A Multiscale Measure for Mixing

George Mathew, Igor Mezić and Linda Petzold *

Department of Mechanical and Environmental Engineering
University of California, Santa Barbara, CA 93106, USA.

Abstract

We present a multiscale measure for mixing that is based on the concept of weak convergence and averages the "mixedness" of an advected scalar field at various scales. This new measure, referred to as the Mix-Norm, resolves the inability of the L^2 variance of the scalar density field to capture small-scale variations when advected by chaotic maps or flows. In addition, the Mix-Norm succeeds in capturing the efficiency of a mixing protocol in the context of a particular initial scalar field, wherein Lyapunov-exponent based measures fail to do so. We relate the Mix-Norm to the classical ergodic theoretic notion of mixing and present its formulation in terms of the power spectrum of the scalar field. We demonstrate the utility of the Mix-Norm by showing how it measures the efficiency of mixing due to various discrete dynamical systems and to diffusion. In particular, we show that the Mix-Norm can capture known exponential and algebraic mixing properties of certain maps. We also analyze numerically the behaviour of scalar fields evolved by the Standard Map using the Mix-Norm.

1 Introduction

Fluid mixing is a critical stage in many engineering applications. Aref [1] has studied the use of chaotic advection to enhance mixing in laminar flows. Books by Ottino [12] and Wiggins [20] address the problem of mixing using concepts and methods of dynamical systems theory. [1], [12] and [20] discuss physical (kinetic) mechanisms for mixing. In spite of this comprehensive study of mixing from the point of view of dynamical systems theory, there is no consensus on how to measure mixing and in particular on how to compare the mixing rates of two different processes. The notion of a mixing measure becomes particularly useful if one is considering the problem of control and optimization of mixing. Previous approaches to this fundamental problem of measurement of mixing include using the Kolmogorov-Sinai entropy of the underlying dynamical system

as an objective for mixing and using the scalar variance of a density field which is being tranported by a dynamical system. Control of mixing using a maximum entropy approach for a prototypical mixing problem was studied in [6]. As the authors themselves point out in [6], the entropy of a dynamical system (given by a spatial integral of the Lyapunov exponents) is independent of the initial fluid configuration. Therefore, if we are interested in "mixing up" a particular initial scalar field in an optimal manner, the maximum entropy approach is not applicable, although one could use ideas like partition entropy which is discussed in [4].

Besides the entropy approach mentioned above, L^p norms have been used when the problem at hand is not purely deterministic, i.e. includes diffusion. In work by Ashwin et al. [3], interesting mixing protocols (combinations of diffusion with permutation operations on phase space) are described. Thiffeault et al. [17] investigate the mixing properties of a map which in an extension of Arnold's Cat Map on the 2-torus combined with diffusion. In [16], Rothstein et al. study the mixing patterns of passive scalars in an electromagnetically driven two-dimensional fluid flow. In [3], [17] and [16] the L^2 and L^∞ norms are used to quantify how far the passive scalar field is from being spatially uniform or homogeneous. In the absence of diffusion, measures based on L^p norms of a scalar density field being transported by a volume-preserving system will not decay. Therefore the L^p norms fail to quantify the "stirring" efficiency of a mixing process accurately because it is insensitive to the small scale structures of the scalar density field generated by the volume-preserving chaotic system. Mostly, this problem of the L^p norm has been ignored because typically there is diffusion associated with the mixing protocols as in [3], [17] and [16].

Other classical approaches, starting from Gibbs' approach, are reviewed in [9]. The key element of the Gibbs' approach is coarse-graining. We extend this idea and the work in [11] to develop a norm based on coarse-graining. This norm called the Mix-Norm is related to the concept of weak convergence and we demonstrate its relation to the classical ergodic theoretic notion of mixing as defined in dynamical systems literature [10], [2] and [14]. The Mix-Norm was motivated by the *mixing variance coefficient* proposed in [19]. The formulation of the Mix-Norm overcomes the deficiences of the two approaches mentioned above. The Mix-Norm depends on the initial fluid configuration and also succeeds in capturing the mixing efficiency of volume-preserving transformations (in the absence of diffusion) wherein the standard L^2 variance fails to do so. In this paper, we discuss the measure with respect to the mixing of advected density fields, but the ideas here can be easily extended to mixing in the context of dispersing solids or droplets. Also, in this paper we have restricted ourselves to periodic domains for ease of presentation and clarity.

The paper is organized as follows. In section 2, we present the basic struc-

tural definition and properties of the Mix-Norm. We discuss its properties as a pseudo-norm induced by an inner product and also its interpretation in terms of the power spectrum of the scalar field. In particular, we show that the Mix-Norm is equivalent to the Sobolev space norm of negative index $s = -\frac{1}{2}$, i.e. the $H^{-\frac{1}{2}}$ norm. In section 3, we discuss the applications of the Mix-Norm. We prove that the Mix-Norm can be used as a metric for checking weak-convergence which is the critical link in justifying its validity as a measure for mixing. We show how the Mix-Norm captures the mixing properties of Arnold's Cat Map and the integrable Standard Map. We numerically explore how the perturbation parameter affects the mixing propeties of the Standard Map using the Mix-Norm. We also give an interepretation of the Mix-Norm in terms of the mixing effectiveness caused by pure diffusion.

2 Structure and Properties of the Mix-Norm

As a motivation, consider the sequence of functions $\{c_m = sin(2m\pi x)\}$. They have the same mean and variance for all m, but there is something fundamentally different about the behaviour of these functions for small m and large m. In particular, the average of $sin(2m\pi x)$ over any open set converges to zero for large m as the functions become more "oscillatory". The most commonly used "global" or average quantities like the mean and the variance of scalar functions do not capture this property. We consider a scalar field to be well-mixed if its averages over arbitrary open sets are uniform. There are many ways of quantifying this idea and in this paper we choose a specific one that coincides well with ideas on mixing in dynamical systems theory ([10], [2], [14]) as well as with intuitive notions with respect to the power spectrum of the scalar field. In contrast to the L^2 norm of a function which is obtained by integrating the square of the function over the whole space, the Mix-Norm is obtained by integrating the square of average values of the function over a dense set of subsets contained in the whole space.

2.1 The Mix-Norm on the circle

First we present the Mix-Norm on the circle and then generalize it to an n-dimensional torus. We parametrize the domain by a non-dimensional distance x ranging from 0 to 1 and μ is the Lebesgue measure. Let $c : S^1 \to \mathbb{R}$. To define the Mix-Norm let

$$d(c, p, s) = \frac{\int_{p-s/2}^{p+s/2} c(x)\mu(dx)}{s} \tag{1}$$

for all $s \in (0, 1)$ and $p \in [0, 1]$. $d(c, p, s)$ is the mean value of the function c within the interval $[p - s/2, p + s/2]$. Then we define

$$\phi(c, s) = \left(\int_0^1 d^2(c, p, s)\mu(dp) \right)^{\frac{1}{2}}. \tag{2}$$

3

$\phi(c, s)$ is the L^2 norm of the averaged function $d(c, ., s)$ for a fixed scale s. Then the Mix-Norm of c is given by

$$\Phi(c) = \left(\int_0^1 \phi^2(c, s) \mu(ds) \right)^{\frac{1}{2}}. \tag{3}$$

The basic idea behind the Mix-Norm is to parametrize all sub-intervals of S^1 and to take the root mean square of the average values of c over these sub-intervals. If $c \in L_{S^1}^2$, the following two limits can be verified.

$$\lim_{s \to 0} \phi(c, s) = \left(\int_0^1 c^2(x) \mu(dx) \right)^{\frac{1}{2}} \tag{4}$$

$$\lim_{s \to 1} \phi(c, s) = \left| \int_0^1 c(x) \mu(dx) \right|. \tag{5}$$

Expressions (4) and (5) are respectively the L^2 norm and absolute value of the mean of the scalar function c, which are two fundamental measures associated with any scalar field (Proofs for these equalities are included in Appendix (B)). Therefore, $\phi(c, s)$ for different values of $s \in (0, 1)$ can be seen as a continuous transition of measures associated with the scalar function c ranging from the L^2 norm to the mean. The Mix-Norm is obtained by the square integral of these measures over all possible scales $s \in (0, 1)$.

2.2 The Mix-Norm on an n-dimensional torus

We consider scalar functions $c : T^n \to \mathbb{R}$ where $T^n = [0, 1]^n$ is an n-dimensional torus. Here again, μ is the Lebesgue measure and the same will be assumed throughout the paper. For notational convenience, we make the following definitions, (All vectors are written in bold font (eg: $\mathbf{x}$) and their respective elements are written in usual font with indices as subscripts (eg: $x_1, x_2, ..$)),

For a given $s \in (0, 1)$ and $\mathbf{p} \in T^n$
$B(\mathbf{p}, s) = \{ \mathbf{y} \in T^n : \|\mathbf{y} - \mathbf{p}\|_2 \leq s/2 \}$.
$VolB(s) =$ Volume of the n-dimensional sphere with radius $s/2 = \mu(B(\mathbf{p}, s))$.
$\chi_{B(\mathbf{p}, s)}$ is the characteristic function on the set $B(\mathbf{p}, s)$.
$$\tag{6}$$

Also, in all the discussions, for any two functions $f, g : T^n \to \mathbb{R}$, the inner product is defined to be

$$\langle f, g \rangle = \int_{T^n} f(x) \overline{g(x)} \mu(dx). \tag{7}$$

To define the Mix-Norm let

4

$$d(c, \mathbf{p}, s) = \frac{\int\limits_{x \in B(\mathbf{p},s)} c(\mathbf{x})\mu(d\mathbf{x})}{VolB(s)} = \frac{\langle c, \chi_{B(\mathbf{p},s)} \rangle}{VolB(s)} \tag{8}$$

for all $s \in (0,1)$ and $\mathbf{p} \in T^n$. $d(c, \mathbf{p}, s)$ is the mean value of the function c within the sphere $B(\mathbf{p}, s)$. Now define

$$\phi(c, s) = \left(\int\limits_{T^n} d^2(c, \mathbf{p}, s)\mu(d\mathbf{p}) \right)^{\frac{1}{2}}$$
$$= (\langle d(c, \cdot, s), d(c, \cdot, s) \rangle)^{\frac{1}{2}}. \tag{9}$$

Just as in the case for the circle, $\phi(c, s)$ is the L^2 norm of the averaged function $d(c, ., s)$ for a fixed scale s. Then the Mix-Norm of c is given by

$$\Phi(c) = \left(\int_0^1 \phi^2(c, s)\mu(ds) \right)^{\frac{1}{2}}. \tag{10}$$

An equality similar to (4) holds true for higher dimensions also, but not equality (5).

2.2.1   The Mix-Variance

The degree of mixedness of a scalar field needs to be measured as a distance from a uniform field which is not necessarily zero. For this purpose, we define the quantity referred to as the Mix-Variance of a scalar field and in fact, it is the Mix-Variance which will be of greater relevance in the context of mixing. Let $\bar{c}$ be the mean of the function. i.e., $\bar{c} = \int_{T^n} c(\mathbf{x})\mu(d\mathbf{x})$. Then we refer to the quantity $\Phi^2(c - \bar{c})$ as the Mix-Variance of c. This is in parallel to the situation where we refer to $\|c - \bar{c}\|_2^2$ as the L^2 variance of c. To avoid any ambiguity, we write all the steps in the definition of the Mix-Variance here.

$$d(c - \bar{c}, \mathbf{p}, s) = \frac{\int\limits_{x \in B(\mathbf{p},s)} c(\mathbf{x})\mu(d\mathbf{x})}{VolB(s)} - \bar{c} = d(c, \mathbf{p}, s) - \bar{c} \tag{11}$$

for all $s \in (0,1)$ and $\mathbf{p} \in T^n$. Now define

$$\phi^2(c - \bar{c}, s) = \int\limits_{T^n} d^2(c - \bar{c}, \mathbf{p}, s)\mu(d\mathbf{p})$$
$$= \int\limits_{T^n} (d(c, \mathbf{p}, s) - \bar{c})^2 \, \mu(d\mathbf{p}). \tag{12}$$

$\phi^2(c - \bar{c}, s)$ is the L^2 "distance" between the averaged function $d(c, ., s)$ at scale s and $\bar{c}$. Then the Mix-Variance of c is given by

$$\Phi^2(c - \bar{c}) = \int_0^1 \phi^2(c - \bar{c}, s)\mu(ds). \tag{13}$$

2.2.2 The Weighted Mix-Norm

A more general version of the Mix-Norm can be considered with a weighting function on the physical space and scale space. Consider a function $w : T^n \times (0,1) \to \Re$ such that $w(\mathbf{p}, s) \geq 0$. Then the Mix-Norm with weight w, $\Phi_w(c)$ is obtained by replacing equation (8) by

$$d_w(c, \mathbf{p}, s) = \frac{\underset{x \in B(\mathbf{p}, s)}{\int} c(\mathbf{x})\mu(d\mathbf{x})}{VolB(s)} \cdot \sqrt{w(\mathbf{p}, s)} \tag{14}$$

and then following through the same steps. The effects of the weighting function on various aspects will be pointed out throughout the paper, but mostly we will be dealing with the case where there is no weighting function.

2.3 Mix-Norm as a pseudo-norm

The Mix-Norm is a pseudo-norm on the space of functions, meaning that it satisfies the following properties. For any $c : T^n \to \mathbb{R}$,

1. $\Phi(c) \geq 0$, and $c = 0 \Rightarrow \Phi(c) = 0$.

2. $\Phi(\lambda c) = | \lambda | \Phi(c)$, where λ is a scalar constant.

3. $\Phi(c_1 + c_2) \leq \Phi(c_1) + \Phi(c_2)$.

A pseudo-norm is different from a norm in that a pseudo-norm can be zero for nonzero functions. In particular, the Mix-Norm is zero for a special class of nonzero functions which have a mean of zero on all sets of nonzero measure, but have a nonzero L^2 norm. This is related to the concept of weak convergence and will be discussed in subsection 3.1. The triangle inequality property of the Mix-Norm follows from the fact that each $\phi(c, s)$ is a pseudo-norm and $\Phi(c)$ is a "summation" of these pseudo-norms. The proof for the triangle inequality property is included in Appendix (A). The introduction of a weighting function in the definition of the Mix-Norm as described before does not change its properties as a pseudo-norm.

The Mix-Norm is bounded for all functions $c \in L^2_{T^n}$. This follows from the fact that the averaging operation in (8) is a contraction from $L^2_{T^n}$ to itself. Moreover, $\Phi(c) \leq \|c\|_2$ for all $c \in L^2_{T^n}$. This will be more obvious in the next subsection, when the connection between the Mix-Norm and the Fourier basis is made.

2.4 The Mix-Norm as an inner product

From the previous sections, one can observe that the Mix-Norm can be written as a triple integral on the unit circle and as a $3n$ integral on an n-dimensional torus. This section shows that the Mix-Norm can be written in a much more compact form as an inner product. Let the linear operator $[D(s)] : L_{T^n}^2 \to L_{T^n}^2$ be defined as follows (all operators will be written in square brackets):

$$[D(s)]c(\mathbf{p}) = \frac{\int\limits_{x \in B(\mathbf{p},s)} c(\mathbf{x})\mu(d\mathbf{x})}{VolB(s)} = \frac{\langle c, \chi_{B(\mathbf{p},s)} \rangle}{VolB(s)}. \tag{15}$$

Now $\phi(c, s)$ can be written as:

$$\phi(c, s) = (\langle [D(s)]c, [D(s)]c \rangle)^{\frac{1}{2}} = (\langle c, [D(s)]^*[D(s)]c \rangle)^{\frac{1}{2}} \tag{16}$$

where $[D(s)]^*$ is the adjoint operator of $[D(s)]$. Then the Mix-Norm $\Phi(c)$ is given by

$$\begin{aligned} \Phi^2(c) &= \int_0^1 \phi^2(c, s)\mu(ds) = \int_0^1 \langle c, [D(s)]^*[D(s)]c \rangle \, \mu(ds) \\ &= \left\langle c, \left[\int_0^1 [D(s)]^*[D(s)]\mu(ds) \right] c \right\rangle = \langle c, [M]c \rangle \end{aligned} \tag{17}$$

where

$$[M] := \int_0^1 [D(s)]^*[D(s)]\mu(ds). \tag{18}$$

By the above definition of $[M]$, we mean that its action on a function c is given as follows:

$$[M]c(\mathbf{x}) = \int_0^1 [D(s)]^*[D(s)]c(\mathbf{x})\mu(ds) \tag{19}$$

We refer to $[M]$ as the Mix-Operator. Note that $[M]$ is a symmetric definite operator by construction and that $[M]$ depends only on the domain under consideration. Thus, the Mix-Norm of any function c can be computed using the inner product of c with $[M]c$. The following discussion presents some interesting facts about the Mix-Operator $[M]$. It turns out that the Fourier basis functions are the eigenfunctions of the Mix-Operator $[M]$. This can be inferred from the fact that [M] belongs to the class of spatially invariant operators for which the Fourier basis functions are always eigenfunctions. Even though the Mix-Norm was defined for real valued functions in the previous discussions, its extension to complex valued functions is straightforward and will be assumed in the following discussions.

Theorem 2.1. *The eigenfunctions of the Mix-Operator $[M]$ defined on the n-dimensional torus T^n are $\{c_{\mathbf{k}}(x) = e^{i2\pi(\mathbf{k}.\mathbf{x})}\}$ for $\mathbf{k} \in \mathbb{Z}^n$ and the corresponding eigenvalues are $\Lambda_{\mathbf{k}} = \int_0^1 K_n^2(s, \mathbf{k})\mu(ds)$, where*

$$K_n(s, \mathbf{k}) = \frac{2^{\frac{(n-2)}{2}} n\Gamma(\frac{n}{2}) J_{\frac{n}{2}}\left(s\pi\|\mathbf{k}\|\|\right)}{(s\pi\|\mathbf{k}\|)^{\frac{n}{2}}} \tag{20}$$

and where

$$
\begin{aligned}
&\Gamma \text{ is the gamma function} \\
&J_{\frac{n}{2}} \text{ is a Bessel function of the first kind} \\
&\|\mathbf{k}\| = \sqrt{k_1^2 + k_2^2 + ... + k_n^2}.
\end{aligned}
\tag{21}
$$

Proof. Motivated by (17), we can define a degenerate inner-product T as $T(c_1, c_2) = \langle c_1, [M]c_2 \rangle$. $T : L^2_{T^n} \times L^2_{T^n} \to \mathbb{R}$ is bilinear and symmetric but degenerate, meaning that there exist nonzero functions $c \in L^2_{T^n}$ such that $T(c, c) = 0$. If we show that $T(c_\mathbf{j}, c_\mathbf{k}) = 0$ when $\mathbf{j} \neq \mathbf{k}$, it follows that $\{c_\mathbf{k}\}$ are indeed the eigenfunctions of $[M]$. Also, the eigenvalues are given by $\Lambda_\mathbf{k} = T(c_\mathbf{k}, c_\mathbf{k})$.

The Fourier basis functions satisfy the following mean-value property over spherical surfaces.

$$\frac{\displaystyle\int_{\mathbf{x} \in S(\mathbf{p}, r)} c_\mathbf{k}(\mathbf{x})\mu(d\mathbf{x})}{Area(S(\mathbf{p}, r))} = \frac{\Gamma(\frac{n}{2}) J_{\frac{(n-2)}{2}}\left(r2\pi\|\mathbf{k}\|\right)}{(r\pi\|\mathbf{k}\|)^{\frac{(n-2)}{2}}} c_\mathbf{k}(\mathbf{p}) \tag{22}$$

where $S(\mathbf{p}, r) = \{x : \|x - p\|_2 = r\}$ and $Area(S(\mathbf{p}, r)) = $ Surface area of the n-sphere. This implies a mean value theorem for the interior of the sphere which is as follows:

$$
\begin{aligned}
[D(s)]c_\mathbf{k}(\mathbf{p}) &= \left(\frac{2^{\frac{(n-2)}{2}} n\Gamma(\frac{n}{2}) J_{\frac{n}{2}}\left(s\pi\|\mathbf{k}\|\right)}{(s\pi\|\mathbf{k}\|)^{\frac{n}{2}}}\right) e^{i2\pi\mathbf{k}\cdot\mathbf{p}} \\
&= K_n(s, \mathbf{k})e^{i2\pi\mathbf{k}\cdot\mathbf{p}}
\end{aligned}
\tag{23}
$$

Details of this derivation are given in Appendix (C). Now, computing the inner product $T(c_\mathbf{j}, c_\mathbf{k})$, we get

$$\langle c_{\mathbf{j}}, [M]c_{\mathbf{k}} \rangle = \left\langle c_{\mathbf{j}}, \int\limits_0^1 [D(s)]^*[D(s)]c_{\mathbf{k}}\mu(ds) \right\rangle$$

$$= \int\limits_0^1 \langle [D(s)]c_{\mathbf{j}}, [D(s)]c_{\mathbf{k}} \rangle \, \mu(ds)$$

$$= \int\limits_0^1 \left[\int\limits_{T^n} K_n(s,\mathbf{j})e^{i2\pi(\mathbf{j}\cdot\mathbf{p})} K_n(s,\mathbf{k})e^{-i2\pi(\mathbf{k}\cdot\mathbf{p})}\mu(d\mathbf{p}) \right] \mu(ds) \qquad (24)$$

$$= \int\limits_0^1 \left[K_n(s,\mathbf{j})K_n(s,\mathbf{k}) \int\limits_{T^n} e^{i2\pi(\mathbf{j}-\mathbf{k})\cdot\mathbf{p}}\mu(d\mathbf{p}) \right] \mu(ds)$$

$$= \begin{cases} 0 & \text{if } \mathbf{j} \neq \mathbf{k}, \\ \int_0^1 K_n^2(s,\mathbf{k})\mu(ds) & \text{if } \mathbf{j} = \mathbf{k}. \end{cases}$$

The result in (24) says that when the Mix-Operator $[M]$ acts on any one of the Fourier modes, the resulting function has no components in any of the other Fourier modes. Therefore it has to be that the Fourier basis functions are eigenfunctions of the Mix-Operator $[M]$.

□

Corollary 2.1. *For a function $c \in L_{T^n}^2$, which has a Fourier representation $c(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} a_{\mathbf{k}}e^{i2\pi(\mathbf{k}\cdot\mathbf{x})}$, the Mix-Norm is given by*

$$\Phi(c) = \left(\sum_{\mathbf{k}\in\mathbb{Z}^n} \Lambda_{\mathbf{k}}|a_{\mathbf{k}}|^2 \right)^{\frac{1}{2}}. \qquad (25)$$

The above corollary follows directly from (24). For the special case when $\mathbf{k} = 0$, $\Lambda_{\mathbf{k}} = 1$. Also, for the special case when $n = 1$, the expression for $\Lambda_{\mathbf{k}}$ reduces to

$$\Lambda_k = \int_0^1 \frac{sin^2(k\pi s)}{(k\pi s)^2}\mu(ds). \qquad (26)$$

The key aspect of the eigenvalues $\Lambda_{\mathbf{k}}$ given in Theorem 2.1 is that they decay monotonically with respect to the argument $\|\mathbf{k}\|$. Figure 1 shows the variation of $\Lambda_{\mathbf{k}}$ with respect to $\|\mathbf{k}\|$ for $n = 1$ and $n = 2$. Moreover, there exist bounded real constants $\mu_1, \mu_2 > 0$ such that

$$\frac{\mu_1}{(1 + (2\pi\|\mathbf{k}\|)^2)^{\frac{1}{2}}} \leq \Lambda_{\mathbf{k}} \leq \frac{\mu_2}{(1 + (2\pi\|\mathbf{k}\|)^2)^{\frac{1}{2}}}. \qquad (27)$$

Details of the above inequality are given in Appendix (D). Therefore the norm

$\|c\|_{H^{-\frac{1}{2}}(T^n)}$ defined as

$$\|c\|_{H^{-\frac{1}{2}}(T^n)} = \left(\sum_{\mathbf{k} \in \mathbb{Z}^n} \frac{1}{(1 + (2\pi\|\mathbf{k}\|)^2)^{\frac{1}{2}}} |a_{\mathbf{k}}|^2 \right)^{\frac{1}{2}}, \tag{28}$$

which is a Sobolev norm of negative index $s = -1/2$, is equivalent to the Mix-Norm. $H^{-\frac{1}{2}}(T^n)$ is the Sobolev space of negative index $s = -1/2$, defined as

$$H^{-\frac{1}{2}}(T^n) = \{c : T^n \to \mathbb{R} : \|c\|_{H^{-\frac{1}{2}}(T^n)} < \infty\}. \tag{29}$$

Therefore, for all $c \in H^{-\frac{1}{2}}(T^n)$,

$$\sqrt{\mu_1}\|c\|_{H^{-\frac{1}{2}}(T^n)} \leq \Phi(c) \leq \sqrt{\mu_2}\|c\|_{H^{-\frac{1}{2}}(T^n)}. \tag{30}$$

n = 1   n = 2
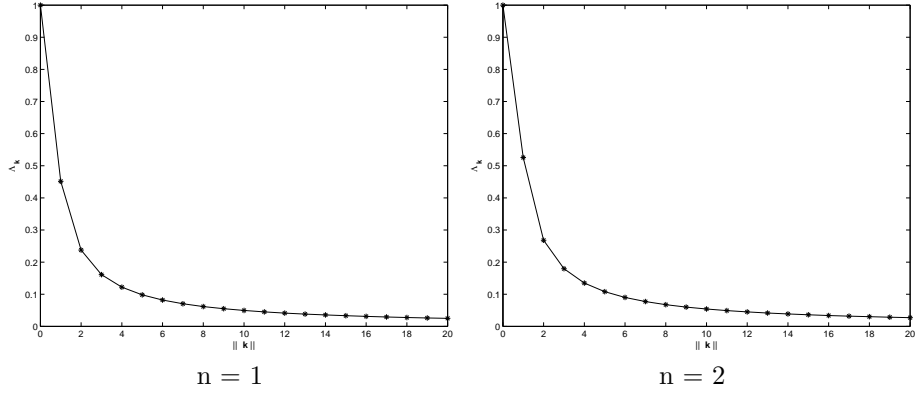
Figure 1: Decay of Eigenvalues of Mix-Operator with respect to magnitude of wave-number vector

The intuition that should be gathered from (25) is that essentially the eigenvalues $\Lambda_{\mathbf{k}}$ act as a weighting on the energy contained in various Fourier modes. The larger the wave number, the smaller the weighting. Therefore, if the energy of a scalar density field is concentrated in small scale(high wave-number) Fourier modes, its Mix-Norm will be very small. In contrast, the L^2 norm is insensitive to the particular Fourier modes within which energy is concentrated due to Parseval's Theorem. The formulation in (25) not only facilitates analysis, but also makes the numerical computation of the Mix-Norm for a given field very efficient via FFT software.

2.4.1   Effects of a weighting function on the Mix-Operator

If we have a weighting function $w : T^n \times (0,1) \to \Re$ with $w(\mathbf{p}, s) \geq 0$, to define the weighted Mix-Norm $\Phi_w(c)$, as described before, we can define the linear operator $[D(s)]_w$ as follows:

$$[D(s)]_w c(\mathbf{p}) = \frac{\underset{x \in B(\mathbf{p},s)}{\int} c(\mathbf{x})\mu(d\mathbf{x})}{VolB(s)} \cdot \sqrt{w(\mathbf{p},s)}. \tag{31}$$

Then the Mix-Operator $[M]_w$ is defined as

$$[M]_w = \int_0^1 [D(s)]_w^* [D(s)]_w \mu(ds). \tag{32}$$

If the weighting function is just a function of the scale variable (i.e., w($\mathbf{p}$, s) = w(s)), then the Fourier basis functions are still the eigenfunctions of the Mix-Operator $[M]_w$. This is because the spatial invariance property of the Mix-Operator without the weighting function is retained. By doing the same calulation as in (24), we get the correponding eigenvalues to be $\Lambda_{\mathbf{k}} = \int_0^1 w(s) K_n^2(s,\mathbf{k})\mu(ds)$. When the weighting function depends on the space variable $\mathbf{p}$, $[M]_w$ is no more spatially invariant and therefore the Fourier basis functions are not necessarily eigenfunctions any more. This can also be inferred from the fact that the calculation in (24) does not follow through anymore.

3   Applications of the Mix-Norm

3.1   The Mix-Norm and weak convergence

To rigorously justify the use of the Mix-Norm as a measure for mixing, we need to use the concept of weak convergence. For a more detailed discussion on the relevance of weak convergence to mixing, refer to [10].

Definition 3.1. *A sequence of functions* $\{c_m\}$, $c_m \in L^2_{T^n}$ *is weakly convergent to* $c \in L^2_{T^n}$ *if*

$$\lim_{m\to\infty} \langle c_m, g \rangle = \langle c, g \rangle \ \text{ for all } g \in L^2_{T^n}. \tag{33}$$

Theorem 3.1. *A sequence of functions* $\{c_m\}$, $c_m \in L^2_{T^n}$ *which is bounded in the* L^2 *norm is weakly convergent to* $c \in L^2_{T^n}$ *if and only if*

$$\lim_{m\to\infty} \Phi(c_m - c) = 0. \tag{34}$$

Proof. First, assume that $\{c_m\}$ converges weakly to c. We need to show that $\lim_{m\to\infty} \Phi(c_m - c) = 0$. By assumption, $\lim_{m\to\infty}\langle c_m, g \rangle = \langle c, g \rangle$ for any $g \in L^2_{T^n}$. In particular, $\lim_{m\to\infty}\langle c_m, \chi_{B(\mathbf{p},s)} \rangle = \langle c, \chi_{B(\mathbf{p},s)} \rangle$ for all $s \in (0,1)$ and $\mathbf{p} \in T^n$. Therefore

$$\begin{aligned}
\lim_{m\to\infty} d(c_m - c, \mathbf{p}, s) &= \lim_{m\to\infty} \frac{\langle c_m - c, \chi_{B(\mathbf{p},s)} \rangle}{VolB(s)} \\
&= \lim_{m\to\infty} \frac{\langle c_m, \chi_{B(\mathbf{p},s)} \rangle - \langle c, \chi_{B(\mathbf{p},s)} \rangle}{VolB(s)} = 0
\end{aligned} \tag{35}$$

11

for all $s \in (0, 1)$ and $\mathbf{p} \in T^n$. Therefore

$$
\begin{aligned}
\lim_{m\to\infty} \phi^2(c_m - c, s) &= \lim_{m\to\infty} \int_{T^n} [d(c_m - c, \mathbf{p}, s)]^2 \, \mu(d\mathbf{p}) \\
&= \int_{T^n} \lim_{m\to\infty} [d(c_m - c, \mathbf{p}, s)]^2 \, \mu(d\mathbf{p}) = 0
\end{aligned}
\tag{36}
$$

for all $s \in (0, 1)$. Now

$$
\begin{aligned}
\lim_{m\to\infty} \Phi^2(c_m - c, s) &= \lim_{m\to\infty} \int_0^1 \phi^2(c_m - c, s) \mu(ds) \\
&= \int_0^1 \lim_{m\to\infty} \phi^2(c_m - c, s) \mu(ds) = 0.
\end{aligned}
\tag{37}
$$

The passage of the limit under the integral sign in both of the above equations is possible due to the boundedness and the pointwise convergence of the respective integrands. Its an application of Lebesgue's dominated convergence theorem.

Now, assume that $\lim_{m\to\infty} \Phi(c_m - c) = 0$. This implies that

$$
\lim_{m\to\infty} \int_0^1 \int_{T^n} d^2(c_m - c, \mathbf{p}, s) \mu(d\mathbf{p})\mu(ds) = 0.
\tag{38}
$$

It follows that for almost every $(\mathbf{p}, s) \in T^n \times (0, 1)$, $\lim_{m\to\infty} d(c_m - c, \mathbf{p}, s) = 0$. i.e.,

$$
\lim_{m\to\infty} \langle c_m, \chi_{B(\mathbf{p},s)} \rangle = \langle c, \chi_{B(\mathbf{p},s)} \rangle \text{ for almost every } (\mathbf{p}, s) \in T^n \times (0, 1).
\tag{39}
$$

Therefore, the set of functions

$$
\mathbf{K} = \{ \chi_{B(\mathbf{p},s)} : \lim_{m\to\infty} d(c_m - c, \mathbf{p}, s) = 0 \}
\tag{40}
$$

is linearly dense in $L^2_{T^n}$. Since $\{c_m\}$ is bounded in the L^2 norm, it follows that $\lim_{m\to\infty} \langle c_m, g \rangle = \langle c, g \rangle$ for any $g \in L^2_{T^n}$. Thus $\{c_m\}$ converges weakly to c.

$\square$

For example, the sequence of functions $\{c_m = sin(m\pi x)\}$ on S^1 converges weakly to zero. Therefore $\lim_{m\to\infty} \Phi(c_m) = 0$ and $\lim_{m\to\infty} [M]c_m = 0$. This is also clear from the fact that $\lim_{m\to\infty} \Lambda_m = 0$.

3.2 Mixing by discrete dynamical systems

In this section, we explain how the Mix-Norm is able to capture the mixing effectiveness of discrete dynamical systems. Of course, all the discussion here

extends easily to continuous-time systems. Roughly speaking, a dynamical system can be considered to be mixing if every portion of the phase space gets spread uniformly throughout the phase space under the action of the dynamical system. If a scalar density field is being transported by a volume-preserving dynamical system, it can be said to be mixing if the mean of the evolving scalar field in all subsets of the phase space becomes closer and closer to the mean of the scalar field over the entire phase space. These were the ideas behind the formulation of the Mix-Norm. Here we connect all of these concepts. In addition, we consider particular discrete dynamical systems on the 2-torus and study their mixing properties using the Mix-Norm. First, we summarize some definitions concerning discrete dynamical systems and mixing. For a more detailed exposition, refer to [10], [2] or [14].

Definition 3.2. *Let $(X, \mathbf{A}, \mu)$ be a measure space. If $S : X \to X$ is a non-singular transformation, the unique operator $[P] : L_X^2 \to L_X^2$ defined by (41) is called the* **Frobenius-Perron operator** *corresponding to S.*

$$\int_A [P]c(x)\mu(dx) = \int_{S^{-1}(A)} c(x)\mu(dx), \textit{for } A \in \mathbf{A} \textit{ and } c \in L_X^2. \qquad (41)$$

The Frobenius-Perron operator $[P]$ is a linear operator that expresses how a scalar density field on the domain evolves with time when advected by a mapping S on the domain. For invertible measure-preserving transformations the Frobenius-Perron operator reduces to (42):

$$[P]c(x) = c(S^{-1}(x)). \qquad (42)$$

Definition 3.3. *Let $(X, \mathbf{A}, \mu)$ be a normalized measure space and $S : X \to X$ be a measure-preserving transformation. S is called* **mixing** *if*

$$\lim_{i \to \infty} \mu(A \cap S^{-i}(B)) = \mu(A)\mu(B) \textit{ for all } A, B \in \mathbf{A} \qquad (43)$$

Theorem 3.2. *Let $S : T^n \to T^n$ be a measure-preserving transformation. Then the following statements are equivalent*

1. *S is mixing;*

2. *The sequence of functions $\{[P]^i c\}$ is weakly convergent to $\bar{c}$ for all $c \in L_{T^n}^2$;*

3. *$\lim_{i \to \infty} \Phi([P]^i c - \bar{c}) = 0$ for all $c \in L_{T^n}^2$,*

where $\bar{c} = \langle c, \chi_{T^n} \rangle$ is the mean of the function c over the whole phase space.

Proof. Here of course, we consider T^n to be equipped with the Lebesgue measure and that S preserves the Lebesgue measure. For the proof on the equivalence of statements (1) and (2), refer to [10]. The equivalence of statements (2) and (3) follows from Theorem 3.1. $\square$

13

Statements (1) and (2) help us to classify transformations as mixing or non-mixing, but don't give an idea of the "mixedness" of a certain scalar density field after a finite number of iterations of the transformation. Statement (3) helps to address this problem and serves to design optimal mixing protocols tailored to a specific initial scalar density field. Statement (2) is also equivalently captured in terms of the so-called correlation function. For two functions f and g (referred to as observables), the correlation function is defined as

$$\varphi_i(f,g) = \left| \int_{T^n} g(\mathbf{x}).f(S^{-i}(\mathbf{x}))\mu(d\mathbf{x}) - \int_{T^n} g(\mathbf{x})\mu(d\mathbf{x}) \int_{T^n} f(\mathbf{x})\mu(d\mathbf{x}) \right| \quad (44)$$

If the map S is mixing, then the correlation function for any two observables must decay to zero. The rate of decay of correlations is considered to be a inidicator of the rate of mixing and their behaviour for various classes of mappings is an active research topic in dynamical systems theory. In this paper, we propose that one needs to study only the rate of decay of the Mix-Variance for a particular initial scalar field to study how efficiently the map S is "mixing it up". For instance if $c = \chi_A$ and if $\Phi([P]^i c - \bar{c}) = O(t(i))$, then it implies that for almost every sphere (i.e, for almost every $(\mathbf{p}, s) \in T^n \times (0,1)$),

$$\langle [P]^i c, \chi_{B(\mathbf{p},s)} \rangle - \langle \bar{c}, \chi_{B(\mathbf{p},s)} \rangle = O(t(i)), \quad (45)$$

which is equivalent to

$$\mu(S^i(A) \cap B(\mathbf{p},s)) - \mu(A)\mu(B(\mathbf{p},s)) = O(t(i)). \quad (46)$$

Here, $t(i)$ could be an exponentially or algebraically decaying function of time and accordingly we would say that S "mixes up" c exponentially or algebraically. We will discuss this using some specific examples in the next subsections. Before we proceed, we discuss the weaknesses of the approaches which were indicated in the Introduction.

3.2.1 Comparison with Traditional Approaches

Lyapunov exponents: Let $S : T^2 \to T^2$ be a differentiable volume-preserving transformation on the 2-dimensional torus. Then the Lyapunov exponent for a given point $\mathbf{x} \in T^2$ under S is given by:

$$\lambda(\mathbf{x}) = \lim_{n \to \infty} \frac{1}{n} log \| D_\mathbf{x} S^n(\mathbf{x}) \| \quad (47)$$

where $D_\mathbf{x} S^n(\mathbf{x})$ is the Jacobian of the nth iterate of the point $\mathbf{x}$ under S. Then the so-called Kolmogorov-Sinai entropy corresponding to the transformation S is given by

$$h(S) = \int_{T^2} \lambda(\mathbf{x})\mu(d\mathbf{x}). \quad (48)$$

The above formula is known as Pesin's formula ([13], [8]). The Kolmogorov-Sinai entropy is a measure of the disorder created by the transformation S. The Lyapunov exponent at the point $\mathbf{x}$ is the maximum average logarithmic expansion rate of an *infinitesimal* circle centred at $\mathbf{x}$. From the point of view of control of mixing, if there was a parameter associated with the transformation S, then one would choose it so as to maximize the Lyapunov exponents.

To illustrate the drawback of using Lyapunov exponents when we have a specific initial scalar field, consider the following simple example. Consider the discrete version of a linear shear flow given by

$$
\left[\begin{array}{c} x_1' \\ x_2' \end{array} \right] = \left[\begin{array}{c} x_1 + x_2 \\ x_2 \end{array} \right] (\text{mod } 1). \tag{49}
$$

A straightforward calculation would show that this map has zero Kolmogorov-

$$
\text{Linear Shear Flow} \qquad c_1(x_1, x_2) = f(x_1) \qquad c_2(x_1, x_2) = g(x_2)
$$
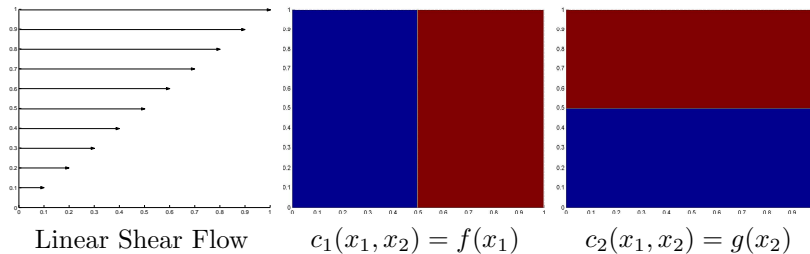
Figure 2: The linear shear flow effects the two different density fields pictured above very differently. Density fields of the form $c_1(x_1, x_2) = f(x_1)$ get "mixed-up" at an algebraic rate, whereas density fields of the form $c_2(x_1, x_2) = g(x_2)$ are invariants of the map.

Sinai entropy (i.e. zero Lyapunov exponents). But, the effects of this map on the two initial scalar fields shown in Figure 2 are very different. It can be shown that for initial distributions of the form $c_1(x_1, x_2) = f(x_1)$, $\Phi([P]^i c_1 - \bar{c}_1)$ will decay to zero when advected by this map, although at an algebraic rate, whereas distrbutions of the form $c_2(x_1, x_2) = g(x_2)$ remain invariant under this map. This is an example of a transformation which has zero entropy, but can "mix-up" certain initial scalar fields. This will be discussed in more detail in a following section. Similarly, there are transformations with non-zero entropy, but may not completely "mix-up" certain initial scalar fields. The bottomline is that the Mix-Norm helps to address the mixing efficacy of a flow with respect to an initial scalar field or fluid configuration whereas it is not directly possible to do the same with information about Lyapunov exponents.

Decay of L^2 variance: The variance of a density field c given by

$$\|c - \bar{c}\|_2^2 = \int_{T^2} (c - \bar{c})^2 \mu(d\mathbf{x}) \tag{50}$$

where $\bar{c} = \langle c, \chi_{T^2} \rangle$, measures how far the density field is from being spatially uniform or homogeneous. In the absence of diffusion, the variance of a density field being advected by an invertible volume-preserving transformation will remain constant. This can be verified as follows:

$$
\begin{aligned}
\int_{T^2} ([P]c(\mathbf{x}) - \bar{c})^2 \mu(d\mathbf{x}) &= \int_{T^2} (c(S^{-1}(\mathbf{x})) - \bar{c})^2 \mu(d\mathbf{x}) \\
&= \int_{S^{-1}(T^2)} (c(\mathbf{y}) - \bar{c})^2 |det(S(\mathbf{y}))| \mu(d\mathbf{y}) \tag{51} \\
&= \int_{T^2} (c(\mathbf{x}) - \bar{c})^2 \mu(d\mathbf{x}).
\end{aligned}
$$

But, by Theorem 3.2, the Mix-Variance of the advected field will decay to zero if the transformation is mixing and the rate at which the Mix-Variance decays gives information about the mixing rate of the underlying transformation. In any realistic physical system, there is going to be some diffusion and therefore one might be tempted to use the variance anyway to measure mixing. But, in a typical mixing process, there is an initial phase during which the variance remains almost constant and after which the variance decays exponentially. It is during this initial phase that stretching and folding of fluid elements occur which eventually facilitate diffusion to act efficiently causing the exponential decay of the variance. Therefore, to optimize this initial phase of stretching and folding or "stirring" one needs some other notion of a mixing measure which the Mix-Variance serves as. In other words, the variance fails to capture the "stirring" effects of a mixing protocol while the Mix-Variance succeeds to do so.

3.2.2 Arnold's Cat Map

Arnold's Cat Map is a volume-preserving diffeomorphism on the 2-torus T^2 given by

$$
\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} (\text{mod } 1) = \mathbf{M}.\mathbf{x}(\text{mod } 1). \tag{52}
$$

It is well known that the Cap Map is mixing in the sense described in Definition 3.3. To prove statement (2) in Theorem 3.2, one needs to prove it only for every element in a complete basis of functions (eg: Fourier basis functions). The Frobenius-Perron operator $[P]$ corresponding to this map when acting on one of the Fourier modes is given explicitly as follows:

$$[P](e^{i2\pi(k_1 x_1 + k_2 x_2)}) = e^{i2\pi((k_1 - k_2)x_1 + (-k_1 + 2k_2)x_2)}. \tag{53}$$

Thus the resulting function is another Fourier mode with wave numbers $(k_1 - k_2, -k_1 + 2k_2)$. We can think of a mapping between the wave numbers as follows:

$$\begin{bmatrix} k_1' \\ k_2' \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \mathbf{M}^{-1}.\mathbf{k}. \tag{54}$$

The matrix $\mathbf{M}^{-1}$ has eigenvalues $\sigma_1 = (3 - \sqrt{5})/2 < 1$, $\sigma_2 = (3 + \sqrt{5})/2 > 1$, and corresponding eigendirections $(1, (\sqrt{5} - 1)/2), (-1, (\sqrt{5} + 1)/2)$. Since the slope of the stable eigendirection is irrational and the wave numbers (k_1, k_2) take only integer values, the magnitudes of the wave numbers (k_1, k_2) increase exponentially under the mapping in (54), for any initial set of wave numbers. This gives an immediate proof that the Cat Map is actually mixing. Moreover, when starting with a Fourier mode with wave numbers (q_1, q_2), the magnitude of the wave numbers of the evolving distibution after i iterations are such that $(|k_1| = O(\sigma_2^i), |k_2| = O(\sigma_2^i))$. In addition, since we know that $\Phi^2(e^{i2\pi(k_1 x_1 + k_2 x_2)}) \approx 1/(1 + 4\pi^2(k_1^2 + k_2^2))^{\frac{1}{2}}$, the Mix-Variance at the i-th iteration, $\Phi_i^2 = O(\sigma_2^{-i})$. Thus, the Cat Map exponentially mixes each one of the Fourier modes and therefore any bounded linear combination of the Fourier modes.

3.2.3 Mixing properties of the Standard Map

We consider the Standard Map $S : T^2 \to T^2$ in the following form:

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + \epsilon sin(2\pi x_1) \\ x_2 + \epsilon sin(2\pi x_1) \end{bmatrix} (\text{mod } 1) = \mathbf{S}(\mathbf{x}). \tag{55}$$

The above map is a volume-preserving diffeomorphism on the 2-torus and was first introduced by Chirikov [18]. Its behaviour changes with the value of ϵ. For values of ϵ close to one, it has been observed to have chaotic properties. Here we study its mixing properties. For the case when $\epsilon = 0$, we get a mapping describing the evolution of the wave numbers just as for the Cat Map, given by

$$\begin{bmatrix} k_1' \\ k_2' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \mathbf{H}.\mathbf{k}. \tag{56}$$

The matrix H is degenerate, meaning that it has only one eigenvalue ($\sigma = 1$) and one eigendirection ($\mathbf{k} = (0, 1)$). Therefore, Fourier modes with wave number $\mathbf{k} = (0, q_2)$ will be fixed points for the Frobenius-Perron operator of the Standard Map when $\epsilon = 0$. Also, when starting with a Fourier mode with wave numbers (q_1, q_2) where $q_1 \neq 0$, the wave numbers of the evolving distibution after i iterations are given exactly as $(k_1 = q_1, k_2 = q_2 - i.q_1)$. Coupled with the fact that $\Phi^2(e^{i2\pi(k_1 x_1 + k_2 x_2)}) \approx 1/(1 + 4\pi^2(k_1^2 + k_2^2))^{\frac{1}{2}}$, we can approximate the Mix-Variance at the i-th iteration as $\Phi_i^2 \approx 1/(1 + 4\pi^2(q_1^2 + (q_2 - i.q_1)^2))^{\frac{1}{2}} = O(i^{-1})$. Therefore, when $q_1 \neq 0$, the Integrable Standard Map algebraically "mixes up" the density. The map can be said to be mixing on the orthogonal complement of the space of functions which depend only on the variable x_2 and the map is not even ergodic.

When the perturbation $\epsilon \neq 0$, energy from one Fourier mode is transferred to more than one Fourier mode, thereby making any kind of analysis harder. We resort to numerics to study the mixing properties of the Standard Map for various ϵ. The computational domain is the square $[0,1] \times [0,1]$ whose sides are being identified (upper with lower, left with right).We discretize the computational domain and let the grid points be (x_1^i, x_2^j) where $x_1^i = i.dx, x_2^j = j.dy$ for $i = 0, 1, .., (N_x - 1)$ and $j = 0, 1, .., (N_y - 1)$ and where $dx = 1/N_x$ and $dy = 1/N_y$. The inverse of the Standard Map is given explicitly as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1' - x_2' \\ x_2' - \epsilon sin(2\pi(x_1' - x_2')) \end{bmatrix} (\text{mod } 1). = \mathbf{S^{-1}}(\mathbf{x}). \tag{57}$$

At the n-th iterate, for each grid point we compute $S^{-n}(x_1^i, x_2^j)$ using (57). Therefore, the discrete version of the Frobenius-Perron operator can be written as

$$\hat{P}^n c(x_1^i, x_2^j) = c_0(S^{-n}(x_1^i, x_2^j)_1, S^{-n}(x_1^i, x_2^j)_2) \tag{58}$$

where c_0 is the initial density field. At each iteration, the Fourier coefficients of the density field were computed using FFT software [7], and the Mix-Variance was computed using (25). Figure 3 shows the decay of the Mix-Variance with time for various values of ϵ when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_2)$. When $\epsilon = 0.0$, there is no mixing and as the value of ϵ is gradually increased, the mixing gets better. Figure 4, shows a plot of $\phi(c - \bar{c}, s)$ versus s corresponding to the evolved density field after 10 and 100 iterations and for various values of ϵ. Figure 5 shows the contour plot of the evolved density field after 5 iterations for $\epsilon = 0.4$ and 0.8.

Figure 6 shows the decay of the Mix-Variance with time when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_1)$. When $\epsilon = 0.0$, there is an algebraic decay of the Mix-Variance as can be observed from the log-log plot in Figure 6. In this case, when ϵ is increased, there is actually a decrease in the mixing efficiency as can be seen for $\epsilon = 0.2, 0.4$ and 0.6. Figures 7 and 8 give a closer look at this behaviour. Figure 7 shows a plot of $\phi(c - \bar{c}, s)$ versus s corresponding to the evolved density field after 10 and 100 iterations. The curve corresponding to $\epsilon = 0.0$ indicates good "mixedness" which is due to the algebraic decay of the Mix-Variance as described earlier. From $\epsilon = 0.0$ to $\epsilon = 0.2$, there is a sharp jump from a well "mixed" field to a badly "mixed" field. For values of ϵ beyond 0.2, there is again a gradual improvement in the mixing efficiency.

Figure 8 shows the corresponding plots for values of ϵ ranging from 0.0 to 0.22 and then 0.24 to 0.4. One can observe a gradual decrease in the "mixedness" of the density field as ϵ increases from 0.0 to 0.22. Then, as ϵ increases from 0.24 to 0.4, there is a gradual increase in the "mixedness" of the density field. One could attribute the initial decline in the mixing performance to the formation of regular islands which prevent the continuous shearing of "fluid" which happenes for the case when $\epsilon = 0.0$. This is validated by the contour plots in Figure 9 which show the evolved density field after 10 iterations for $\epsilon = 0.0, 0.1$ and 0.2.
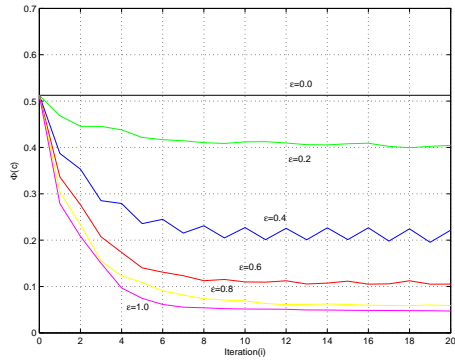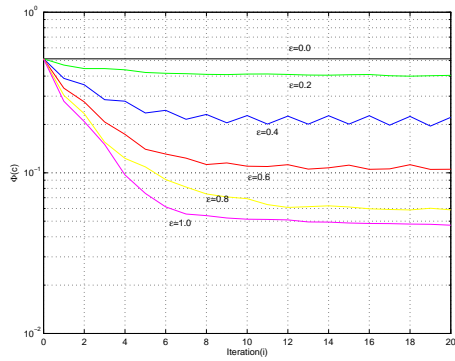
18
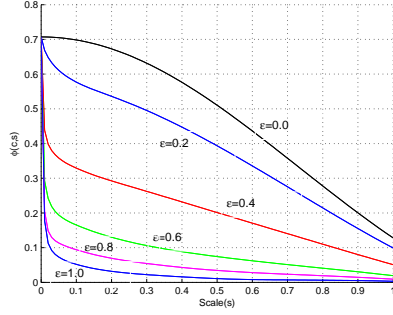
Mix-Variance versus Iteration
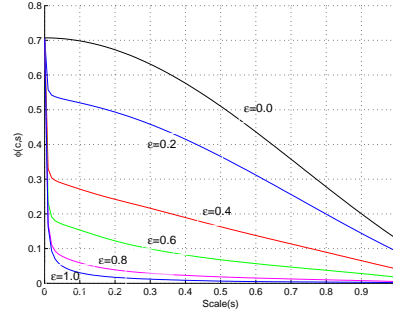

Linear-log plot of Mix-Variance versus Iteration
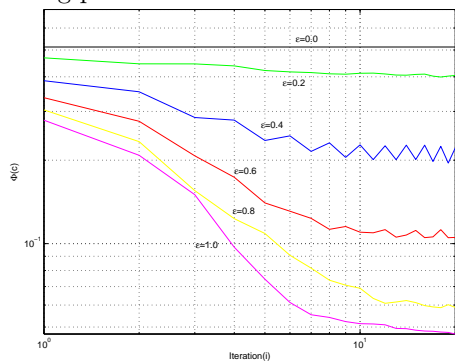
Log-log plot of Mix-Variance versus Iteration

Figure 3: Mixing properties of the Standard Map. Decay of Mix-Variance with respect to iteration for different values of ϵ when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_2)$.

19

After 10 iterations

After 100 iterations

Figure 4: Plot of $\phi(c, s)$ versus s for the evolved density fields after 10 and 100 iterations of the Standard Map and for various values of ϵ, when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_2)$.

$\epsilon = 0.4$

$\epsilon = 0.8$
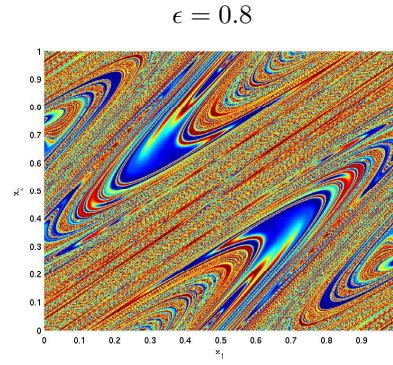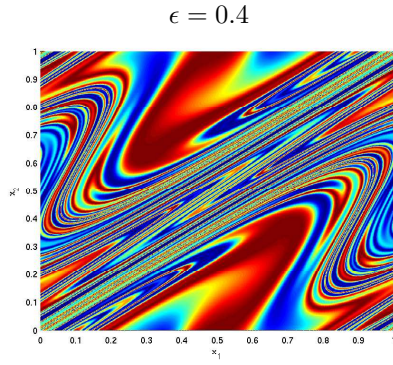
Figure 5: Contour plot of density field after 5 iterations of the Standard Map for $\epsilon = 0.4$ and $\epsilon = 0.8$, when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_2)$.

20

A precise mathematical explanation for this behaviour is beyond the scope of this paper.

3.3 Mixing by diffusion

Consider the diffusion equation on T^n with a diffusivity rate $D > 0$.

$$\frac{\partial c}{\partial t} = D \sum_{i=1}^{n} \frac{\partial^2 c}{\partial x_i^2} = D\Delta c. \tag{59}$$

First, note that the Mix-Operator $[M]$ and the Laplacian operator Δ have the same eigenfunctions (i.e., $e^{i2\pi(\mathbf{k}.\mathbf{x})}$). A Fourier mode with wavenumber $\mathbf{k}$ decays exponentially at a rate of $4\pi^2\|\mathbf{k}\|^2$ under pure diffusion. Therefore, if one were to arrange the Fourier modes in descending order of their amplitude when acted upon by pure diffusion over a unit time interval, one would get the same order as when arranging them in descending order of their eigenvalues when acted upon by the Mix-Operator.

Now, assume an initial density field $c(\mathbf{x}, 0)$ on T^n, whose average is given by

$$\bar{c} = \int_{\mathbf{x} \in T^n} c(\mathbf{x}, 0)d\mathbf{x}. \tag{60}$$

Let the distribution at time t be $c(\mathbf{x}, t) \geq 0$. In work by Ashwin et al. [3] the L^2 and L^∞ norms are used to measure mixing. Also, they define the time to 95% mixing, t_{95} to be the smallest $t > 0$ such that

$$\|c(\mathbf{x}, t) - \bar{c}\|_\alpha \leq 0.05 \tag{61}$$

where α refers to the norm used. Note that t_{95} is a function of the initial distribution, the norm chosen and the mixing protocol. Assume that $c(\mathbf{x}, 0)$ has a Fourier expansion

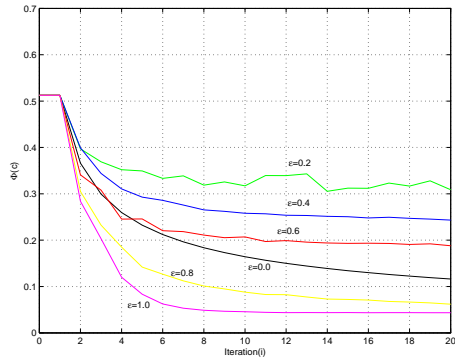$$c(\mathbf{x}, 0) = \sum_{\mathbf{k} \in \mathbb{Z}^n} a_{\mathbf{k}} e^{2\pi i(\mathbf{k}.\mathbf{x})}, \tag{62}$$

where $a_{\mathbf{k}} = \int c(\mathbf{x}, 0)e^{-2\pi i(\mathbf{k}.\mathbf{x})}d\mathbf{x}$. The distribution at time t is given by

$$c(\mathbf{x}, t) = \sum_{\mathbf{k} \in \mathbb{Z}^n} a_{\mathbf{k}} e^{2\pi i(\mathbf{k}.\mathbf{x})} e^{-4\pi^2\|\mathbf{k}\|^2 Dt}. \tag{63}$$
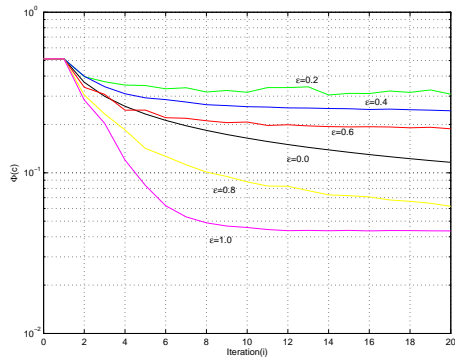
An upper bound for t_{95} corresponding to the L^2 norm can be computed as in [3] to be

$$t_{95} \leq \frac{\log(20\|c(\mathbf{x}, 0) - \bar{c}\|_2)}{4\pi^2 D}. \tag{64}$$
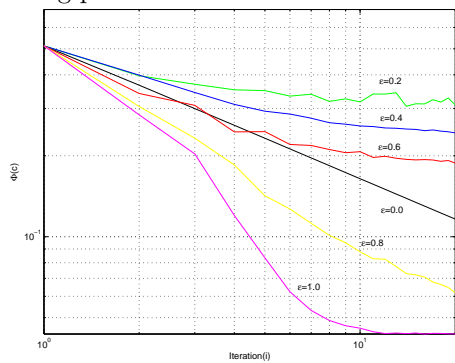
Doing the same computation corresponding to the Mix-Norm we obtain
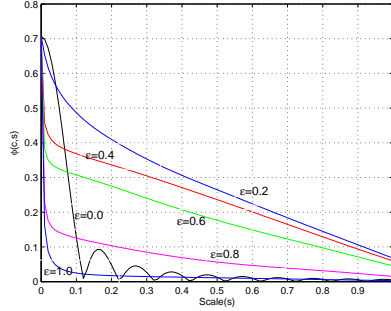
21

Mix-Variance versus Iteration

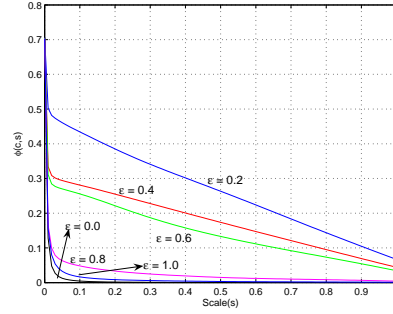Linear-log plot of Mix-Variance versus Iteration

Log-log plot of Mix-Variance versus Iteration

Figure 6: Mixing properties of the Standard Map. Decay of Mix-Variance with respect to iteration for different values of ϵ when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_1)$.
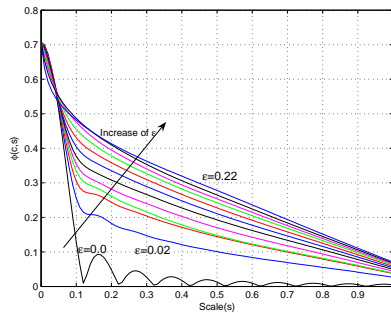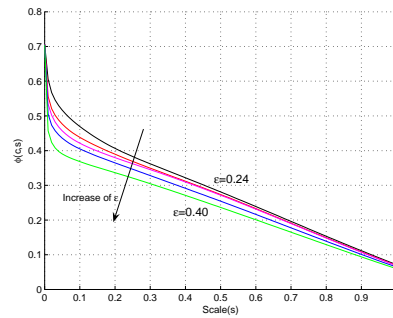
After 10 iterations

After 100 iterations

Figure 7: Plot of $\phi(c, s)$ versus s for the density fields after 10 and 100 iterations of the Standard Map and for various values of ϵ, when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_1)$.

For $\epsilon = 0.0, 0.02, ..., 0.22$

For $\epsilon = 0.24, 0.28, ..., 0.40$

Figure 8: Plot of $\phi(c, s)$ versus s for the density fields after 10 iterations of the Standard Map and for various values of ϵ, when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_1)$.

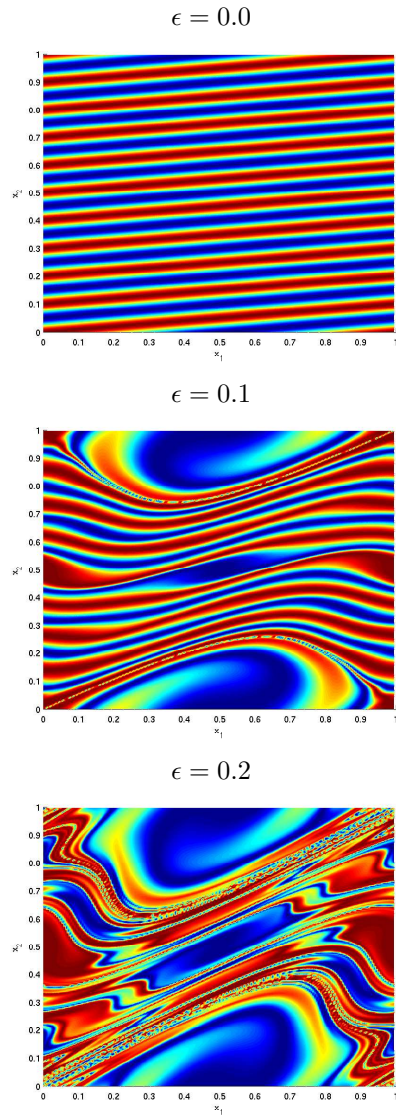$$\epsilon = 0.0$$

$$\epsilon = 0.1$$

$$\epsilon = 0.2$$

Figure 9: Contour plot of density field after 10 iterations of the Standard Map for $\epsilon = 0.0, 0.1$ and 0.2, when starting with an initial density field of $c_0(x_1, x_2) = cos(2\pi x_1)$.

$$\Phi^2(c(\mathbf{x}, t) - \bar{c}) = \sum_{\mathbf{k} \neq 0} \Lambda_{\mathbf{k}} |a_{\mathbf{k}}|^2 e^{-8\pi^2 \|\mathbf{k}\|^2 Dt}$$

$$\leq e^{-8\pi^2 Dt} \sum_{\mathbf{k} \neq 0} \Lambda_{\mathbf{k}} |a_{\mathbf{k}}|^2 = e^{-8\pi^2 Dt} \Phi^2(c(\mathbf{x}, 0) - \bar{c}). \tag{65}$$

Thus an upper bound for t_{95} corresponding to the Mix-Norm can be found as

$$t_{95} \leq \frac{\log(20\Phi(c(\mathbf{x}, 0) - \bar{c}))}{4\pi^2 D}. \tag{66}$$

The difference in the estimates obtained for t_{95} corresponding to the Mix-Norm compared to that in (64) can be demonstrated as follows. Consider initial distributions of the form $c(x, 0) = c_m = 1 + sin(2m\pi x)$ on the circle. $\|c_m - 1\|_2$ is a constant for all m whereas $\Phi(c_m - 1)$ is small for large m. According to (66), diffusion will achieve almost perfect mixing in much less time for large m, whereas the estimate in (64) does not depend on the magnitude of m. In general, if the initial distribution has strong components in high wave number Fourier modes, then the diffusion process will achieve mixing very quickly, which is reflected in the estimate obtained using the Mix-Norm. However, note that (66) does not directly say anything about the decay of the L^2 variance. A more useful estimate corresponding to the L^2 norm may probably be obtained using a combination of the Mix-Variance and L^2 variance of the initial distribution. If the ratio between $\Phi(c_m - 1)$ and $\|c_m - 1\|_2$ is very small, it necessarily implies that there exist strong components of high wave number Fourier modes in the initial distribution. The smaller this ratio, the faster the diffusion process will be. Unfortunately, it is not straightforward to derive any estimates for t_{95} corresponding to the L^2 norm just from this ratio.

4   Conclusions

A multiscale measure for quantifying mixing which is particularly useful for studying mixing of chaotic flows and maps has been presented. Its properties as a pseudo-norm induced by an inner product and its connection with the Fourier spectrum have been discussed. Its validity as a measure for mixing has been rigorously justified through the notion of weak convergence. Although we derived the mixing rates for a couple of classical examples and performed some numerical experiments for the Standard Map, the real usefulness of this measure is to the problem of designing mixing protocols tailored to mix a specific initial density field. This problem can be framed as an optimal control problem and is being currently pursued. Such a notion of a measure for mixing can also be very useful for optimizing the design of mixing devices. This new measure has been successfully used for optimization of mixing in microdevices [11] and has provided valuable insights.

In an interesting paper [15], Protas et al. have studied issues of optimal flow control in spaces other than L^2. Since we have shown equivalence of the

Mix-Norm to the $H^{-\frac{1}{2}}$ norm, it is natural to pursue optimal control problems in this context. The equivalence of these two norms raises the following question - why not just use the $H^{-\frac{1}{2}}$ norm instead of the Mix-Norm? We think that it is perfectly fine to use the $H^{-\frac{1}{2}}$ norm or for that matter any Sobolev space norm of negative index instead of the Mix-Norm. But, we wish to convey in this paper that by defining a norm based on the intuitive notion of averaging function values over spherical sets and by weighing the average values over spherical sets of all radii equally, we get a norm equivalent to the $H^{-\frac{1}{2}}$ norm. Such an intuition did not exist before and even Sobolev space norms of negative index have not been usefully applied for control and optimization of mixing. Note that by weighing average values over spheres of different radii differently, we would not necessarily have equivalence with the $H^{-\frac{1}{2}}$ norm. Also note that, by defining the Mix-Norm based on averaging over sets of different shapes (say rectangular sets), we would not necessarily have equivalence with any Sobolev space norm. Future work includes the generalization of the Mix-Norm to arbitrary domains and also its application it as a means of comparing the statistical properties of dynamical systems.

A Proof for triangular inequality property of the Mix-Norm

Proof. Let $c_1, c_2 : T^n \to \Re$. We need to prove that $\Phi(c_1 + c_2) \leq \Phi(c_1) + \Phi(c_2)$. Clearly,

$$d(c_1 + c_2, \mathbf{p}, s) = d(c_1, \mathbf{p}, s) + d(c_2, \mathbf{p}, s). \tag{67}$$

Now,

$$
\begin{aligned}
\phi^2(c_1 + c_1, s) &= \int_{T^n} d^2(c_1 + c_2, \mathbf{p}, s)\mu(dp) \\
&= \int_{T^n} \left[d(c_1, \mathbf{p}, s) + d(c_2, \mathbf{p}, s)\right]^2 \mu(dp) \\
&= \int_{T^n} \left[d^2(c_1, \mathbf{p}, s) + d^2(c_2, \mathbf{p}, s) + 2d(c_1, \mathbf{p}, s)d(c_2, \mathbf{p}, s)\right] \mu(dp) \\
&= \phi^2(c_1, s) + \phi^2(c_2, s) + 2 \int_{T^n} \left[d(c_1, \mathbf{p}, s)d(c_2, \mathbf{p}, s)\right] \mu(dp).
\end{aligned}
\tag{68}
$$

Applying the Cauchy-Schwarz inequality to the above equation,

$$\phi^2(c_1 + c_2, s) \le \phi^2(c_1, s) + \phi^2(c_2, s) + 2\sqrt{\left(\int_{T^n} d^2(c_1, \mathbf{p}, s)\mu(dp)\right)\left(\int_{T^n} d^2(c_2, \mathbf{p}, s)\mu(dp)\right)}$$
$$= \phi^2(c_1, s) + \phi^2(c_2, s) + 2\phi(c_1, s)\phi(c_2, s). \tag{69}$$

It follows that

$$\Phi^2(c_1 + c_2) \le \Phi^2(c_1) + \Phi^2(c_2) + \int_0^1 2\phi(c_1, s)\phi(c_2, s)\mu(ds). \tag{70}$$

Applying the Cauchy-Schwarz inequality once again, we get

$$\Phi^2(c_1 + c_2) \le \Phi^2(c_1) + \Phi^2(c_2) + 2\sqrt{\left(\int_0^1 \phi^2(c_1, s)\mu(ds)\right)\left(\int_0^1 \phi^2(c_2, s)\mu(ds)\right)}$$
$$= \Phi^2(c_1) + \Phi^2(c_2) + 2\Phi(c_1)\Phi(c_2)$$
$$= (\Phi(c_1) + \Phi(c_2))^2. \tag{71}$$

$\square$

B Proof for Equations (4) and (5)

Consider the delta sequence

$$\delta_s(x) = \begin{cases} 0 \text{ for } x < -\frac{s}{2} \\ \frac{1}{s} \text{ for } -\frac{s}{2} \le x \le \frac{s}{2} \\ 0 \text{ for } x > \frac{s}{2} \end{cases} \tag{72}$$

Then for $c \in L^2_{S^1}$, we have

$$d(c, p, s) = \int_0^1 c(x)\delta_s(p - x)\mu(dx) = c * \delta_s(p) \tag{73}$$

i.e., the function $d(c, ., s)$ is the convolution of c and δ_s. For continuous functions c, since

$$\lim_{s \to 0} d(c, p, s) = \lim_{s \to 0} c * \delta_s(p) = c * \delta(p) = c(p), \tag{74}$$

it is straightforward to see that

$$\lim_{s \to 0} \phi(c, s) = \lim_{s \to 0} \left(\int_0^1 d^2(c, p, s)\mu(dp)\right)^{\frac{1}{2}} = \left(\int_0^1 \lim_{s \to 0} d^2(c, p, s)\mu(dp)\right)^{\frac{1}{2}}$$
$$= \left(\int_0^1 c^2(p)\mu(dp)\right)^{\frac{1}{2}}. \tag{75}$$

27

The proof for any $c \in L_{S^1}^2$ is as follows. Let $\hat{c}$, $\hat{d}(c,.,s)$ and $\hat{\delta}_s$ be the Fourier transforms of c, $d(c,.,s)$ and δ_s respectively. i.e.,

$$\hat{c}(k) = \int_0^1 c(x)e^{-i2\pi kx}\mu(dx). \tag{76}$$

Then, we have

$$\hat{d}(c,k,s) = \hat{c}(k).\hat{\delta}_s(k) \tag{77}$$

Now, by Parseval's identity

$$\int_0^1 d^2(c,p,s)\mu(dp) = \sum_{k\in\mathbb{Z}} \hat{d}^2(c,k,s) = \sum_{k\in\mathbb{Z}} \hat{c}^2(k).\hat{\delta}_s^2(k) \tag{78}$$

Since,

$$\lim_{s\to 0} \hat{\delta}_s(k) = \int_0^1 \delta(x)e^{-i2\pi kx}\mu(dx) = 1, \tag{79}$$

we have

$$\lim_{s\to 0} \int_0^1 d^2(c,p,s)\mu(dp) = \sum_{k\in\mathbb{Z}} \hat{c}^2(k) \lim_{s\to 0}\hat{\delta}_s^2(k) = \sum_{k\in\mathbb{Z}} \hat{c}^2(k) = \int_0^1 c^2(x)\mu(dx). \tag{80}$$

The proof for Equation 5 follows easily from noting that

$$\lim_{s\to 1} d(c,p,s) = \frac{\int_{p-1/2}^{p+1/2} c(x)\mu(dx)}{1} = \int_0^1 c(x)\mu(dx) = \bar{c}. \tag{81}$$

Therefore

$$\lim_{s\to 1} \phi(c,s) = \lim_{s\to 1} \left(\int_0^1 d^2(c,p,s)\mu(dp)\right)^{\frac{1}{2}} = \left(\int_0^1 \lim_{s\to 1} d^2(c,p,s)\mu(dp)\right)^{\frac{1}{2}}$$
$$= \left(\int_0^1 (\bar{c})^2\mu(dp)\right)^{\frac{1}{2}} = |\bar{c}| = \left|\int_0^1 c(x)\mu(dx)\right|. \tag{82}$$

C Derivation for the eigenvalues of the Mix-Operator

The eigenfunctions of the Laplacian operator or solutions of the Helmholtz equation

$$\Delta f + \lambda f = 0, \tag{83}$$

satisfy the following mean value property over spherical surfaces [5].

$$\frac{\int\limits_{\mathbf{x}\in S(\mathbf{p},r)} f(\mathbf{x})d\mathbf{x}}{Area(S(\mathbf{p},r))} = \frac{\Gamma(\frac{n}{2})J_{\frac{(n-2)}{2}}(r\sqrt{\lambda})}{(r\sqrt{\lambda}/2)^{\frac{(n-2)}{2}}} f(\mathbf{p}), \tag{84}$$

where $S(\mathbf{p}, r) = \{x : \|x - p\|_2 = r\}$ and $Area(S(\mathbf{p}, r)) = $ Surface area of the n-sphere. The Fourier basis functions $\{c_{\mathbf{k}}(\mathbf{x}) = e^{i2\pi\mathbf{k}\cdot\mathbf{x}}\}$ being solutions of the Helmholtz equation on the torus domain for $\lambda = 4\pi^2\|\mathbf{k}\|^2$, we have

$$\frac{\int\limits_{\mathbf{x}\in S(\mathbf{p},r)} c_{\mathbf{k}}(\mathbf{x})d\mathbf{x}}{Area(S(\mathbf{p},r))} = \frac{\Gamma(\frac{n}{2})J_{\frac{(n-2)}{2}}(r2\pi\|\mathbf{k}\|)}{(r\pi\|\mathbf{k}\|)^{\frac{(n-2)}{2}}}c_{\mathbf{k}}(\mathbf{p}) \tag{85}$$

This implies a mean value theorem for the interior of the sphere thereby giving

$$[D(s)]c_{\mathbf{k}}(\mathbf{p}) = \frac{1}{VolB(s)}\int\limits_0^{s/2}\frac{\Gamma(\frac{n}{2})J_{\frac{(n-2)}{2}}(r2\pi\|\mathbf{k}\|)e^{i2\pi(\mathbf{k}\cdot\mathbf{p})}Area(S(\mathbf{p},r))}{(r\pi\|\mathbf{k}\|)^{\frac{(n-2)}{2}}}dr \tag{86}$$

where $VolB(s) = $ Volume of the n-dimensional sphere with radius $s/2$. Substituting the formula for $Area(S(\mathbf{p}, r)) = \frac{2\pi^{\frac{n}{2}}r^{(n-1)}}{\Gamma(\frac{n}{2})}$, we get

$$[D(s)]c_{\mathbf{k}}(\mathbf{p}) = \frac{e^{i2\pi(\mathbf{k}\cdot\mathbf{p})}}{VolB(s)}\int\limits_0^{s/2}\frac{J_{\frac{(n-2)}{2}}(r2\pi\|\mathbf{k}\|)2\pi^{\frac{n}{2}}r^{(n-1)}}{(r\pi\|\mathbf{k}\|)^{\frac{(n-2)}{2}}}dr \tag{87}$$

Multiplying and dividing the integrand by $(r2\pi\|\mathbf{k}\|)^{\frac{n}{2}}$, we get

$$[D(s)]c_{\mathbf{k}}(\mathbf{p} = \frac{e^{i2\pi(\mathbf{k}\cdot\mathbf{p})}}{VolB(s)(2\pi)^{\frac{(n-2)}{2}}(\|\mathbf{k}\|)^{(n-1)}}\int\limits_0^{s/2}J_{\frac{(n-2)}{2}}(r2\pi\|\mathbf{k}\|)(r2\pi\|\mathbf{k}\|)^{\frac{n}{2}}dr$$

$$= \frac{e^{i2\pi(\mathbf{k}\cdot\mathbf{p})}}{VolB(s)(2\pi)^{\frac{n}{2}}(\|\mathbf{k}\|)^n}\int\limits_0^{s\pi\|\mathbf{k}\|}J_{\frac{(n-2)}{2}}(y)y^{\frac{n}{2}}dy. \tag{88}$$

Using the derivative identity $\frac{d}{dx}[x^m J_m(x)] = x^m J_{(m-1)}(x)$ and using the formula for $VolB(s) = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}\frac{(s/2)^n}{n}$ we get

$$[D(s)]c_{\mathbf{k}}(\mathbf{p}) = \frac{e^{i2\pi(\mathbf{k}\cdot\mathbf{p})}}{VolB(s)(2\pi)^{\frac{n}{2}}(\|\mathbf{k}\|)^n}\left[(s\pi\|\mathbf{k}\|)^{\frac{n}{2}}J_{\frac{n}{2}}(s\pi\|\mathbf{k}\|)\right]$$

$$= \left(\frac{2^{\frac{(n-2)}{2}}n\Gamma(\frac{n}{2})J_{\frac{n}{2}}(s\pi\|\mathbf{k}\|)}{(s\pi\|\mathbf{k}\|)^{\frac{n}{2}}}\right)e^{i2\pi\mathbf{k}\cdot\mathbf{p}} \tag{89}$$

$$= K_n(s, \mathbf{k})e^{i2\pi\mathbf{k}\cdot\mathbf{p}}.$$

D   Proof for the Inequality in (27)

For $\mathbf{k} \neq 0$, the following holds true:

$$\Lambda_{\mathbf{k}} = \int_0^1 \frac{2^{n-2}n^2\Gamma^2(\frac{n}{2})J_{\frac{n}{2}}^2(s\pi\|\mathbf{k}\|)}{(s\pi\|\mathbf{k}\|)^n}ds = \frac{2^{n-2}n^2\Gamma^2(\frac{n}{2})}{\|\mathbf{k}\|\pi}\int_0^{\|\mathbf{k}\|\pi}\left(\frac{J_{\frac{n}{2}}(y)}{y^{n/2}}\right)^2 dy$$

$$\leq \frac{2^{n-1}n^2\Gamma^2(\frac{n}{2})}{2\pi\|\mathbf{k}\|}\int_0^\infty\left(\frac{J_{\frac{n}{2}}(y)}{y^{n/2}}\right)^2 dy = \frac{\gamma_n}{2\pi\|\mathbf{k}\|},$$

$$(90)$$

where

$$\gamma_n = 2^{n-1}n^2\Gamma^2\left(\frac{n}{2}\right)\int_0^\infty\left(\frac{J_{\frac{n}{2}}(y)}{y^{n/2}}\right)^2 dy. \tag{91}$$

Now, for any $a \geq 2$, (in fact for any $a \geq \sqrt{\left(1 + \frac{1}{4\pi^2}\right)}$), the following holds true:

$$\Lambda_{\mathbf{k}} \leq \frac{\gamma_n}{2\pi\|\mathbf{k}\|} = \frac{a\gamma_n}{((a2\pi\|\mathbf{k}\|)^2)^{\frac{1}{2}}} \leq \frac{a\gamma_n}{(1 + (2\pi\|\mathbf{k}\|)^2)^{\frac{1}{2}}}. \tag{92}$$

Therefore, by choosing a sufficiently large, we can make $\mu_2 = a\gamma_n \geq \Lambda_0 = 1$. Therefore,

$$\Lambda_{\mathbf{k}} \leq \frac{\mu_2}{(1 + (2\pi\|\mathbf{k}\|)^2)^{\frac{1}{2}}} \text{ for all } \mathbf{k}. \tag{93}$$

Again, for $\mathbf{k} \neq 0$ and $0 < \epsilon < \pi$, the following holds true

$$\Lambda_{\mathbf{k}} = \frac{2^{n-2}n^2\Gamma^2(\frac{n}{2})}{\|\mathbf{k}\|\pi}\int_0^{\|\mathbf{k}\|\pi}\left(\frac{J_{\frac{n}{2}}(y)}{y^{n/2}}\right)^2 dy \geq \frac{2^{n-1}n^2\Gamma^2(\frac{n}{2})}{2\pi\|\mathbf{k}\|}\int_0^\epsilon\left(\frac{J_{\frac{n}{2}}(y)}{y^{n/2}}\right)^2 dy$$

$$= \frac{\beta_n(\epsilon)}{2\pi\|\mathbf{k}\|} \geq \frac{\beta_n(\epsilon)}{(1 + (2\pi\|\mathbf{k}\|)^2)^{\frac{1}{2}}}.$$

$$(94)$$

where

$$\beta_n(\epsilon) = 2^{n-1}n^2\Gamma^2\left(\frac{n}{2}\right)\int_0^\epsilon\left(\frac{J_{\frac{n}{2}}(y)}{y^{n/2}}\right)^2 dy. \tag{95}$$

Clearly, we can choose ϵ small enough so that $\mu_1 = \beta_n(\epsilon) \leq \Lambda_0 = 1$. Therefore,

$$\Lambda_{\mathbf{k}} \geq \frac{\mu_1}{(1 + (2\pi\|\mathbf{k}\|)^2)^{\frac{1}{2}}} \text{ for all } \mathbf{k}. \tag{96}$$

References

[1] H. Aref. Stirring by chaotic advection. *Journal of Fluid Mechanics*, 143:1–21, 1984.

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

[3] Peter Ashwin, Matthew Nicol, and Norman Kirkby. Acceleration of one-dimensional mixing by discontinuous mappings. *Physica A*, 310:347–363, 2002.

[4] D. Betz. *Physical mechanisms of mixing*. PhD thesis, Univ. California, Santa Barbara, 2001.

[5] R. Courant and D. Hilbert. *Methods of Mathematical Physics*. Interscience, New York, 1962.

[6] Domenico D'Alessandro, Mohammed Dahleh, and Igor Mezić. Control of mixing: A maximum entropy approach. *IEEE Transactions on Automatic Control*, 44:1852–1864, 1999.

[7] M. Frigo and S. G. Johnson. Fftw: An adaptive software architecture for the fft. In *Proc. ICASSP*, volume 3, pages 1381–1384, 1998.

[8] A. Katok and J. M. Strelcin. *Invariant manifolds, entropy and billiards–Smooth maps with singularities*, volume 1222 of *Lecture Notes in Mathematics*. Springer-Verlag, New York, 1980.

[9] T. S. Krasnopolskaya, V. V. Meleshko, G. W. M. Peters, and H. E. H. Meijer. Mixing in stokes flow in an annular wedge cavity. *Eur. J. Mech. B.Fluids*, 18:793–822, 1999.

[10] A. Lasota and M. Mackey. *Chaos, Fractals and Noise*, volume 97 of *Applied Mathematical Sciences*. Springer, Berlin, 1994.

[11] G. Mathew, I. Mezić, R. Serban, and L. Petzold. Characterization of mixing in an active micromixer. Preprint. Dept. Mechanical and Environmental Eng., Univ. California, Santa Barbara, 2003.

[12] J. M. Ottino. *The Kinematics of Mixing: Stretching, Chaos and Transport*. Cambridge University Press, Cambridge, England, 1989.

[13] Ya. B. Pesin. Characteristic lyapunov exponents and smooth ergodic theory. *Russian Math. Surveys*, 32(4):55–114, 1977.

[14] K. Petersen. *Ergodic Theory*, volume 2. Cambridge Studies in Advanced Mathematics, 1983.

[15] B. Protas, T. Bewley, and G. Hagen. A computational framework for the regularization of adjoint analysis in multiscale pde systems. *Journal of Computational Physics*, 195(1):49–89, 2004.

[16] D. Rothstein, E. Henry, and J. P. Gollub. Persistent patterns in transient chaotic fluid mixing. *Nature*, 401:770–772, Oct 1999.

[17] Jean-Luc Thiffeault and Stephen Childress. Chaotic mixing in a torus map. *Chaos*, 13(2):502–507, 2003.

[18] Chirikov B. V. A universal instability of many-dimensional oscillator systems. *Physics Reports*, 52:264–379, 1979.

[19] Marion Volpert, Carl Meinhart, Igor Mezić, and Mohammed Dahleh. Modeling and numerical analysis of mixing in an actively controlled micromixer. In *HEFAT 2002, 1st International Conference on Heat Transfer, Fluid Mechanics and Thermodynamics*, Kruger Park, South Africa, April 2002.

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