [34]. Of the samples, we select the protocol that achieves the smallest mix norm. This process will be referred to as the Random Optimal (RO) algorithm.

Figure 3.5(a) shows the distribution of $\|\rho^T\|_{-1}$ for some two million randomly chosen amplitude vectors with the fixed energy (3.2) using the Gaussian initial profile (3.1). Almost all of the trials, 94.8%, perform better than the constant protocol, and the best case achieves the norm $\|\rho^T\|_{-1} = 0.0545$, which is 3.35 times smaller than the constant protocol. We will refer to this vector as the "random optimal protocol", even though it is certainly not truly optimal.

The evolution of $\|\rho^t\|_{-1}$ is shown in Fig. 3.5(b) for the constant protocol, the RO protocol, and the worst performing solution of the random trials, which we call "RW". Though the seminorm for the RW case decreases on the first step, it remains essentially constant after that. Aside from the first step of the map, the optimal trial significantly outperforms the constant case.

The twenty most successful amplitude vectors are shown in Fig. 3.6. For these vectors, the largest magnitudes of a_t occur primarily in steps $t = 2$ or 3; the first step and all following steps have smaller amplitudes. This suggests a simple scheme: concentrate the energy in the vertical shear in a small number of key steps of the map. We investigate this next.

3.1.4 Sparse Stirring

Since the most successful protocols found at random have energy concentrated in a small number of map steps, we can modify the optimization scheme to assume sparsity in a. Fixing $d \leq T$ to be the number of nonzero a_t, we first randomly select d values from the set $\{1, 2, \ldots, T\}$

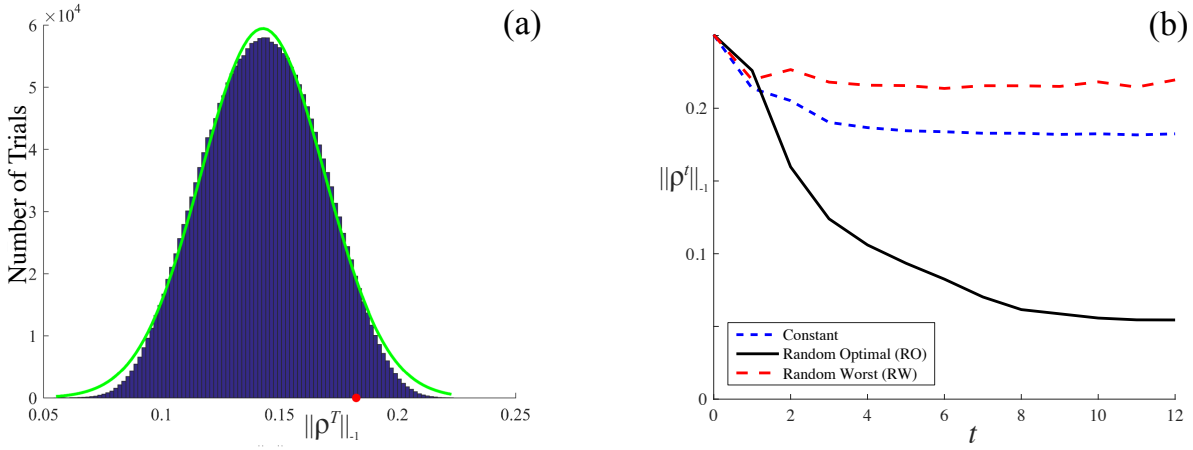

Figure 3.5: (a) Histogram of $\|\rho^T\|_{-1}$ for $\sim 2.2 \times 10^6$ trials of the RO algorithm applied to the standard map (2.8) with $E_C = T + 0.03$, $T = 12$, and initial condition (3.1). The green curve indicates the best fit normal distribution with mean 0.1429 and standard deviation 0.03704. The red point indicates the norm achieved for the constant case $(a_0, \ldots, a_0)$. (b) Comparison between the mix-norms as a function of time for the constant, RO, and RW protocols.

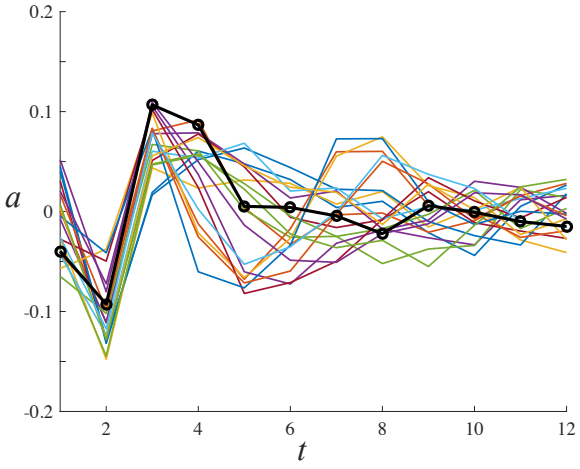

Figure 3.6: Amplitude vectors for the best twenty protocols found from the RO algorithm described in Fig. 3.5. The black curve indicates the RO protocol and a movie of the spatial evolution of this case is at `http:/amath.colorado.edu/faculty/jdm/movies/ChirikovMC.mp4`

to correspond to the nonzero a_t, and then perform the random optimization as described above, but with $n = d$. For $d = 1$, there are simply $2 \times 12 = 24$ discrete choices, since exactly one of the $a_t = \pm\sqrt{0.03}$. For $d > 1$, in addition to the discrete selection of $\binom{12}{d}$ shearing times, we must choose sparse vectors on $\mathbb{S}^{d-1}$ to fix the energy. Consequently the number of trials should increase

with d. The best amplitude vectors that we found for each case are shown in Fig. 3.7; the optimal protocol for $d = 1$ corresponds to shearing on the second step with $a_2 = -\sqrt{0.03} \approx -0.173$. The performance of the optimal protocols as d varies is shown in Fig. 3.8; note that $\|\rho^t\|_{-1}$ does not not depend significantly on d. Moreover, the optimal sparse protocols achieve even a slightly smaller mix-norm than the full set of random trials. This shows that the latter technique did not find the true optimum, even with more than two million trials.

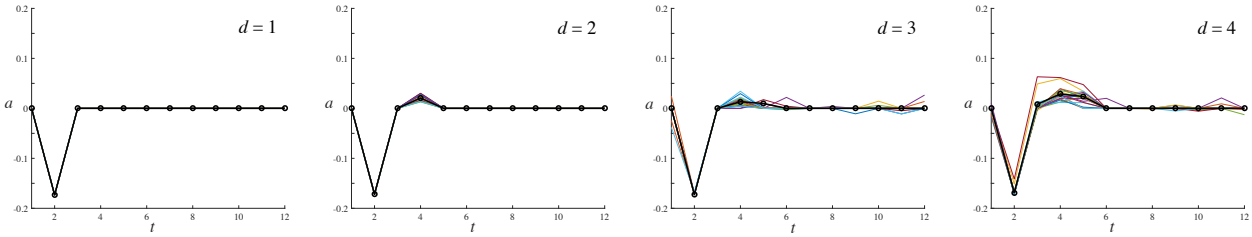

Figure 3.7: Amplitude vectors of the best ($d = 1$) and the best twenty ($d > 1$) protocols found from the sparse RO method. The open circles/black line indicates the optimal protocol for each case.

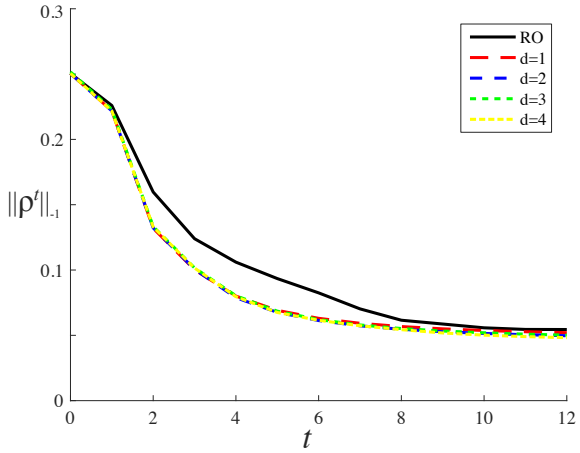

Figure 3.8: Comparison of $\|\rho^t\|_{-1}$ for the random optimal protocol with the sparse optimal protocols for $d = 1, 2, 3, 4$.

If we think of the composite map $F_T = f_T \circ \cdots \circ f_1$ as a single, area-preserving map, for $d = 1$

and $a_2 \neq 0$ it has the form

$$F_T(x,y) = \begin{cases} x + y + (T-1)[y - a_2 \sin(2\pi(x+y))] \\ y - a_2 \sin(2\pi(x+y)) \end{cases} \tag{3.4}$$

when $T \geq 2$. Defining new variables $(\xi, \eta) = (x+y, Ty)$, this reduces to a single Chirikov standard map in (ξ, η) with effective amplitude Ta_2. Several phase portraits of the map (3.4) are shown in Fig. 3.9. As T increases the elliptic regions for this map decrease in size in contrast to the regularity of the constant amplitude case which has the same total energy. Indeed, when $T = 12$ the dynamics are nearly uniformly chaotic since, $12|a_2| \approx 2.08 \gg 0.972/(2\pi) \approx 0.155$, which is the threshold for global chaos for the standard map [37]. This is one reason why the sparse case is such an effective stirrer.

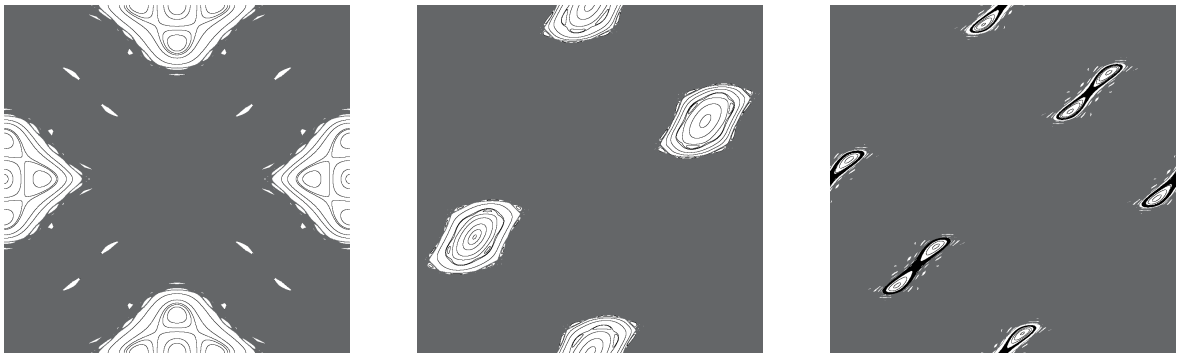

Figure 3.9: Phase portraits of the map (3.4) for $T = 2$, 3, and 4 with $a_2 = -0.1732$. For $T \geq 6$ (not shown), the dynamics appears to be nearly uniformly chaotic, except for small period-six island chains along lines of slope -1.

The sparse amplitude vectors which performed the worst are shown in Fig. 3.10. In general, the sparse amplitude vectors, a, that maximize the final mix-norm have $a_2 > 0$, the opposite of the optimal sparse. These protocols vary more with d than the optimal protocols did, but the difference in the final mix-norms is small. For the worst $d = 1$ case $\|\rho^T\|_{-1} = 0.2041$, which is less than that of the full RW case shown in Fig. 3.5. A sparse amplitude vector is not as effective in inhibiting mixing as one that is more nearly constant.

To further understand stirring in the $d = 1$ sparse case, we can observe the effects of the

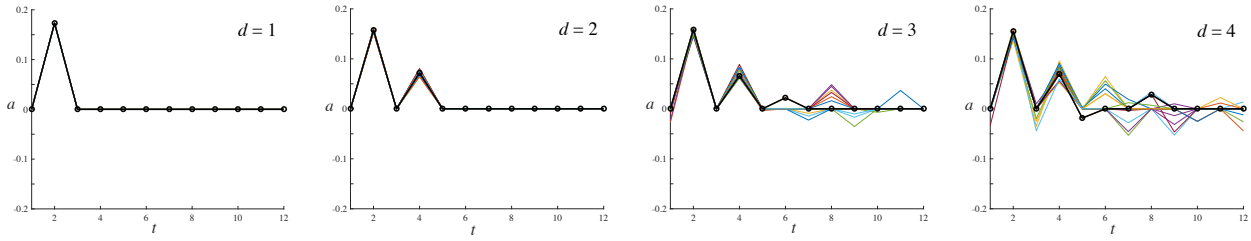

Figure 3.10: Amplitude vectors for the worst ($d = 1$) and the worst twenty ($d > 1$) stirring protocols up to $d = 4$. The open circles/black line indicates the worst protocol for each case.

horizontal shear in Fourier space. When $a = 0$, (2.20) gives

$$\hat{\rho}_m^t = \sum_{k\in\mathbb{Z}} \hat{\rho}_{k,m_1+m_2}^{t-1} J_{m_1-k}(0) = \hat{\rho}_{m_1,m_1+m_2}^{t-1}.$$

This implies that each Fourier coefficient has its y-wavenumber shifted by m_1 and, in particular, that the coefficients $\hat{\rho}_{0,m_2}^t$ remain unchanged. The effect of applying the map with $a = 0$ is that the Fourier amplitudes with $m_1 > 0$ are shifted up in y-wavenumber; thus, for a smooth initial condition, the amplitudes of the low modes become asymptotically small. The effect of one step of horizontal shear is shown in real space and Fourier space in Fig. 3.11. Of course the $m_1 = 0$ amplitudes are unaffected, so an optimal protocol must use the nonzero a_t to minimize the magnitude of these coefficients. Conversely, a poor protocol must accomplish the opposite.

Let us now consider the effect of a single nonzero a_t on the amplitudes with zero y wave number. Since the horizontal shear has no effect on these modes, we may write the updated coefficients in terms of the initial vertically shearing step. Using (2.20) the coefficients $\hat{\rho}_{0,m_2}$ are given by

$$\hat{\rho}_{0,m_2}^t = \sum_{k\in\mathbb{Z}} \hat{\rho}_{k,m_2}^{t-1} J_k(-2\pi m_2 a_t).$$

After one step with $a_0 = 0$ (Fig. 3.11), the largest Fourier coefficients from the initial condition (3.1) with $m_1 = 0$ are $\hat{\rho}_{0,\pm1} \approx -0.1204$. For a nonzero a_2, the $(0, 1)$ mode becomes

$$\hat{\rho}_{0,1}^2 \approx \hat{\rho}_{-2,1}^1 J_{-2}(-2\pi a_2) + \hat{\rho}_{-1,1}^1 J_{-1}(-2\pi a_2) + \hat{\rho}_{0,1}^1 J_0(-2\pi a_2) + \hat{\rho}_{1,1}^1 J_1(-2\pi a_2) + \hat{\rho}_{2,1}^1 J_2(-2\pi a_2), \quad (3.5)$$

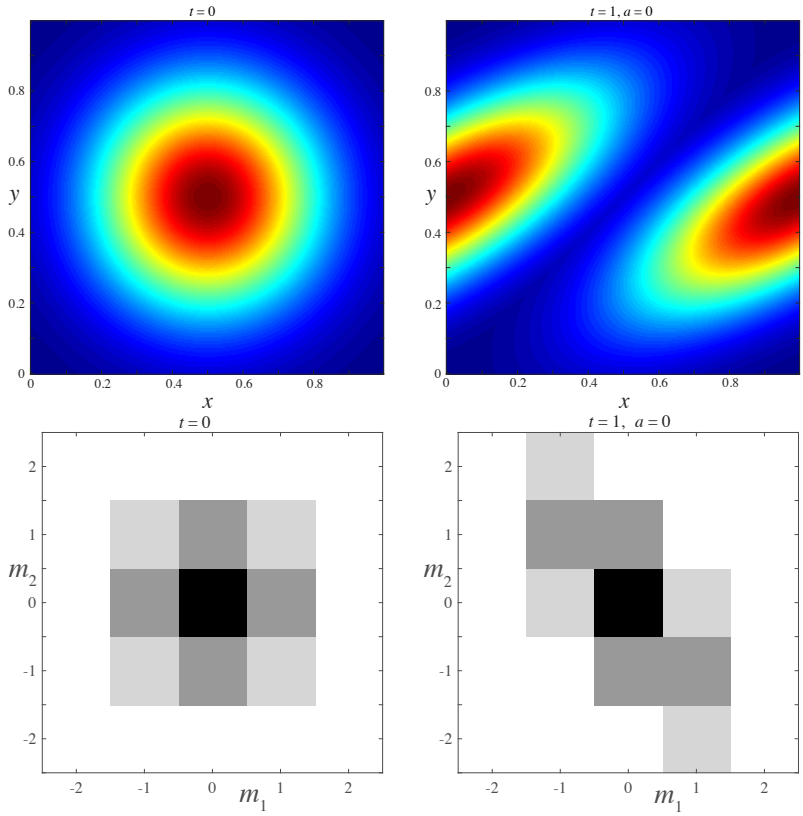

Figure 3.11: Evolution of $\rho^t(x, y)$ after one step of shear ($a = 0$). Top row shows evolution in real space. Bottom row shows the evolution of magnitudes of Fourier coefficients of degree up to 2. One step of shear results in a vertical shift of the Fourier coefficients, except for those corresponding to x wavenumber $= 0$, which remain constant.

where we truncate at the second mode number. For the optimal case, $a_2 = -0.1732$, (3.5) gives

$$\begin{aligned}\hat{\rho}^2_{0,1} &\approx (-0.0008)(0.1340) + (-0.1204)(-0.4675) + (-0.1204)(0.7251) \\ &\quad + (-0.0008)(0.4675) + (0.0000)(0.1340) \\ &\approx -0.0315.\end{aligned}$$

For this negative a_2, the Bessel functions J_{-1} and J_0 have opposite signs. This causes the largest terms to nearly cancel, causing the value $\hat{\rho}^2_{0,1}$ to decrease by a factor of 4. This can be seen in Fig. 3.12, which shows the progression of the density in both physical space (top row) and Fourier space (bottom row). The larger Fourier amplitudes in the figure are shown as darker squares, and moving from the $t = 1$ column, to the $t = 2$ column, shows that the $(0, \pm 1)$ mode amplitudes

decrease significantly. The following steps of shear then push the other large coefficients, the $(\pm 1, \mp 2)$ modes in the figure, to larger y-wavenumbers, so that the magnitude of the low Fourier modes diminishes significantly after just the four time steps shown.

By contrast, the $d = 1$ RW protocol has the opposite effect. When $a_2 = +0.1732$, (3.5) gives

$$\begin{aligned}\hat{\rho}^2_{0,1} &\approx (-0.0008)(0.1340) + (-0.1204)(0.4675) + (-0.1204)(0.7251) \\ &\quad + (-0.0008)(-0.4675) + (0.0000)(0.1340) \\ &\approx -0.1433\end{aligned}$$

The sign change causes this mode amplitude to grow in magnitude, resulting in a clumping of density that cannot be undone by the remaining steps of purely horizontal shear. This is illustrated in Fig. 3.13, which shows the first four steps of the sparse RW protocol in physical and Fourier space.

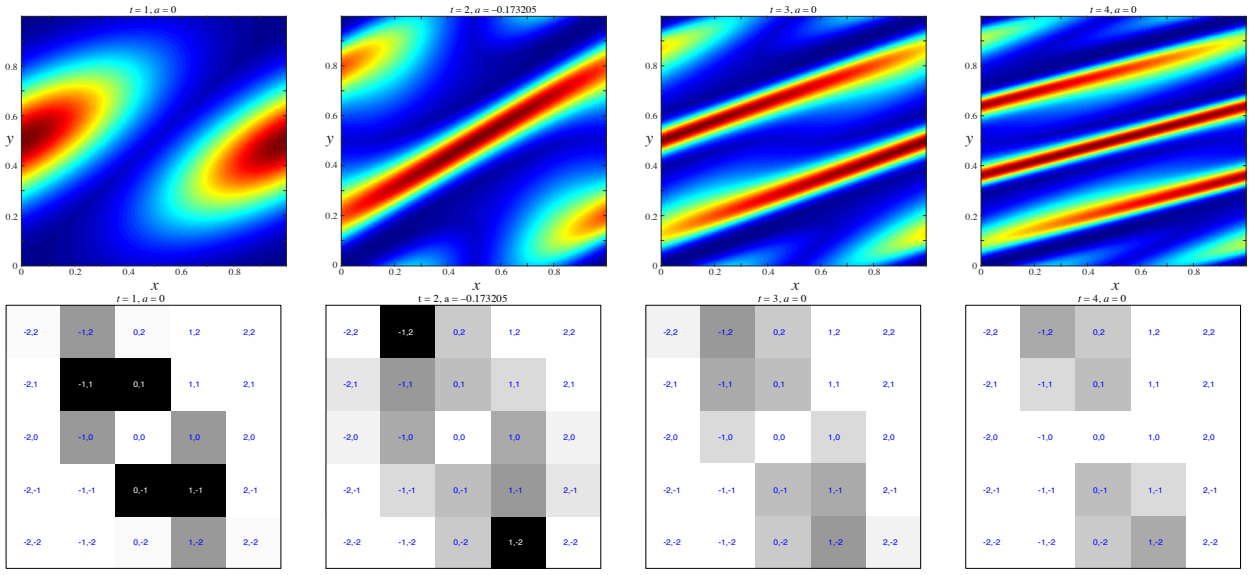

Figure 3.12: Evolution of $\rho^t(x, y)$ for the $d = 1$ sparse RO protocol of Fig. 3.7. Top row shows the evolution in physical space. Bottom row shows the evolution of magnitudes of Fourier coefficients with mode numbers up to 2; the value of the invariant $(0, 0)$ mode, $\hat{\rho}_{0,0}$ set to zero. The numbered pairs indicate the wave numbers and the greyscale indicates the coefficient magnitude, black being the maximum and white being zero. For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/ChirikovSparse.mp4`.

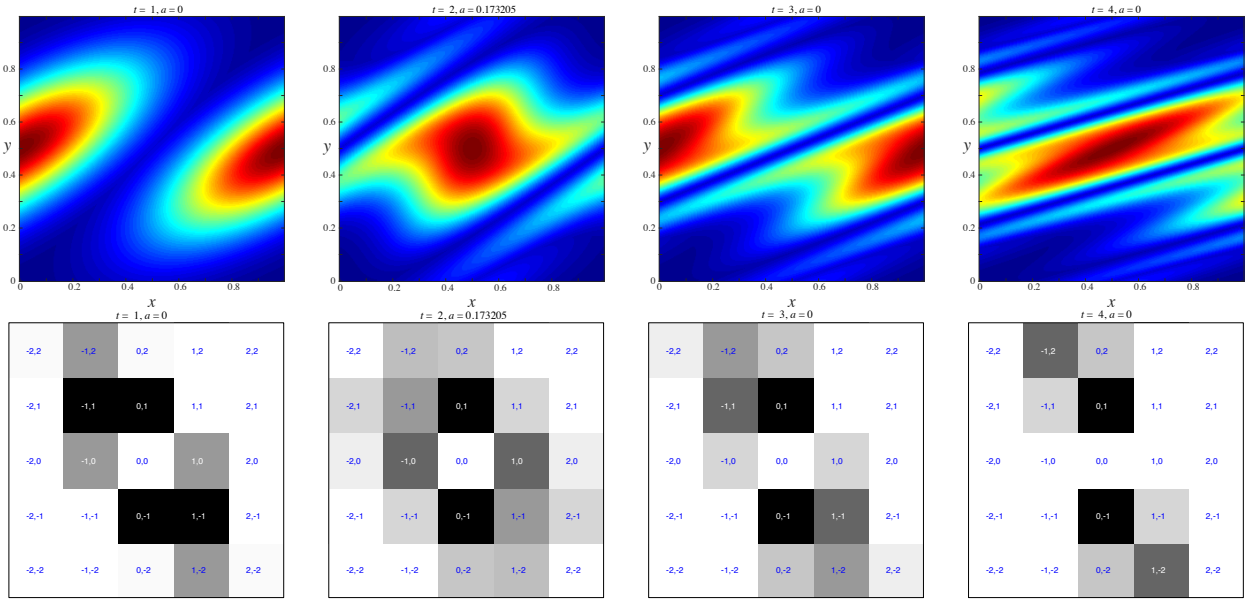

Figure 3.13: Evolution of $\rho^t(x, y)$ for the $d = 1$ sparse RW protocol. The eight panels are otherwise equivalent to those in Fig. 3.12. For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/ChirikovSparseBad.mp4`.

3.1.5 Other Initial Conditions

Given this intuition for optimizing mixing for the Gaussian initial density (3.1), we now determine if the sparse protocol is still effective for other initial conditions. We study initial densities constructed by randomly superimposing periodic bump functions of the form

$$f(x, y; C, \varepsilon, u, v) = Ce^{-\varepsilon \sin^2[\pi(x-u)] - \varepsilon \sin^2[\pi(y-v)]}, \tag{3.6}$$

so that the center of the bump can range over the torus. Adding a set of sinusoidal perturbations, we construct the density by defining

$$\rho^0(x, y) = \sum_{k=1}^{K} f(x, y; C_k, \varepsilon_k, u_k, v_k) + \sum_{l=1}^{L} C_l \sin(2\pi n_l \cdot z - \phi_l). \tag{3.7}$$

We randomly choose the centers $(u, v) \in [0, 1)^2$, the widths $\varepsilon \in [0.1, 5]$, and K and L from 1 to 5. The wavenumbers n_l ranged over $\{-10, 10\}^2$ and the phases ϕ_l over $[0, 2\pi)$. The amplitudes C_l of the sine terms are kept smaller than 20% of the maximum bump amplitude (to create a noisy effect), and each of the function (3.7) is scaled so that $\|\rho^0\|_{-1} = 1$. Examples of density profiles of

the form (3.7) are shown in Fig. 3.14

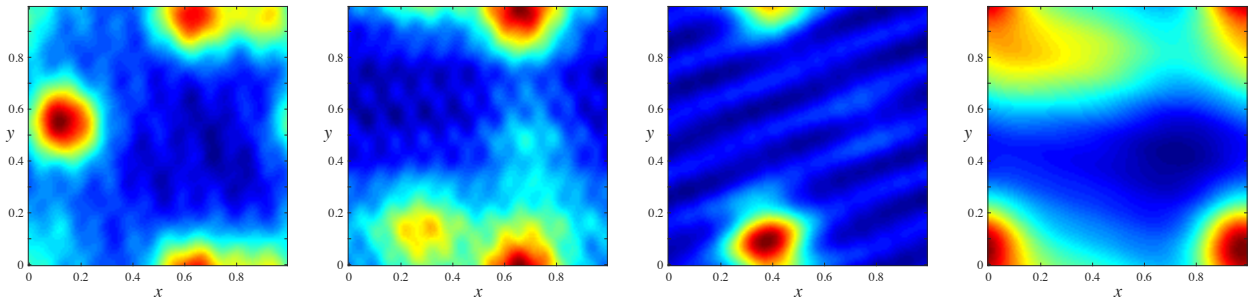

Figure 3.14: Density profiles of the form (3.7), with the randomly chosen parameters $(u, v) \in [0, 1)^2$, $\varepsilon \in [0.1, 5]$, $K, L \in \{1, 5\}$, $n_l \in \{-10, 10\}^2$, $\phi_l \in [0, 2\pi]$, C_l less that 20% of the maximum bump amplitude. Each density profile is scaled such that $\|\rho^0\|_{-1} = 1$.

The question becomes: Is a $d = 1$ sparse protocol, in which we apply all the energy at a single step, still nearly optimal for varying initial conditions? If this is the case, then the optimal $d = 2$ case, in which we divide the energy between two steps, should resemble a $d = 1$ protocol, in that the energy at one step should be much larger than that at the second. For each of the 1200 initial conditions, we found the best $d = 2$ protocol from 10^4 random trials with the energy (3.2). Figure 3.15(a) shows a histogram of step times t_{max} at which the larger of the two nonzero amplitudes, $|a_{t_{max}}|$, was applied. The majority of these optimal protocols have $t_{max} \leq 2$: as for (3.1) the largest amplitude is used early in the twelve applications of the map. Figure 3.15(b) shows the fraction of protocols for which $a^2_{t_{max}} > rE$; that is, the fraction of trials for which the largest amplitude step contains more than a fraction r of the available energy. Nearly three-quarters of protocols apply more than 70% of the available energy in one step. This supports the conclusion that the scheme applying a large shear early in the evolution is nearly optimal for the majority of initial conditions.

3.2 Harper Map

The Harper map (2.9) differs from the standard map in that the horizontal shear is nonlinear with the additional amplitude parameter b. The mixing protocol is now determined by two param-

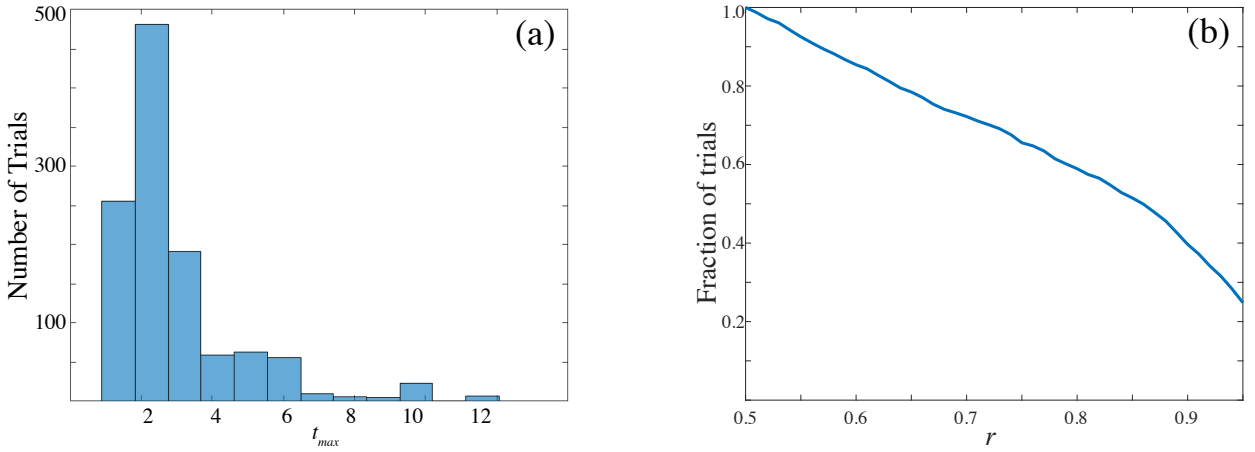

Figure 3.15: (a) Histogram of time at which largest $|a_t|$ occurs for $d = 2$ RO protocols. (b) The fraction of the RO protocols that satisfy $a^2_{t_{max}} > rE$ as a function of r. We perform 10^4 trials to find the $d = 2$ RO protocols for 1200 different initial conditions of the form (3.7) on a grid of size $N = 128$ with (3.2).

eter vectors a and b so that the dimensionality of the optimization problem is doubled, allowing for more flexibility. Moreover, since b_t can vary, the energy (2.11) in the horizontal shear is no longer fixed as it was for (2.12). In this section, we search for an optimal mixing protocol by means of the random optimization method and also develop other methods that require less computation.

For most of the results in this section, we again use the Gaussian density profile (3.1), the time constraint $T = 12$, and fix

$$E_H = \|a\|^2 + \|b\|^2 = 0.5.$$

Unless otherwise specified, we use a 301×301 grid, setting $M = 150$.

3.2.1 Fixed Points

Whenever $ab \neq 0$, the Harper map has exactly four fixed points, $(0,0)$, $(0,\frac{1}{2})$, $(\frac{1}{2},0)$ and $(\frac{1}{2},\frac{1}{2})$. For small a and b the map can be approximated by the Hamiltonian system

$$(\dot{x},\dot{y}) = (\partial_y H, -\partial_x H), \quad H(x,y) = -\tfrac{a}{2\pi}\cos(2\pi x) - \tfrac{b}{2\pi}\cos(2\pi y). \tag{3.8}$$

Contours of this Hamiltonian are shown in Fig. 3.16 for various parameters. For the flow, the stability of the fixed points depends only on the relative signs of a and b. When a and b are of the

same sign, the points $(0,0)$ and $(\frac{1}{2},\frac{1}{2})$ are elliptic and $(0,\frac{1}{2})$ and $(\frac{1}{2},0)$ are hyperbolic. When a and b are of opposite sign, the stabilities are reversed. Note that $(\frac{1}{2},\frac{1}{2})$, the center of the domain, is elliptic when $ab > 0$ and hyperbolic when $ab < 0$. When $|a| > |b|$ there is more vertical shearing; conversely, when $|b| > |a|$, the horizontal shear is dominant.

The dynamics of the map (2.6) with (2.9)—for fixed parameter values—is similar to the Hamiltonian flow when $|a|$, $|b| \ll 1$, except for chaotic layers near the separatrices; two examples are shown in Fig. 3.17. For the map, the elliptic fixed points undergo a period doubling bifurcation at $|ab| = \pi^{-2}$, becoming hyperbolic with reflection. However, for the total energy levels that we will choose, none of the shear amplitudes will be large enough for this to occur.

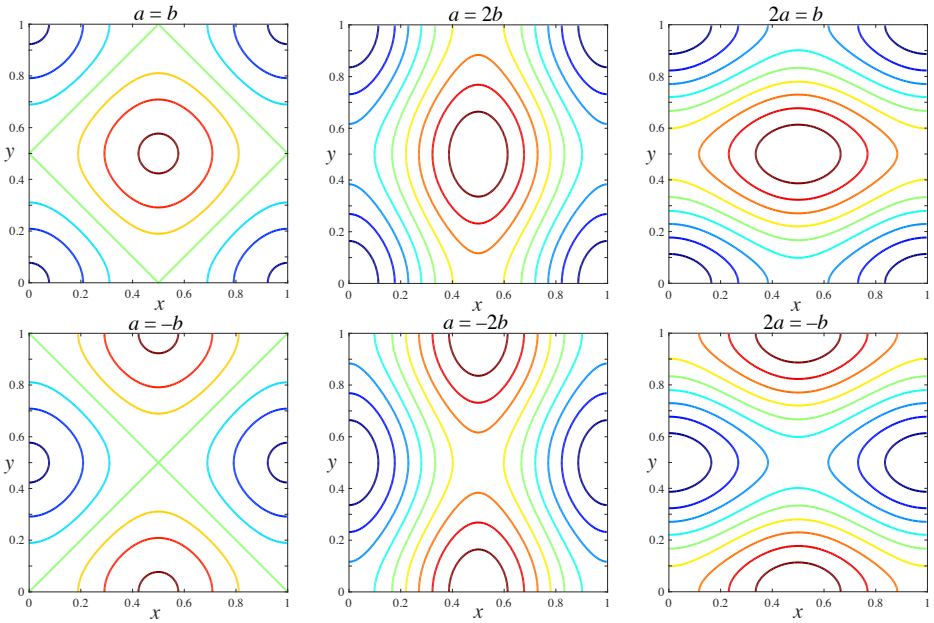

Figure 3.16: Level sets of the Hamiltonian (3.8) with various relative values of (a, b).

3.2.2 Constant Shear Amplitude

If the shears have constant amplitudes $|a_t| = |b_t| = \gamma_o$, then the energy constraint (2.11) implies that $\gamma_o = \sqrt{E_H/2T}$. It would seem to be counterproductive to have successive signs of the horizontal (or vertical) shears oscillate, so we will consider here two characteristic cases,

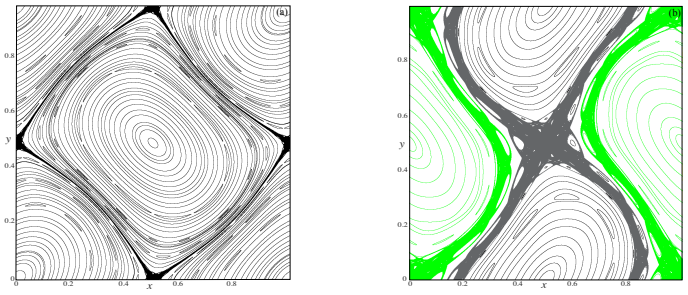

Figure 3.17: Phase portraits of the Harper map (2.9) with constant amplitudes, (a) $a_t = b_t = 0.1443$ and (b) $(a_t, b_t) = (-0.211881, 0.15726)$. With the relatively small amplitudes in (a) the evolution resembles the Hamiltonian flow shown in Fig. 3.16, with a thin chaotic separatrix layer. For the larger amplitudes in (b) there is more chaos.

$a_t = b_t = \gamma_o$, shown in the linked movie (`http:/amath.colorado.edu/faculty/jdm/movies/HarperConst.mp4`), and $a_t = -b_t = \gamma_o$, shown in the linked movie (`http:/amath.colorado.edu/faculty/jdm/movies/HarperConstNeg.mp4`) The contours of the Hamiltonian (3.8) are also shown in these movies to indicate the approximate dynamics of the next map to be applied—indeed as Fig. 3.17(a) shows, these contours reasonably approximate the dynamics since the amplitude, $\gamma_o = 0.1443$, is small when $E_H = 0.5$.

Since the initial density is centered at $(\frac{1}{2}, \frac{1}{2})$, which is elliptic when a and b have the same sign, and the standard deviation of the Gaussian profile (3.1) is 0.2236, which is not too large, the dominant dynamics seen for $a_t = b_t$ is a sheared rotation about the center which results in very little mixing, as shown in Fig. 3.18. By contrast, when a and b have opposite signs, the point $(\frac{1}{2}, \frac{1}{2})$ is a saddle, so the density is stretched along the unstable manifold until it reaches the saddle at $(0, 0) \equiv (1, 1)$, where it divides and then moves back towards $(\frac{1}{2}, \frac{1}{2})$ along its stable manifolds, as shown in Fig. 3.19.

Figure 3.20 shows the decay of the mix-norm for these two cases and various values of the maximum mode number, M. As with the standard map, the decrease in the mix-norm slows as t increases. As the grid size increases, the values of $\|\rho^T\|_{-1}$ converge—little difference is seen when $M > 25$, well below our standard resolution.

Among the four possible constant amplitude cases, $(a_t = \pm 0.1443, b_t = \pm 0.1443)$ for $E_H =$

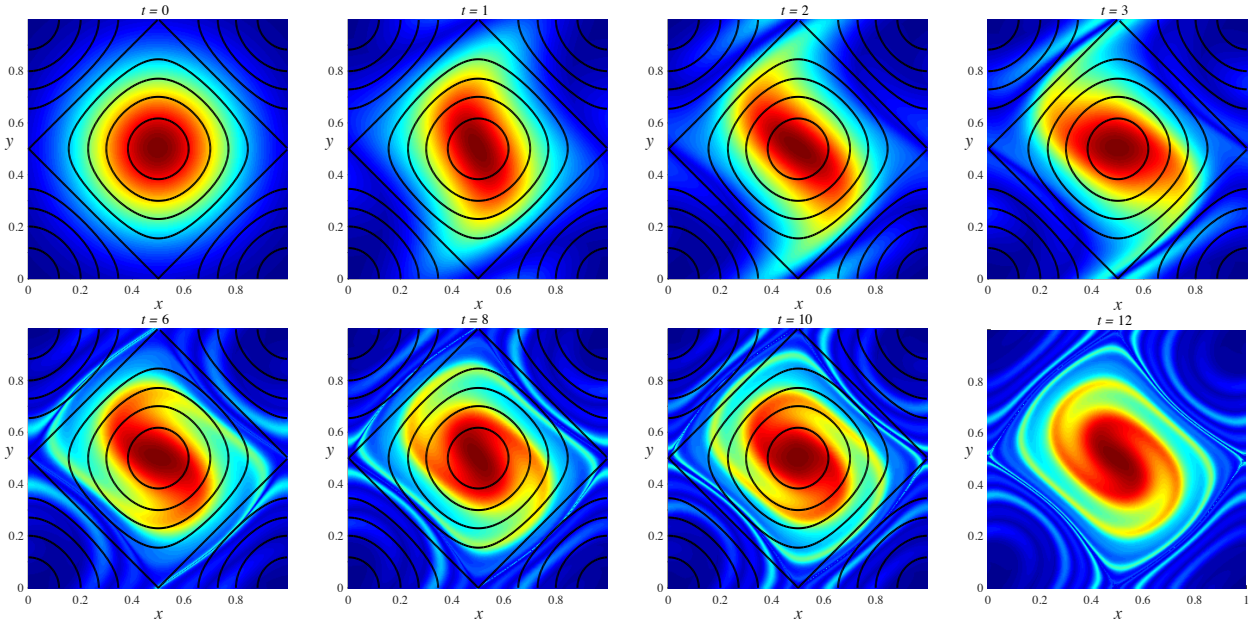

Figure 3.18: Evolution of $\rho^t(x, y)$ for the Harper map (2.9) with initial condition (3.1) for $t = 0, 1, 2, 3$ (top row) and $t = 6, 8, 10, 12$ (bottom row) with constant strengths $a_t = b_t = 0.1443$. Contours correspond to the Hamiltonian (3.8). For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/HarperConst.mp4`.

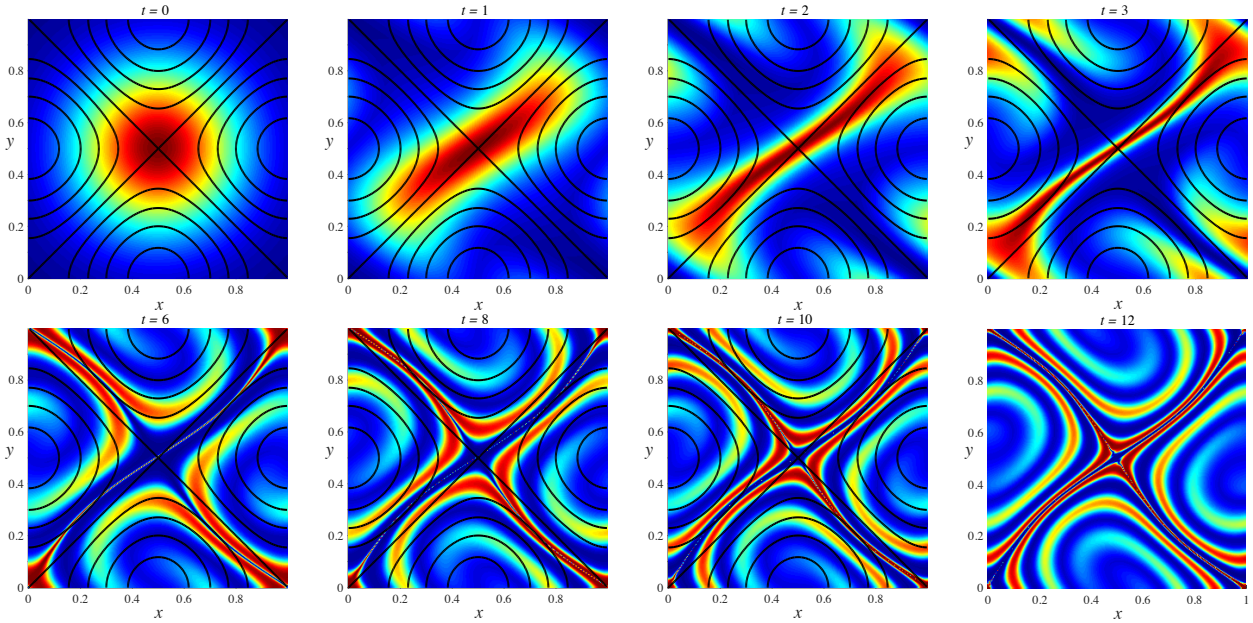

Figure 3.19: Evolution of $\rho^t(x, y)$ for the Harper map (2.9) with initial condition (3.1) for $t = 0, 1, 2, 3$ (top row) and $t = 6, 8, 10, 12$ (bottom row) with constant strengths $a_t = -b_t = 0.1443$. Contours correspond to the Hamiltonian (3.8). For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/HarperConstNeg.mp4`.

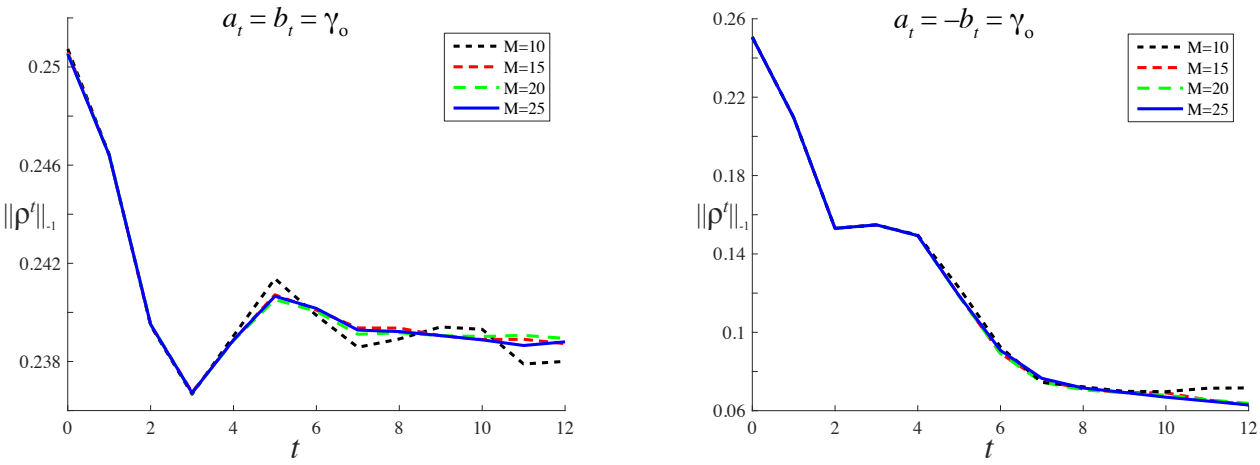

Figure 3.20: Decay of $\|\rho^t\|_{-1}$ for the Harper map (2.9) with initial condition (3.1) for twelve iterates with constant strength $a_t = b_t = 0.1443$ (left) and $a_t = -b_t = 0.1443$ (right). The values in the legend denote the maximum mode number, M, with grid size $N = 2M + 1$.

0.5, the optimal mixing protocol for the centered Gaussian (3.1) corresponds to the opposite sign case, achieving $\|\rho^T\|_{-1} = 0.0636$. Of course, since the initial condition is symmetric about the fixed point, flipping both signs results in the same Sobolev norm. This raises the question: will the opposing amplitude stirring scheme be optimal for different energy budgets? The mix-norm as function of constant amplitudes $(a_t, b_t) = (a, b)$ is shown in Fig. 3.21. It is interesting that though the optimal protocols for both large and small energy have $a \approx -b$, there is an intermediate regime where this scheme is not optimal.

3.2.3 Random Optimal Protocol

Computing the optimal mixing protocol for the Harper map requires choosing the $n = 2T$ dimensional vector (a, b) on the energy sphere

$$\|a\|^2 + \|b\|^2 = E_H$$

that has the smallest $\|\rho^T\|_{-1}$. Following the same procedure as §3.1.3, Fig. 3.22 shows a histogram of the mix-norm for our standard Gaussian initial density. Note that the two constant amplitude protocols (orange and red dots in the figure) correspond to near extremes of the range of $\|\rho^T\|_{-1}$.

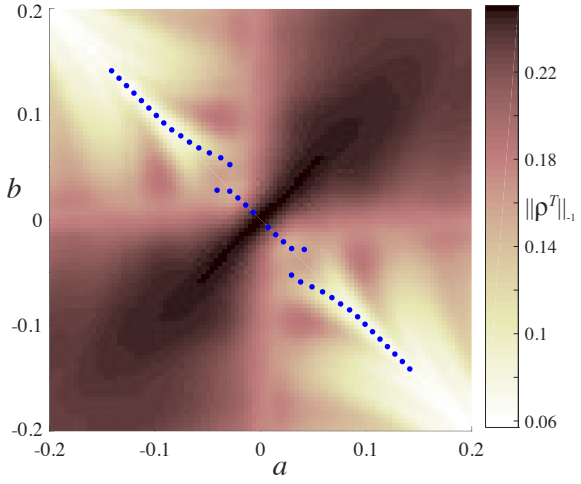

Figure 3.21: Level sets of constant mix-norm for $(a_t, b_t) = (a, b)$ with $T = 12$ and initial condition (3.1) for a grid of amplitudes in the square $-0.2 \leq a, b \leq 0.2$. The color bar shows the value of the mix-norm: dark colors correspond to less mixing, and the optimal case is white. The blue points indicate the optimal protocol as a function of energy $E_H = T(a^2 + b^2)$.

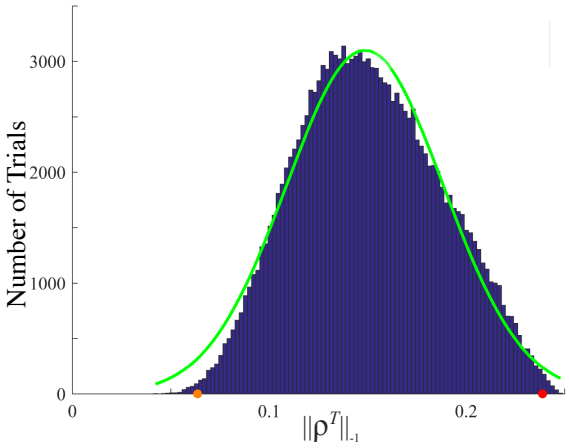

Figure 3.22: Distribution of $\|\rho^T\|_{-1}$ after $\sim 1.4\times10^6$ trials of the random optimization method applied to the Harper map (2.9) with $T = 12$ and initial condition (3.1). The green curve indicates the best fit normal distribution with mean 0.1486 and standard deviation 0.05636. The red and orange dots indicate $\|\rho^T\|_{-1}$ for the constant amplitude elliptic, Fig. 3.18, and hyperbolic, Fig. 3.19, cases respectively.

The evolution of the density for the best protocol from 1.7×10^6 trials is shown in Fig. 3.23; again we call this the RO protocol. Note that for this protocol a_1 and b_1 have opposite signs, causing the initial density to be stretched along the unstable manifold of the point $(\frac{1}{2}, \frac{1}{2})$; this

is as expected from our earlier discussion of the constant case. It is interesting that the density evolution in Fig. 3.23 exhibits not only the stretching seen in the opposite amplitude hyperbolic case, of Fig. 3.19, but also considerable folding, the second hallmark of chaos.

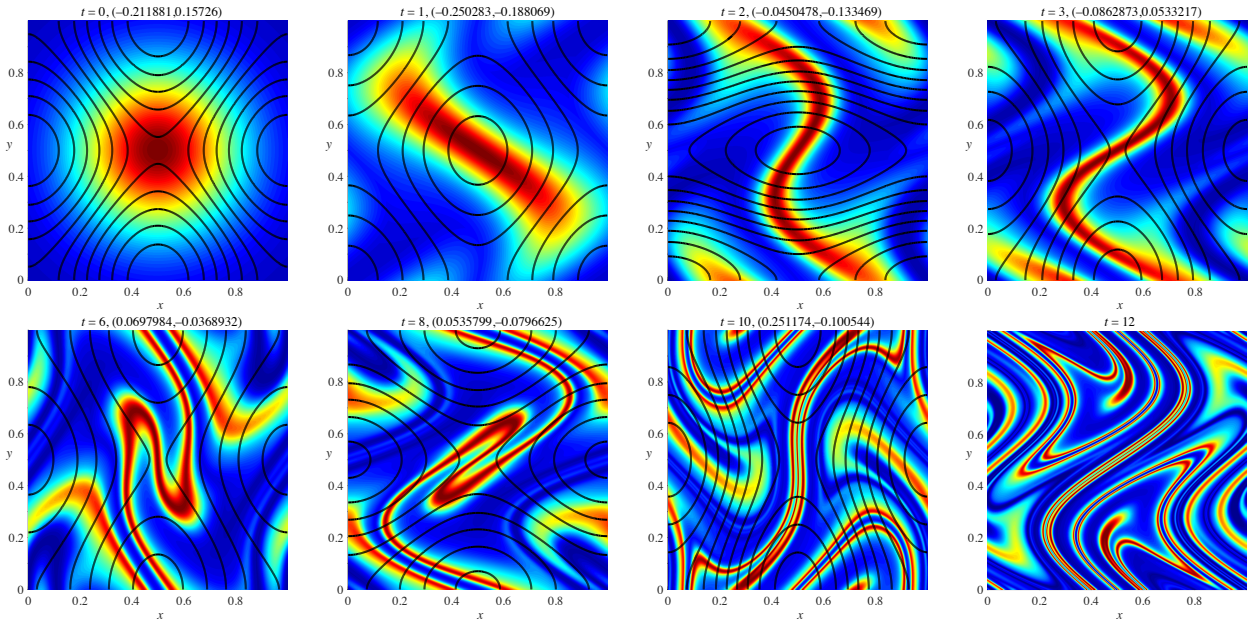

Figure 3.23: Evolution of $\rho^t(x, y)$ for the Harper map (2.9) with initial condition (3.1) for $t = 0, 1, 2, 3, 6, 8, 10, 12$ with best protocol of $\sim 1.7 \times 10^7$ random trials. Overlaid on each frame, except the last, are the contours of the Hamiltonian (3.8) for the *upcoming* map step. For a movie of all 12 steps, see `http:/amath.colorado.edu/faculty/jdm/movies/HarperMC.mp4`.

It is also important that the amplitudes of the RO protocol do not have equal magnitudes. Indeed, for several of the steps ($t = 5, 8, 10$, and 12) one of the amplitudes is at least three times smaller than the other, so that the dynamics are nearly a horizontal or vertical shear (Recall that the Hamiltonian contours for these steps are overlaid on the previous density.). In §3.2.5, we exploit this by examining stirring protocols that allow only one of the shear amplitudes to be nonzero for each step.

In addition, the energy of the shears in the RO protocol varies considerably with time. Defining the one-step energy by

$$E_t^{RO} = a_t^2 + b_t^2, \tag{3.9}$$

we see in Fig. 3.24 that most of the energy is expended in steps $t = 1$, 2, 10, and 11.

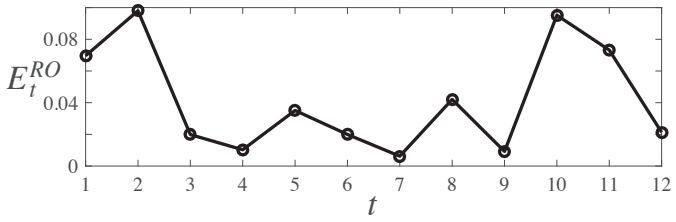

Figure 3.24: Energy profile of the RO solution for initial density (3.1), $E_H = 0.5$, and $T = 12$.

The evolution of $\|\rho^t\|_{-1}$ for the RO protocol is shown in Fig. 3.25(a); this figure also includes the two constant amplitude cases from §3.2.2. As seen in Table 3.1, at $t = 12$, the RO protocol just outperforms the hyperbolic constant case, showing that it is possible to do better than just stretching the density along the manifolds of the central fixed point. It is interesting, however, that the constant amplitude case has lower mix-norm when $4 < t < 9$. During this period, as can be seen in Fig. 3.23, the RO protocol is setting up improved stirring for the last three steps by folding the density contours; this is accomplished using disparate amplitudes of the two shears.

Scheme	$\|\rho^T\|_{-1}/\|\rho^0\|_{-1}$	% Protocols better
Elliptic Constant	0.953	99.74
CS SW	0.277	0.46
CS FS	0.271	0.37
Hyperbolic Constant	0.254	0.09
CS Full	0.217	0.02
SW	0.160	0.00
RO	0.145	0.00
CS RO	0.133	0

Table 3.1: Mix norms for the Harper map with initial condition (3.1), energy $E_H = 0.5$, and total steps $T = 12$ for the optimization schemes discussed in §3.2.2-3.2.5. The initial condition has $\|\rho^0\|_{-1} = 0.2505$. The last column gives the percentage of protocols from the random trials of Fig. 3.22 that perform better, i.e., have a smaller mix-norm.

As the energy varies, so does the achievable minimum mix-norm. The optimal protocol as a function of energy gives a Pareto frontier, shown in Fig. 3.25(b). This figure was constructed using 6.5×10^5 trials for the initial condition (3.1) spread over 20 values of $E_H \in [0, 2]$. Also shown,

for each energy, are the mix-norms for the worst trial and for the constant amplitude, hyperbolic protocol. Note that the latter protocol appears to be optimal when $E_H < 0.3$; however, for larger energies the RO protocol beats the constant case.

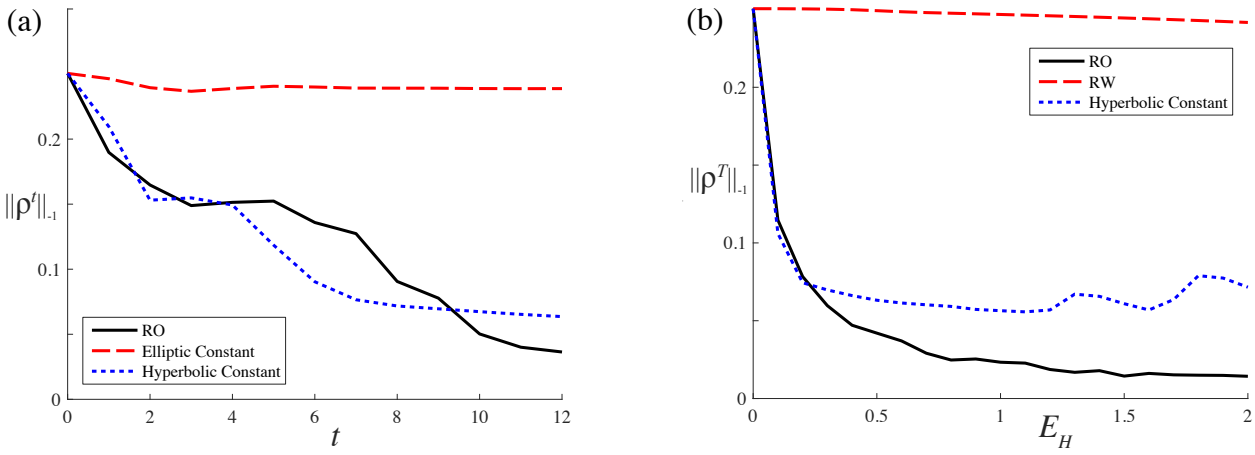

Figure 3.25: (a) Mix-norm as a function of time for the RO protocol for the Harper map (2.9) with $E_H = 0.5$, solid (black) curve. Also shown are the elliptic constant, $a_t = b_t = \gamma_o$, (dashed red curve) and the hyperbolic constant, $a_t = -b_t = \gamma_o$ (dotted blue curve) protocols of §3.2.2. (b) Pareto frontier, showing the attainable mix-norm as a function of energy for the Harper map, solid (black) curve. The dashed (red) curve shows the worst trial, RW, and the mix-norm achieved by the hyperbolic constant protocol is shown as the dotted (blue) curve.

3.2.4 Stepwise Minimization

The decay of the mix-norm for the RO protocol shown in Fig. 3.25(a) is not monotone. Though $\|\rho\|_{-1}$ decreases sharply during the first three steps, it appears to increase slightly during steps four and five, finally decreasing for the remaining steps. This is odd, since one might expect the optimal twelve-step protocol to minimize the mix-norm at each step as much as possible. Instead of considering the time-history as a whole, could it be as effective to choose a_t and b_t independently to maximize the decrease in the mix-norm from step $t-1$ to step t? This approach corresponds to the "short time horizon" optimization method used in [10, 25, 26]. For these studies, which assumed $(a_t, b_t) = (0, \tau)$ or $(\tau, 0)$ for a fixed $\tau > 0$, the one-step optimization performed nearly as well as multistep look-ahead protocols.

For simplicity, we will look for a protocol that has same stepwise values of the energy as the RO protocol. Thus, if the amplitudes (a_t, b_t) for the RO protocol at time t have energy (3.9), then we will search on this circle for the optimal one-step protocol: for each time t, we apply (1.1) to find

$$(a_t, b_t) = \text{argmin}\{\|\rho^{t-1} \circ f_{a_t,b_t}^{-1}\|_{-1} \ : \ a_t^2 + b_t^2 = E_t^{RO}\}. \tag{3.10}$$

Recall that the time evolution of the energy E_t^{RO} was shown in Fig. 3.24. Since (3.10) requires succession of a one-dimensional searches at each time, it is much quicker than the full random optimization. However, the restriction to a known energy distribution means that we will not be looking for more general protocols. We call this new stepwise minimization protocol "SW".

Figure 3.26 illustrates the stepwise minimization by comparing the RO protocol of §3.2.3 to a SW protocol at three times. For this illustration, we used the density ρ^{t-1} obtained by the RO protocol as the initial state for the SW minimization, shown in the second column. So to obtain the stepwise minimization at t, we perform $t-1$ steps of the RO protocol and then use (3.10). The amplitudes (a_1, b_1) for the SW protocol are essentially the same as those for RO–this can be seen in the third column which shows $\|\rho^t\|_{-1}$ as a function of (a, b). The RO amplitudes, the green dot in this panel, and the selected SW amplitudes nearly overlap.

This is also true at $t = 3$, though there is a larger difference, presumably because none of the selected random trials landed on the best point. However the difference at $t = 2$ is nontrivial: the sign of b_2 flips between the RO and SW cases. The central point $(\frac{1}{2}, \frac{1}{2})$ is elliptic for the RO amplitudes, whereas it is a saddle for the SW amplitudes. This supports our hypothesis that the RO protocol, by looking ahead, does not minimize the mix-norm at every step: for this step it has emphasized folding instead of stretching.

Fully implementing (3.10) (without reinitializing to the RO solution at every step) for the initial density (3.1) gives $\|\rho^t\|_{-1}$ shown in Fig. 3.27(a). This figure also shows the mix-norm of the RO protocol from Fig. 3.25 for comparison. Though the overall performance of the two protocols is about the same at $t = 12$, the SW protocol reaching $\|\rho^T\|_{-1} = 0.0401$, the RO protocol does

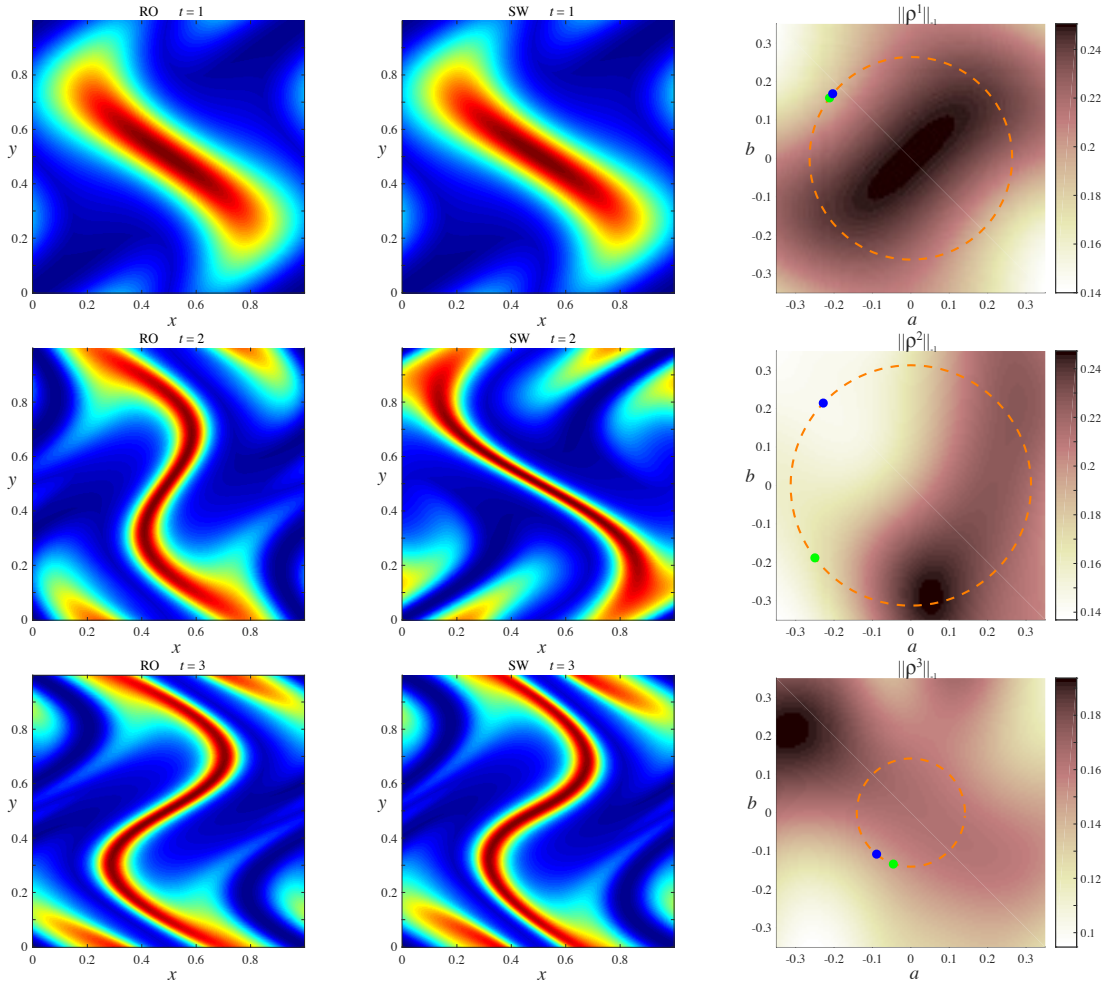

Figure 3.26: Comparison of RO and SW protocols of the Harper map for $t = 1, 2, 3$ using the initial condition (3.1). The first column shows $\rho^t(x, y)$ for first three steps of the RO protocol, and the second gives that of the stepwise protocol applied to the RO state ρ^{t-1}. The color scale shows the mix-norm as a function of (a, b) for step t. The dotted circle is (3.9), from which the parameters for the SW protocol are selected. The green point on this circle corresponds to the RO and the blue point to the SW parameters.

slightly outperform the stepwise minimizer. It seems that an optimal mixing scheme cannot simply consider each step of the map independently, but must instead forgo a one-step improvement to setup the density for future steps of the map.

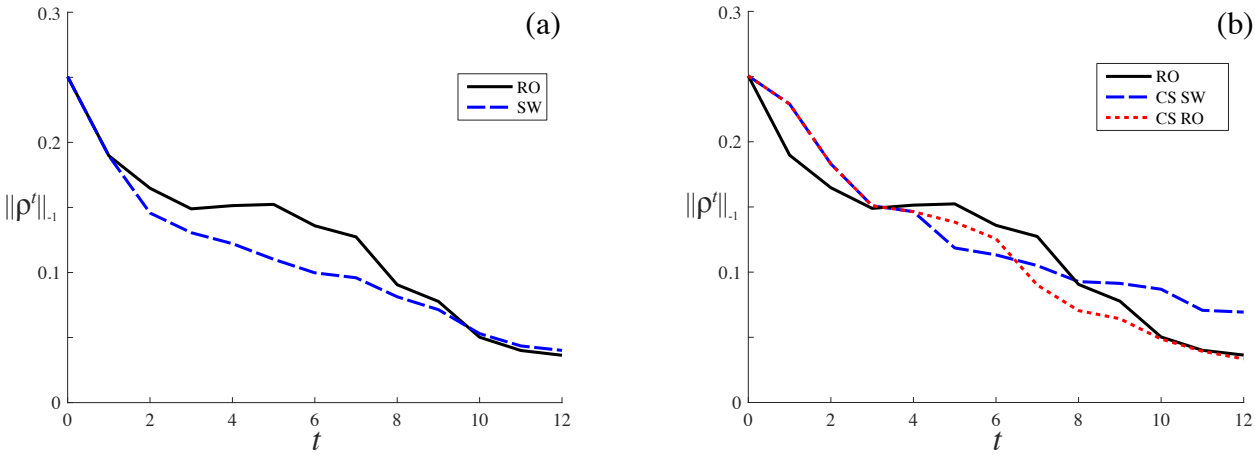

Figure 3.27: (a) Progression of $\|\rho^t\|_{-1}$ for the RO and SW protocols described in Figure 3.26. For a movie of the SW protocol steps, see `http:/amath.colorado.edu/faculty/jdm/movies/HarperSW.mp4`. (b) Progression of $\|\rho^t\|_{-1}$ for the RO, CS SW, and CS RO protocols. The latter used $\sim 4 \times 10^5$ trials. The two constant shear protocols are identical until $t = 3$.

3.2.5 Horizontal or Vertical Shearing

For many of the time steps of the RO protocol, as illustrated in Fig. 3.23, one of the shear amplitudes is much smaller than the other, giving either primarily horizontal or vertical shearing. In this section we investigate protocols in which there is only one nonzero shear at each step. To further simplify the problem, we assume that the energy per step is constant. Defining the amplitude $\gamma_o = \sqrt{E_H/T}$, there are now just four options for the parameter pair at a given step: $(a, b) \in \{(\gamma_o, 0), (-\gamma_o, 0), (0, \gamma_o), (0, -\gamma_o)\}$.

To find an optimal stirring protocol with the new constraints is easier than a full RO since there are only four choices for parameters at every step. In fact, except for $t = 1$, only three of these can decrease the mix-norm. For example if $(a_1, b_1) = (\gamma_o, 0)$, then the choice $(a_2, b_2) = (-\gamma_o, 0)$ simply undoes the previous step, so it cannot be an optimal choice. Therefore, only the three options $(\gamma_o, 0), (0, \gamma_o), (0, -\gamma_o)$ need be considered at $t = 2$. In this way, there are $4 \times 3^{T-1}$ distinct shear protocols to consider, i.e., $\sim 7 \times 10^5$ when $T = 12$. The protocol found from randomly sampling this discrete space is compared to the RO protocol in Fig. 3.27(b). The optimal shear

protocol (labeled "CS RO" for Constant Shear Random Optimal in the figure) just beats the full RO protocol, as shown in Table 3.1.

An even simpler method is to do a stepwise minimization, as in §3.2.4, now imposing the single-shear restriction. In this case there are only $4 + (T-1)3$ choices, i.e., 40 when $T = 12$. The optimal result in this case gives the mix-norm curve labelled CS SW (Constant Shear SW) in Fig. 3.27(b). This simplified protocol achieves $\|\rho^T\|_{-1} = 0.0693$, which is not comparable to the better performing protocols. Nevertheless, it still performs better than 99.6% of the randomly sampled amplitude vectors shown in Fig. 3.22. Note that, for this specific initial density profile, the hyperbolic constant scheme is slightly better (see Table 3.1); however, its effectiveness depends strongly on the localization of (3.1) near the saddle point.

To gain some insight into how to choose the direction of shear based on the current density distribution, we can consider what the optimal one-step parameter choice would be when the density can be described by a single sinusoid. Consider the initial condition

$$\rho^0(x, y) = 2\cos(2\pi n \cdot z - \phi) \tag{3.11}$$

for wave vector n, and phase shift ϕ. The resulting Fourier expansion will have only two nonzero terms, both of modulus one. Using this initial condition and fixing the amplitude $\gamma_o = 0.2041$, we compute the mix-norm $\|\rho^1\|_{-1}$ as a function of mode numbers for 200 randomly chosen phase shifts. Figure 3.28 shows a density plot of the number of times choosing $a \neq 0$ results in better one-step mixing than choosing $b \neq 0$. The results are remarkably simple: whenever $|n_1| < |n_2|$, we should choose $a \neq 0$, and if $|n_1| > |n_2|$ we should choose $b \neq 0$. The lines $|n_1| = |n_2|$ are ambiguous.

To understand why this is true, consider the simple case of (3.11) with $\phi = 0$, so that the Fourier expansion will contain the only nonzero terms $\hat{\rho}^0_n = \hat{\rho}^0_{-n} = 1$. Recalling (2.21), the Fourier coefficients at $t = 1$ will be

$$\hat{\rho}^1_{m_1,m_2} = J_{m_1+n_1}(-2\pi n_2 a)J_{-n_2-m_2}(2\pi m_1 b) + J_{m_1-n_1}(2\pi n_2 a)J_{n_2-m_2}(2\pi m_1 b).$$

When $b = 0$ and a is nonzero, the only terms that survive are those that satisfy $m_2 = \pm n_2$ so that

$$\hat{\rho}^1_{k,\pm n_2} = J_{m_1 \mp n_1}(\pm 2\pi n_2 a), \quad (b = 0),$$

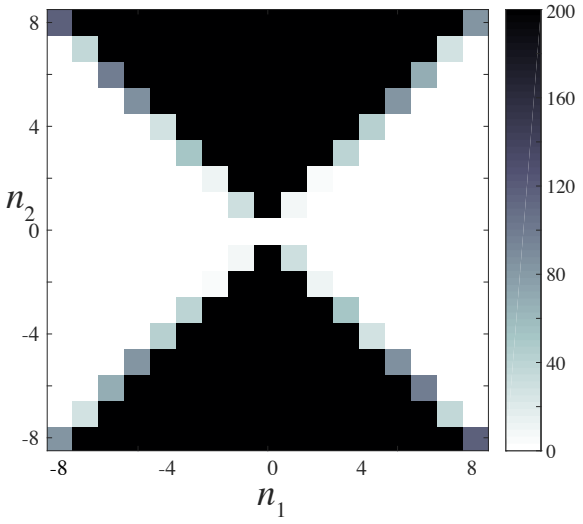

Figure 3.28: Density plot of the number of times choosing $a \neq 0$ results in better one step mixing than choosing $b \neq 0$ with respect to mode numbers (n_1, n_2) for initial condition (3.11) with 200 random choices of phase shift ϕ. Here the shear amplitudes, $|a|$ or $|b|$, are fixed to $\gamma_0 = 0.2041$, but varying γ_0 within reasonable bounds does not meaningfully change the results.

for $k \in \mathbb{Z}$. On the other hand, if $a = 0$ and b is nonzero, we are left with the terms in which $m_1 = \pm n_1$, giving

$$\hat{\rho}^1_{\pm n_1,k} = J_{\pm n_2 - m_2}(\pm 2\pi n_1 b), \quad (a = 0).$$

Since the H^{-1} norm (2.13) has the wavenumber in the denominator, the implication is that we would then want to choose $b = 0$ when $|n_1| < |n_2|$, since the resulting terms $\hat{\rho}_{k,\pm n_2}$ will have higher combined wave number than the terms $\hat{\rho}_{\pm n_1,k}$. The Bessel function terms will also tend to be smaller, at least when the shears have comparable amplitude and the mode numbers are large. This supports the more general results of the numerical experiments.

To apply this insight to more general density profiles which have a full Fourier spectrum, we define the partial sums

$$S_< = \sum_{|n_1|<|n_2|} \frac{|\hat{\rho}_{n_1,n_2}|^2}{n_1^2 + n_2^2}, \qquad S_> = \sum_{|n_1|>|n_2|} \frac{|\hat{\rho}_{n_1,n_2}|^2}{n_1^2 + n_2^2}.$$

The shear choice is then determined by the relative sizes of these partial norms: if $S_<$ is smaller we choose $(|a|, |b|) = (0, \gamma_o)$; otherwise, if $S_>$ is smaller we choose $(|a|, |b|) = (\gamma_o, 0)$.

To make this a complete protocol, we must also select the signs of the shears. Note that successive steps $(a_t, 0)$ to $(0, b_{t+1})$ are equivalent to one step with amplitudes (a_t, b_{t+1}), so that the choice of sign is equivalent to the selection of the location of elliptic and saddle points of the map. The opposite switch, from $(0, b_t)$ to $(a_{t+1}, 0)$, does not give a single Harper map of the form (2.9), but is instead equivalent to its inverse with the signs of both parameters switched. Thus the locations of the elliptic and saddle cases are still the same as those shown in Fig. 3.16.

If the density is concentrated in a specific location, then it seems reasonable that a sign choice that makes the nearest fixed point a saddle will beat the sign choice that makes it elliptic. Indeed, for the Gaussian centered at $(\frac{1}{2}, \frac{1}{2})$, it was important to choose opposite signs to allow this fixed point to be hyperbolic, since making it elliptic would result in virtually no mixing.

To determine the relationship between the location of density peaks and the optimal sign choice, we can do a simple one-step experiment with various localized initial densities—we use the initial state (3.6), varying the center (u, v) of the profile. Figure 3.29 shows the best sign choice as a function of (u, v). The color at a given point indicates the sign choice that most often resulted in the best one-step mixing under randomly chosen values of the width parameter ε. As expected, if the density is localized near $(\frac{1}{2}, \frac{1}{2})$, we should choose a and b of opposite sign. This generalizes: If the density is localized near any fixed point, the relative signs should be chosen to make that fixed point hyperbolic. The regions in which we choose a given sign combination form nearly perfect squares around each fixed point. To understand why this is, consider the approximate eigenvalues and eigenvectors given by the Hamiltonian (3.8). The fixed point $(\frac{1}{2}, \frac{1}{2})$ has eigenvalue-eigenvector pairs

$$\lambda = \pm 2\pi\sqrt{-ab}, \quad \mathbf{v} = \left(\mp\frac{b}{\sqrt{-ab}}, 1\right).$$

Now assume that we are using the constant shearing scheme such that $|a| = |b| = \gamma_o$. Also assume that the density is concentrated near $(\frac{1}{2}, \frac{1}{2})$ so that we at least know we want $a = -b$ to make the point hyperbolic. This causes the eigenpairs to reduce to

$$\lambda = \pm 2\pi\gamma_o, \quad \mathbf{v} = (\mp\text{sign}(b), 1).$$

The vector $(\text{sign}(b), 1)$ corresponds to the stable manifold and $(-\text{sign}(b), 1)$ to the unstable. The closer the concentration of density is to a hyperbolic fixed point, the more it gets stretched along the unstable manifold of that point. Ideally, the density should be localized to straddle the stable manifold so that it gets pushed towards the fixed point and then spread out along the two branches of the unstable manifold. This implies that if the concentration is located in the 1st or 3rd quadrants, relative to $(\frac{1}{2}, \frac{1}{2})$, we should choose $b > 0$ to make the local stable manifold have positive slope. By contrast, if the concentration lies in the 2nd or 4th quadrants, we choose $b < 0$ to make the local stable manifold have negative slope. Performing a similar analysis for each of the fixed points implies the results obtained experimentally in Fig. 3.29.

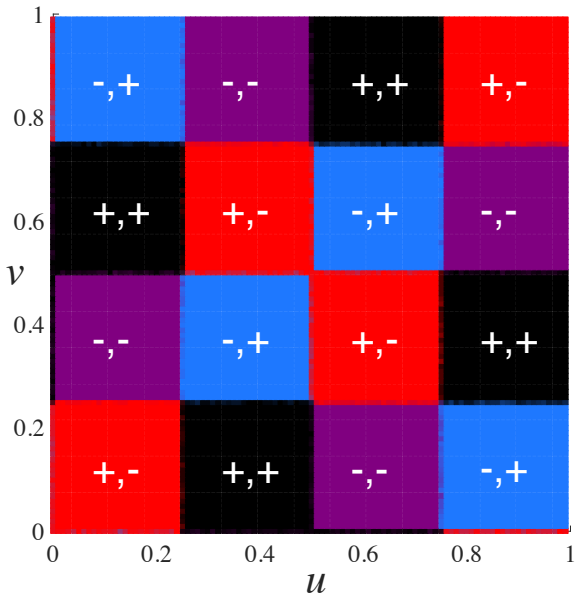

Figure 3.29: The intensity of the color at a given point (u, v) is proportional to the percentages of sign choices that yielded the best mixing over different ε ranging from 0.1 to 5 using the initial density (3.6).

From this we can extract an algorithm to determine effective sign choices. We denote the 16 regions in Fig. 3.29 by $R_1, R_2, \ldots, R_{16}$, and define the indicator functions

$$\chi_i(x, y) = \begin{cases} 1, & (x, y) \in R_i \\ 0, & (x, y) \notin R_i \end{cases}.$$

For each region we compute a partial mix-norm

$$S_i = \|\rho \cdot \chi_i\|_{-1}, \quad i = 1, 2, \ldots, 16.$$

Under the hypothesis that region with the largest S_i dominates the need for mixing, we pick the signs of (a, b) for this region according to Fig. 3.29.

A combination of the Fourier and sign algorithms gives a full protocol choice for the multistep shear problem. We attempt two different schemes. The first allows for the signs of (a, b) to change between steps of the map if the location of the density has changed. We refer to this as the "CS Full" method. The second scheme chooses signs for (a, b) based solely on the initial condition, maintaining this sign choice throughout the protocol. We refer to this as the "CS FS" (Fixed Signed) method. the CS Full method has the advantage of tracking the density and switching the fixed points accordingly. However, the CS FS method allows for more consistency in the direction of stirring.

In the case of the initial condition (3.1), the CS Full and CS FS algorithms produce the same protocol, depicted by Fig. 3.30, which is almost qualitatively identical to the Hyperbolic constant case shown in Fig. 3.19, in that it alternates between steps of $(0, \gamma_o)$ and $(-\gamma_o, 0)$. Figure 3.31 compares the decay of the mix norms for the RO, SW, and CS FS protocols. The CS FS protocol matches the SW solution in performance, though both are significantly outperformed by the RO protocol.

The CS Full algorithm performs better when we limit the instances in which a sign change can occur. Switching the signs, and thereby changing the direction of stirring, often causes a momentary increase in the mix-norm. Performing a sign change when it is not strictly necessary can negatively affect the results. However, forgoing a sign change can cause the mix-norm to level off. Recall that the RO solution outperformed the Hyperbolic Constant scheme because it sacrificed some steps for folding, which allowed for more effective stretching later. Experimentally, we find that CS Full performs the best when we only switch signs if the difference between the two largest S_i values is more than 10% of the total mix-norm.

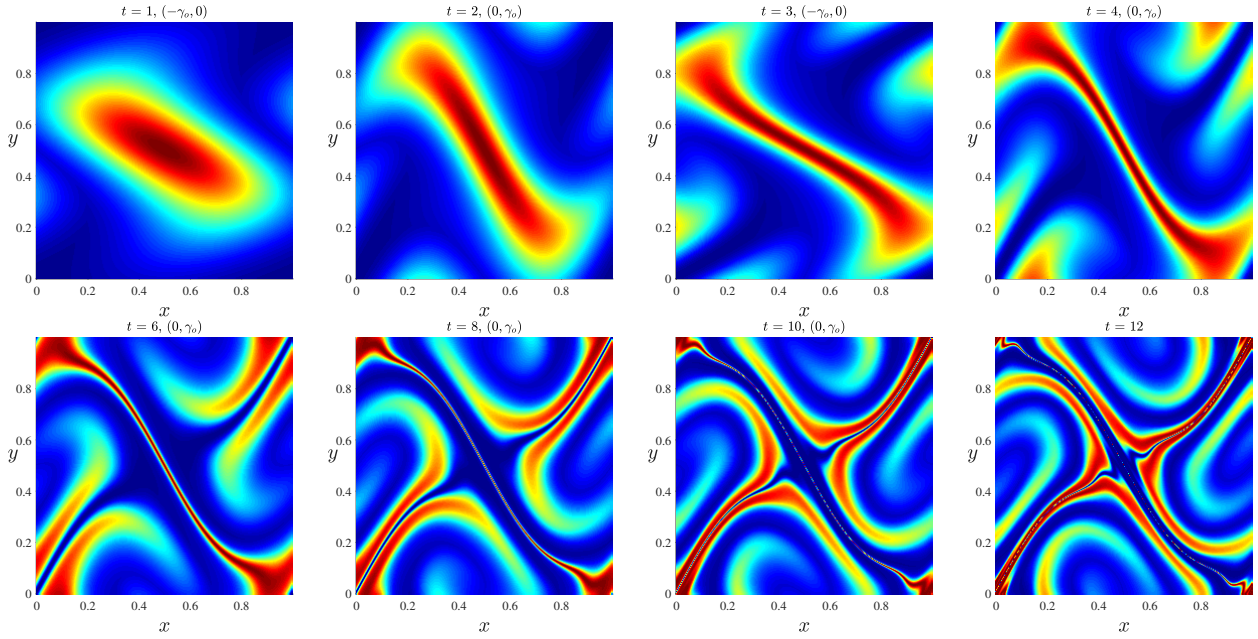

Figure 3.30: Evolution of $\rho^t(x, y)$ for the stirring protocol produced by CS Full and CS FS: alternating steps of $(0, \gamma_o)$ and $(-\gamma_o)$, $\gamma_o =\approx 0.204$, with initial condition (3.1) for $t = 1, 2, 3, 4$ (top row) and $t = 6, 8, 10, 12$ (bottom row) on a 128×128 grid.

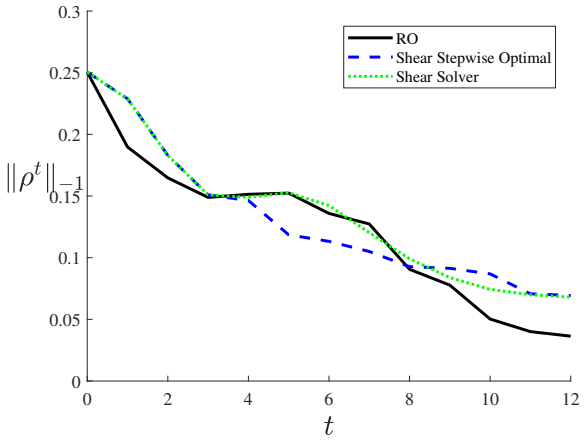

Figure 3.31: Decay of $\|\rho^t\|_{-1}$ for the RO protocol, the SW minimizing protocol, and the CS FS protocol shown in Fig. 3.30.

Figure 3.32 shows the performance of the algorithm over 10^5 random choices of initial conditions (3.7) normalized such that $\|\rho^0\|_{-1} = 1$. The minimum, mean, and standard deviation of $\|\rho^T\|_{-1}$ for each distribution depicted in Fig. 3.32 can be found in Table 3.2. The scheme with sign

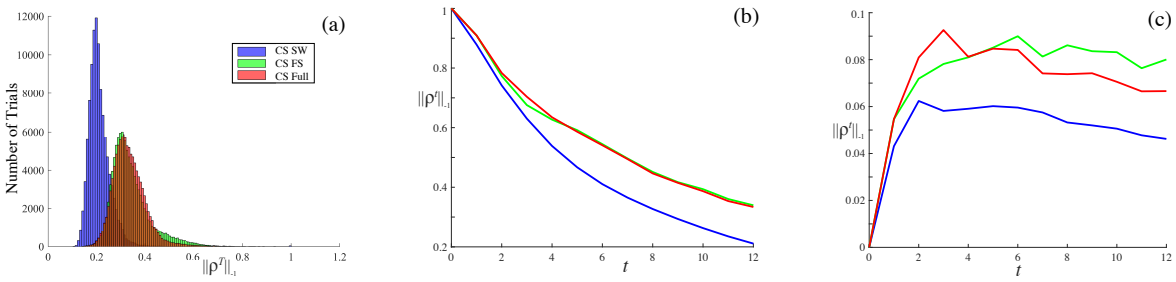

Figure 3.32: (a) Distribution of $\|\rho^T\|_{-1}$ using the three shearing techniques, for random choices of (3.7) where $(u, v) \in [0, 1)^2$, $\varepsilon \in [0.1, 5.0]$, and $n \in \{-10, 10\}^2$. We use energy $E_H = 0.5$ and $T = 12$. (b) Means and (c) standard deviations of $\|\rho^t\|_{-1}$ for the CS SW, CS Full, and CS FS solvers.

switching beats out the one with fixed signs slightly in terms of mean final mix-norm. However, the SW minimizer outperforms both analytic solvers. The standard deviation of mix-norms using the sign switching methods is smaller than that for fixed signs. Indeed, the distribution of mix-norms for the fixed sign method has a right tail that extends nearly to $\|\rho^T\|_{-1} = 1$, suggesting that for some initial conditions, the solver performs extremely poorly and results in virtually no mixing. Using sign switching helps to reduce the size of this tail.

	$\|\rho^T\|_{-1}$		
	Optimal	Mean	Std. Dev.
Elliptic Constant	0.1191	0.4905	0.1924
CS SW	0.1085	0.2110	0.0462
CS FS	0.1449	0.3395	0.0801
CS Full	0.1449	0.3338	0.0666

Table 3.2: Optimal, mean and standard deviations of the mix-norms at time $T = 12$ for the Harper map from the distributions of Fig. 3.32(a). The initial conditions are normalized such that $\|\rho^0\|_{-1} = 1$.

Chapter 4

Results of Blinking Vortex Map

In this chapter, we present results for the blinking vortex model described in §2.4. There are multiple parameter choices we can manipulate to find the sequence of blinking vortex maps $\{H_{v_t,\tau_t,s_t}\}$, defined by (2.29), for $n = 1, \ldots, N$ that optimizes the mix norm on the disk (2.36). These include, N, the total number of vortex switches, v_t, the position of the vortex at step t, τ_t, the time vortex v_t is activated, and s_t, the direction that vortex v_t is spinning. For all results, we use the initial Gaussian profile,

$$\rho^0(z) = C \exp\left(-\varepsilon \|z - z_o\|^2\right), \quad z_o = (0,0), \quad \varepsilon = 5, \quad C \approx 8.9574. \tag{4.1}$$

The normalizing coefficient C is chosen such that $\|\rho^0\|_{-1} = 1$. Figure 4.1 shows the evolution of the density profile for the initial condition (4.1) with $N = 3$ and stirring protocol vectors

$$v = (0.5, -0.6 + 0.6, -0.2 - i0.3), \quad \tau = (0.1, 0.5, 0.02), \quad s = (1, -1, 1). \tag{4.2}$$

The vortices at $t = 1, 3$ rotate counterclockwise, and the vortex at $t = 2$ rotates clockwise. The second vortex is activated for the longest amount of time, the first vortex activated for a medium amount of time, and the third vortex for a very short amount of time. The size of the "spirals" is proportional to the amount of time a given vortex is activated.

We consider two different optimization problems. The first will be referred to as the Roaming Vortex Problem (RV), in which N, t_t, and s_t are held fixed and the position v_t is allowed to vary to any point in the unit disk at each step. For the numerical results, we fix N at 12, τ_t at 0.1, and s_t at 1, so that the vortices always spin counterclockwise. The numerical methods and results of this

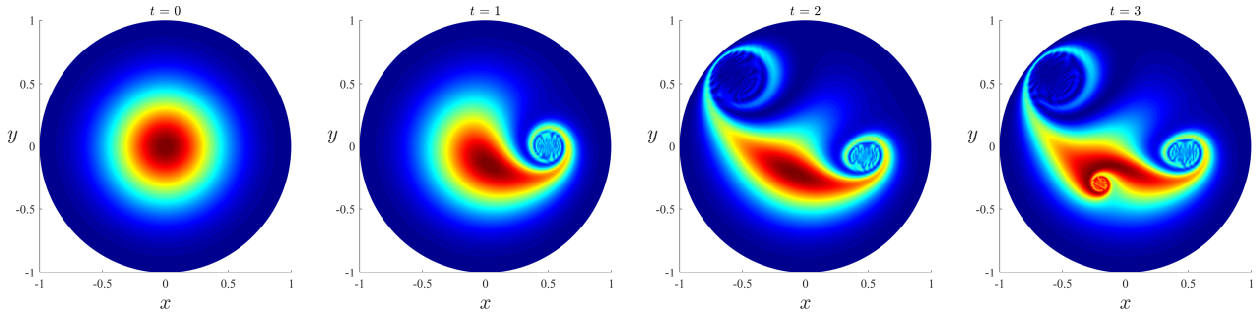

Figure 4.1: Evolution of ρ^t for $t = 0, \ldots, 3$ with initial density (4.1), $N_r = N_\theta = 128$, and blinking vortex protocol vectors (4.2).

problem are discussed in §4.1. The second problem considered is called the Four Vortex Problem (FV). For this problem, we consider only four possible vortex locations at $\frac{1}{2}$, $\frac{i}{2}$, $-\frac{1}{2}$, and $-\frac{i}{2}$. We hold the total time T that the vortices are activated fixed and force the time τ_t to remain constant for each vortex so that $\tau_t = T/N$, where we allow the total number of vortex switches N to vary. We also allow s_t to vary so that there is a total of eight possibilities at each step: four vortex position choices, and two vortex direction choices. For the numerical results, we hold T fixed at 1. The methods and results of the Four Vortex Problem are discussed in §4.2.

4.1 Roaming Vortex Problem

Unlike the periodic shear maps discussed in Chapter 4, the vortex model (2.25) is a more concentrated stirring mechanism. The stirring action of the Chirikov (2.8) and Harper (2.9) maps affected every point in the domain, whereas the vortex stirring action causes a swirling of points nearby, but has virtually no affect on points far away from the vortex position. Recalling (2.28), we know that $\frac{d\phi}{dt}$ is proportional to $\frac{1}{p^2}$. Therefore, for the vortex position to be chosen optimally, it must be placed in a way so that it has the maximum disruption of the most "unmixed" region of the domain. There are multiple parts to this choice that make it difficult. The first is, how does one identify the most unmixed region of the domain; where is it located and what is its size and shape? Secondly, once the most unmixed region is identified, how does one choose the vortex location? It

is clear that the vortex should be chosen near to this region, but shouldn't be placed at its center, since this will result in the concentrated density being rotated and sheared, but not distributed. Figure 4.2 illustrates the effect of vortices at different distances from the unmixed region. Fig. 4.2 (a) shows a vortex too close to the center of the unmixed region, only resulting in a 0.2% decrease in the mix norm (2.36). Fig. 4.2 (b) shows a vortex closer to the optimal distance, resulting in a 5.4% decrease mix norm. Fig. 4.2 (c) shows a vortex too far from the unmixed region, resulting in a 3.6% decrease in mix norm. Since the stirring action of the vortex is so concentrated, choosing the right location at every step is essential to deriving an effective mixer.

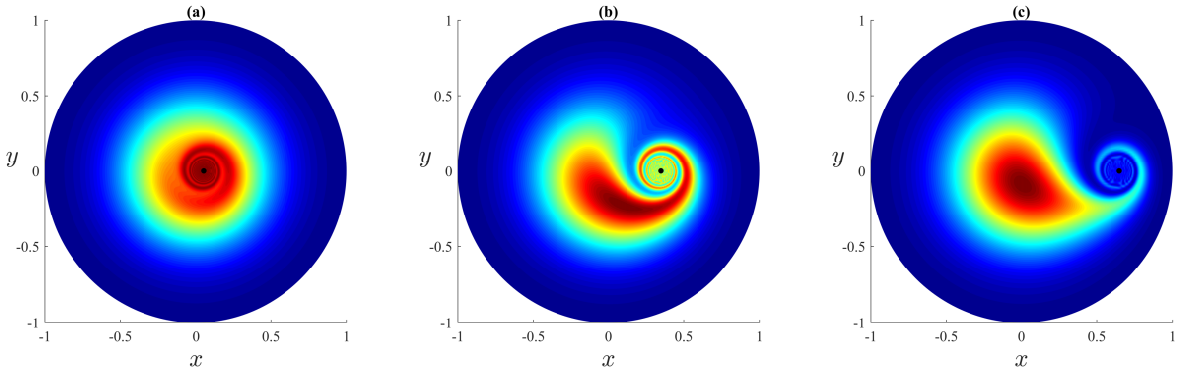

Figure 4.2: Initial Gaussian density profile (4.1) with $N_r = N_\theta = 128$ subjected to vortices at different distances: (a) $v = 0.05$, $\|\rho\|_{-1}$ defined by (2.36) decreases by 0.2%, (b) $v = 0.35$, $\|\rho\|_{-1}$ decreases by 5.4%, (c) $v = 0.65$, $\|\rho\|_{-1}$ decreases by 3.6%.

One way to determine the optimal vortex locations would be to systematically compute the one-step mix norm (2.36) for each vortex location and choose the one that minimizes the norm. However, the mix norm is a more expensive calculation on the disk than the box $[0, 1)^2$ because it requires the computation of a Bessel Fourier transform, recall §2.4.1. For the angular direction, we can use the fast Fourier transform, but for the radial part, we us the trapezoid rule for numerical integration, which is considerably slower. Computing many instances of the mix norm to find every v_t is not a reasonable approach. Instead, we use a cheaper method that fits an ellipse to the most unmixed region, and then finds a near optimal vortex location that best spreads the points the contained in the ellipse.

To locate the unmixed region of the density profile ρ^t at step t, we consider the difference between local medians and the global median. We use the median as opposed to the mean to avoid local medians being skewed by a few outlier density points. Given a uniform grid of N_z points and a disc radius ϱ at each point z_i, $i = 1, \ldots, N_z$, we compute the quantity

$$A_i = \underset{z \in B_\varrho(z_i)}{\text{median}} \left[\rho(z) - \mu\right],$$

where μ is the median of ρ^t and $B_\varrho(z)$ is the open ball of radius ϱ centered at z. We choose the center of the unmixed region z_C to be the point that corresponds to the maximum A_i value and satisfies the condition

$$\min_{z \in B_{\frac{1}{N_r}}(z_i)} \left[\rho(z)\right] \geq \frac{1}{2} \max\left[\rho^t\right].$$

Recall from §2.4.1 that $\frac{1}{N_r}$ is the interval size of the radial discretization. This is to ensure that that we do not choose the point to be the center in a region that is not empty, and not too granulated.

Once the location of the most unmixed region is identified, the next step is to determine the shape and size of the region. This is done by using the covariance to fit an ellipse to the region. For all points $z_i = (x_i, y_i)$, $i = 1, \ldots, N_{z_C,\varrho}$ contained in $B_\varrho(z_C)$, we compute the 2×2 covariance matrix M_{cov}:

$$M_{\text{cov}} = \begin{bmatrix} \sum_{i=1}^{N_{z_C,\varrho}} \rho^t(z_i)(x_i - x_C)^2 & \sum_{i=1}^{N_{z_C,\varrho}} \rho^t(z_i)(x_i - x_C)(y_i - y_C) \\ \sum_{i=1}^{N_{z_C,\varrho}} \rho^t(z_i)(x_i - x_C)(y_i - y_c) & \sum_{i=1}^{N_{z_C,\varrho}} \rho^t(z_i)(y_i - y_C).^2 \end{bmatrix}.$$

The matrix M_{cov} has the two eigenpairs (λ_1, w_1) and (λ_2, w_2), where λ_1 is the largest eigenvalue. The eigenfunctions w_1 and w_2 give the directions of major and minor axes of the ellipse respectively. To determine the lengths of the axis, the vectors are scaled according to

$$|w_{\text{maj}}| = \sqrt{\chi^2_{2,\alpha}\lambda_1}, \quad |w_{\text{min}}| = \sqrt{\chi^2_{2,\alpha}\lambda_2},$$

where $\chi_{2,\alpha}$ refers to the two-dimensional chi squared value and α is the confidence level of a local Gaussian, so that $100(1-\alpha)\%$ of ρ^t defined over $B_\varrho(z_C)$ is contained within the fitted ellipse. Define Φ to be the angle between w_{min} and the x-axis. Then the fitted ellipse $\mathcal{E}_{z_C,\varrho,\alpha}$ has the

parametrization

$$z(s) = z_c + \begin{bmatrix} \cos\Phi & -\sin\Phi \\ \sin\Phi & \cos\Phi \end{bmatrix} \begin{bmatrix} |w_{\min}|\cos s \\ |w_{\mathrm{maj}}|\sin s \end{bmatrix}, \quad s \in [0, 2\pi].$$

Figure 4.3 shows the initial density profile

$$\rho^0(r,\theta) = \exp\left[-r^2\left(\frac{1}{a^2}\cos^2\theta + \frac{1}{b^2}\sin^2\theta\right)\right], \tag{4.3}$$

with $a = 0.5$ and $b = \frac{1}{4}a = 0.125$, along with the fitted covariance ellipse $\mathcal{E}_{z_C,\varrho,\alpha}$ with $z_C = 0$, $\varrho = 0.5$, and $\alpha = 0.01$, so that the ellipse contains 99% of the density within the ball $B_{0.5}(0)$.

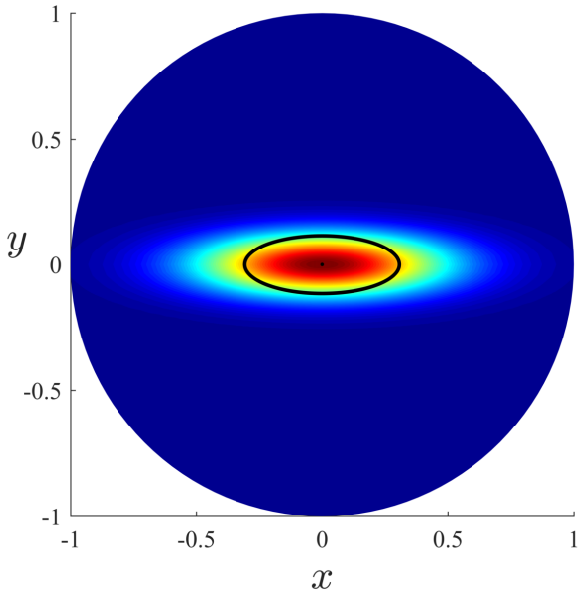

Figure 4.3: Density profile (4.3) with $(a, b) = (0.5, 0.125)$ with fitted covariance ellipse $\mathcal{E}_{z_C,\varrho,\alpha}$ with $z_C = 0$, $\varrho = 0.5$. The black point marks z_C and the black curve marks $\mathcal{E}_{z_C,\varrho,\alpha}$.

With an ellipse fitted to the unmixed region, the third and final step is to determine the location for the vortex v_{t+1} to give the optimally mixed density profile at the following step ρ^{t+1}. We find the vortex v that optimizes the one-step variance of the points contained in $\mathcal{E}_{z_C,\varrho,\alpha}$. Define $E_{z_C,\varrho,\alpha}$ to be the set of points on the grid contained within $\mathcal{E}_{z_C,\varrho,\alpha}$. Recalling the definition (1.3) and the map (2.29), we make the claim that the quantity $\mathrm{Var}(H_{t+1}(E_{z_C,\varrho,\alpha}))$ will be nearly maximal for the vortex that optimizes the one-step mix norm $\|\rho^{t+1}\|_{-1}$, because maximizing the variance means

we are allowing for the largest possible "spreading-out" of the unmixed region. Numerically, we find that choosing a confidence level of $\alpha = 0.25$, such that the covariance ellipse contains 75%, or about one standard deviation, of the density of the unmixed region, gives similar vortex locations that maximize the variance and minimize the mix norm. Figure 4.4 compares the vortex solution that minimizes the mix norm of $\rho^1(r,\theta) = \rho^0(H_{v_1}^{-1}(r,\theta))$ with ρ^0 defined by (4.3), and the vortex solution that maximizes the quantity

$$\mathrm{Var}(H_1(E_{z_C,\varrho,\alpha})). \tag{4.4}$$

The parameters of the initial condition are $(a,b) = (0.5, 0.125)$. The variance (Fig. 4.4(a)) and the mix norm (Fig. 4.4(b)) appear to be inversely related: where one plot has a local maximum, the other has a local minimum, and vice versa. This supports our claim that maximizing the variance of the points within the unmixed region will come close to minimizing the mix norm. The distance between the variance maximizing vortex $v = -0.2381 - 0.1429i$ and the mix norm minimizing vortex $v = -0.3333 - 0.2381i$ is 0.1347. With an initial mix norm of 0.0543, the optimal one-step mix norm value is 0.0448, and the vortex that maximizes the variance achieves a mix norm of 0.0461, which is only a 2% increase of the true optimal. This indicates that, should the covariance ellipse be a good fit for the most unmixed section of density, we can use the much cheaper variance computation to approximate the vortex location that will minimize the mix norm. By symmetry, there are two optimal positions, assuming the ellipse is far enough from the boundary. In some cases there is only one because the reflected point falls outside of the disk.

The steps above can be summarized as follows:

(1) Given a grid of N_z points and radius ϱ, search through the N_z points to find the center z_C of the unmixed region.

(2) Given the point z_C, search radius ϱ, and confidence level α, find the covariance elliptic region $E_{z_C,\varrho,\alpha}$ centered at z_C containing the specified proportion of density.

(3) Given a grid of N_v possible vortex positions, and elliptic region $E_{z_C,\varrho,\alpha}$, determine a near

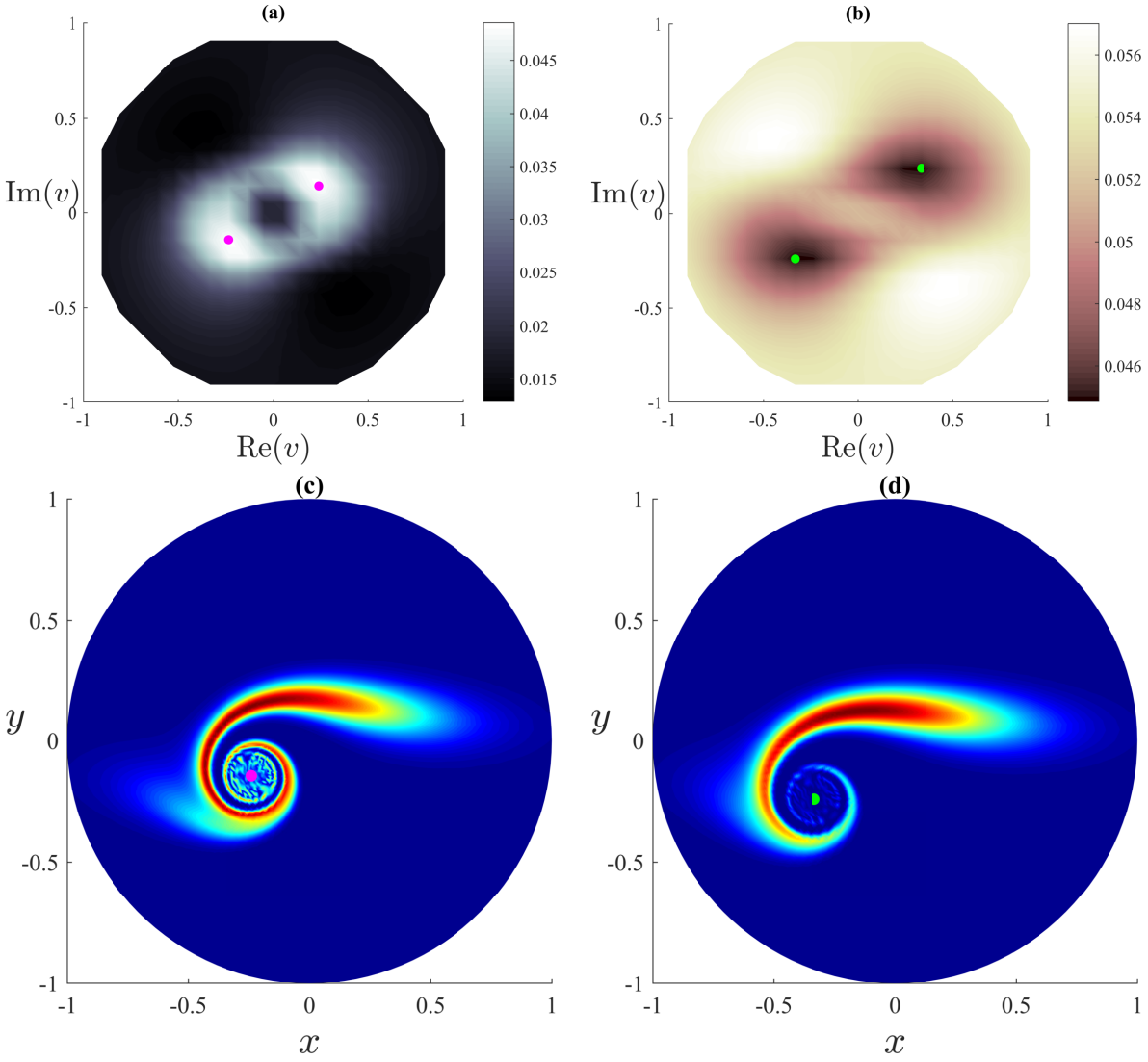

Figure 4.4: Optimizing one-step mixing of the RV Problem for an elliptic unmixed region with initial condition (4.3) and $(a, b) = (0.5, 0.125)$ depicted by Fig. 4.3. Plot (a) shows the variance (4.4). Plot (b) shows the mix norm ρ^1 with respect to v. Plot (c) shows the density profile ρ^1 after applying the variance optimizing vortex. Plot (d) shows ρ^1 after applying the mix norm optimizing vortex. The purple point marks the variance optimizing vortex $v = -0.2381 - 0.1429i$. The green point marks the mix norm optimizing vortex $v = -0.3333 - 0.2381i$.

optimal choice of vortex location v that maximizes the quantity (4.4).

At this point, it should be clear why the RV algorithm described above will be problematic. For one thing, it depends on the choices of the parameters N_z, ϱ, α, and N_v. If any of these parameters are chosen poorly, it could compromise the result. The parameters may need to change

as the number of stirring steps increases, because the length scales of uniform regions of density will decrease as the density becomes more mixed up. In addition, the algorithm is based on a number of assumptions and approximations. For example, it depends on the assumption that we are able to approximate the unmixed region by fitting an ellipse, which may not be the case if the unmixed region is oddly shaped. We also rely on the assumption that maximizing the variance of the positions of the discrete points contained within the fitted ellipse gives a similar result as minimizing the mix norm, which will result in at least some error. Figure 4.5 shows the evolution of density with initial condition (4.1) normalized such that $\|\rho^0\|_{-1} = 1$ and the RV algorithm with a constant choice of parameters: $N_z = 1005$, $\varrho = 0.2$, $\alpha = 0.75$, and $N_v = 88$. The algorithm performs poorly, achieving a mix norm of 0.8551, only about a 15% decrease. Because the parameters ϱ and α stay constant, the sizes of the fitted ellipses stay about the same despite the length scales getting finer with increasing t.

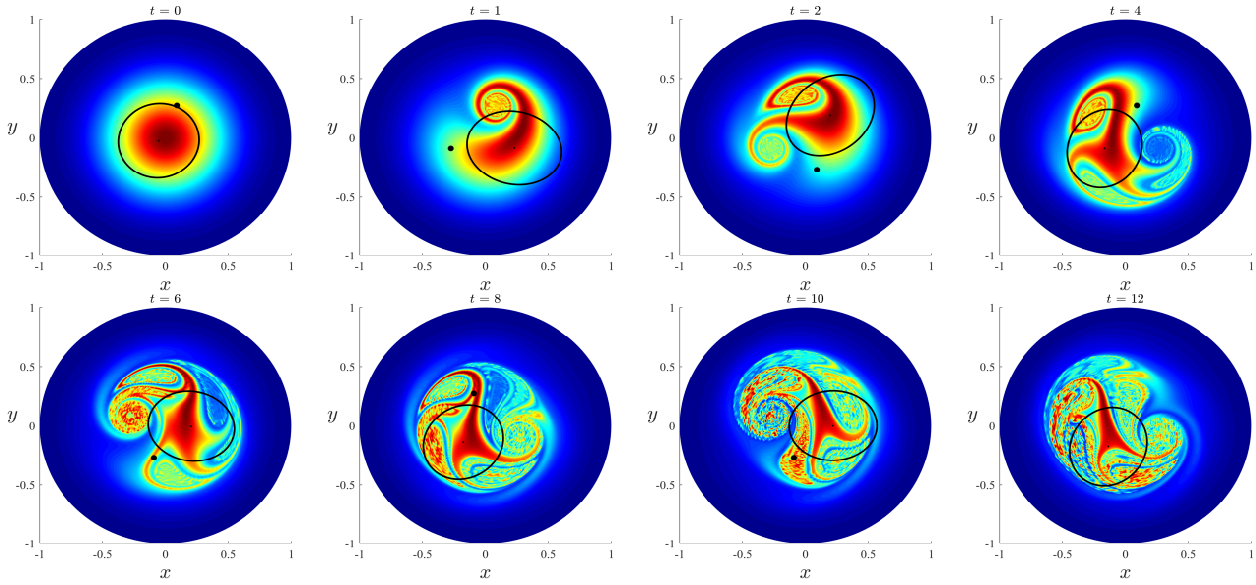

Figure 4.5: Evolution of ρ^t for $t = 0, 1, 2, 4, 6, 8, 10, 12$ with initial condition (4.1) and $N_r = N_\theta = 128$. Parameters of the RV algorithm are $N_z = 1005$, $\varrho = 0.2$, $\alpha = 0.75$, and $N_v = 88$.

Since it seems the algorithm could be improved by allowing more flexibility in the ellipse size, we allow some of the parameters to change with t. Figure 4.6 shows the evolution of density with

the same parameters as Fig. 4.5, but with ϱ decaying and α growing according to

$$\varrho(t) = 0.5\exp(-0.1t), \quad \alpha(t) = 1 - 0.5\exp(-0.5t). \tag{4.5}$$

With the varying parameters, the performance of the RV algorithm improves, achieving a mix norm of 0.6050, about a 40% decrease. By allowing the ellipse to be larger at low n, the vortex targets features at larger scales, then focuses in on smaller scale structures as t increases and the ellipses shrink, allowing for more effective mixing.

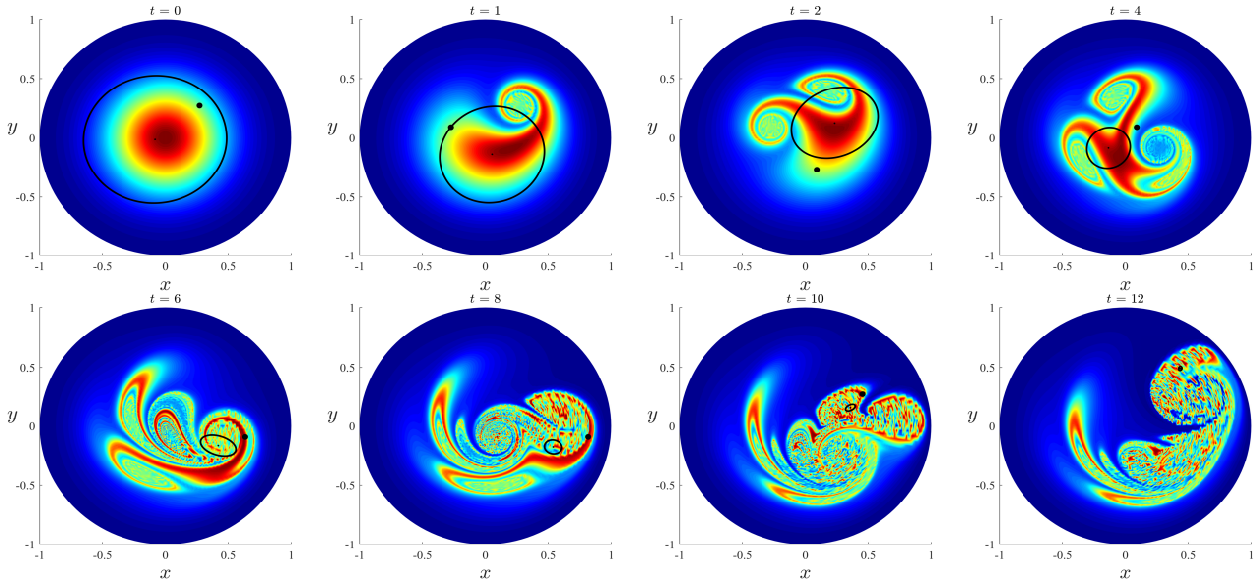

Figure 4.6: Evolution of ρ^t for $t = 0, 1, 2, 4, 6, 8, 10, 12$ with initial condition (4.1) and $N_r = N_\theta = 128$. Parameters of the RV algorithm are $N_z = 1005$, $N_v = 88$ and ϱ and α vary according to (4.5).

The mix norm evolution for the RV algorithm with constant and nonconstant parameters is compared in Figure 4.7. The mix norm decays vary slowly in the case of the constant parameters, even increasing at $t = 4, 6, 8$. The version with the varying parameters decays much faster and more consistently. However, there still seem to be flaws in the method, as the mix norm does not decrease sharply until $t = 6$, and stays mostly constant at $t = 7, 9$. Though there appears to be good mixing where the density is localized, the algorithm struggles to distribute the density to the entire domain, as it tends to stay mostly centralized near the origin. This is especially clear in the final plot of Fig. 4.5, where the majority of the density is clumped into a disc. There are two

possible explanations for this. One is that the algorithm itself is designed to work locally rather than globally, and therefore will be biased towards stirring where the most density is located. This approach is inherently flawed because mixing is a global process, and trying to address it as a local process is an oversimplification.

The other explanation for the density clumping effect is that the algorithm only considers one step in the future. It may be more optimal to forgo the largest decrease in mixing at the current step in order to setup the density profile for more effective mixing in future steps. In the next section, we will consider a stirring scheme that takes a more global approach to the vortex model.

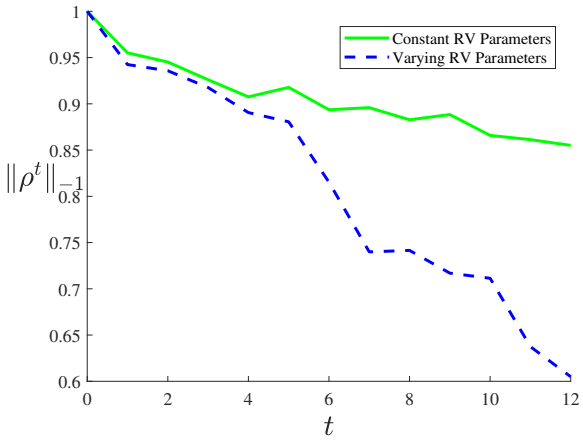

Figure 4.7: Decay of $\|\rho^t\|_{-1}$ for the RV stirring protocols with constant parameters (Fig. 4.5) and nonconstant parameters (Fig. 4.6).

4.2 Four Vortex Problem

The Roaming Vortex Problem was challenging because of too much freedom in choice of vortex location at each step. The Four Vortex Problem addresses this issue by only allowing four fixed locations for a vortex to appear: $\frac{1}{2}$, $\frac{i}{2}$, $-\frac{1}{2}$, and $-\frac{i}{2}$. For this problem we will allow a choice of either counterclockwise ($s = 1$) and clockwise ($s = 1$) spin for each vortex. Thus, at a given step t there are eight choices of stirring action: four choices of vortex location v_t and two choices

of rotation direction s_t. The eight stirring actions will be labeled as follows:

$$\left\{\begin{array}{c} +\mathrm{r} \\ +\mathrm{t} \\ +\mathrm{l} \\ +\mathrm{b} \\ -\mathrm{r} \\ -\mathrm{t} \\ -\mathrm{l} \\ -\mathrm{b} \end{array}\right\} = \left\{\begin{array}{cc} v=\frac{1}{2}, & s=1 \\ v=\frac{i}{2}, & s=1 \\ v=-\frac{1}{2}, & s=1 \\ v=-\frac{i}{2}, & s=1 \\ v=\frac{1}{2}, & s=-1 \\ v=\frac{i}{2}, & s=-1 \\ v=-\frac{1}{2}, & s=-1 \\ v=-\frac{i}{2}, & s=-1 \end{array}\right\}, \tag{4.6}$$

where r, t, l, and b, stand for right, top, left, and bottom respectively. The value ω_t will be used to denote the v_t and s_t pair label from (4.6). We will allow the total number of vortex switches N to vary, but hold the total time T the vortices are activated to remain constant. We will also hold the time between vortex switches constant, so that $\tau_t = \frac{T}{N}$.

For a given choice of N, the total possible stirring protocols will be 8^N. When searching for the optimal protocol, we can narrow down some of these choices. For example, it is obvious that the optimal protocol will not have $v_{n+1} = v_t$ and $s_{n+1} = -s_t$, i.e. it will never be the case that $(\omega_t, \omega_{t+1}) = (+\mathrm{r}, -\mathrm{r}), (-\mathrm{r}, +\mathrm{r}), (+\mathrm{t}, -\mathrm{t}), \ldots$ That would put a vortex in the same location as the previous step, but spinning in the opposite direction, which is the same as applying the inverse map and canceling out the stirring action. This means that there are eight choices for the first vortex and only seven for all following vortices, reducing the number of possible protocols to $8(7^{N-1})$. For all results we will use the initial density profile (4.1). The vortices are symmetric with respect to this initial condition, so the initial choice of vortex location and rotation is trivial. We shall always choose $\omega_1 = +r$ with $v_1 = \frac{1}{2}$ and $s_1 = 1$. The one-step action of this vortex is shown in Fig. 4.8.

We are left with only 7^{N-1} protocols to search through. Therefore, finding the true optimal protocol is computationally feasible for small N. Figure 4.9 shows the optimal stirring protocols for $N = 2, \ldots, 8$. There are many interesting things to note about the vortex choices. One is that a vortex will never be chosen twice in a row, suggesting that an optimal scheme must change

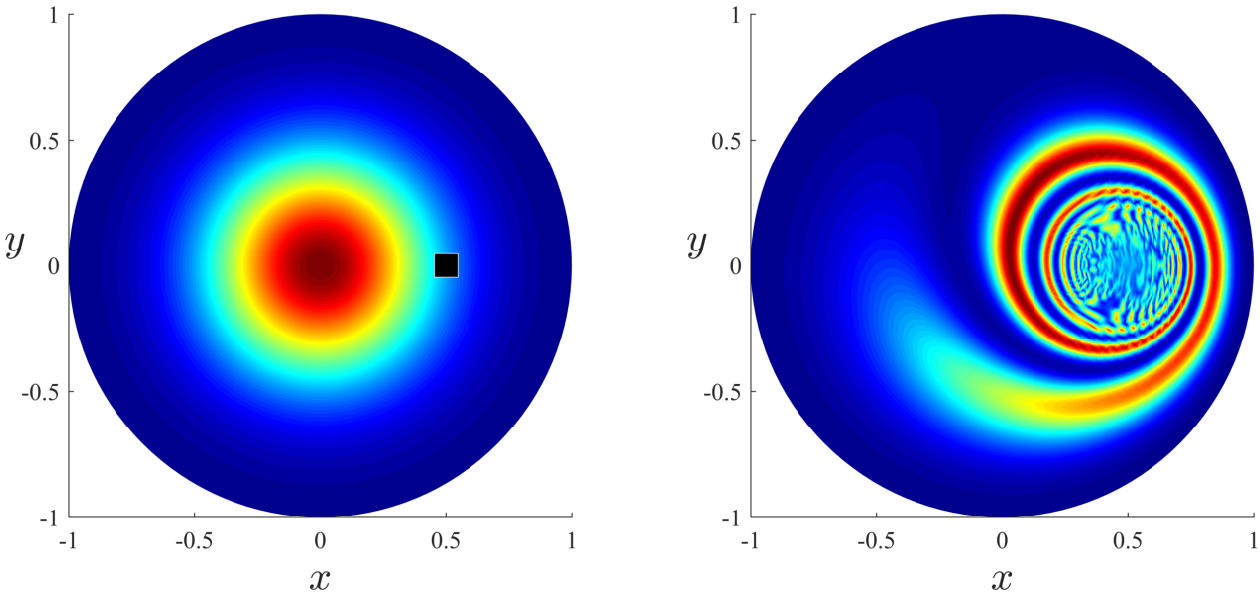

Figure 4.8: One-step evolution of initial density (4.1) with default first vortex choice $\omega_1 = +r$ and $N_r = N_\theta = 2^8$. Plot $t = 0$ shows ρ^0 with a square marking the vortex $v_1 = \frac{1}{2}$. Plot $t = 1$ shows ρ^1 after vortex at v_1 was applied counterclockwise.

the vortex location at every opportunity. This seems reasonable, since the initial activation of a vortex accomplishes a large shift in density, but continuing to activate that same vortex will only accomplish a spiraling effect near the vortex and have little impact on overall mixing. Instead, the vortex locations should be alternated in order to continue moving the unmixed regions throughout the domain.

Another important observation of Fig. 4.9 is that opposite vortices share a rotation direction that differs from that of the adjacent vortices. Figure 4.10 illustrates why this scheme outperforms the other six choices for relative rotation choice for the four vortices. To create these plots, we activate the vortex locations in the same order as the optimal $N = 4$ protocol, but vary the relative rotation directions. The common feature in the less optimal schemes is two adjacent vortices that have the same rotation direction. Anytime two neighboring vortices stir in the same direction, the majority of density gets shifted to that quadrant of the disc, causing it to be less effected by the action of the other vortices. Keeping adjacent vortices of opposite sign will ensure that the density

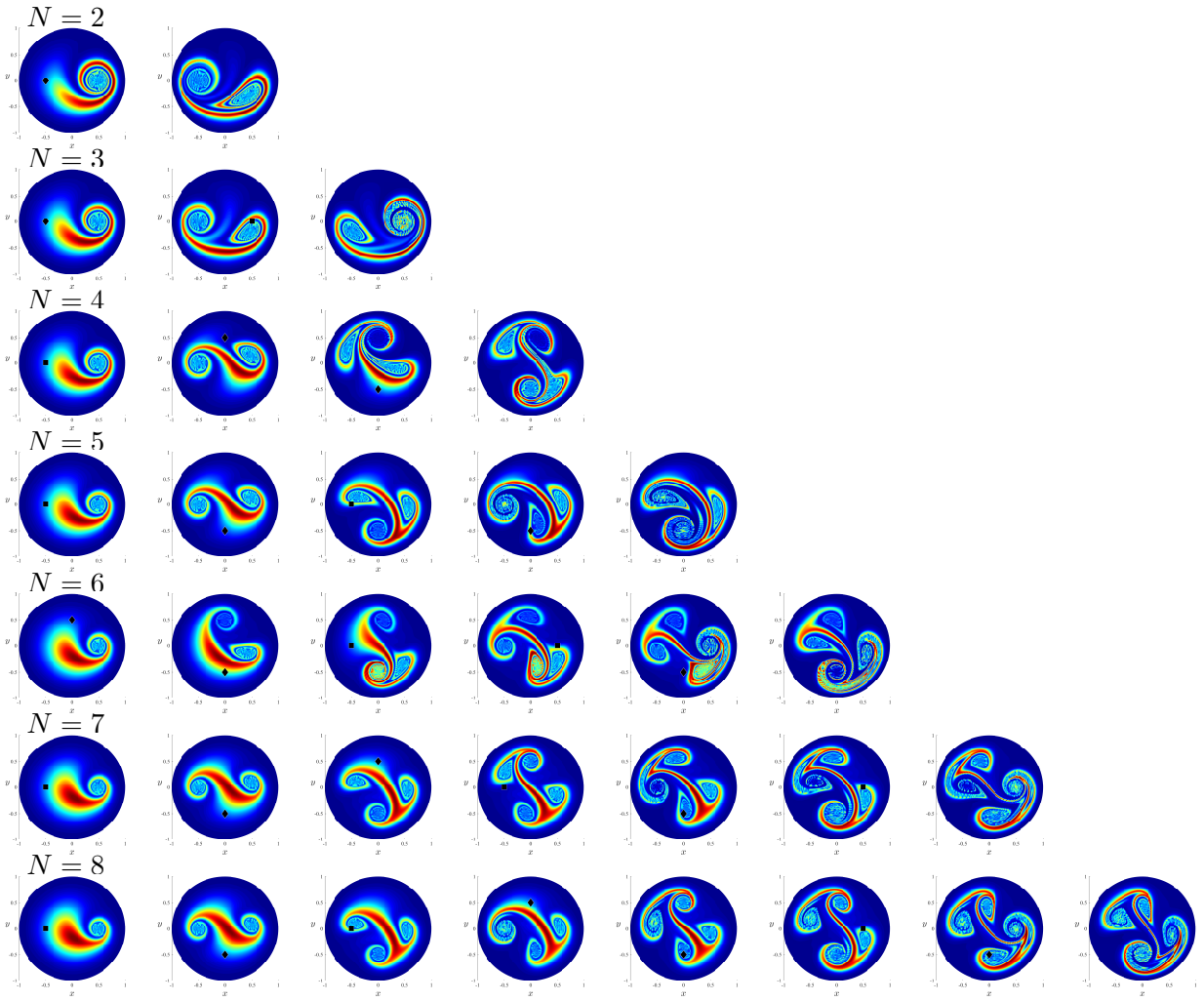

Figure 4.9: Evolution of ρ^t for the optimal stirring protocols of the FV problem for $N = 2, \ldots, 8$ with initial condition (4.1) and $N_r = N_\theta = 2^7$. Squares mark where a vortex will be activated counterclockwise, diamonds mark where a vortex will be activated clockwise.

remains centralized enough that all four vortices remain effective.

The shape of the final plots for the rows corresponding to $N = 2$ and $N = 3$ of Fig. 4.9 are very similar, since they both result from the vortex choices $\omega = +\mathrm{r}, +\mathrm{l}$. Since two vortex locations are activated, we see the formation of two "bulbs". For the $N = 4$ case, we see four bulbs because all four vortex locations are used. The $N = 5$ solution only uses three of the vortices. The rightmost plots for $N = 6, 7, 8$ appear very similar to $N = 4$, with the same four bulb shape. Figure 4.11 (a)

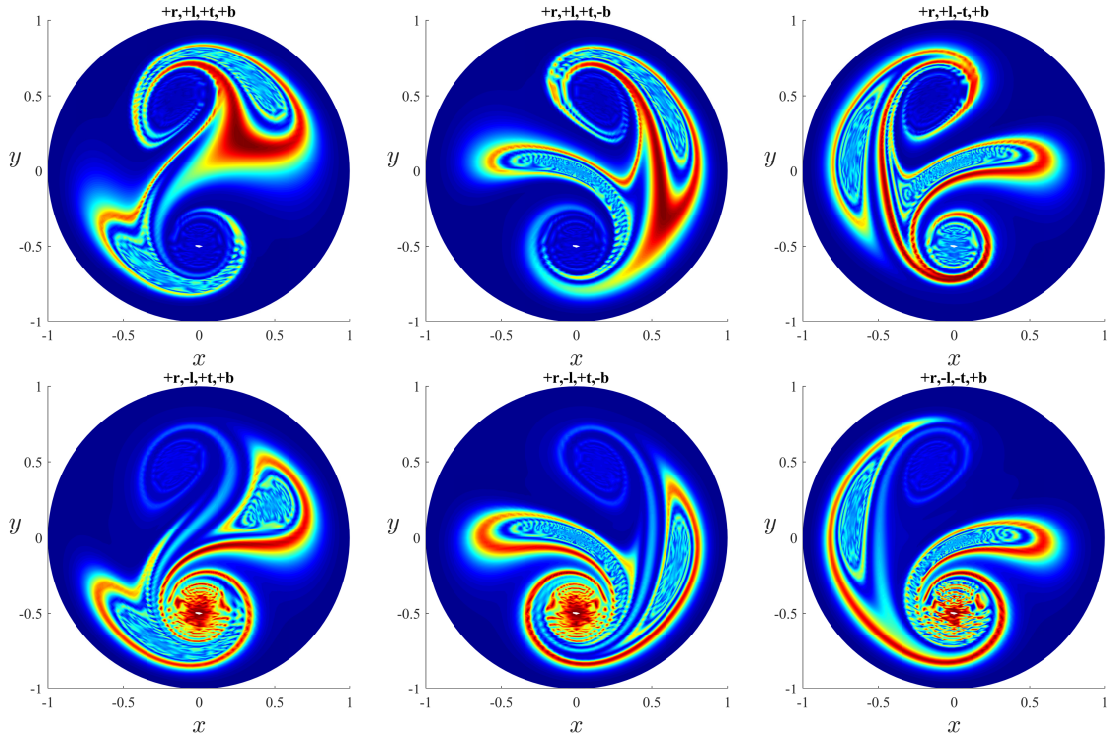

Figure 4.10: ρ^4 after applying the same v_t as the optimal 4-step FV scheme (Third row of Fig. 4.9), with all possible other choices of rotation direction s_t.

compares the decay of mix norm for the optimal $N = 1, \ldots, 8$ solutions. Figure 4.11 (b) shows how $\|\rho^N\|_{-1}$ decays with respect to N, quickly at first, but then it begins to level off. This suggests that allowing more and more vortex switches may not cause a notable decrease in mix norm.

To get an indication of how the mix norm will decay for larger N, we consider the case where we allow N to tend towards infinity. At large N, the discrete action of the blinking vortices resembles a Hamiltonian flow in which we are applying all four vortices simultaneously at a quarter strength. The Hamiltonian of (2.25) can be easily derived to be

$$H(z) = -s \log \left| \frac{z - v}{z - \frac{1}{\bar{v}}} \right|. \tag{4.7}$$

For the infinite four vortex case, the Hamiltonian becomes

$$H_4(z) = -\frac{1}{4} \left(s_{\frac{1}{2}} \gamma_{\frac{1}{2}} \log \left| \frac{z - \frac{1}{2}}{z - 2} \right| + s_{\frac{i}{2}} \gamma_{\frac{i}{2}} \log \left| \frac{z - \frac{i}{2}}{z - i2} \right| + s_{-\frac{1}{2}} \gamma_{-\frac{1}{2}} \log \left| \frac{z + \frac{1}{2}}{z + 2} \right| + s_{-\frac{i}{2}} \gamma_{-\frac{i}{2}} \log \left| \frac{z + \frac{i}{2}}{z + 2i} \right| \right), \tag{4.8}$$

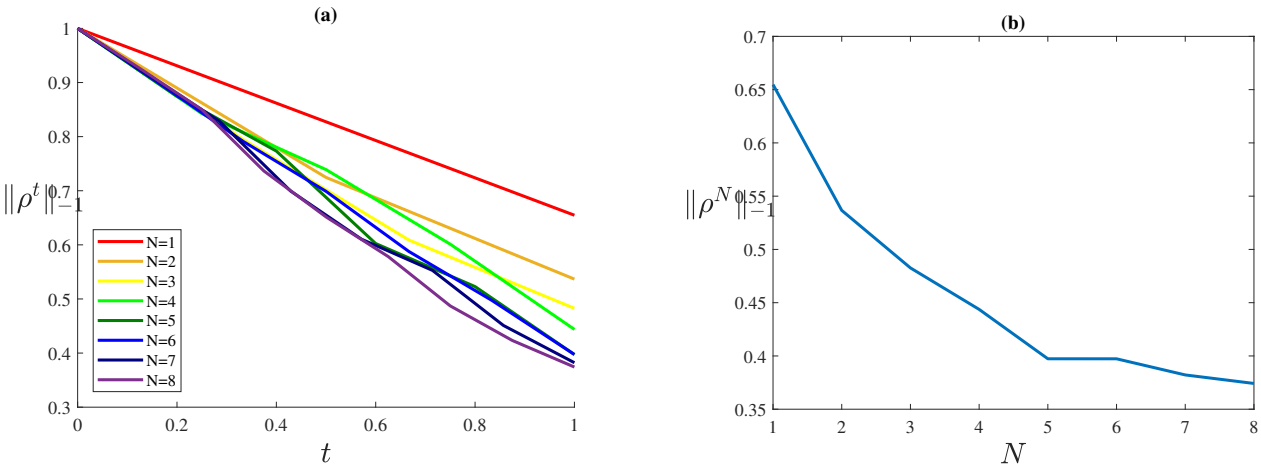

Figure 4.11: Mix norms of the FV optimal stirring protocols for $N = 1, \dots 8$ shown in Figures 4.8 and 4.9. Plot (a) shows the decay of $\|\rho^t\|_{-1}$ with respect to t . Plot (b) shows the decay of $\|\rho^N\|_{-1}$ with respect to N.

where γ_v specifies the proportion of time that vortex v is activated. We require that $\gamma_v \in [0, 1]$ and that

$$\gamma_{\frac{1}{2}} + \gamma_{\frac{i}{2}} + \gamma_{-\frac{1}{2}} + \gamma_{-\frac{i}{2}} = 1. \tag{4.9}$$

We make the assumption that the sign s_v of a given vortex remains constant, since switching the direction of rotation would result in "unstirring" and would be inefficient. Once again we set $s_{\frac{1}{2}} = 1$ to be the default. This leaves the following eight choices for the four stirring directions:

$$\left\{ s_{\frac{1}{2}}, s_{\frac{i}{2}}, s_{-\frac{1}{2}}, s_{-\frac{i}{2}}) \right\} = \left\{ \begin{array}{c} 1,1,1,1 \\ 1,1,1,-1 \\ 1,-1,1,1 \\ 1,1,-1,-1 \\ 1,-1,1,-1 \\ 1,-1,-1,1 \\ 1,-1,-1,-1 \end{array} \right\}. \tag{4.10}$$

We use a Random Optimal algorithm, similar to the RO algorithm described in §3.1.3, by randomly sampling vectors $(\gamma_{\frac{1}{2}}, \gamma_{\frac{i}{2}}, \gamma_{-\frac{1}{2}}, \gamma_{-\frac{i}{2}})$ from the section of four dimensional hyperplane (4.9), such that

$\gamma_v \in [0,1]$, selecting randomly from the eight sign choices (4.10) and choosing the proportion and sign vectors that minimize $\|\rho^\infty\|_{-1}$. We approximate the infinite case by setting N to be a large number, and activate each vortex for time $\gamma_v \frac{T}{N}$. The proportions of time the vortices are activated for each time step must match the proportions of the total time T. This uniform application of the vortices is necessary to simulate the infinite case, since it should be as if we are activating all four vortices at once. We continue to fix T at 1, and we set $N = 100$.

Figure 4.12 shows the results of the Random Optimal algorithm applied for 10^4 trials. The near optimal solution produced by the algorithm is $\gamma = (0.47, 0.4, 0.01, 0.11)$ and $s = (1, -1, 1, -1)$, which achieves a mix norm of 0.4129. The relative direction choices echo that of the optimal solution from Fig. 4.9. What is surprising is that the scheme favors two of the vortices heavily, barely using the other two. In the case of the schemes from Fig. 4.9, it was best to activate all vortices relatively equally so as to keep the density centralized. However, in the infinite case, it seems more optimal to weight the density towards one half of the domain where two vortices can have the most effect.

The most important takeaway from the Roaming Vortex Problem was that a local approach to a global problem will not yield good results, and too much freedom in parameter choices makes it difficult to produce any valid conclusions. The Four Vortex Problem addressed these issues with a simplified model in which we limited the possible locationss of the vortices, but introduced new elements of rotation direction and number of vortex switches. The optimized stirring protocols of this scheme were far more successful than those of the RV Problem, able to achieve a mix norm of about 0.37 in the case of the 8 step protocol as opposed to a mix norm of 0.6 in the case of RV solution with varying parameters. It became clear that it is not enough to use the action of the vortices to induce good one-step mixing, but it also important to use them to keep the density positioned in a way so that other vortices can be most effective in future steps.

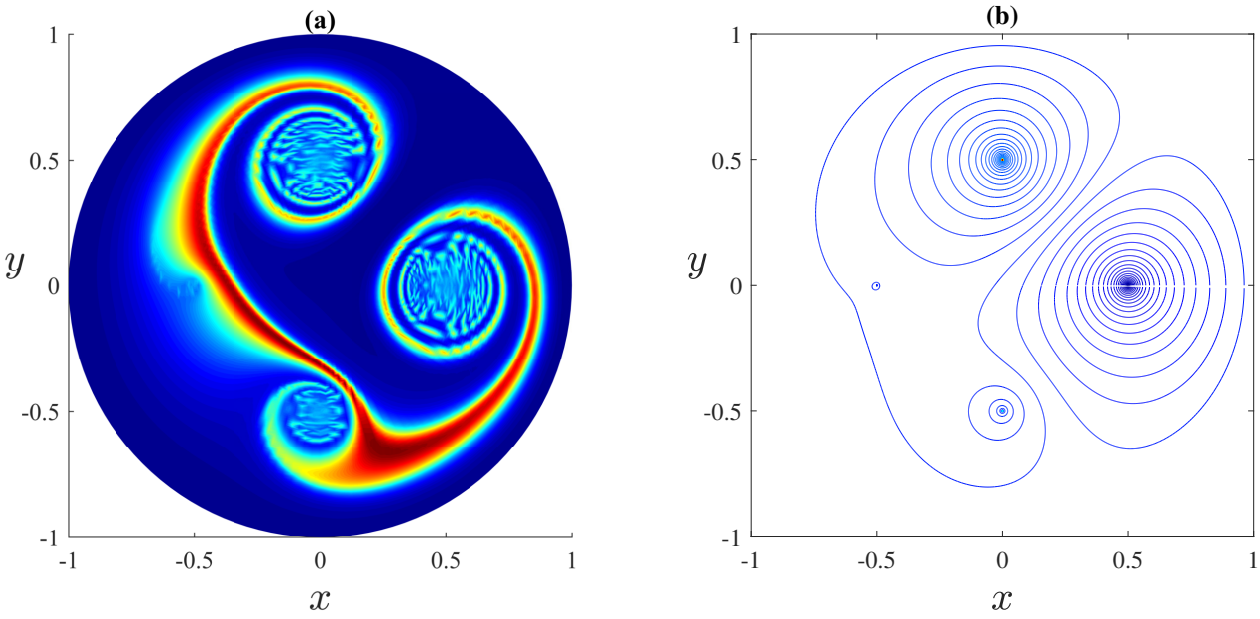

Figure 4.12: Results of the Random Optimal algorithm applied with 10^4 trials to find the proportion vector $(\gamma_{\frac{1}{2}}, \gamma_{\frac{i}{2}}, \gamma_{-\frac{1}{2}}, \gamma_{-\frac{i}{2}})$ which satisfies $\gamma_v \in [0,1]$ and (4.9) and the sign vector $(s_{\frac{1}{2}}, s_{\frac{i}{2}}, s_{-\frac{1}{2}}, s_{-\frac{i}{2}})$ which minimize $\|\rho^{100}\|_{-1}$ with $T = 1$ and $\tau = \frac{1}{100}$. The vectors produced by the algorithm are $\gamma = (0.47, 0.4, 0.01, 0.11)$ and $s = (1, -1, 1, -1)$. Plot (a) shows the density profile after the scheme is applied. Plot (b) shows the contours of the Hamiltonian (4.8).

Chapter 5

Conclusions

We investigated the design of mixers by studying the optimization of stirring behavior of a two-dimensional device modeled by a finite number of nonlinear shears, under a constraint of fixed energy, and with a finite number of blinking vortex models, varying vortex location, rotation direction, and number of "blinks" between vortices. The minus-one Sobolev seminorm, one example of a "mix-norm", provides a cost-effective way to measure mixing efficiency without explicitly computing diffusion. We computed the results of stirring by evaluating the Perron-Frobenius operator on a uniform grid, assuming that the initial density profile is known analytically. This allowed us to avoid the aliasing errors and numerically induced diffusivity of other methods.

One of the most important insights from the results of Chapter 3 is that effective mixing is highly dependent on the behavior of the lower order Fourier modes: in order to construct a good stirring scheme, one must target the low-order modes with the highest amplitudes. For the first model, a composition of Chirikov standard maps, concentrating the nonlinear shear energy in a small number of steps was shown to give an efficient mixing scheme. For this map, the horizontal shear is assumed fixed, and using the energy for vertical shearing early in the protocol takes maximal advantage of the shift in Fourier coefficients to higher mode numbers. It is probable that this approach would also maximize the effects of any nonzero diffusivity, since the striations are formed early in the evolution.

For the Harper map model, both vertical and horizontal shears can be controlled, allowing for more diverse stirring mechanisms and more interesting behaviors. A simplified scheme assumed

that each step corresponds to either a horizontal or a vertical shear with constant energy. Within this reduced parameter space, we devised an algorithm that combines insight on the action of the shears in Fourier space with knowledge of the character of the fixed points, to produce a near-optimal stirring protocol. The method was shown to remain effective when applied to a variety of initial density profiles.

Table 3.1 compares the optimal stirring protocols produced by the different methods when applied to the initial condition (3.1), and Table 3.2 compares the overall performance of solvers for varying initial conditions (3.7). Our results suggest that both consistency and change—contradictory requirements—are necessary to produce an optimal stirring protocol. While the former leads to stretching near saddle points, the latter can result in enhanced folding. Both were required to obtain the optimal protocol for the Harper map. With only stretching, the mix-norm tends to level off; but with too much folding, stirring can be reversed, causing the mix-norm to grow. An optimal protocol will tend to use hyperbolic behavior if a smooth portion of the density is concentrated near a fixed point—if a nearby fixed point is elliptic, mixing efficiency can be decreased even though there will be (weak) shearing. Though it is clear that both stretching and folding are required to design an effective mixer, it is less clear how to best combine these processes to optimize mixing. Observing stretching and folding in a wider variety of mixing problems may lead to more general results about how one should balance the two mechanisms to induce mixing.

The shearing models studied have their fixed points at fixed locations. The blinking vortex model studied in Chapter 4 allows for the fixed points of the map to change locations at every step. Advection created by the shearing problems was more uniform, but in the case of vortices the advection was more localized. This offered a more complex investigation with more choices of parameters. Since the complexity of the problem was too large, we limited the parameter choices in two ways: by means of the Roaming Vortex Problem and the Four Vortex Problem.

The parameters where chosen in the Roaming Vortex Problem by fixing the rotation direction, time, and number of vortex switches, and allowing the location of the vortex to be free. We devised a scheme in which we tracked the movement of the most unmixed region and chose the vortex location

to optimize the variance of this region. The algorithm was problematic because it depended on many numerical parameters that determined how to identify the location, size, and shape of the unmixed region. The main issue was that, as the density became more mixed, the length scales shrank, requiring these numerical parameters to change as the number of steps increased. In addition, the localized approach was not conducive for spreading the density throughout the domain. The algorithm achieved modest success, allowing for a 40% decrease in mix norm. Experiments with the RV Problem made it clear that it would be worthwhile to simplify the complexity of the model, so as to allow for a more focused approach.

To remedy the numerical problems that arose from the RV Problem, we devised the Four Vortex Problem by confining the vortices to four possible locations, holding the total time fixed, and allowing a choice of number of vortex switches and vortex rotation direction. This confined the possible stirring protocols to a finite set, which could be searched through exhaustively to find the true optimal protocol for a given number of vortex switches. The main result of the experiments with the FV model was that the most effective protocols oriented the stirring direction of the vortices so as to keep the density centralized and allowed all vortex locations to remain effective as the mixing progressed.

A key observation that arose from both the shear maps and the blinking vortex map was that, to devise an effective finite-time stirring scheme, it is not enough to try to optimize the mixing at every step. This is evidenced by the fact that algorithms that only looked one step in the future, the SW minimizer in the case of the Harper map, and the Roaming Vortex algorithm in the case of the blinking vortex model, performed worse than algorithms that took a global approach, i.e. the full Random Optimal algorithm for the Harper map and the Four Vortex algorithm for the blinking vortex model. Schemes that were most effective included steps of "setup" in addition to steps of mixing, i.e., steps that were devoted to reorienting the density, so that future steps of mixing could be more effective. More generally, any mixer being applied to a fluid with little diffusion most use a combination of stretching and folding actions. As a future avenue of study, it would be worthwhile to investigate more general results about what ratio of setup and mixing steps is required for a

mixer to be effective, and how to identify which steps should do which.

Though interesting, the problems studied in this investigation are simplified two-dimensional stirring models. It would be worthwhile to study models that are more physically applicable. In particular, it would be interesting to study three dimensional models such as the Kenics Static Mixer and the rotated arc mixer discussed in Chapter 1 [22, 42]. Our investigation resulted in the design of adaptive mixing algorithms that changed stirring parameters as they went along based on the current density profile. It would be interesting to attempt to devise such adaptive schemes for these three dimensional problems and attempt to understand what sort of physical mechanisms could implement the adaptive scheme. The challenge for three-dimensional problems would be to devise an efficient model, rather than resorting to solving the full Navier-Stokes equation.

Bibliography

[1] M. Abramowitz and I.A. Stegun. Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables. Dover Publications, Inc., New York, 1974.

[2] H. Aref. Stirring by chaotic advection. J. Fluid Mech., 143:1–21, 1984. http://dx.doi.org/10.1017/S0022112084001233.

[3] V.I. Arnold and A. Avez. Ergodic Problems of Classical Mechanics. Benjamin, New York, 1968.

[4] R. Artuso, G. Casati, F. Borgonovi, L. Rebuzzini, and I. Guarneri. Fractal and dynamical properties of the kicked Harper model. Int. J. Mod. Phys. B, 08(03):207–235, 1994. http://dx.doi.org/10.1142/S0217979294000099.

[5] R. Burton and R.W. Easton. Ergodicity of linked twist maps. In Z. Nitecki and C. Robinson, editors, Global Theory of Dynamical Systems, volume 819 of Springer Lecture Notes in Math., pages 35–49. Springer-Verlag, Berlin, 1980.

[6] S. Cerbelli, V. Vitacolonna, A. Adrover, and M. Giona. Eigenvalue eigenfunction analysis of infinitely fast reactions and micromixing regimes in regular and chaotic bounded flows. Chem. Eng. Sci., 59(11):2125–2144, 2004. http://dx.doi.org/10.1016/j.ces.2004.02.005.

[7] Q. Chen, I. Dana, J.D. Meiss, N. Murray, and I.C. Percival. Resonances and transport in the sawtooth map. Physica D, 46(2):217–240, 1990. http://www.sciencedirect.com/science/article/pii/016727899090037P.

[8] Q. Chen and J.D. Meiss. Flux, resonances and the devil's staircase for the sawtooth map. Nonlinearity, 2(2):347–356, 1989. http://iopscience.iop.org/0951-7715/2/2/009.

[9] B.V. Chirikov. A universal instability of many-dimensional oscillator systems. Phys. Rep., 52(5):263–379, 1979. http://dx.doi.org/10.1016/0370-1573(79)90023-1.

[10] L. Cortelezzi, A. Adrover, and M. Giona. Feasibility, efficiency and transportability of short-horizon optimal mixing protocols. J. Fluid Mech., 597:199–231, 2008. http//dx.doi.org/10.1017/S0022112007009676.

[11] I.J. Couchman and E.C. Kerrigan. Control of mixing in a stokes' fluid flow. J. of Process Control, 20(10):1103–1115, 2010. http://doi.org/10.1016/j.jprocont.2010.06.014.

[12] D. D'Alessandro, M. Dahleh, and I. Mezić. Control of mixing in fluid flow: A maximum entropy approach. IEEE Trans. Automat. Control, 44(10):1852–1863, 1999.

[13] P.V. Danckwerts. The definition and measurement of some characteristics of mixtures. Appl. Sci. Res. A, A3(4):279–296, 1952. http://dx.doi.org/10.1007/BF03184936.

[14] P.V. Danckwerts. Theory of mixtures and mixing. Research, 6:355–361, 1953.

[15] K.M. Frahm and D. L. Shepelyansky. Ulam method for the chirikov standard map. Eur. Phys. J. B, 76(1):57–68, 2010. http://dx.doi.org/10.1140/epjb/e2010-00190-6.

[16] J. G. Franjione, C.-W. Leong, and J. M. Ottino. Symmetries within chaos: A route to effective mixing. Physics of Fluids A: Fluid Dynamics, 1(11):1772–1783, 1989. http://dx.doi.org/10.1063/1.857504.

[17] J.G. Franjione and J.M. Ottino. Symmetry concepts for the geometric analysis of mixing flows. Phil. Trans. R. Soc. Lond. A, 338(1650):301–323, 1992. http://dx.doi.org/10.1098/rsta.1992.0010.

[18] G. Froyland. Statistically optimal almost-invariant sets. Physica D, 200(3-4):205–219, 2005. http://dx.doi.org/10.1016/j.physd.2004.11.008.

[19] G. Froyland, C. Gonzalez-Tokman, and T.M. Watson. Optimal mixing enhancement by local perturbation. SIAM Review, 58(3):494–513, 2016. http://dx.doi.org/10.1137/15M1023221.

[20] 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. http://dx.doi.org/10.1016/j.physd.2009.03.002.

[21] M. Funakoshi. Chaotic mixing and mixing efficiency in a short time. Fluid Dynam. Res., 40(1):1–33, 2008. http://dx.doi.org/10.1016/j.fluiddyn.2007.04.004.

[22] O. S. Galaktionov, P. D. Anderson, G. W. M. Peters, and H. E. H. Meijer. Analysis and optimization of Kenics static mixers. Intern. Polymer. Processing, XVIII(2):138–150, 2003.

[23] S. Gibout, Y. Le Guer, and E. Schall. Coupling of a mapping method and a genetic algorithm to optimize mixing efficiency in periodic chaotic flows. Comm. Nonl. Sci and Num. Sim., 11(3):413–423, 2006. http://dx.doi.org/10.1016/j.cnsns.2004.11.006.

[24] 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. Phys. Fluids, 24(7):073603, 2012. http://dx.doi.org/10.1063/1.4738598.

[25] O. Gubanov and L. Cortelezzi. Towards the design of an optimal mixer. J. Fluid Mech., 651:27–53, 2010. http://dx.doi.org/10.1017/S0022112009993806.

[26] O. Gubanov and L. Cortelezzi. On the cost efficiency of mixing optimization. J. Fluid Mech., 692:112–136, 2012. http://dx.doi.org/10.1017/jfm.2011.498.

[27] G. Haller. Lagrangian coherent structures. Annual Review of Fluid Mechanics, 47:137–162, 2013. http://dx.doi.org/10.1146/annurev-fluid-010313-141322.

[28] K.-Y. Hsiao, C.-Y. Wu, and Y.-T. Huang. Fluid mixing in a microchannel with longitudinal vortex generators. Chem. Eng. J., 235:27–36, 2014. http://dx.doi.org/10.1016/j.cej.2013.09.010.

[29] C.-P. Jen, C.-Y. Wu, Y.-C. Lin, and C.-Y. Wu. Design and simulation of the micromixer with chaotic advection in twisted microchannels. Lab on a Chip, 3(2):77–81, 2003. http://dx.doi.org/10.1039/B211091A.

[30] P. G. M. Kruijt, O. S. Galaktionov, P. D. Anderson, G. W. M. Peters, and H. E. H. Meijer. Analyzing mixing in periodic flows by distribution matrices: Mapping method. AIChE J., 47(5):1005–1015, 2001. http://dx.doi.org/10.1002/aic.690470507.

[31] A. Lasota and M. Mackey. Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics. Springer-Verlag, New York, 2nd edition, 1994.

[32] Z. Lin, J.-L. Thiffeault, and C.R. Doering. Optimal stirring strategies for passive scalar mixing. J. Fluid Mech., 675:465–476, 2011. http://dx.doi.org/10.1017/S002211201000563X.

[33] M. Liu, F.J. Muzzio, and R.L. Peskin. Quantification of mixing in aperiodic chaotic flows. Chaos, Solitons and Fractals, 4(6):869–893, 1994. http://dx.doi.org/10.1016/0960-0779(94)90129-5.

[34] G. Marsaglia. Choosing a point from the surface of a sphere. The Annals of Mathematical Statistics, 43(2):645–646, 1972. http://dx.doi.org/10.1214/aoms/1177692644.

[35] G. Mathew, I. Mezić, S. Grivopoulos, U. Vaidya, and L. Petzold. Optimal control of mixing in stokes fluid flows. J. Fluid Mech., 580:261–281, 2007. https://doi.org/10.1017/S0022112007005332.

[36] G. Mathew, I. Mezić, and L. Petzold. A multiscale measure for mixing. Physica D, 211(1-2):23–46, 2005. http://dx.doi.org/10.1016/j.physd.2005.07.017.

[37] J.D. Meiss. Symplectic maps, variational principles, and transport. Rev. Mod. Phys., 64(3):795–848, 1992. http://dx.doi.org/10.1103/RevModPhys.64.795.

[38] R.A. Mitchell and J.D. Meiss. Designing a finite-time mixer: Optimizing stirring for two-dimensional maps. in Press, 2017. http://adsabs.harvard.edu/abs/2017arXiv170105620M.

[39] S. Shinohara. The threshold for global diffusion in the kicked Harper map. Phys. Lett. A, 298(5):330–334, 2002. http://dx.doi.org/10.1016/S0375-9601(02)00579-0.

[40] M. K. Singh, O. S. Galaktionov, H. E. H. Meijer, and P. D. Anderson. A simplified approach to compute distribution matrices for the mapping method. Computers & Chemical Eng., 33(8):1354–1362, 2009. http://dx.doi.org/10.1016/j.compchemeng.2009.01.021.

[41] T. H. Solomon, S. Tomas, and J.L. Warner. Role of lobes in chaotic mixing of miscible and immiscible impurities. Phys. Rev. Lett., 77:2682–2685, 1996. https://doi.org/10.1103/PhysRevLett.77.2682.

[42] M.F.M. Speetjens, E.A. Demissie, G. Metcalfe, and H.J.H. Clercx. Lagrangian transport characteristics of a class of three-dimensional inline-mixing flows with fluid inertia. Phys. Fluids, 26(11):113601, 2014. http://dx.doi.org/10.1063/1.4901822.

[43] R. S. Spencer and R. M. Wiley. The mixing of very viscous liquids. J. Colloid Science, 6(2):133–145, 1951. http://dx.doi.org/10.1016/0095-8522(51)90033-5.

[44] R. Sturman, S.W. Meier, J.M. Ottino, and S. Wiggins. Linked twist map formalism in two and three dimensions applied to mixing in tumbled granular flows. J. Fluid Mech., 602:129–174, 2008. http://dx.doi.org/10.1017/S002211200800075X.

[45] R. Sturman, J.M. Ottino, and S. Wiggins. The Mathematical Foundations of Mixing: The Linked Twist Map as a Paradigm in Applications : Micro to Macro, Fluids to Solids, volume 22 of Cambridge monographs on applied and computational mathematics. Cambridge University Press, Cambridge, 2006.

[46] R. Sturman and J.-L. Thiffeault. Lyapunov exponents for the random product of two shears. ArXiv e-prints, 2017. https://arxiv.org/abs/1706.03398.

[47] J.-L. Thiffeault. Stretching and curvature of material lines in chaotic flows. Physica D, 198(3-4):169–181, 2004. http://dx.doi.org/10.1016/j.physd.2004.04.009.

[48] J-L. Thiffeault. Using multiscale norms to quantify mixing and transport. Nonlinearity, 25(2):R1–R44, 2012. http://dx.doi.org/10.1088/0951-7715/25/2/R1.

[49] S.M. Ulam. A Collection of Mathematical Problems, volume 8 of Interscience Tracts in Pure and Applied Mathematics. Interscience, New York, 1960.

[50] F. Városi, T.M. Antonsen, and E Ott. The spectrum of fractal dimensions of passively convected scalar gradients in chaotic fluid flows. Phys. Fluids, A3(5):1017–1028, 1991. http://dx.doi.org/10.1063/1.858081.