




Identifying invariant ergodic subsets and barriers to mixing by cutting and shuffling: Study in a birotated hemisphere

Thomas F. Lynn 


Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, Illinois 60208, USA

Julio M. Ottino 

*Department of Chemical and Biological Engineering, Department of Mechanical Engineering,
Northwestern University, Evanston, Illinois 60208, USA
and Northwestern Institute on Complex Systems (NICO), Northwestern University, Evanston, Illinois 60208, USA*

Paul B. Umbanhowar 

Department of Mechanical Engineering, Northwestern University, Evanston, Illinois 60208, USA

Richard M. Lueptow *

*Department of Mechanical Engineering, Department of Chemical and Biological Engineering,
Northwestern University, Evanston, Illinois 60208, USA
and Northwestern Institute on Complex Systems (NICO), Northwestern University, Evanston, Illinois 60208, USA*



(Received 17 July 2019; revised manuscript received 25 October 2019; published 10 January 2020)

Mixing by cutting and shuffling can be mathematically described by the dynamics of piecewise isometries (PWI), higher dimensional analogs of one-dimensional interval exchange transformations. In a two-dimensional domain under a PWI, the exceptional set, \bar{E} , which is created by the accumulation of cutting lines (the union of all iterates of cutting lines and all points that pass arbitrarily close to a cutting line), defines where mixing is possible but not guaranteed. There is structure within \bar{E} that directly influences the mixing potential of the PWI. Here we provide computational and analytical formalisms for examining this structure by way of measuring the density and connectivity of ε -fattened cutting lines that form an approximation of \bar{E} . For the example of a PWI on a hemispherical shell studied here, this approach reveals the subtle mixing behaviors and barriers to mixing formed by invariant ergodic subsets (confined orbits) within the fractal structure of the exceptional set. Some PWIs on the shell have provably nonergodic exceptional sets, which prevent mixing, while others have potentially ergodic exceptional sets where mixing is possible since ergodic exceptional sets have uniform cutting line density. For these latter exceptional sets, we show the connectivity of orbits in the PWI map through direct examination of orbit position and shape and through a two-dimensional return plot to explain the necessity of orbit connectivity for mixing.

DOI: [10.1103/PhysRevE.101.012204](https://doi.org/10.1103/PhysRevE.101.012204)

I. INTRODUCTION

The mathematical foundation of cut-and-shuffle mixing is the piecewise isometry (PWI), which cuts a domain into pieces that are rearranged to reform the original domain [1–6]. Mixing by cutting and shuffling arises in several natural systems such as granular materials [7–12], valved fluid flows [13,14], and imbricate thrust faults in geology [15–17]. In these systems and others, cutting and shuffling is not the only mixing mechanism (chaotic advection and diffusion can also be present), but it is presently the least understood. Here we present techniques to investigate the mixing of a particular class of hemispherical PWIs related to the mixing of granular particles in a spherical tumbler [8–10,18]. Our goal is not to study the wide range of PWIs, which can occur in one, two, and three dimensions (or higher) [6], but instead to develop computational and analytical approaches

to identify invariant ergodic subsets that result in barriers to mixing. And, more importantly, the results we will describe are largely independent of the choice of (almost everywhere invertible) PWI.

PWIs act isometrically (i.e., as a distance-preserving transformation) on each of a finite number of pieces, or “atoms,” P_i of a domain. For this paper, PWI dynamics are examined near the boundaries of these pieces where the map is discontinuous. Following previous work [10,11,19,20], the PWI here is defined as a multivalued map acting on the power set (the set of subsets, which allows multivalued boundaries) such that the different, but overlapping, boundaries of the pieces of the domain are allowed to have different actions under the map. This is to say that, under the map, atoms that touch each other retain a copy of their mutual boundary under the map. This is an identical treatment to our previous papers [10,20], although it was not emphasized in the arguments presented there. Atom boundaries are essential elements of some of the arguments presented in this paper, and it is a point that

*r-lueptow@northwestern.edu

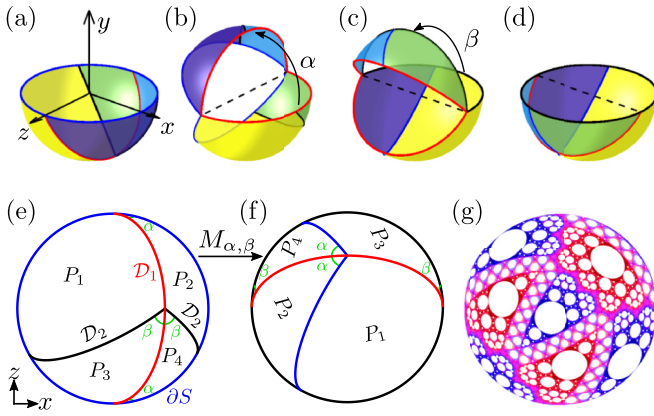


FIG. 1. (a–d) Illustration of the rotations that define the hemispherical PWI studied here. (a) Initial condition of the hemispherical shell (S). (b) Rotation about the z axis by α (arrow shows direction only). (c) Rotation about the x axis by β . (d) Reformed S . (e), (f) Equivalent PWI for the ($\alpha = 57^\circ$, $\beta = 57^\circ$) protocol results in four atoms, P_1 , P_2 , P_3 , P_4 , with cutting lines \mathcal{D}_1 and \mathcal{D}_2 . (g) Accumulation of pre- and postimages of ε -fattened ($\varepsilon = 10^{-3}$) cutting lines \mathcal{D} after $N = 20000$ iterations, viewed from the negative y axis, approximating the true exceptional set \bar{E} as $\bar{E}_{\varepsilon,N}$ using color to indicate cutting line density as explained in Sec. III. A Lambert azimuthal equal area projection is used [29].

requires special care. Other treatments define dynamics on cutting lines such that the map is not multivalued [2,3,5,21–25], and single-valued cutting lines are sufficient for some arguments here (the computational methods are ultimately indifferent to the treatment of the boundaries).

The specific system we study as an example is inspired by a physically realizable spherical tumbler half-filled with granular matter and rotated in a periodic sequence about two perpendicular axes [8,26,27], though the approach described here to identify barriers to mixing could be applied to any PWI. In the equivalent hemispherical shell PWI example we consider here [9–11,18,20,22,23,26–28], the lower unit hemispherical shell, $S = \{\mathbf{x} = (x, y, z) \in \mathbb{R}^3 : \|\mathbf{x}\| = 1, y \leq 0\}$, in Fig. 1(a) is rotated by angle $\alpha \in [0, \pi)$ about the z axis and then by angle $\beta \in [0, \pi)$ about the x axis, as shown in Figs. 1(b) and 1(c) to remake S in Fig. 1(d). The equator provides the cutting action as the pieces of the hemisphere that are flipped across the boundary separate from neighboring points. This rotational procedure, i.e., Figs. 1(a)–1(d), motivates the equivalent PWI map $M_{\alpha,\beta}$, referred to as M except where α , β need to be specified, corresponding to the *protocol* (α , β) shown in Figs. 1(e) and 1(f).

In this PWI, the hemispherical domain S is split into at most four atoms, $\{P_1, P_2, P_3, P_4\}$, as shown in Fig. 1(e), which are mutually disjoint but overlap at their shared boundaries, \mathcal{D}_1 and \mathcal{D}_2 . These atoms are rearranged according to the rotations in Figs. 1(a)–1(d) as shown in Fig. 1(f). Repeating the PWI, i.e., cutting and shuffling over and over, reveals the *singular set*, E , which is defined as $E = \bigcup_{n=0}^{\infty} M^n(\mathcal{D}_1 \cup \mathcal{D}_2)$ and is the accumulation of cuts produced by the PWI map M (all multivalued points of the map and the union of forward iterates of cutting lines) [19]. As previously defined, the singular set and its limit points define the closure of E called

the *exceptional set*, \bar{E} , which is E and all points that pass arbitrarily close to E under the PWI map.¹

For generic protocols, it has been conjectured that \bar{E} is a fat fractal [30,31], and there is strong numerical evidence that \bar{E} has nonzero measure for almost all protocols [1,2,10,23,32] (that is, a finite fraction of the domain passes arbitrarily close to the cutting lines under the PWI map) except for those protocols that produce polyagonal tilings [11,20]. Instead of an exact representation of \bar{E} , an N -iterate approximation [20] to the exceptional set using ε -fattened cutting lines, $\bar{E}_{\varepsilon,N}$, is used here; for the PWI in Figs. 1(e) and 1(f), this approximation is shown in Fig. 1(g). An obvious feature of the exceptional set is the open circular regions, or *cells*, which are periodic nonmixing regions that rotate about an internal elliptic periodic point and are never cut by the cutting lines of the PWI [10,11,20,22,23,25,28]. The fractional coverage of the hemispherical shell's area contained in the associated fat fractal is denoted Φ and its limiting value is denoted Φ_∞ [10,11,20,28]. A key point is that while the exceptional set indicates where mixing is possible, mixing is not guaranteed within the exceptional set [20].

The purpose of this paper is to investigate invariant ergodic subsets within \bar{E} that do not intermix. By ergodic set, we mean that every *orbit* (i.e., the set of all pre- and postimages of a single point, e.g., $X = \{M^n(\mathbf{x}) : -\infty < n < \infty\}$ is the orbit of \mathbf{x}) from the set either exists in a lower dimensional set (zero measure, e.g., a finite number of periodic points or curves) or fills out the entire ergodic set (full measure); a set that is only partially filled by an orbit is not ergodic, but the subset that is filled may be ergodic on its own [33]. In this case, minimally invariant sets (orbits under the PWI and their closures) with nonzero measure are also ergodic due to the dynamics of PWIs [31]. In other words, an ergodic set is one in which, under the PWI map, an orbit passes arbitrarily close to every point within the set infinitely often (Poincaré recurrence) and with equal frequency (with respect to an invariant measure) throughout the set. The closure of every orbit that has nonzero measure under the PWI is invariant under the map and defines an ergodic subset of the domain (due to the dynamics of PWIs). If the exceptional set \bar{E} can be split into multiple, separate ergodic subsets, then there are additional barriers to mixing within the exceptional set. We present tools to examine invariant subsets in \bar{E} and definitively determine that \bar{E} is not always ergodic and instead contains many ergodic subsets for some PWIs.

Our goal in doing this is to answer two specific questions about the physical nature of solid mixing by cutting and shuffling. First, do more cuts necessarily generate more mixing? We demonstrate that this is not the case, and, in fact, variation in the density of cuts at different locations throughout the

¹The limit points of a set define all the points for which there is a sequence of points within the set that approach the limit point with arbitrary closeness. In this case, E is the set of all points \mathbf{x} for which the minimum distance between $\{M^n(\mathbf{x}) : -\infty < n < \infty\}$, i.e., the *orbit* of \mathbf{x} , and \mathcal{D} is exactly zero; $\min_n d(M^n(\mathbf{x}), \mathcal{D}) = 0$ for a distance metric d . The limit points of E are all points for which the infimum of this distance is zero, but this distance is never exactly zero; $\inf_n d(M^n(\mathbf{x}), \mathcal{D}) = 0$ and $d(M^n(\mathbf{x}), \mathcal{D}) > 0$.

domain is an indicator of barriers to mixing. Second, do nontrivial cutting actions necessarily produce any mixing at all? We demonstrate, by use of a modified recurrence plot, that there exist invariant subsets of cut regions which indicate barriers to mixing, potentially barring global mixing. Both of these results are likely to be of importance for practical applications of solids mixing.

II. MATHEMATICS OF PIECEWISE ISOMETRIES

Mathematically, a PWI on a metric, measure space S is a map $\Phi : \mathcal{P}(S) \rightarrow \mathcal{P}(S)$, where $\mathcal{P}(S)$ denotes the power set of S (the set of subsets of S). S consists of a finite number of *tiles* or *atoms* $\{P_1, \dots, P_m\}$ with nonempty interiors that form a tiling of S , and Φ consists of isometries on these atoms (metric-preserving maps):

$$\phi_k : P_k \rightarrow S, \quad 1 \leq k \leq m.$$

Although not a requirement for PWI in general, the PWI studied here is almost everywhere invertible such that $\{\phi_1(P_1), \dots, \phi_m(P_m)\}$ is also a tiling of S , i.e., $S = \bigcup_{k=1}^m \phi_k(P_k) = \bigcup_{k=1}^m P_k$, and, as such, both tilings, $\{P_1, \dots, P_m\}$ and $\{\phi_1(P_1), \dots, \phi_m(P_m)\}$ overlap only on their boundaries, ∂P_k and $\partial \phi_k(P_k)$ respectively, which are measure zero. The single PWI map, $\Phi : \mathcal{P}(S) \rightarrow \mathcal{P}(S)$, is the combination of these isometries,

$$\Phi(\mathbf{x}) := \phi_k(\mathbf{x}), \quad \text{when } \mathbf{x} \in P_k, \quad 1 \leq k \leq m,$$

which is multivalued and discontinuous on the overlapping boundaries of P_k . We use \mathcal{D} to denote the collection of these multivalued points (which are discontinuities in the PWI), specifically $\mathcal{D} = \bigcup_{i \neq j} P_i \cap P_j$. The red cutting line \mathcal{D}_1 (due to the first rotation by α) and the black cutting line \mathcal{D}_2 (due to the second rotation by β) together form this multivalued, zero measure set in Figs. 1(e) and 1(f). In many definitions [2,3,5,21–25], PWIs are defined on \mathcal{D} in such a way as to avoid multivalued mappings, but orbits originating from this set, \mathcal{D} , and nearby are used in this paper to understand the mixing regions adjacent to E . Atoms that share a boundary each retain a copy under the map, which will later be described as each retaining a *side* of the discontinuity.

The multivalued map for the hemispherical shell PWI is formally defined as a rotation for each atom independently,

$$M_{\alpha,\beta}(\mathbf{x}) = \{R_i(\mathbf{x}) : \mathbf{x} \in P_i\}, \quad (1)$$

where the four rotation isometries R_i for each P_i are

$$R_1(\mathbf{x}) = R_\beta^x R_\alpha^z \mathbf{x}, \quad (2)$$

$$R_2(\mathbf{x}) = R_\beta^x R_{\alpha+\pi}^z \mathbf{x}, \quad (3)$$

$$R_3(\mathbf{x}) = R_{\beta+\pi}^x R_\alpha^z \mathbf{x}, \quad (4)$$

$$R_4(\mathbf{x}) = R_{\beta+\pi}^x R_{\alpha+\pi}^z \mathbf{x}, \quad (5)$$

and R_θ^a is a rotation about axis a by angle θ [11,20]. Rotations are applied right to left such that rotation about the z axis is first and rotation about the x axis is second. Additional rotations by π are included where the specified atom crosses the (periodic) equator.

This PWI is orientation preserving (free of reflections) and almost everywhere invertible except for the points that map to the equator after either rotation (atom boundaries) which map to two locations. The pre-image of the points that encounter the equator during the rotation procedure in Fig. 1 are labeled \mathcal{D}_1 (first rotation about the z axis) and \mathcal{D}_2 (subsequent rotation about the x axis) in Fig. 1(c). Since $M\mathcal{D}_2 = \partial S$, the same blue color is assigned to ∂S , \mathcal{D}_2 , and their pre- and postimages in later figures to aid visualization. Likewise, the pre- and postimages of \mathcal{D}_1 are red in later figures. Figure 1(g) shows $\tilde{E}_{0.001,20000}$ ($N = 20\,000$ iterations of cutting lines given half-width $\varepsilon = 0.001$), for the $(57^\circ, 57^\circ)$ protocol, an approximation of the exceptional set, \bar{E} , which is the accumulation of all possible pre- and postimages of the cutting lines using this red-blue coloring viewed from the negative y axis and flattened to a disk using a Lambert azimuthal equal-area projection [29]. This projection preserves the relative areas of features and is used in Sec. III B to evenly sample across the shell (an evenly spaced grid is projected onto the hemisphere) [20].

The hemispherical shell, S , can be partitioned into distinct invariant sets with respect to the PWI. $O \subset S$ is the collection of periodic islands, maximally open neighborhoods about central periodic points, which almost always, with the exception of degenerate protocols without periodic islands, has positive Lebesgue measure, i.e., two-dimensional (2D) area. Periodic islands are never cut by the PWI. In contrast, E is the countable union of pre- and postimages of the cutting lines, \mathcal{D} , and, as a collection of great circle arcs, has zero Lebesgue measure. E can contain periodic points, but since M is treated as a multivalued map on \mathcal{D} , it may be that only one of the multiple images of a point in E is actually periodic. \bar{E} , the closure of E , is called the *exceptional set*. While E may contain periodic points, the *remainder* of \bar{E} outside of E , $E' = \bar{E} \setminus E$, necessarily cannot contain any periodic points [32].² Since E has zero measure as a collection of thin arcs, all of the measure associated with a fat fractal \bar{E} is from E' . These three sets, E , E' and O are all invariant under M (and, of course, $\bar{E} = E \cup E'$ is trivially invariant as the union of two invariant sets).³

In general, O is not minimally invariant (containing no smaller invariant subset), as each periodic cell can only map to cells of the same size [11]. For any given protocol, it is not clear whether \bar{E} is minimally invariant (containing no smaller

²Take periodic orbit X (such that every $\mathbf{x} \in X$ is a periodic point) and the distance metric $d(\mathbf{a}, B) = \inf_{\mathbf{b} \in B} d(\mathbf{a}, \mathbf{b}) \geq 0$ which is the minimum distance between a point \mathbf{a} and a set B . If $d(\mathbf{x}, \mathcal{D}) > 0$ for all $\mathbf{x} \in X$, then there exists a $\delta > 0$ such that $d(\mathbf{x}, \mathcal{D}) > \delta$ for all $\mathbf{x} \in X$ and, therefore, there exists a neighborhood around each $\mathbf{x} \in X$ constituting a periodic cell with nonzero measure such that $X \subset O$. If $d(\mathbf{x}, \mathcal{D}) = 0$ for some $\mathbf{x} \in X$, then there is at least one point $\mathbf{y} \in X$ such that $d(\mathbf{y}, \mathcal{D}) = 0$ and $\mathbf{y} \in \mathcal{D}$, which implies $X \subset E$. Therefore, there are no periodic points in E' .

³ E is invariant by definition as it is the smallest set containing \mathcal{D} that is invariant. O is also invariant as the image of any periodic cell is another cell. $E' = S \setminus (O \cup E)$ must then also be invariant. If it were not invariant, there would exist $\mathbf{x} \in E'$ such that $M(\mathbf{x}) \notin E'$, i.e., $M(\mathbf{x}) \in E \cup O$, which violates the invariance of E or O .

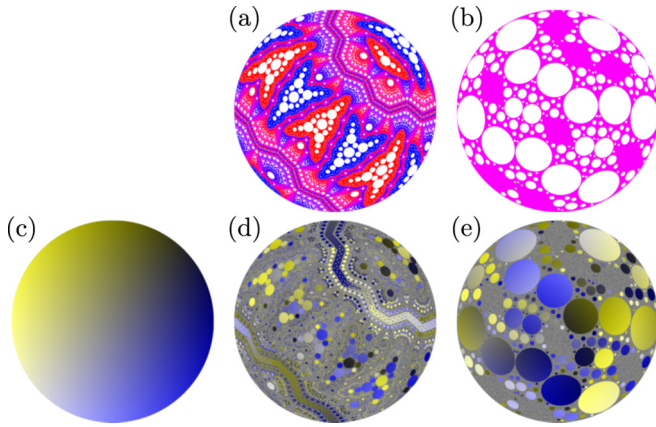


FIG. 2. Approximate exceptional sets, $\tilde{E}_{0.001, 2 \times 10^4}$, for (a) $(45^\circ, 45^\circ)$ and (b) $(57^\circ, 32.75^\circ)$ protocols viewed from below using a Lambert equal area projection (see Sec. III for explanation of how the hue is computed). Corresponding mixing of initial condition in (c) for (d) $(45^\circ, 45^\circ)$ and (e) $(57^\circ, 32.75^\circ)$ after 2×10^4 iterations of $M_{\alpha, \beta}$.

invariant subsets within them), but the existence of distinct, positive measure, minimally invariant subsets in \tilde{E} introduces barriers to the mixing induced by the PWI. Orbits in \tilde{E} can be approximated by nearby trajectories in E , but due to roundoff error, no numerically evaluated orbits will ever enter E .

With this background, we begin by examining an approximation to the natural invariant measure of \tilde{E} and the numerical considerations in representing orbits via color coding in Sec. III. In Sec. IV we examine how images of \mathcal{D} return to \mathcal{D} using a type of return plot and draw conclusions about invariant subsets within \tilde{E} that can create barriers to mixing. Again, our goal is not to attempt to characterize the infinite range of possible PWIs on the hemispherical shell, which has been done previously [10,11,20,28], but instead to develop approaches to understand the mixing characteristics of PWIs in general by considering a few specific hemispherical shell PWIs as examples.

III. COLOR CODING INVARIANT SUBSETS

Previous studies on the hemispherical PWI [11,20] used a correlation [10] between the measure of \tilde{E} (2D area) and the *intensity of segregation* due to Danckwerts [34], a measure of mixing, to compare the degree to which two different protocols mix. Some PWIs have ergodic exceptional sets [35], some PWIs produce weak mixing within their exceptional sets [36], and other PWIs have exceptional sets containing separate invariant subsets that do not mix between each other [37,38]. In general, it cannot be predicted whether a given PWI will have an ergodic exceptional set, and, as a result, the measure of \tilde{E} is not always an accurate indicator of mixing. Compare, for example, the protocols $(45^\circ, 45^\circ)$ and $(57^\circ, 32.75^\circ)$, shown in Figs. 2(a) and 2(b), respectively, both of which have exceptional sets that cover roughly 41% of S , i.e., $\Phi_\infty \approx 0.41$ [10,11,20]. For both protocols, the N -iterate approximation to \tilde{E} , $\tilde{E}_{\varepsilon, N}$, shown in Figs. 2(a) and 2(b) is formed by using $N = 2 \times 10^4$ iterations of fattened cutting lines with width $2\varepsilon = 0.002$ (ε on either side of the line).

Using this approach, ε -fattened cutting lines subject to the PWI completely cover \tilde{E} in a finite number of iterations [20].

It is illustrative to compare the mixing for these two protocols. The continuously varying initial condition shown in Fig. 2(c) is partially mixed by both protocols as shown in Figs. 2(d) and 2(e). As expected, nonmixing cells do not mix with the rest of the domain, but, surprisingly, the $(45^\circ, 45^\circ)$ protocol in Fig. 2(d) generates noncircular regions inside of $\tilde{E}_{\varepsilon, N}$ that do not mix with the rest of the domain. Previous measurements of fractional coverage [10,11,20] count all of \tilde{E} as mixing, yet there appear to be regions within \tilde{E} that are isolated from one another. The most obvious example is the zigzag band across the upper right and lower left of Fig. 2(d). Some parts of this band are collections of periodic cells that travel together, but more interesting is that the parts of this band outside of these cells are completely self-contained, indicating a separate invariant subset within \tilde{E} . This zigzag band appears to be an interval exchange transform embedded into the PWI [38]. An interval exchange transformation is the one-dimensional (1D) version of cutting and shuffling in which a line segment (or other 1D object) is split into pieces that are reordered [39–45]. A closer examination of the $(45^\circ, 45^\circ)$ protocol in Sec. IV reveals many other self-contained invariant subsets in \tilde{E} , each of which represents an isolated mixing region and a barrier to mixing that also indicates \tilde{E} is not ergodic as a whole. In contrast, the $(57^\circ, 32.75^\circ)$ protocol in Fig. 2(e) produces good mixing (gray regions), except for the unmixed cells corresponding with the white regions in Fig. 2(b). Thus, this protocol does not appear to have any structure outside of the nonmixing islands, which suggests that there are no smaller invariant subsets in \tilde{E} .

These examples show that although the fractional coverage Φ_∞ indicates what fraction of the domain is outside of nonmixing cells, it fails to indicate the degree of mixing within \tilde{E} . Hence, the fractional coverage of \tilde{E} only indicates how much mixing is possible, not whether it occurs. In previous work [1,3,5,10,11,20,23], exceptional sets for the hemispherical PWI, as well as other PWIs, have been constructed without considering the dynamics within them. However, in the process of generating the approximation $\tilde{E}_{\varepsilon, N}$, the local amount of cutting in the exceptional set, which will be subsequently defined as the *density*, ρ , can be approximated by layering ε -fattened (finite width 2ε) cutting lines on top of one another. Variation in this ε -fattened cutting line density reveals the subtle nonmixing structures within \tilde{E} , as described in this section.

Two conditions must be true for an exceptional set to be completely ergodic on its own, that is, without separate ergodic subsets. First, since the PWI is composed of isometries (which are nondistorting), the entire ergodic exceptional set must encounter the cutting lines, \mathcal{D} , with equal frequency or density (i.e., the local distribution of cuts is uniform). That is, if one point encounters a cutting line more often than another, the two points must lie on different orbits (and therefore different invariant sets) within the exceptional set. Second, almost any point in an ergodic exceptional set (with the exception of measure zero orbits, e.g., periodic points) necessarily has an orbit dense in the entire exceptional set which includes the initial cutting lines, \mathcal{D} , such that any point in \tilde{E} will have a trajectory that falls arbitrarily close to every point along the

entire length of \mathcal{D} , and, in conjunction with the first condition, encounters all of \mathcal{D} with equal frequency or density. These two conditions motivate the following measurement of the cutting line density.

A. Cutting line density

Cutting line *density* is an invariant measure of how often a region of the domain is cut. Using the ε -fattened cutting line approximation [20], this can be thought of as layering paint along each cutting line such that the paint thickness, when normalized by the number of cuts painted, represents the local density of cutting lines.

Before formally defining the cutting line density ρ , we address some mathematical preliminaries. A mathematical *ball* is used to consider neighborhoods of points (a point plus points in the region surrounding it) such that given a point $\mathbf{x} \in S$, the closed ball of radius $r > 0$, $B_r(\mathbf{x}) = S \cap \{\mathbf{z} \in S : \|\mathbf{x} - \mathbf{z}\| \leq r\}$, contains all points \mathbf{z} within distance r of point \mathbf{x} .⁴ This is generalized to the closed r -neighborhood of a set A which is defined as $B_r(A) = S \cap \bigcup_{\mathbf{x} \in A} B_r(\mathbf{x}) = S \cap \bigcup_{\mathbf{x} \in A} \{\mathbf{z} \in S : \|\mathbf{x} - \mathbf{z}\| \leq r\}$ and likewise is the set of all points \mathbf{z} within a distance r of A .

The ε -fattened cutting lines of the PWI are defined as $\mathcal{D}_{1,\varepsilon} = B_\varepsilon(\mathcal{D}_1)$ and $\mathcal{D}_{2,\varepsilon} = B_\varepsilon(\mathcal{D}_2)$ for some small $\varepsilon > 0$. Note that the 2D Lebesgue measure ($\mathcal{L}^2(\cdot)$, or area) of the overlap is $\mathcal{L}^2(\mathcal{D}_{1,\varepsilon} \cap \mathcal{D}_{2,\varepsilon}) > 0$ for $\varepsilon > 0$ because there is in fact a small region of overlap at the intersections of the cutting lines $\mathcal{D}_{1,\varepsilon}$ and $\mathcal{D}_{2,\varepsilon}$. The ε -fattened cutting lines are used to define an ε -fattened E [20], which is called $E_\varepsilon = \bigcup_{n=-\infty}^{\infty} M^n \mathcal{D}_\varepsilon = B_\varepsilon(E)$ where $\mathcal{D}_\varepsilon = \mathcal{D}_{1,\varepsilon} \cup \mathcal{D}_{2,\varepsilon}$, that approximates \bar{E} (contains E and its limit points as well as extra points due to the ε -fattened cutting lines) since in the limit as $\varepsilon \rightarrow 0$, $E_\varepsilon \rightarrow \bar{E}$ (extra points are reduced and only E and its limit points remain) [20,46]. It is not possible to apply infinite iterations of M to completely construct E_ε , so N iterations are used to find an estimate of E_ε : $\bar{E}_{\varepsilon,N} = \bigcup_{n=0}^N M^n \mathcal{D}_\varepsilon$. There always exists a sufficiently large N such that the approximation $\bar{E}_{\varepsilon,N}$ composed of the ε -fattened cutting lines completely covers \bar{E} [20], and this estimate is increasingly accurate as $\varepsilon \rightarrow 0$ and $N \rightarrow \infty$.

Before defining the cutting line density, we define the fraction, F , of iterates for which \mathbf{x} is contained in some set $A \subset S$ as

$$F(\mathbf{x}, A) = \lim_{n \rightarrow \infty} \frac{\text{No.}\{\mathbf{x} \in M^i(A) : 0 \leq i < n\}}{n} \quad (6)$$

$$= \lim_{n \rightarrow \infty} \frac{\text{No.}\{M^{-i}(\mathbf{x}) \in A : 0 \leq i < n\}}{n}, \quad (7)$$

where ‘‘No.’’ denotes the counting measure such that the numerator is the count of iterates of \mathbf{x} in A [47]. The ‘‘moving-set’’ (M is applied to A) definition, Eq. (6), is equivalent to the more useful (when considering orbits) ‘‘moving-point’’ definition in Eq. (7). Equation (7) does not include limit points

⁴The geodesic distance on the unit sphere is used here. Geodesic distance, d_G , and Euclidean distance, d_E , on the unit sphere are related by $d_E^2 = \sin^2(d_G) + [1 - \cos(d_G)]^2$ and are equivalent in the limit as d_E or $d_G \rightarrow 0$.

arbitrarily close to the measurement set A , which makes it a poor measure for orbits that are arbitrarily close to A . For orbits arbitrarily close, but not contained in A , F is always 0. To include these limit points, the density of the orbit starting at \mathbf{x} in a set A is defined as

$$\rho(\mathbf{x}, A) = \lim_{r \rightarrow 0} F(\mathbf{x}, B_r(A)), \quad (8)$$

such that limit points are included in this measurement. For attracting maps, this density is identical to the *natural invariant measure* of the set A under said attracting map [47–49].⁵

For any $A \subset S$ it is straightforward to show that $\rho(\mathbf{x}, A) = \rho(M^n \mathbf{x}, A)$ for any integer n , i.e., ρ is invariant under the map and constant along the orbit of any \mathbf{x} in the domain. By extension, if $C \subset S$ is an ergodic subset, then $\rho(\mathbf{z}, D)$ will be identical for every $\mathbf{z} \in C$ and any subset $D \subset S$ as a direct result of ergodicity, i.e., ρ is constant throughout an ergodic set.

The sets of interest for measuring the density of the exceptional set are the ε -fattened cutting lines that generate the fat fractal. In fact, $\lim_{\varepsilon \rightarrow 0} \rho(\mathbf{x}, \mathcal{D}_\varepsilon)$ is the natural invariant density in \mathcal{D} [47,49] but is zero for all nonperiodic \mathbf{x} since the PWI has no attractors.

Finally, the density of iterates of $\mathcal{D}_{1,\varepsilon}$ at a point $\mathbf{x} \in S$ is $\rho(\mathbf{x}, \mathcal{D}_{1,\varepsilon})$ (or equivalently, the density of the orbit of \mathbf{x} in $\mathcal{D}_{1,\varepsilon}$) and the density of $\mathcal{D}_{2,\varepsilon}$ is likewise $\rho(\mathbf{x}, \mathcal{D}_{2,\varepsilon})$. As the areas of $\mathcal{D}_{1,\varepsilon}$ and $\mathcal{D}_{2,\varepsilon}$ scale linearly with ε in the limit of small ε and the map is composed of isometries, $\rho(\mathbf{x}, \mathcal{D}_{1,\varepsilon})$ and $\rho(\mathbf{x}, \mathcal{D}_{2,\varepsilon})$ are also proportional to ε in the limit $\varepsilon \rightarrow 0$. As a consequence, density measurements are normalized by ε to more easily compare densities across values of ε . The error in this measurement is discussed in Appendix A. Normalizing in this way gives, for nonperiodic orbits, the almost ε -invariant quantity ρ/ε , which, as discussed in Sec. V and Appendix B, is closely connected to the area of the corresponding orbit in the limit as $\varepsilon \rightarrow 0$.

We consider two representations of this cutting line density. First, a coloring method based on measurements of cutting line density is used to create a colored exceptional set in which the color intensity represents the cutting line density, as described next. Second, a modified recurrence plot, called a *return plot*, that captures localized density throughout the approximate exceptional set is constructed and analyzed in Sec. IV, revealing barriers to mixing and confined trajectories within \bar{E} .

B. Color coding the exceptional set

With this background, we consider the cutting line density for an example protocol (45°, 45°) after 2×10^4 iterations.

⁵The natural invariant density is a global property (any starting position \mathbf{x} will produce the same value) for attracting maps describing the normalized fraction of ‘‘time’’ a typical orbit spends in A and can be used as a measure on the attractor for attracting maps. Since PWIs are nonattracting maps, this density can be measured throughout the exceptional set and will produce different values for starting points \mathbf{x} that are in different invariant sets. The natural invariant density is therefore not a global property of the map [e.g., $\rho(\mathbf{x}, E) = 0$ for all $\mathbf{x} \in O$ since E and O are distinct invariant sets, $E \cap O = \emptyset$].

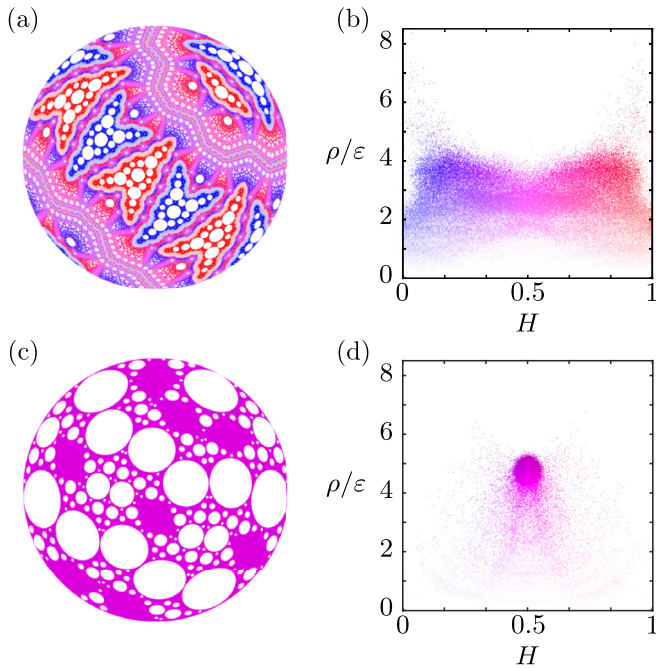


FIG. 3. Area-preserving projection of (a) the $(45^\circ, 45^\circ)$ protocol and (c) the $(57^\circ, 32.75^\circ)$ protocol after $N = 2 \times 10^4$ iterations using a cutting line width of $\varepsilon = 0.001$ as viewed from below. (a), (c) Both hue (relative density) and lightness (density) of cutting lines. (b), (d) Scatter plot of hue, H , and cutting line density, ρ/ε , for colors shown in (a) and (c) respectively.

Applying the PWI repeatedly, the cutting lines, \mathcal{D}_1 and \mathcal{D}_2 , are cut and shuffled around the domain forming a part of E , which has cutting lines overlapping only at individual points. With ε -fattened cutting lines, cutting line intersections are small patches instead of individual points, that is $\mathcal{L}^2(\mathcal{D}_\varepsilon \cap M^t \mathcal{D}_\varepsilon) > 0$. This is what contributes to density (individual points do not contribute to density). These small overlapping areas are categorized in two ways to make the overall coloring shown in Fig. 3(a): the total density (of both fattened cutting lines) is used for lightness and the relative density of the two different cutting lines is used for hue. The hue measurement alone is shown in Fig. 2(a). The different hues in Fig. 2(a) demonstrate that some parts of the domain are cut mostly (or even exclusively) by \mathcal{D}_1 (red), others are cut mostly by \mathcal{D}_2 (blue), and still others are cut by both \mathcal{D}_1 and \mathcal{D}_2 (magenta).

When the absolute density is also included as a lightness with the hue, the image in Fig. 3(a) is produced. White areas show cells in which there are no cutting lines at all (and thus no cutting line density). The lightness of color in the ε -fattened exceptional set in Fig. 3(a) reflects the total cutting line density, $\rho(\mathbf{x}, \mathcal{D}_\varepsilon)/\varepsilon$. Vivid color represents a high cutting line density (many cuts), whereas pale color indicates low cutting line density (few cuts).

Thus, in visualizing the exceptional set, two different color dimensions are used: lightness based on the total cutting line density, and hue based on the relative densities of the original, red-blue cutting lines. An HSL [hue, saturation (constant), lightness] color space is used to show both dimensions together. The lightness at a point $\mathbf{x} \in S$ is defined as the total

density of the two cutting lines,

$$L(\mathbf{x}) = \rho(\mathbf{x}, \mathcal{D}_{1,\varepsilon}) + \rho(\mathbf{x}, \mathcal{D}_{2,\varepsilon}). \quad (9)$$

As noted earlier, since $\mathcal{L}^2(\mathcal{D}_{1,\varepsilon} \cap \mathcal{D}_{2,\varepsilon}) > 0$ for $\varepsilon > 0$, this lightness is not exactly density of \mathcal{D}_ε , $L(\mathbf{x}) \neq \rho(\mathbf{x}, \mathcal{D}_\varepsilon)$, because $L(\mathbf{x})$ double counts the region at the intersection of ε -fattened cutting lines. However, since $\mathcal{L}^2(\mathcal{D}_{1,\varepsilon} \cap \mathcal{D}_{2,\varepsilon})$ is roughly equal to $2\varepsilon^2[\csc(\alpha)\sec(\alpha) + \csc(\beta)\sec(\beta)]$ for small ε , the two measurements are similar, $L(\mathbf{x}) \sim \rho(\mathbf{x}, \mathcal{D}_\varepsilon)$, in the limit of small ε for $\alpha, \beta \gg 0$ (see Appendix A for more discussion on the error incurred by this overlap). Hence, we use the symbols ρ and L interchangeably to mean $L(\mathbf{x}) \approx \rho(\mathbf{x}, \mathcal{D}_\varepsilon)$. For small values of α or β , this similarity breaks down, because the overlap of \mathcal{D}_1 and \mathcal{D}_2 is large. For an ergodic set, the lightness will be uniform throughout since every point returns to the cutting lines with equal frequency.

The second color dimension is the hue. Hue is measured as a fraction of interactions with one of the two cutting lines out of the total interactions with both cutting lines,

$$H(\mathbf{x}) = \frac{\rho(\mathbf{x}, \mathcal{D}_{1,\varepsilon})}{\rho(\mathbf{x}, \mathcal{D}_{1,\varepsilon}) + \rho(\mathbf{x}, \mathcal{D}_{2,\varepsilon})}, \quad (10)$$

which is undefined when both $\rho(\mathbf{x}, \mathcal{D}_{1,\varepsilon})$ and $\rho(\mathbf{x}, \mathcal{D}_{2,\varepsilon})$ are zero (i.e., inside periodic cells). The range of $H(\mathbf{x})$ is $[0, 1]$, from blue to red. For an ergodic set, the hue is uniform throughout the set since every point encounters either cutting line with equal frequency.

As mentioned above, an ergodic subset $B \subset S$ has the same density, $\rho(\mathbf{x}, B)$ for all $\mathbf{x} \in B$ for any sample set C . Equal densities in an ergodic set imply the same coloring according to this visualization method. Although identical color cannot be used to verify that a subset is indeed ergodic, differences in color can be used to show that subsets are necessarily not ergodic. The scatter plot of hue and lightness in Fig. 3(b) clearly shows that the $(45^\circ, 45^\circ)$ protocol is not ergodic since color is not uniform in either hue or lightness. Similar color coding has been used by Ashwin and Goetz [37] to color code periodic cells based on their return frequency to a particular atom of the PWI, $\rho(\mathbf{x}, P_i)$.

The hue of the $(45^\circ, 45^\circ)$ protocol, calculated using Eq. (10) and shown in Fig. 2(a), reveals large arrowhead like structures that are exclusively red or blue that do not mix with the rest of the domain. From the lightness, combined with the hue in Fig. 3(a), it is clear that there are even regions within these arrowheads that are not mixing within themselves, notably their lighter borders have a low cutting line density and do not mix with their darker cores, which have a different (higher) cutting line density. Furthermore, recall that Fig. 2(d) shows that the $(45^\circ, 45^\circ)$ protocol has invariant subsets within \tilde{E} that hinder mixing, such as the zigzag bands across the upper right and lower left. When using the normalized density in Fig. 3(a), these zigzag bands are a light magenta indicating low cutting line density bands that separate darker regions of high cutting line density in the exceptional set from one another.

For a protocol to have an ergodic \tilde{E} , there are conditions on what the coloring of $\tilde{E}_{\varepsilon,N}$, the approximate exceptional set can be. Ergodicity implies that dynamics within \tilde{E} should have no bias in proximity to the two cutting lines. Since the average color of \mathcal{D} is a magenta halfway between red

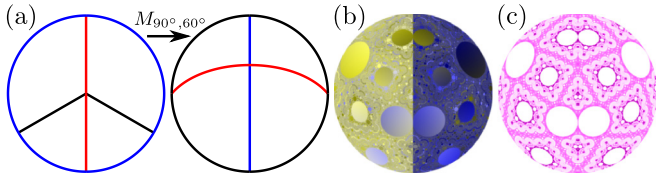


FIG. 4. (a) The $(90^\circ, 60^\circ)$ protocol shown from below using an equal-area projection to a disk. (b) Using the initial condition in Fig. 2(c), mixed state after 2×10^4 iterations. (c) Hue and lightness visualization of cutting lines for $\varepsilon = 0.001$ after 2×10^4 iterations of the PWI.

and blue, the resulting stacked cutting lines should be this same average hue of magenta [$H(\mathbf{x}) = 0.5$, $\mathbf{x} \in \bar{E}$]. Further, since M is a PWI and therefore area preserving, all regions within \bar{E} should pass within ε of cutting lines with an equal frequency. If \bar{E} is ergodic, $\bar{E}_{\varepsilon, N}$ should be a solid magenta color without variation in intensity everywhere inside of \bar{E} (periodic cells will have single colored boundaries resulting from an intersection between E_ε and O). This does not mean that a single magenta-colored $\bar{E}_{\varepsilon, N}$ implies ergodicity, but it does mean that if any portion of $\bar{E}_{\varepsilon, N} \cap \bar{E}$ is not a single magenta color, the protocol definitively produces a nonergodic exceptional set. The distribution of colors using hue and lightness for the $(45^\circ, 45^\circ)$ protocol in Fig. 3(b) indicates that neither hue nor cutting line density are uniform throughout the exceptional set, each varying substantially, and therefore the exceptional set for the $(45^\circ, 45^\circ)$ protocol is not ergodic.

The differences in $\bar{E}_{\varepsilon, N}$ for the $(45^\circ, 45^\circ)$ protocol shown in Fig. 3(a) and the $(57^\circ, 32.75^\circ)$ protocol shown in Fig. 3(c) are obvious. The lightness of cutting line intersections in Fig. 3(c) shows no distinguishable pattern and is nearly uniform in density. This does not imply that \bar{E} is ergodic, but it does not rule out ergodicity as it did for $(45^\circ, 45^\circ)$. The hue of overlapped cutting lines in Figs. 2(b) and 3(c) also shows no discernible pattern within \bar{E} , showing the entire hemispherical shell is uniformly cut by both \mathcal{D}_1 and \mathcal{D}_2 throughout except, of course, in the cells where there are no cuts. Although it cannot necessarily be concluded that \bar{E} is ergodic, it is likely that \bar{E} for the $(57^\circ, 32.75^\circ)$ protocol has fewer ergodic subsets than for $(45^\circ, 45^\circ)$. We conclude that mixing within the exceptional set in the $(57^\circ, 32.75^\circ)$ protocol is quite different from that for the $(45^\circ, 45^\circ)$ protocol, despite both exceptional sets occupying nearly the same fraction of the hemisphere, due to the apparent absence of separate invariant subsets and the barriers to mixing they create within the exceptional set in the former. This is emphasized by the tight distribution of hue and cutting line density in the scatter plot in Fig. 3(d).

Unfortunately, neither hue, $H(\cdot)$, nor density, ρ/ε , as defined here completely identifies invariant subsets in \bar{E} . To illustrate this, consider the $(90^\circ, 60^\circ)$ protocol shown in Fig. 4. This protocol has coincident red and blue cutting lines, shown in Fig. 4(a) with a red cutting line down the vertical diameter for the initial cutting lines and a blue cutting line down the vertical diameter. This creates poor mixing as shown in Fig. 4(b), but this also makes separation of cutting lines by hue as defined here impossible as shown in Fig. 4(c), since every red cutting line maps to a blue cutting line and $\rho(\mathbf{x}, \mathcal{D}_{1, \varepsilon}) = \rho(\mathbf{x}, \mathcal{D}_{2, \varepsilon})$. Notice that all points in $\bar{E}_{\varepsilon, N}$ in Fig. 4(c) have the

same magenta hue, and the only variation is due to changes in cutting line density. This protocol produces a left-right barrier to mixing due to the coincident cutting lines [20] that is obvious in Fig. 4(b) but not reflected in the hue-lightness visualization, Fig. 4(c). In order to differentiate invariant subsets in protocols like this, more information about orbits within \bar{E} is needed, as described in Sec. IV.

IV. CATEGORIZING INVARIANT SUBSETS USING RECURRENCE

The coloring scheme in Sec. III B based on different colors for cutting lines \mathcal{D}_1 and \mathcal{D}_2 is useful for understanding which cutting lines are responsible for different parts of \bar{E} , but this red-blue partitioning is somewhat coarse. Instead of considering just \mathcal{D}_1 and \mathcal{D}_2 separately (by color), the cutting lines can be decomposed into smaller pieces and ultimately into a continuum of points with densities that indicate their interactions.

Every intersection point of iterated cutting lines, $M^n \mathcal{D}$, with the original cutting lines, \mathcal{D} , is a recurrent point from \mathcal{D} to \mathcal{D} . Points $\mathbf{x} \in \mathcal{D}$ and $M^n(\mathbf{x})$ are called recurrent if $M^n(\mathbf{x}) \in \mathcal{D}$. Each recurrent point necessarily occupies the same orbit under the PWI. For most protocols, these recurrent points appear to be dense in \mathcal{D} , but there are only a countably infinite number of them,⁶ meaning almost all of \mathcal{D} is not exactly recurrent. This concept of recurrence can be expanded to the ε -fattened cutting lines to approximate the exact return structure of the cutting lines. Orbits in \bar{E} can then be grouped by which points in \mathcal{D} produce the same orbit. This allows invariant subsets in \bar{E} to be uniquely identified by the segments of \mathcal{D} they intersect.

A. Parameterization of cutting lines

In order to separate orbits in \bar{E} by their intersections with \mathcal{D} , every point in \mathcal{D} needs a unique identifier. For the hemispherical shell PWI, atom boundaries that make up \mathcal{D} are great circle arcs with finite length. This allows a simple parametrization along atom boundaries to uniquely identify each element of \mathcal{D} . For almost all protocols, it is sufficient to parametrize along the length of just the red and black lines in Fig. 1(e) or just the red and blue lines in Fig. 1(f) whose total length is 2π . However, some protocols, such as $(90^\circ, 60^\circ)$ shown in Fig. 4(a), have a pronounced “sidedness” associated with the coincident cutting lines. This is evident in the separation between the two sides of the vertical midline, where the cutting lines are coincident, in Fig. 4(c), indicating that each side of the same cutting line interacts only with its own side of the domain. Because of this, a parametrization should account for the different atoms on either side of the cutting lines. To be clear, the two “sides” of a cutting line are simply the two sets of dynamics assigned to the cutting line, one for each neighboring atom P_k . This sidedness is

⁶Recurrent points can only arise due to intersected cutting lines of which there are only ever a countably infinite number. That is, although there are an infinite number of cutting lines in E , they can be ordered according to the iteration of the map that generates them.

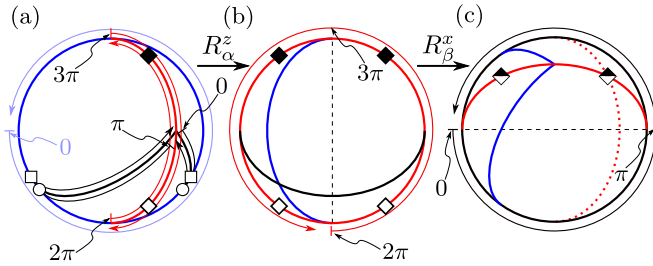


FIG. 5. The $(45^\circ, 45^\circ)$ PWI's cutting lines' (a) initial position, (b) position after rotation about the z axis by α , and (c) position after both rotations; shown orthographically from the negative y axis. The black cutting line (and blue boundary) in (a) is parameterized by $\theta \in [0, 2\pi)$ in (c). The red cutting line is parameterized by $\theta \in [2\pi, 4\pi)$ in (b). The parametrization of the initial black cutting line in (a) is broken by the equator, and the two pairs of matching points for the parametrization are labeled with circles and squares, respectively. The upper right half-filled diamond in (c) is a recurrent point lying on the initial red cutting line (dotted red) and is the combined postimage of the filled and unfilled diamonds in (a).

also prominent for any protocol that produces polygonal cells whose sides return exactly to \mathcal{D} , including those that produce polygonal tilings, which are composed entirely of periodic segments [11,20].

The $(45^\circ, 45^\circ)$ protocol is used as an example to demonstrate the parametrization of cutting lines, as shown in Fig. 5(a). To parametrize along the initial red and black cutting lines in a manner that accounts for the sidedness of the cutting lines, the parametrization is specified based on the orientation, with respect to the rotation procedure in Fig. 1, that places the cutting line on the hemispherical shell equator. Cutting lines are not multivalued when they lie on the equator, which makes measurements accounting for sidedness natural. Working backwards from the final orientation of the cutting lines after one iteration in Fig. 5(c), the black cutting line is parameterized first from $\theta \in [0, 2\pi)$ starting at a point coinciding with the second rotation axis [outer black circle in Fig. 5(c)] such that inverting the rotation about the second axis collapses this parametrization onto the black cutting line in Figs. 5(a) and 5(b). This collapse is why, for protocols without explicitly coincident cutting lines, a sided parametrization often leads to redundant parameter values. The red cutting line is treated similarly and is parameterized according to its orientation when it lies on the equator [outer red circle in Fig. 5(b)] from $\theta \in [2\pi, 4\pi)$ which similarly collapses to the single cutting line in Figs. 5(a) and 5(c). This parametrization defines the arc-length parametrization of \mathcal{D} which is referred to as $s(\theta): [0, 4\pi) \rightarrow \mathcal{D}$. Note that the equatorial circular arc $\theta \in [0, 2\pi)$ in Fig. 5(c) collapses to the two sides of the black cutting line in Fig. 5(a) with two matching points in the parametrization (broken by the PWI) labeled by a pair of circles or squares, and that the equatorial circular arc $\theta \in [2\pi, 4\pi)$ in Fig. 5(b) collapses to the two sides of the red cutting line in Fig. 5(a). This parametrization is not unique as any parametrization could be chosen for \mathcal{D} , but this one is selected for its connection to the single-valued orientation of the original cutting lines and because it is an arc-length parametrization.

The necessity for a sided or single-valued parametrization is made clear by examining the points along \mathcal{D} labeled by filled and unfilled diamonds in Fig. 5. These two points in Fig. 5(a) map to two opposite points on the equator in Fig. 5(b) resulting in four total points. The sided parametrization allows a distinction between these split points to be made before the split occurs. Using a single-valued, or sided, parametrization allows the tracking of all four points even when two of them have collapsed onto one another as in Figs. 5(a) and 5(c). Note that the upper right point in Fig. 5(c) has returned to the original cutting line (the dotted red line) such that the sided parametrization breaks down for the next iteration at this point (since it is multivalued for yet another iteration). Despite this breakdown at a single point, the sided parametrization provides additional information about how the cutting lines move subject to the PWI mapping.

B. Return plots

The cutting line parametrization provides a mechanism for grouping recurrent points in \mathcal{D} , that is, points that start in \mathcal{D} and return back to \mathcal{D} . Orbits may be trapped within an invariant set, and by separating out points that are confined to certain invariant sets, barriers to mixing within cut regions can be located. In the physical tumbler, orbits confined to cells do not mix with the rest of the domain [26]. Other barriers to mixing also seem to exist that, despite being cut by the map, produce confined orbits that do not actually contribute to the overall mixing within the exceptional set.

For example, one recurrent point for $(45^\circ, 45^\circ)$ after one application of the PWI is indicated with a half-filled diamond in the upper right of Fig. 5(c), but there are recurrent points wherever cutting lines intersect, including the boundary (which is always treated using the same parametrization as the initial black cutting line). Since recurrent points are marked by cutting line intersections, an “intersection plot” or “return plot,” similar to a recurrence plot for time series [50–52], can be made where the horizontal and vertical axes are both $\theta \in [0, 4\pi)$. If the point specified by $s(\theta_1)$ returns to \mathcal{D} at a point specified by $s(\theta_2)$, then a return plot will be populated at both (θ_1, θ_2) and (θ_2, θ_1) , since recurrence is symmetric. This also means that the entire $\theta_1 = \theta_2$ line will be populated, since every point in \mathcal{D} is trivially recurrent upon itself. Exact intersections (up to machine precision) can be computed to create an exact intersection map, which is essentially a first return plot for \mathcal{D} , by applying M to segments of \mathcal{D} and finding the exact locations of their intersection with other parts of \mathcal{D} , which breaks each segment into smaller segments, using the parametrization in Fig. 5 to exactly identify points.

Exact recurrence measured in this way has shortcomings. First, the plot will only ever be sparse since cutting line intersections are countable. Additionally, due to the multivalued nature of the discontinuity, any sort of sidedness of cutting lines (neighboring atoms having discontinuous dynamics at the cutting line) cannot be detected computationally using exact recurrence. This is because a segment returning exactly to the original cutting line, either in whole or at its endpoint, will not be cut [an example where segments return exactly at their endpoints is the $(90^\circ, 60^\circ)$ protocol shown in Fig. 4], and numerical comparison of two floating point zeros (used

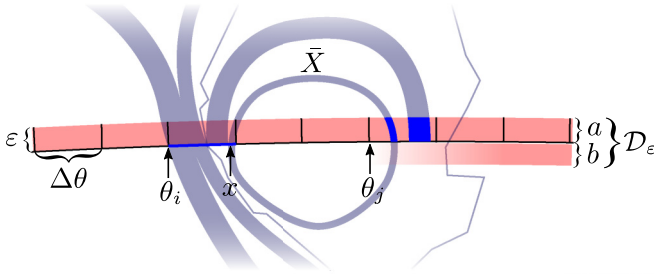


FIG. 6. Schematic diagram of a fat, two-sided (a and b) cutting line in red with example orbits in light gray. The cutting line is split into bins of length $\Delta\theta$ in which local orbit density is measured. These example orbits are not representative of all possible orbits in an actual PWI.

to detect intersections) is impossible. Additionally, as cutting line segments are continually split into smaller and smaller segments, the number of cutting line segments to be tracked (in order to find intersections) appears to grow with the square of the number of iterations, N^2 , for protocols that do not result in polygonal tilings.⁷ This results in a constantly growing memory requirement that quickly makes the process of making an exact return plot intractable.

Instead of an exact return plot, an approximate return plot is used based on ε -fattened cutting lines. Points along the cutting lines are iterated using the PWI, and, if they return to \mathcal{D}_ε , they approximately return to the closest point in \mathcal{D} . The cutting lines are split into a finite number, T , of bins $[\theta_i, \theta_i + \Delta\theta)$ where $\theta_i, \theta_i + \Delta\theta \in [0, 4\pi)$, and each returns to the ε -fattened cutting lines is placed in a bin within a 2D grid of starting bins and return bins. Figure 6 illustrates a sided, ε -fattened cutting line (horizontal with red shading to show the ε -fattening) to demonstrate how the cutting line intersects with some example orbits. The cutting line is segmented into $\Delta\theta$ -long bins in which the local density of an orbit can be measured.

To create the return plot, points are initially placed along \mathcal{D} in the i -bins $s([\theta_i, \theta_i + \Delta\theta))$, which are assigned to the horizontal axis of the return plot. These points then increment the corresponding return j -bin, $[\theta_j, \theta_j + \Delta\theta)$, along the vertical axis when they return within ε of the cutting line as they are iterated by the PWI. This localized return density, called ρ_{ij} , represents how often a point starting in the bin $[\theta_i, \theta_i + \Delta\theta)$ returns to the bin $[\theta_j, \theta_j + \Delta\theta)$. For example, point x in Fig. 6 will fill out the orbit \bar{X} under the PWI, and every return to the θ_j bin (dark blue region of \bar{X} on the right) updates the ρ_{ij} measurement and indicates that the two bins are connected by an orbit. This can also be thought of as the fractional area of orbits from the θ_i bin that intersects the θ_j bin as shown in Fig. 6. Several points to the left of point x in bin θ_i also intersect bin θ_j following a different orbit, while other points further to the left of point x in bin θ_i follow orbits that do not intersect bin θ_j . Also recall that the ε -fattened cutting line

has two sides; points on opposite sides of the cutting line may follow different orbits and intersect different bins. Because of this, initial points are placed a small distance $\delta \ll \varepsilon$ from the cutting line to simplify this distinction in the computational analysis.

Because this local density links a cutting line segment and an ε -wide bin, there is an inherent, but subtle, asymmetry in the resulting local density (only disappearing in the $\varepsilon \rightarrow 0$ limit). The local density is written mathematically as the cumulative fraction of all orbits from the θ_i bin that intersect the θ_j bin,

$$\rho_{ij} = \int_{\theta_i}^{\theta_i + \Delta\theta} \rho(\mathbf{s}(t), \{\varepsilon \times [\theta_j, \theta_j + \Delta\theta)\}) dt, \quad (11)$$

where $\{\varepsilon \times [\theta_j, \theta_j + \Delta\theta)\}$ is the θ_j bin constructed by the ε expansion (to the left in the direction of increasing θ) of the cutting line segment making up the entire θ_j bin as shown in Fig. 6.

Figure 7(a) shows the resulting return plot for the $(45^\circ, 45^\circ)$ protocol. We plot the value of ρ_{ij} normalized by both the cutting line half-width, ε , and the bin length, $\Delta\theta$. Orbits where points return frequently to a small number of bins (possibly localized near the cutting line) result in highly skewed local density ρ_{ij} values [e.g., black dots at the centers of the dark squares in Fig. 7(a)]. The skewed data are transformed to $\log_{10}(\frac{\rho_{ij}}{\varepsilon\Delta\theta} + 1)$ to better reveal the less frequent return regions, which are lighter (due to orbits occupying more space away from the cutting lines). Note that the earlier measurement of hue used to color the exceptional set in Sec. III is exactly this measurement with only two bins: one bin for all the points along the red cutting line, $[0, 2\pi)$, and one bin for all the points along the blue cutting line, $[2\pi, 4\pi)$, indicated on the axes of Fig. 7(a) with adjacent blue and red lines.

The return plot is reminiscent of the adjacency matrix of a weighted graph, with each entry relating the weight of connectivity between two infinitesimal nodes along the cutting lines.⁸ The weights in a traditional weighted graph adjacency matrix indicate the difficulty of traveling between nodes. Here the return plot indicates the reciprocal of this difficulty, i.e., the ease of travel between two “nodes,” a value of zero indicating travel is impossible. Questions about the shortest orbit given some diffusion on the scale of ε , which is built into the return plot, between any two parts of the cutting line can be answered. A measurement of a time-to- ε -connectivity between regions of the cutting line, which comes from this perspective, could indicate the rate of mixing within the system, but these analyses are beyond the scope of this study.

For Fig. 7(a), each $[\theta_i, \theta_i + \Delta\theta)$ bin is seeded with 10^3 points uniformly arranged along the bin. In order to maintain sidedness, the initial points are placed a distance $\delta \ll \varepsilon$ away from either side of the cutting line. The error in position

⁷This is a preliminary result that is backed up by evidence from the described method not detailed here, but beyond the scope of this work.

⁸Any introductory text on graph theory (e.g., Ref. [53]) should include a discussion of weighted graphs and shortest path problems as well as an algorithm due to Dijkstra for computing the shortest path [58].

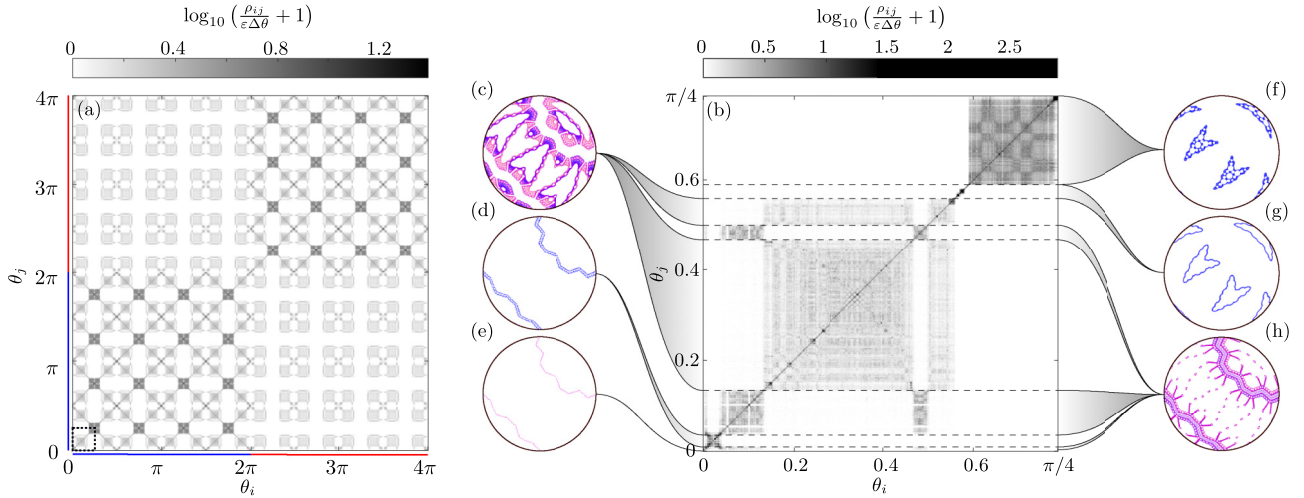


FIG. 7. (a) Approximate return plot for the $(45^\circ, 45^\circ)$ protocol as a 2D histogram of ρ_{ij} , return to bin i from bin j ; $\Delta\theta = \pi \times 10^{-3}$. (b) Detail of (a) from the dashed black box in the lower left; $\Delta\theta = \pi/16 \times 10^{-3}$. (a), (b) 2×10^6 points are initialized $\delta = 10^{-9}$ from the cutting line, iterated $N = 2 \times 10^6$ times using a cutting line width $\varepsilon = 10^{-5}$. The $\theta \in [0, 2\pi]$ range represents \mathcal{D}_2 [blue (black) cutting line] and the $\theta \in [2\pi, 4\pi]$ range represents \mathcal{D}_1 (red cutting line). (c)–(h) Various regions of the exceptional set corresponding to portions of the return plot.

after N iterations for the double precision numbers grows like $N \times (2 \times 10^{-16})$ where 2×10^{-16} is roughly machine precision (sometimes called machine epsilon). As a direct result of this error, it is not possible to place points exactly on a cutting line. The plots shown in the following figures use $N = 2 \times 10^6$ iterations such that the error in position after all iterations is roughly 4×10^{-10} . Setting δ smaller than this value allows for points to potentially switch sides of the initial cutting line at some point during the application of the PWI, so $\delta = 10^{-9}$ is used for the following figures. See Appendix A for further discussion of the error introduced and mitigated by δ . Since sidedness often adds a symmetry to protocols without coincident cutting lines such as $(45^\circ, 45^\circ)$, the approximate return plot in Fig. 7(a) is symmetrical across $\theta = \pi$ within the $[0, 2\pi)$ range and across $\theta = 3\pi$ within the $[2\pi, 4\pi)$ range in both axes. There is also a symmetry across $\theta_i = \theta_j$, because an orbit connecting θ_i to θ_j also connects θ_j to θ_i [there is a subtle asymmetry when using ε -fat cutting lines that is invisible in Fig. 7(a); see Appendix A].

How the orbits throughout the domain are connected with one another directly affects mixing. Figure 7(a) shows approximately where \bar{E} is connected and, likewise, where it is disconnected. If ρ_{ij} is nonzero (gray or black), there exists at least one orbit in \bar{E} that begins in $[\theta_i, \theta_i + \Delta\theta)$ along the cutting line and either intersects or passes within ε of the $[\theta_j, \theta_j + \Delta\theta)$ bin. If ρ_{ij} is zero, white in Fig. 7(a), then no such orbit exists and bins i and j are disconnected from one another such that the orbits within each are necessarily in different invariant subsets of \bar{E} . White stripes indicate that the corresponding orbit is disconnected from a large part of \bar{E} . Disconnected regions directly impact the mixing dynamics by containing the mixing to a region smaller than the entire exceptional set.

The larger the value of ρ_{ij} [darker in Fig. 7(a)], the more often an orbit returns to that bin. If an orbit spends many iterations returning to the same cutting lines, it is not spending as many iterations exploring the domain. As such, a high value of ρ_{ij} indicates the associated orbit is small or short. Similarly,

large orbits that spread throughout the domain return to the cutting lines with a much lower frequency (lighter color). We refer to filled regions in the same horizontal (row) or vertical (column) position together as *blocks* due to their rectangular shape. Each block has orbits that return close to everywhere else in the block, indicating potential mixing. Since mixing can only occur between connected regions, disconnected blocks represent disparate mixing regions.

Figure 7(a) shows that some regions (invariant sets) are completely disconnected from the rest of \bar{E} , as indicated by the white stripes in the return plot. For example, in the lower left [corresponding to \mathcal{D}_2 , $\theta \in [0, 2\pi]$] and upper right [corresponding to \mathcal{D}_1 , $\theta \in [2\pi, 4\pi]$] the dark square blocks with empty vertical and horizontal stripes around them correspond directly with the blue and red arrowheads in Fig. 3(a). These are regions that only interact with one of the two cutting lines and are isolated from the rest of \bar{E} .

The return plot in Fig. 7(a) can be used to split the exceptional set into its constituent invariant sets. To explain, consider a detail of the lower left corner of Fig. 7(a), which is shown in Fig. 7(b) for $\theta \in [0, \pi/4]$. A close examination of Fig. 7(b) demonstrates how the repeating section shown in the figure approximately decomposes the exceptional set for the $(45^\circ, 45^\circ)$ protocol into different invariant sets. The thin curve at the center of the zigzag in the exceptional set, shown in Fig. 7(e), intersects $(\theta_i, \theta_j) = (0, 0)$ in the domain shown in Fig. 7(b). Immediately surrounding this zigzag curve is a caterpillar-like shell in Fig. 7(h) and a region outside the arrowheads and the zigzag in Fig. 7(c). Some invariant sets only intersect the blue cutting line \mathcal{D}_2 . A small region inside the caterpillar-like zigzag is distinct from the central zigzag curve, as shown in Fig. 7(d). The boundary of the blue arrowheads in Fig. 7(g) is actually a line of small dark blocks, each representing a thin invariant set that outlines an arrowhead, as shown in Fig. 7(g) surrounding the blue arrowhead core in Fig. 7(f). Finally, there is a separate set intersecting $(\theta_i, \theta_j) = (\pi/4, \pi/4)$ (not shown) which is a point of tangency with a circular cell and produces a series of very

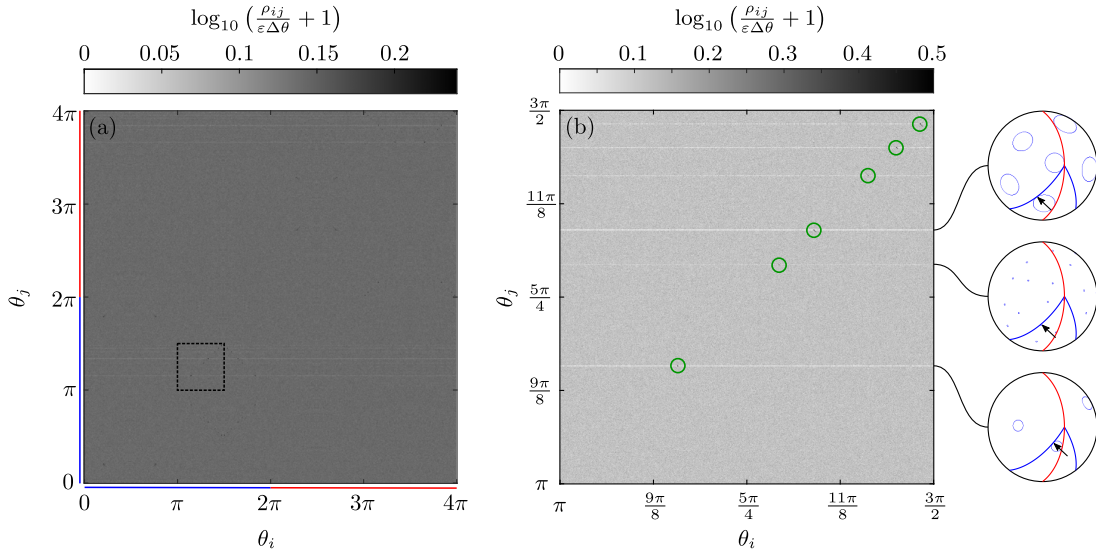


FIG. 8. (a) Approximate return plot for the $(57^\circ, 32.75^\circ)$ protocol as a 2D histogram of ρ_{ij} , return to bin i from bin j ; $\Delta\theta = \pi \times 10^{-3}$. (b) Detail of (a) from the dashed black box in the lower left; $\Delta\theta = \pi/8 \times 10^{-3}$. Some of the dark spots along the diagonal are circled in green with the corresponding cells responsible for the horizontal white lines and dark spots shown on the right. The arrows indicate the point of tangency with the circular cells responsible for the depletion of area. (a), (b) 2×10^6 points are initialized $\delta = 10^{-9}$ from the cutting line, and iterated for $N = 2 \times 10^6$ iterations using a cutting line width $\varepsilon = 10^{-5}$. The $\theta \in [0, 2\pi)$ range represents \mathcal{D}_2 [blue (black) cutting line] and the $\theta \in [2\pi, 4\pi)$ range represents \mathcal{D}_1 (red cutting line).

small circles throughout the exceptional set. The associated short orbits that return frequently to the cutting line create the extreme skew in the return plot, which in turn requires the use of the log scale in the figures.

None of the gray blocks associated with the $(45^\circ, 45^\circ)$ protocol in Fig. 7(b) have a uniform ρ_{ij} throughout the block. The existence of nonuniform patterns within blocks indicates that these blocks are not the smallest invariant sets, although different sets may be so intertwined with one another that separating them in this fashion is impossible. This is obvious in the upper right of Fig. 7(b), where the cutting line density has an almost checkerboard pattern suggesting that different parts of the arrowhead in Fig. 7(f) correspond to different intensities in the block. On the other hand, the combination of two blocks is evident in the color variation within Fig. 7(c), demonstrating how blocks may correspond to intertwined invariant sets that are close but not overlapping.

In order for a mixing block to exist, an orbit in the block must intersect a nonzero length of the cutting line. Combined with an occupation of a nonzero width away from the cutting line in an orbit, this means that any mixing block must have orbits that individually have nonzero area on the hemispherical shell (finite length and width). It appears that a necessary condition for mixing is that individual orbits occupy a finite area, as it allows orbits to interact with one another while thin (zero-area) curves that do not overlap anywhere do not allow mixing between them. Although a large exceptional set can be constructed using thin curves (see Appendix C), there will be no mixing if orbits do not create the area for mixing to occur.

With this decomposition of the $\theta_i, \theta_j \in [0, \pi/4]$ region in mind, the entirety of the structure in Fig. 7(a) can be decomposed into similar invariant sets. Invariant sets intersecting both \mathcal{D}_1 and \mathcal{D}_2 from $[0, \pi/4] \times [0, \pi/4]$ are repeated throughout the entire $[0, 4\pi] \times [0, 4\pi]$ space. Invariant sets

intersecting just the blue cutting line, \mathcal{D}_2 , are repeated only in the lower left, $[0, 2\pi) \times [0, 2\pi)$. The invariant sets intersecting just the red cutting line, \mathcal{D}_1 , are repeated only in the upper right, $[2\pi, 4\pi) \times [2\pi, 4\pi)$, but are identical in their return structure to those that intersect just \mathcal{D}_2 .

The return plot for $(45^\circ, 45^\circ)$ corresponds with the structure evident when invariant subsets are color coded as in Fig. 3. In contrast, the return plot for $(57^\circ, 32.75^\circ)$, shown in Fig. 8(a), has no distinct blocks that would indicate distinct invariant subsets. This is consistent with the previous observations that there are no distinct invariant subsets within \bar{E} in Fig. 3(c). In fact, if \bar{E} for $(57^\circ, 32.75^\circ)$ is truly ergodic, then Fig. 3(c) would be uniform except for the nonmixing cells. However, upon close inspection of a small section of the return plot for the $(57^\circ, 32.75^\circ)$ protocol in Fig. 8(b), faint white horizontal lines are visible which result from the proximity of orbits to a large circular periodic cell. In addition, faint dark dots, inside the superimposed circles, are also visible along each horizontal white line at the intersection with the $\theta_i = \theta_j$ diagonal. These dark spots are from points placed δ away from \mathcal{D} that fall inside a periodic cell. The horizontal white lines result from the decreased width of \bar{E} immediately next to a cell. Essentially, the cell occupies some of the \mathcal{D}_ε sample region, which creates a light region since points are less likely to be in this small region compared with the full or nearly full width regions elsewhere along the cutting line. White vertical lines from points placed in a cell are noticeably absent since there are several hundred points seeded in each horizontal bin. Since $\delta \ll \varepsilon$, the vertical stripes produced by placement inside a cell (due to δ) are smaller (and not visible) than those produced by the intersection of a cell with \mathcal{D}_ε (due to ε). Thus, the trapping of points within cells is not the dominant effect producing these aberrations, but rather it is the depletion of area of $\mathcal{D}_\varepsilon \cap \bar{E}$ due to adjacent

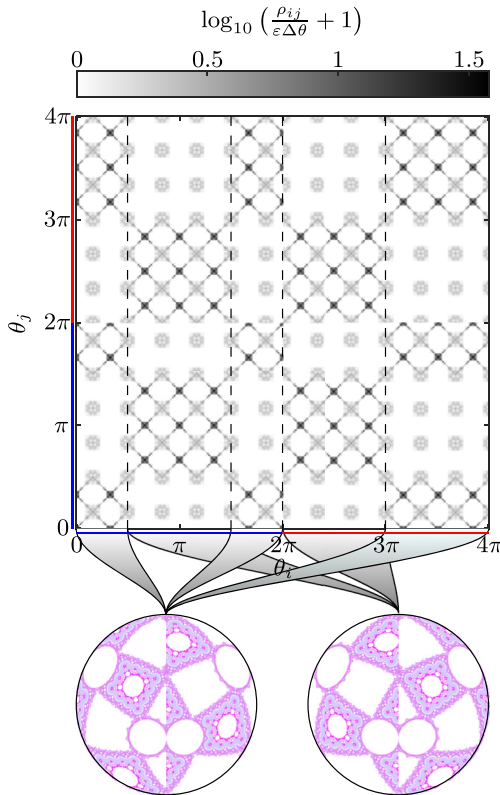


FIG. 9. Return plot ($\delta = 10^{-9}$) for $(90^\circ, 60^\circ)$ protocol showing ρ_{ij} , using 2×10^6 points, $N = 2 \times 10^6$ iterations, and $\varepsilon = 10^{-5}$.

circular cells which makes returning to cell-adjacent regions less likely.

Recall that this “sided” parametrization of the cutting lines was chosen to specifically deal with cases where cutting lines are coincident, such as the $(90^\circ, 60^\circ)$ protocol shown in Fig. 4. The corresponding return plot is shown in Fig. 9. In this case, $\theta \in [2\pi, 3\pi)$ and $\theta \in [3\pi, 4\pi)$ are two “sides” of the same cutting line, each corresponding to different atoms for the PWI dynamics. These two sides act in a similar fashion to the two separate cutting lines in the $(45^\circ, 45^\circ)$ protocol. For example, the symmetry across $\theta = 3\pi$ for $\theta \in [2\pi, 4\pi)$ evident for the $(45^\circ, 45^\circ)$ protocol is absent in Fig. 9. The two distinct regions (each with the same vertical or horizontal pattern) can also be used to separate the exceptional set into two portions, as shown below the return plot. These two portions are identical when reflected about the vertical diameter, consistent with the left-right barrier to mixing shown in Fig. 4(b). Not only is there a barrier to mixing across the vertical midline, but there are also completely separate regions contributing to each half of \bar{E} that are not captured by the red and blue coloring for the cutting lines used in Fig. 4(c). Based on the return plot, it becomes clear that the red and blue cutting line assignment for color is naive in this case and does not capture the separable mixing sets. Separable mixing sets in \bar{E} [like those for the $(45^\circ, 45^\circ)$ protocol in Fig. 7(b)] are not a generic property. Instead, they occur for some protocols but not for others. However, separable mixing sets are a generic property of protocols with exclusively coincident cutting lines (polygonal tilings) as shown in Appendix D.

We have described an easily separable protocol that has large regions with no returns at all (white) in the return plot indicating barriers to mixing, $(45^\circ, 45^\circ)$, and a seemingly inseparable protocol that has no large regions without returns indicating few barriers to mixing, $(57^\circ, 32.75^\circ)$. An intermediate case is the $(57^\circ, 57^\circ)$ protocol, first shown in Fig. 1(g), for which only 0.4% of the return plot in Fig. 10(a) is empty (white), but which still has many barriers to mixing due to separate invariant subsets within the exceptional set. This is made apparent by the obvious pattern in this protocol’s return plot in Fig. 10(a). The comparison between the $(57^\circ, 57^\circ)$ protocol and the other two we have examined is imperfect because the fractional coverage for the $(57^\circ, 57^\circ)$ protocol is $\Phi_\infty \approx 0.33$, smaller than the coverage of the other two. Nevertheless, the $(57^\circ, 57^\circ)$ protocol has orbits that are not equally dense in their returns to the cutting lines, indicating that orbits are simultaneously very close to one another and still separate.

To illustrate, the $[0, \pi/4] \times [0, \pi/4]$ region of Fig. 10(a) is shown in detail and approximately decomposed into different invariant subsets in Fig. 10(b). Although there is a pattern within the return plot, it is essentially blurred or smudged such that the more precise separation done in Fig. 7(b) for the $(45^\circ, 45^\circ)$ protocol is not possible here. Instead, some of the more “fuzzy” regions can be approximately separated. Figures 10(c) and 10(d) approximate the invariant sets that compose the center of the arrowhead structures visible in Fig. 1(g). That the colors are very faint is a consequence of almost all orbits reaching almost all of the exceptional set with a low density even when each orbit has a region of high density. In order to better show these more disperse, faint orbits, the colors used for the orbits in Figs. 10(c)–10(j) are twice as dark as those used in Fig. 7(b). Figures 10(e) and 10(f) show regions of slightly concentrated blue and wide faint regions of red, exploring the boundary between the major blue and red arrowheads. The regions in Fig. 10(g) and 10(h) explore most of the exceptional set uniformly other than a difference in blue and red density. Figure 10(i) shows a region surrounding the zigzag curve shown in Fig. 10(j), similar to the zigzag in Fig. 7(e) for the $(45^\circ, 45^\circ)$ protocol. Thus, the $(57^\circ, 57^\circ)$ protocol return plot represents an intermediate between the easily separable $(45^\circ, 45^\circ)$ protocol and the seemingly impossible-to-separate $(57^\circ, 32.75^\circ)$ protocol in which almost all orbits visit most of the exceptional set with uneven density.

The return plot contains both information about the area and connectivity of the potentially ergodic region of the PWI, \bar{E} . In an abstract sense, the return plot defines the connectivity of a complex network between different segments of the cutting lines. This return plot analysis is naturally extendable to any area-preserving PWI with finite perimeter atoms (to allow parametrization) and could have applications in the analysis of other area preserving dynamical systems. We conjecture that a metric on the shapes within the return plot in conjunction with the fractional coverage Φ_∞ would create a complete measure of mixing for any given PWI of this type. In the following section, we show that Φ_∞ can be extracted from density measurements such as the return plot. Variation within the return plot seems to be an indicator of distinct mixing regions that form barriers to overall mixing. The amount of

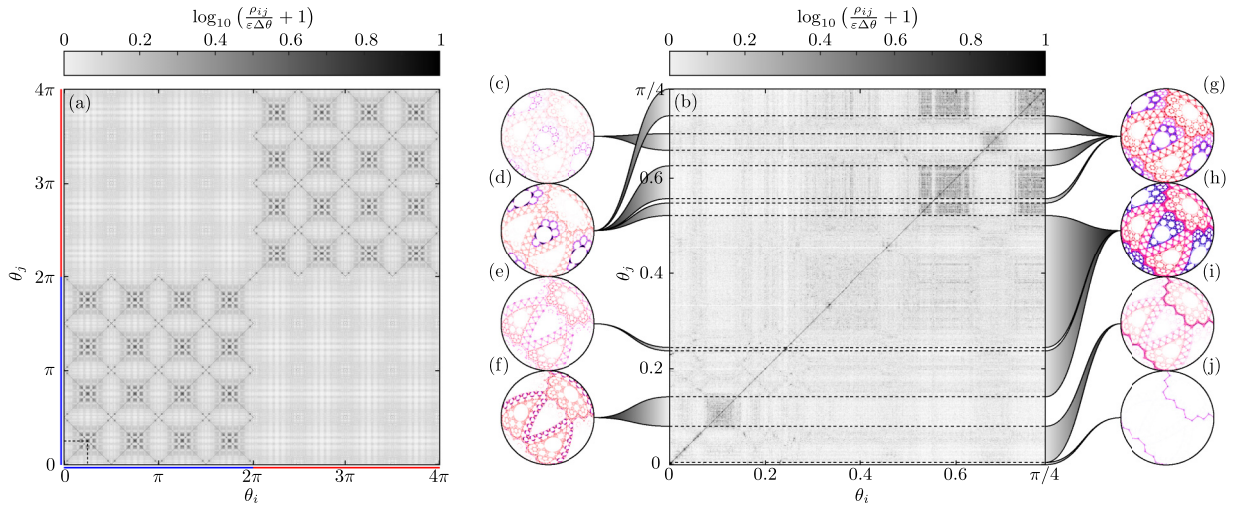


FIG. 10. (a) Approximate return plot for the $(57^\circ, 57^\circ)$ protocol as a 2D histogram of ρ_{ij} , return to bin i from bin j ; $\Delta\theta = \pi \times 10^{-3}$. (b) Detail of (a) from the dashed black box in the lower left; $\Delta\theta = \pi/16 \times 10^{-3}$. (a), (b) 2×10^6 points are initialized $\delta = 10^{-9}$ from the cutting line, iterated $N = 2 \times 10^6$ times using a cutting line width $\epsilon = 10^{-5}$. (c)–(j) Various regions of the exceptional set corresponding to portions of the return plot.

white space in the return plot also is an indicator of the lack of connectivity of the exceptional set that indicates the barriers to mixing within \bar{E} . A measurement of the variation in density or the amount of white space (disconnectedness) could, in conjunction with a measurement of the area of \bar{E} , be used to quantify the mixing under any given protocol.

V. ESTIMATING MIXING METRICS FROM CUTTING LINE DENSITY VALUES

We now return to the fractional coverage of the fat fractal exceptional set, Φ_∞ [10,11], which is the normalized area of the fat fractal, to show that the density measure described in Sec. III can be used to estimate the area of the fat fractal, and hence the value of Φ_∞ , from measurements of ρ along the cutting lines. One of the more interesting features of the exceptional set is that a 2D fat fractal is generated from 1D cutting lines. Since the PWI is composed of isometries, this means that all information (without distortion) about the exceptional set, specifically the cutting line density everywhere, must be contained within the original 1D, ϵ -fattened cutting lines themselves, and applying the PWI over many iterations simply spreads this information to the rest of the exceptional set. With this in mind, the cutting line density $\rho(\mathbf{x}, \mathcal{D}_\epsilon)$ at any point along \mathcal{D}_1 or \mathcal{D}_2 is a measure of how that particular point on the cutting line spreads through the domain under the PWI.

Since the PWI map is measure preserving, $\rho(\mathbf{x}, A)$ can be extended from a simple counting measure fraction to a fraction of how much of the orbit $X = \bigcup_{i=-\infty}^{\infty} M^i(\mathbf{x})$ is in an area A , i.e., a fraction of measures $\rho(\mathbf{x}, A) = \mu(A \cap X)/\mu(X)$. The Lebesgue measure (area) of an orbit X from point $\mathbf{x} \in \mathcal{D}_\epsilon$, where \mathcal{D} is any 1D curve such as the cutting lines (or any subset of the cutting lines), is

$$\mathcal{L}^2(X) = \int_{\mathcal{D}_\epsilon \cap X} \frac{1}{\rho(\mathbf{x}, \mathcal{D}_\epsilon)} dx. \quad (12)$$

This integral is specifically over a 2D region such that if the dimension of the orbit is smaller than two, this integral

evaluates to zero. Since $\rho(\mathbf{x}, \mathcal{D}_\epsilon)$ is invariant under the map, $1/\rho(\mathbf{x}, \mathcal{D}_\epsilon)$ is also invariant. Intuitively, the integral bounds measure the size of the intersection of X and \mathcal{D}_ϵ and the value of $1/\rho(\mathbf{x}, \mathcal{D}_\epsilon)$ measures how many copies of this intersection are used to create the orbit X .

By taking the union of all orbits from \mathcal{D}_ϵ , which is precisely the ϵ -fattened exceptional set E_ϵ , the area of the ϵ -fattened exceptional set is

$$\mathcal{L}^2(E_\epsilon) = \int_{\mathcal{D}_\epsilon} \frac{1}{\rho(\mathbf{x}, \mathcal{D}_\epsilon)} dx. \quad (13)$$

The intersection of \mathcal{D} and any individual orbit X is no longer needed as the intersection of all orbits from \mathcal{D}_ϵ with itself is simply \mathcal{D}_ϵ . In this way, the integral evaluates the expected value of the number of copies of \mathcal{D}_ϵ needed to get the total area and multiplies this by the area of \mathcal{D}_ϵ to get the total area.

In the limit of small ϵ , the region of integration approaches \mathcal{D} which is 1D. With this in mind, the integral in Eq. (13) can be approximated by a 1D integral over \mathcal{D} multiplied by width ϵ ,

$$\mathcal{L}^2(\bar{E}) \approx \lim_{\epsilon \rightarrow 0} \int_{\mathcal{D}_a} \frac{\epsilon}{\rho(\mathbf{x}, \mathcal{D}_\epsilon)} dx + \lim_{\epsilon \rightarrow 0} \int_{\mathcal{D}_b} \frac{\epsilon}{\rho(\mathbf{x}, \mathcal{D}_\epsilon)} dx, \quad (14)$$

where \mathcal{D}_a and \mathcal{D}_b are the two sides of the cutting line as shown in Fig. 6, each side is assigned the dynamics of one of the two overlapping atoms, P_a or P_b . The value of $\lim_{\epsilon \rightarrow 0} \epsilon/\rho(\mathbf{x}, \mathcal{D}_\epsilon)$ is an abstraction of how much of the orbit is spread away from the measurement set. Note that for periodic orbits, $\lim_{\epsilon \rightarrow 0} \rho(\mathbf{x}, \mathcal{D}_\epsilon) \neq 0$ such that $\lim_{\epsilon \rightarrow 0} \epsilon/\rho(\mathbf{x}, \mathcal{D}_\epsilon)$ is always 0. To arrive at Φ_∞ , this measure simply needs to be normalized by the area of domain. Since the $\epsilon \rightarrow 0$ limit is not directly computable, an appropriately small ϵ can be used to approximate $\mathcal{L}^2(\bar{E})$. More thorough derivations of Eqs. (12)–(14) as well as additional intuition regarding density and area are given in Appendix B.

The parametrization in Fig. 5 provides a framework for applying the integral expression for the area of the exceptional

set given in Eq. (14), such that this integral becomes

$$\mathcal{L}^2(\bar{E}) \approx \lim_{\varepsilon \rightarrow 0} \int_0^{4\pi} \frac{\varepsilon}{\rho(\mathbf{s}(\theta), \mathcal{D}_\varepsilon)} d\theta. \quad (15)$$

The previously used metric for the fractional coverage of the exceptional set was $\Phi_\infty = \mathcal{L}^2(\bar{E})/2\pi$, where 2π is the area of the unit hemispherical shell [10,11,20]. The density along the cutting lines $\rho(\mathbf{s}(\theta), \mathcal{D}_\varepsilon)$ is, in general, not continuous, so it is convenient to recast the integral as an average (denoted with angled brackets) of the integrand over θ to evaluate Φ_∞ ,

$$\Phi_\infty \approx \frac{1}{2\pi} \lim_{\varepsilon \rightarrow 0} \int_0^{4\pi} \frac{\varepsilon}{\rho(\mathbf{s}(\theta), \mathcal{D}_\varepsilon)} d\theta \quad (16)$$

$$\approx \lim_{\varepsilon \rightarrow 0} \left\langle \frac{2\varepsilon}{\rho(\mathbf{s}(\theta), \mathcal{D}_\varepsilon)} \right\rangle. \quad (17)$$

From the return plots in Sec. IV, $\rho(\mathbf{s}(\theta), \mathcal{D}_\varepsilon)$ can be approximated by integrating (summing) any corresponding column (or row) in the return plot, which reflects the total return to (or from) any part of the cutting line, i.e., summing across row j approximates $\rho(\mathbf{s}(\theta \in [\theta_i, \theta_i + \Delta\theta]), \mathcal{D}_\varepsilon) \approx \sum_j \rho_{ij}$ and the fractional coverage would be $\Phi_\infty \approx \lim_{\varepsilon \rightarrow 0} (2\varepsilon (\sum_i \rho_{ij})^{-1})_j$. For instance, the $(57^\circ, 32.75^\circ)$ protocol has a value of $\Phi_\infty \approx 0.41$ using direct measurement [20] and $\Phi_\infty \approx 0.41$ using Eq. 17 with $\varepsilon = 10^{-5}$, $\delta = 10^{-9}$, $N = 2 \times 10^6$ iterations, averaged across 2×10^6 different positions along the cutting lines.

VI. CONCLUSIONS

We have described two different representations of the internal dynamics in the exceptional set of the hemispherical PWI by creating an invariant measure, ρ , which we refer to as *cutting line density*. The first representation is easily shown everywhere throughout the domain through a color coding that is generated by interactions with the two different cutting line discontinuities in the PWI. The second is the natural extension of this color coding, treating the cutting lines as a continuum, and measuring the return dynamics at each point along them to create a *return plot*. Although the color coding can show a snapshot of the dynamics everywhere in the exceptional set, the return plot shows only dynamics near the original cutting lines. Due to the nature of the construction of the PWI exceptional set, the dynamics at the cutting line are representative of the dynamics everywhere in the exceptional set.

These two representations of the cutting line density, color coding and return plots, indicate that more cuts do not indicate more mixing, rather an even distribution of cuts is required for mixing and variation in cutting line densities actually indicates barriers to mixing. Additionally, sufficient variation in the return dynamics of the cutting line is required to generate mixing. Without this variation, no amount of cuts can produce mixing, even if the domain is cut with an equal density of cuts. Ultimately, the existence and detection of confined mixing regions in the spherical tumbler mixer is aided by understanding the role of cutting and shuffling in creating trapped orbits. Large, invariant regions in the PWI, such as the arrowheads in the $(45^\circ, 45^\circ)$ protocol, as shown

in Fig. 2(a), are essentially extensions of the nonmixing cells contained in them.

The density of cutting lines in the hemispherical shell PWI can be used as an indicator of nonergodic behavior within the exceptional set. Although the fractional coverage of the exceptional set, Φ_∞ , is correlated with the degree of mixing [10,11], the existence of mixing within the exceptional set is entirely dependent on the dynamics within it. The cutting line density-based coloring techniques described in Sec. IIIB can be used in most situations to rule out (but not to definitively show) ergodicity within an exceptional set with minimal effort.

Return plots described in Sec. IV essentially project 2D orbits into a 1D space (the parameterized cutting lines), which necessarily exists for this PWI since the 1D set of cutting lines \mathcal{D} sweeps out and generates the exceptional set. Separating out orbits using the return plot can reveal nonmixing regions within the exceptional set as well as rule out ergodicity. A measurement of the white space (disconnectedness) of the return plot or the variation of the local densities within the plot, combined with the fractional coverage of the exceptional set, indicates not only how much of the domain can mix, but also accounts for how mixing within this space is occurring. Return plots can also be used to determine cutting line densities from which Φ_∞ can be approximated. Consequently, return plots enable relevant mixing information to be contained in a single representation that expresses both the size and connectivity of the exceptional set.

Just as cells only interchange with other cells within their periodic orbit [11,23,25,28,32], inside the exceptional set points only mix with one another if they are connected by at least one orbit. The colored exceptional set and the return plot are both visual representations of the connectedness of all the orbits within the exceptional set. The color-based classification of the exceptional set allows easy identification of invariant subsets, while the return plot provides a detailed deconstruction of where these subsets appear in the exceptional set. The return plot image is similar to that of an adjacency matrix in a weighted graph system if such a graph had infinitely many nodes [53].

The methods presented in this paper to examine mixing, connectedness, and barriers to both mixing and connectedness can be immediately applied to any 2D, finite domain, almost everywhere invertible PWI (especially different geometries [6]) and even PWI systems of other dimensions (although a 2D return plot may not always be possible). Notably, these methods could be easily applied to simpler PWIs such as a digital filter (a square geometry) [1,2]. Apart from the change in geometry, little changes in the analysis. The ability to construct and measure the connectivity of the exceptional set for any such system provides an effective means of evaluating the mixing, and the coloring methods can be applied to immediately identify the existence of invariant subsets by inspection.

ACKNOWLEDGMENTS

We gratefully acknowledge Lachlan D. Smith and Christian Gorski for insightful discussions. This research was based

upon work supported by the National Science Foundation under Grant No. CMMI-1435065.

APPENDIX A: DISCUSSION OF ERROR

There are four main sources of error in the methods presented here: error due to numerical precision, error due to δ -placement away from cutting lines (for the return plot and measurements of ρ), error due to ε -fattened cutting lines, and error due to the number of iterations, N , being finite.

Numerical precision affects all numerically advanced maps such as the PWI studied here. Error from finite precision calculations manifest as accumulated errors in position, or drift, that scales like N times the machine epsilon (in this case, roughly 2×10^{-16}). Better precision in position is not possible using double-precision data types since this error is scaled by the position of the tracked point, which is often $O(1)$. These types of numerical errors could influence the underlying dynamics of the map by creating spurious new periodic points in the numerical system or destroying actual periodic orbits through drift. The scale of these types of errors is mostly irrelevant to the study here (other error effects are much larger in comparison) because the accumulated drift after 2×10^6 iterations is only about 4×10^{-10} .

δ -placement away from the cutting lines is introduced to ensure that the aforementioned numerical drift does not cause a calculated point to incorrectly drift across the cutting line it is placed near. The value in this paper, $\delta = 10^{-9}$, is chosen to account for an $N \times$ (machine epsilon) $= 4 \times 10^{-10}$ amount of drift. The numerical drift can still cause errant cutting line crossings at later positions, but a positive δ guarantees that points do not immediately collapse onto one another (removing the sided nature of the cutting lines). If a point is advected under the map and then advected under the reverse map, this δ -placement should keep the original point and the iterated point in the same starting atom. The biggest consequence of δ -placement is the potential placement of points inside of a cell instead of inside the exceptional set. The fraction of points errantly placed inside of a cell is incalculable (it depends on the details of \bar{E}), but points placed within cells can only occupy thin circular orbits. These thin circular orbits contribute a δ -fattening to the exceptional set when taken together. In addition to this fattening, they can also exclude parts of the exceptional set; the exceptional set is only guaranteed to be arbitrarily close to the cutting lines, not necessarily arbitrarily close to a line δ from the cutting lines. Any invariant sets within the exceptional set not intersecting this line are not measured. Hence, the error from δ changes values of $\rho(x, A)$ by shifting x by δ . The balance of unmeasured invariant sets and a measured fattening is difficult to determine. However, the error due to δ -placement is potentially orders of magnitude smaller than the error incurred from ε -fattening of the cutting line.

The error due to ε -fattening introduces an additional, adjacent region, where points can “return” to the cutting line, and creates a border within the collection of periodic islands, O , where points in open cells appear to belong to the exceptional set. In the colored exceptional set [e.g., Fig. 1(g)], these regions appear as distinct invariant sets, which confuses any measurement of ergodicity (which would only have a single

invariant set within \bar{E}). In the return plot, this ε -fattening causes a vertical smearing effect since orbits can return to other parts of the cutting lines without actually reaching them. This smearing is not uniform, as evidenced by the horizontal white bands, discussed with respect to Fig. 8, which result from the overlap with the circular cells in O . For direct measurements of Φ , the fattening of the cutting lines includes the entirety of any cells of radius $\leq \varepsilon$ and the boundary of all cells with larger radii. For the density-based calculation of Φ , error is not primarily dependent on ε but rather on the placement of points δ from the cutting line and the spacing between points (an accurate measurement would require values of ρ continuously along the cutting line). As outlined in Appendix B, the size and shape of the region used for measurement of ρ is not as important as the intersection with invariant sets, and the δ -placed line is solely responsible for this. However, there is error in measuring the size of the fattened cutting line (since the line lies on a sphere, the cutting line is a great circle, but the ε -fattened border is not) and overlap between the two cutting lines. This error can be quantified exactly, although the exact overlap is not easily formulated.⁹ The approximate area of the fattened cutting line is $4\pi\varepsilon$ while the true area is $4\pi \sin(\varepsilon)$ minus the overlap of the two fattened cutting lines which is asymptotically $2\varepsilon^2[\csc(\alpha)\sec(\alpha) + \csc(\beta)\sec(\beta)]$ for small ε . The overlap is the dominant error source since $\varepsilon - \sin(\varepsilon) = O(\varepsilon^3)$ and the overlap is $O[\varepsilon^2(\frac{1}{\alpha} + \frac{1}{\beta})]$ accounting for the \csc terms which are larger when α or β are small. The error due to ε changes the value of $\rho(x, A)$ by introducing overlap at the cutting line intersection (double counting) but is still small, $O(\varepsilon^2)$, for most of the (α, β) protocol space.

Error due to finite iterations is subtle since it depends sensitively on the irrationality of measured quantities. Irrationality here concerns rotations and measurements that are not “close” to rational numbers that are “early” in Cantor’s ordering of the rationals. For example, a cell that rotates internally by an amount that is nearly $1/3$ will appear as a triangle for much longer than a cell that rotates internally by an amount that is nearly $1/300$, hence the $1/300$ rotation may appear more irrational since more of the boundary of the cell is initially explored. Quantifying this irrationality is subjective and dependent on the values of ε and N . If, in the previous example, the rotation is orders of magnitude closer to $1/300$ for the second cell than $1/3$ of the first, which one appears more irrational may switch after some N . These are errors in the construction of $\bar{E}_{\varepsilon, N}$, but there are also errors in measuring ρ . If ρ at some point is truly irrational, then there is no sufficient N at which ρ can be measured exactly. In this case, the true value of ρ can be considered the average of a random variable, i.e., $\rho(x, A)$ is the probability of the orbit of x being in A , for which one can consider the standard error of the mean, which is $\sqrt{\rho(1-\rho)/N}$. The \sqrt{N} dependence would be a dominant effect if not for the influence of the actual value of ρ , which is small for nonperiodic orbits and small ε . The error

⁹The exact value of this overlap is $4 \int_0^{\frac{\pi}{2}} 4 - 4\sqrt{1 - \varepsilon^2 \csc(\beta + \phi)^2} - 4\sqrt{1 - \varepsilon^2 \csc(\alpha + \phi)^2} d\phi$ when $\varepsilon \leq \min(\sin(\alpha), \sin(\beta))$.

due to finite N changes the value of $\rho(\mathbf{x}, A)$ by introducing error in the actual measurement of ρ rather than an error in orbit (δ) or orbit intersection with the overlapped region (ε). Additionally, the error in measurements of ρ as a result of using a finite number of iterations manifests itself as a sparse return plot, and, as ε is decreased, the number of iterations required to offset this sparsity increases.

With all this said, the dominant error term is unclear since the various nonnumerical errors depend on the geometry of the exceptional set which varies continuously through the (α, β) protocol space. Nevertheless, our experience in considering a range of protocols with the N , δ , and ε values used here, gives us confidence that the errors in the current analysis are inconsequential to the conclusions about invariant mixing regions within the exceptional set.

APPENDIX B: EXTRACTING THE MEASURE OF THE EXCEPTIONAL SET FROM CUTTING LINE DENSITY MEASUREMENTS

Previous papers [10,11,20] have examined the correlation between mixing and the size of the exceptional set (the fat fractal where mixing is possible). It can be shown that the size of the exceptional set is related to the measured density of cutting lines along the initial cutting lines. Intuitively, density can increase only if cutting lines are not spreading to other parts of the hemisphere, but instead stack on themselves such that $1/\rho$ is, abstractly, a measure of how cutting line “area” is spread through the hemisphere. Since density is easy to measure for a single point in a set, using density to approximate the measure of the exceptional set is useful.

For a measure-preserving map $M : S \rightarrow S$, define the density of the orbit from \mathbf{x} in set A as in Eq. (8), and define the closure of a set C as $\bar{C} = \lim_{r \rightarrow 0} B_r(C)$ such that the following are equivalent:

$$\begin{aligned} \rho(\mathbf{x}, A) &= \lim_{r \rightarrow 0} F(\mathbf{x}, B_r(A)) & (8) \\ &= F(\mathbf{x}, \bar{A}) \\ &= \lim_{n \rightarrow \infty} \frac{\text{No.}\{M^i(\mathbf{x}) \in \bar{A} : 0 \leq i < n\}}{n} \\ &= \lim_{n \rightarrow \infty} \frac{\text{No.}\{M^i(\mathbf{x}) \in \bar{A} : -n \leq i < n\}}{2n}, & (B1) \end{aligned}$$

where “No.” is the counting measure and ρ measures the fraction of iterates that an orbit spends in \bar{A} . The forward iterates are sufficient for the map in this paper, but the backward iterates can be included as shown previously [10,20]. To be explicit, the measurable space here is (S, \mathcal{E}) where S is the domain and \mathcal{E} is the invariant σ -algebra with respect to M . Then ρ is the conditional probability of A with respect to the invariant σ -algebra (i.e., the possible combinations of minimally invariant sets under the map). That is, if $\mathbf{x} \in X$ where X is a minimally invariant set, $X \in \mathcal{E}$, such as an orbit of the map, $X = \bigcup_{n=-\infty}^{\infty} M^n(\mathbf{x})$,

$$\rho(\mathbf{x}, A) = P_{\mathbf{x}}(A|X) = \frac{P_{\mathbf{x}}(A \cap X)}{P_{\mathbf{x}}(X)}, \quad (B2)$$

where $P_{\mathbf{x}}(A|X)$ is the conditional probability measure such that given a random iterate of \mathbf{x} under the map, $P_{\mathbf{x}}(A|X)$

is the probability that this iterate is in A . This probability measure only makes sense if $P_{\mathbf{x}}(X) \neq 0$, so we define the symbol $P_{\mathbf{x}}(\cdot)$ with respect to \mathbf{x} such that $P_{\mathbf{x}}(X)$, given X is the orbit of \mathbf{x} , is always 1. In an inexact way, we extend this definition to measures other than $P_{\mathbf{x}}(\cdot)$ and allow $\rho(\mathbf{x}, A) = \lim_{r \rightarrow 0} \mu[B_r(A \cap X)]/\mu[B_r(X)]$ where $\mu(\cdot)$ is a higher dimensional measure matching the dimension of the ball $B_r(\cdot)$ such that $\mu[B_r(\cdot)]$ is not zero or ∞ . We ultimately utilize the definition $\rho(\mathbf{x}, A) = \mu(A \cap X)/\mu(X)$ for intermediate steps, although this is inexact.

This density function is constant along orbits, i.e., it is measurable with respect to the invariant σ -algebra. Take $\mathbf{y} \in A$ in minimally invariant set Y such that $\rho(\mathbf{y}, A) \neq 0$, and therefore, since ρ is constant throughout orbits,

$$\mu(Y) = \frac{\mu(A \cap Y)}{\rho(\mathbf{y}, A)} \quad (B3)$$

$$= \frac{1}{\rho(\mathbf{y}, A)} \int_{A \cap Y} d\mu \quad (B4)$$

$$= \int_{A \cap Y} \frac{1}{\rho(\mathbf{y}, A)} d\mathbf{y}. \quad (B5)$$

Equation (B5) is the generalization of Eq. (12). Define the *streak* of set A , \mathcal{A} , under the map as the union of all orbits to or from A and that intersect A at least once (i.e., the orbit of the set A), \mathcal{E}_i , such that

$$\mathcal{A} = \bigcup_{\mathbf{x} \in A} \bigcup_{n=-\infty}^{\infty} M^n(\mathbf{x}) = \bigcup_{\mathbf{x} \in A} \mathcal{E}_i. \quad (B6)$$

\mathcal{A} is likewise the union of all minimally invariant sets, \mathcal{E}_i , from the invariant σ -algebra, \mathcal{E} , that intersect A , i.e., every $\mathcal{E}_i \cap A \neq \emptyset$ for all $\mathcal{E}_i \in \mathcal{E}$. Likewise, \mathcal{A} is the smallest collection of minimally invariant sets $\bigcup \mathcal{E}_i$ such that $A \subset \bigcup \mathcal{E}_i$. To be explicit, we choose minimally invariant sets from the invariant σ -algebra, \mathcal{E} , that intersect A and label these \mathcal{E}_i . There may be uncountably many minimally invariant sets, \mathcal{E}_i , that compose \mathcal{A} . Then, $\mu(\mathcal{A})$ can be written as

$$\mu(\mathcal{A}) = \sum_i \mu(\mathcal{E}_i). \quad (B7)$$

$\mu(\mathcal{A})$ is the “summation” of $\mu(\mathcal{E}_i)$ for all minimally invariant sets \mathcal{E}_i that make up \mathcal{A} . Since $\bigcup_i A \cap \mathcal{E}_i = A$ and all \mathcal{E}_i are mutually disjoint, this implies

$$\mu(\mathcal{A}) = \sum_i \mu(\mathcal{E}_i) \quad (B8)$$

$$= \sum_i \int_{A \cap \mathcal{E}_i} \frac{1}{\rho(\mathbf{y}, A)} d\mathbf{y} \quad (B9)$$

$$= \int_A \frac{1}{\rho(\mathbf{y}, A)} d\mathbf{y}. \quad (B10)$$

Equation (B10) is the generalization of Eq. (13). This intuitively says that the density information within A is the reciprocal of how A is advected through the domain. If density is “high” at some \mathbf{x} , the associated orbit is mostly contained within A , while if density is “low” at some \mathbf{x} , the associated orbit is mostly outside of A , spreading out the streak of A .

By referring to ρ as a density, we acknowledge that it is analogous to a physical mass density. Consider that A

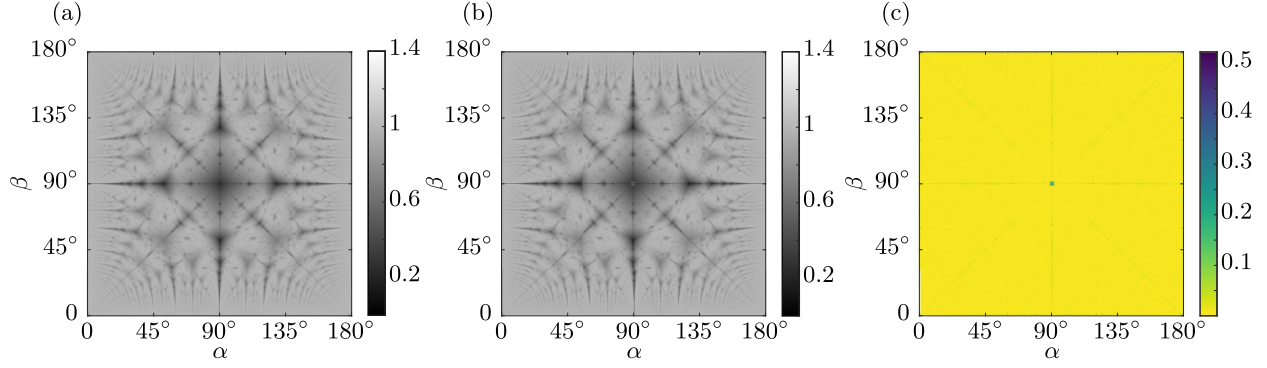


FIG. 11. A comparison of Φ values computed by (a) direct measurement of the exceptional set as in previous work [20] and (b) integrating the cutting line density. The absolute value of the difference between the two is shown in (c). For most protocols, the difference is near zero. Also note that $\Phi > 1$, which is impossible, for some protocols near the boundaries in (b) while values above 1 are impossible using the method in (a). Both methods use $\varepsilon = 1 \times 10^{-3}$ wide cutting lines for $N = 20000$ iterations.

contributes a certain amount of “mass” to \mathcal{A} with each iteration. This contribution is constant since the map M is measure preserving (all iterates of A have the same “mass”). This means that each individual orbit from a point in A is given the same amount of mass. The density is then the distribution of this mass, and, since density is necessarily constant along orbits, the area of each orbit can be calculated, i.e., $\rho = \text{mass}/\text{area}$. Note that although some orbits are given more mass with each iteration than others depending on how much the orbit overlaps the cutting line, this additional mass is taken into account by a proportional increase in ρ .

This can be applied directly to fattened cutting lines \mathcal{D}_ε such that

$$\rho(\mathbf{x}, \mathcal{D}_\varepsilon) = \frac{\mu(\mathcal{D}_\varepsilon \cap X)}{\mu(X)}. \quad (\text{B11})$$

Note that the streak of \mathcal{D}_ε is, for infinite iterations, the fattened exceptional set E_ε , and, for finitely many iteration N , this streak is the approximation to E_ε , $\tilde{E}_{\varepsilon,N}$. Note that $\lim_{\varepsilon \rightarrow 0}(\lim_{N \rightarrow \infty} \tilde{E}_{\varepsilon,N}) = \lim_{\varepsilon \rightarrow 0} E_\varepsilon = \bar{E}$ [20]. Using the 2D Lebesgue measure instead of $\mu(\cdot)$, the 2D measure of the approximate exceptional set is

$$\mathcal{L}^2(\tilde{E}_{\varepsilon,N}) \approx \mathcal{L}^2(E_\varepsilon) = \int_{\mathcal{D}_\varepsilon} \frac{1}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)} dx. \quad (\text{B12})$$

In practice, measuring $\rho(\mathbf{x}, \mathcal{D}_\varepsilon)$ cannot be done exactly, such that this integral is closer to value of $\tilde{E}_{\varepsilon,N}$ than E_ε .

In the limit of small ε , the fattened cutting line approaches a thin, 1D, cutting line such that $\mathcal{L}^2(\mathcal{D}_\varepsilon) \sim 2\varepsilon \mathcal{L}^1(\mathcal{D})$ (ε contribution from either side of the cutting line). Using this, the integral is, as $\varepsilon \rightarrow 0$, asymptotically,

$$\lim_{\varepsilon \rightarrow 0} \mathcal{L}^2(E_\varepsilon) = \lim_{\varepsilon \rightarrow 0} \int_{\mathcal{D}_\varepsilon} \frac{1}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)} dx \quad (\text{B13})$$

$$\approx \lim_{\varepsilon \rightarrow 0} \int_{\mathcal{D}} \frac{2\varepsilon}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)} dx. \quad (\text{B14})$$

Since we have used a multivalued cutting line in this paper, getting values for ρ along the cutting line is actually not possible. Instead, consider the boundary of the fattened cutting line $\partial \mathcal{D}_\varepsilon$ as $\varepsilon \rightarrow 0$, which is essentially the two sides of the cutting line, which we call \mathcal{D}_a and \mathcal{D}_b (the order does not matter here,

only that a and b are different sides). The integral should then be split into a sided cutting line (each side of width ε), such that

$$\mathcal{L}^2(\bar{E}) \approx \lim_{\varepsilon \rightarrow 0} \int_{\mathcal{D}} \frac{2\varepsilon}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)} dx \quad (\text{B15})$$

$$\approx \lim_{\varepsilon \rightarrow 0} \int_{\mathcal{D}_a} \frac{\varepsilon}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)} dx + \lim_{\varepsilon \rightarrow 0} \int_{\mathcal{D}_b} \frac{\varepsilon}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)} dx. \quad (\text{B16})$$

Since the parametrization in Sec. IV A accounts for the sidedness of the cutting lines already, this can be written succinctly as

$$\mathcal{L}^2(\bar{E}) = \lim_{\varepsilon \rightarrow 0} \int_0^{4\pi} \frac{\varepsilon}{\rho(\mathbf{s}(\theta), \mathcal{D}_\varepsilon)} d\theta. \quad (\text{B17})$$

If the regions of intersection of the cutting lines and a particular orbit are known, then the measure of the particular orbit can be computed,

$$\begin{aligned} \mathcal{L}^2(\bar{X}) &\approx \lim_{\varepsilon \rightarrow 0} \int_{\mathcal{D}_a \cap \bar{X}} \frac{\varepsilon}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)} dx \\ &+ \lim_{\varepsilon \rightarrow 0} \int_{\mathcal{D}_b \cap \bar{X}} \frac{\varepsilon}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)} dx \quad (\text{B18}) \\ &\approx [\mathcal{L}^1(\mathcal{D}_a \cap \bar{X}) + \mathcal{L}^1(\mathcal{D}_b \cap \bar{X})] \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\rho(\mathbf{x}, \mathcal{D}_\varepsilon)}. \end{aligned} \quad (\text{B19})$$

For many orbits, this is zero, but any orbit that is dense in a 2D set will evaluate as the measure of the 2D set.

To illustrate application of this method, the fractional coverage of the hemispherical shell, $\Phi = \mathcal{L}^2(\bar{E})/2\pi$, was computed using 20000 iterations of protocols within the $(\alpha, \beta) \in [0, 180^\circ] \times [0, 180^\circ]$ space using the method presented here [Fig. 11(b)] as well a previously used method [Fig. 11(a)] that directly measures a fattened exceptional set [20]. Both use $\varepsilon = 1 \times 10^{-3}$ fattened cutting lines. The direct measurement method samples the fattened cutting line at 1024 evenly spaced points across the shell. The cutting line density measurement uses $\delta = 1 \times 10^{-6}$ and measures $\rho(\mathbf{x}, \mathcal{D}_\varepsilon)$ at 2000 evenly spaced points along the cutting lines. For most protocols, the difference between the two, shown

in Fig. 11(c), is near zero, but has large deviations in some regions from the direct measurement method. The large difference near $(90^\circ, 90^\circ)$ could result from orbits being incomplete after 20 000 iterations. The direct measurement method is less sensitive to incomplete orbits as area accumulates nearly uniformly as iterations increase. The density measurement requires that many circuits be taken around an orbit in order for the value of ρ to converge near its true value. For most protocols, this happens quickly (i.e., within 1000 iterations), but near $(90^\circ, 90^\circ)$, convergence is very slow, which overestimates ρ and results in an incorrect estimate of Φ . A more concerning phenomenon occurs near the boundaries when approximating Φ using cutting line density and cutting lines begin to significantly overlap with one another (overlap is roughly $2\varepsilon^2[\csc(\alpha)\sec(\alpha) + \csc(\beta)\sec(\beta)]$), which results in an estimate for Φ that is larger than 1. The only remedy for this is to use a smaller value of ε and a larger number of iterates. Since cutting lines are so close, it becomes challenging to separate the two cutting lines from one another and the estimate breaks down.

The main diagonals in Fig. 11 contain protocols that have additional symmetries [11] and create cells that scale more gradually than protocols off the diagonal. For these protocols, the estimate of Φ using the cutting line density is more accurate than the direct measurement, which handles smaller cells sizes poorly, of which there are many along the diagonal.

APPENDIX C: ANALYTIC DENSITY EXAMPLE

For some trivial protocols, the density of cutting lines can be analytically defined. Take the example of the $(\phi, 0)$ protocol where $\phi/\pi \notin \mathbb{Q}$ (i.e., irrational). This is an example of irrational rotation that creates an exceptional set dense in S such that there are no cells, i.e., $O = \emptyset$ where O is the set of cells as defined in Sec. I, and every orbit is dense in an arc across the hemisphere in the x direction (rotation about the z axis does not change z values) with arc-length $l(z) = \pi\sqrt{1-z^2}$. For this example, we ignore the cutting line at the equator (due to $\beta = 0$, rotation about the second axis) since it generates no cutting lines. Each orbit is also ergodic such that the density $\rho(\mathbf{x}, \mathcal{D}_\varepsilon)$ for any point $\mathbf{x} = (x, y, z) \in S$, given cutting line width $\varepsilon > 0$ (on both sides), is the fraction of $l(z)$ that intersects \mathcal{D}_ε , which is roughly¹⁰ of length 2ε due to contributions from both separate sides of the cutting line. Therefore, using sided cutting lines in this analytic example, the cutting line density at any point $\mathbf{x} \in S$ is

$$\rho(\mathbf{x}, \mathcal{D}_\varepsilon) = \begin{cases} \frac{2\varepsilon}{l(z)} & \text{if } 2\varepsilon < l(z) \\ 1 & \text{if } 2\varepsilon \geq l(z) \end{cases}. \quad (\text{C1})$$

Note that in the limit of $\varepsilon \rightarrow 0$, $\rho \rightarrow 0$ everywhere except at the poles where ρ is always 1. Here, it is immediately clear that $\rho \sim \varepsilon$ because ρ can easily be thought of as a fraction of an ergodic orbit, and we have used this in the explicit definition of ρ . The normalized density, i.e., ρ/ε , is

simply $2/l(z)$. In this example, the arc-length of any orbit is closely connected to the normalized density. Density, although symmetric about $z = 0$, is different for each orbit indicating that all of the domain is confined along infinitely thin orbit curves that do not mix. Rewriting $l(z)$ so that it can be parameterized by arc-length along \mathcal{D} using $\phi \in [0, 2\pi)$ for a sided parametrization of the single cutting line gives

$$l(\phi) = \pi|\sin(\phi)|, \quad (\text{C2})$$

and the corresponding fractional coverage using Eq. (14) normalized by the area of the unit hemisphere,

$$\Phi = \frac{\mathcal{L}^2(\bar{E})}{2\pi} = \frac{1}{2\pi} \int_0^{2\pi} \frac{\varepsilon}{\rho} d\phi = \frac{1}{2\pi} \int_0^{2\pi} \frac{l(\phi)}{2} d\phi \quad (\text{C3})$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{\pi|\sin(\phi)|}{2} d\phi = 1, \quad (\text{C4})$$

which indicates the entire hemisphere is covered, as expected.

When $\phi/\pi \in \mathbb{Q}$ (i.e., a rational rotation) for the same $(\phi, 0)$ protocol, we expect that $\Phi = 0$ which is, indeed, the case. First, let $\phi/\pi = p/q$ such that q is the periodicity of the entire domain. In this case, there are exactly q cutting lines of zero width comprising \bar{E} . Then, for small $\varepsilon > 0$ (with some caveats near the poles that we will ignore since they disappear in the limit of $\varepsilon \rightarrow 0$), the cutting line density at any point $\mathbf{x} \in S$ is

$$\rho(\mathbf{x}, \mathcal{D}_\varepsilon) = \begin{cases} 1/q, & \text{if } \mathbf{x} \in E_\varepsilon \\ 0, & \text{if } \mathbf{x} \notin E_\varepsilon \end{cases}, \quad (\text{C5})$$

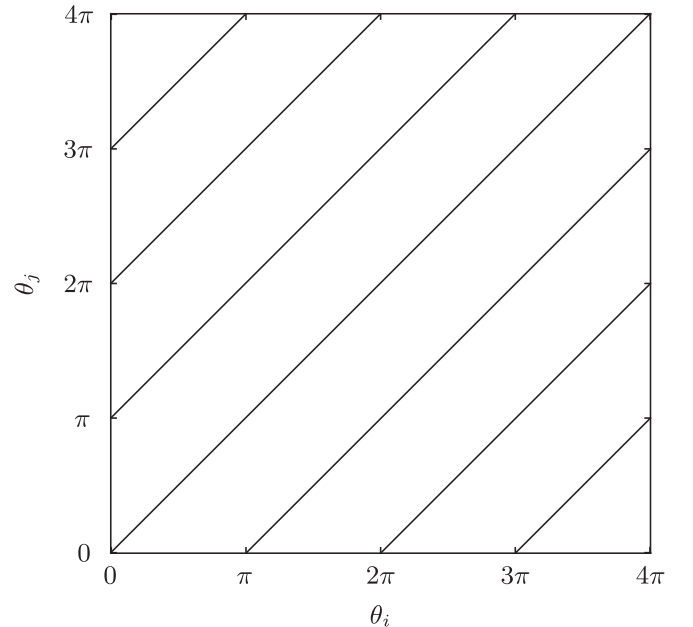


FIG. 12. Exact return plot for the $(90^\circ, 90^\circ)$ protocol. This is a protocol with only coincident cutting lines (i.e., a polygonal tiling of the sphere) resulting in a periodic interval exchange transform (IET) along the cutting line.

¹⁰There are some caveats with regard to the curvature of orbit arc versus the curvature of \mathcal{D} [e.g., the true length is $2\arcsin(\varepsilon)$], but these disappear as $\varepsilon \rightarrow 0$.

where we do not double count multivalued points that exactly return to the cutting line. Each orbit in \bar{E} can then be assigned the “arc-length” ε/ρ as before, which is simply εq in this case, and the fractional coverage is

$$\Phi = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\pi} \int_0^{2\pi} \varepsilon q d\phi = \lim_{\varepsilon \rightarrow 0} \varepsilon q = 0, \quad (\text{C6})$$

as expected.

The values of $\Phi = 1$ for irrational rotations and $\Phi = 0$ for rational rotations are completely unsurprising, but we use this trivial example to demonstrate the concept of density. The closed curves that are the orbits in the irrational case contribute no coverage individually, but together, they cover the hemisphere. For rotation about a single axis, neither the irrational rotation nor the rational rotation cause any mixing since there are no orbits that intersect \mathcal{D} in more than a finite number of points, and it seems that a condition for mixing in the exceptional set is the existence of orbits that contribute

nonzero coverage individually (which would necessarily intersect \mathcal{D} in some nonzero length region).

APPENDIX D: PURELY IET PROTOCOLS

A protocol where E forms a polygonal tiling (in this case, \bar{E} is equal to E since there are no limit points of E) [20], such as the $(90^\circ, 90^\circ)$ protocol, consists exclusively of coincident (overlapping or repeating) cutting lines and is essentially an interval exchange transform (IET). As such, it is a 1D cut-and-shuffle system [36,39–45,54–57] on the cutting lines, and the exact return plot can be constructed as shown in Fig. 12. For a protocol of this type, the return plot is constructed out of separate diagonal lines consisting of neighboring periodic points along the cutting lines. This results in a continuum of separate invariant sets (which are just collections of periodic points) that makes the entire exceptional set nonmixing and the return plot completely free from mixing blocks.

-
- [1] P. Ashwin, W. Chambers, and G. Petkov, *Int. J. Bifurcation Chaos* **07**, 2603 (1997).
- [2] P. Ashwin, *Phys. Lett. A* **232**, 409 (1997).
- [3] A. Goetz, *Ill. J. Math.* **44**, 465 (2000).
- [4] R. Adler, B. Kitchens, and C. Tresser, *Ergod. Theory Dyn. Syst.* **21**, 959 (2001).
- [5] A. Goetz and G. Poggiaspalla, *Nonlinearity* **17**, 1787 (2004).
- [6] L. D. Smith, P. B. Umbanhowar, R. M. Lueptow, and J. M. Ottino, *Phys. Rep.* **802**, 1 (2019).
- [7] R. Sturman, *Adv. Appl. Mech.* **45**, 51 (2012).
- [8] G. Juarez, R. M. Lueptow, J. M. Ottino, R. Sturman, and S. Wiggins, *EPL* **91**, 20003 (2010).
- [9] G. Juarez, I. C. Christov, J. M. Ottino, and R. M. Lueptow, *Chem. Eng. Sci.* **73**, 195 (2012).
- [10] P. P. Park, T. F. Lynn, P. B. Umbanhowar, J. M. Ottino, and R. M. Lueptow, *Phys. Rev. E* **95**, 042208 (2017).
- [11] L. D. Smith, P. P. Park, P. B. Umbanhowar, J. M. Ottino, and R. M. Lueptow, *Phys. Rev. E* **95**, 062210 (2017).
- [12] P. B. Umbanhowar, R. M. Lueptow, and J. M. Ottino, *AIChE J.* **59**, 3237 (2013).
- [13] S. W. Jones and H. Aref, *Phys. Fluids* **31**, 469 (1988).
- [14] L. D. Smith, M. Rudman, D. R. Lester, and G. Metcalfe, *Chaos* **26**, 023113 (2016).
- [15] S. E. Boyer and D. Elliott, *AAPG Bull.* **66**, 1196 (1982).
- [16] T. H. Bell, *Nature (London)* **304**, 493 (1983).
- [17] R. W. Butler, *J. Struct. Geol.* **4**, 239 (1982).
- [18] P. P. Park, Ph.D. thesis, Department of Engineering Science and Applied Mathematics, Northwestern University (2016).
- [19] B. Kahng, *Chaos* **19**, 023115 (2009).
- [20] T. F. Lynn, L. D. Smith, J. M. Ottino, P. B. Umbanhowar, and R. M. Lueptow, *Phys. Rev. E* **99**, 032204 (2019).
- [21] A. Goetz, *Nonlinearity* **14**, 205 (2001).
- [22] A. Scott, C. Holmes, and G. Milburn, *Physica D (Amsterdam)* **155**, 34 (2001).
- [23] A. Scott, *Physica D (Amsterdam)* **181**, 45 (2003).
- [24] P. Ashwin, M. Nicol, and N. Kirkby, *Physica A (Amsterdam)* **310**, 347 (2002).
- [25] A. Goetz, in *Fractals in Graz 2001* (Birkhäuser, Basel, 2003), pp. 135–144.
- [26] Z. Zaman, M. Yu, P. P. Park, J. M. Ottino, R. M. Lueptow, and P. B. Umbanhowar, *Nat. Commun.* **9**, 3122 (2018).
- [27] M. Yu, P. B. Umbanhowar, J. M. Ottino, and R. M. Lueptow, *Phys. Rev. E* **99**, 062905 (2019).
- [28] P. P. Park, P. B. Umbanhowar, J. M. Ottino, and R. M. Lueptow, *Chaos* **26**, 073115 (2016).
- [29] J. P. Snyder, *Map Projections—A Working Manual* (U.S. GPO, Washington, DC, 1987).
- [30] J. Farmer, E. Ott, and J. A. Yorke, *Physica D (Amsterdam)* **7**, 153 (1983).
- [31] D. K. Umberger and J. D. Farmer, *Phys. Rev. Lett.* **55**, 661 (1985).
- [32] H. Bruin, A. Lambert, G. Poggiaspalla, and S. Vaienti, *Chaos* **13**, 558 (2003).
- [33] K. D. Trinh, in *Mathematics for Industry* (Springer Japan, Tokyo, 2014), pp. 297–309.
- [34] P. V. Danckwerts, *Appl. Sci. Res.* **3**, 279 (1952).
- [35] B. Kahng, *Ergod. Theory Dyn. Syst.* **22**, 483 (2002).
- [36] A. Avila and G. Forni, *Ann. Math.* **165**, 637 (2007).
- [37] P. Ashwin and A. Goetz, *SIAM J. Appl. Dyn. Syst.* **4**, 437 (2005).
- [38] P. Ashwin, A. Goetz, P. Peres, and A. Rodrigues, *Ergod. Theory Dyn. Syst.*, doi: 10.1017/etds.2018.112.
- [39] M. Yu, P. B. Umbanhowar, J. M. Ottino, and R. M. Lueptow, *Int. J. Bifurcation Chaos* **26**, 1630038 (2016).
- [40] C. F. Novak, *J. Mod. Dyn.* **3**, 379 (2009).
- [41] A. Katok, *Isr. J. Math.* **35**, 301 (1980).
- [42] H. Hmili, *Discret. Contin. Dyn. S.* **27**, 1079 (2010).
- [43] M. Keane, *Isr. J. Math.* **26**, 188 (1977).
- [44] H. Masur, *Ann. Math.* **115**, 169 (1982).
- [45] M. Wang and I. C. Christov, *Phys. Rev. E* **98**, 022221 (2018).
- [46] C. Grebogi, S. W. McDonald, E. Ott, and J. A. Yorke, *Phys. Lett. A* **110**, 1 (1985).
- [47] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An Introduction to Dynamical Systems* (Springer, New York, NY, 1996).
- [48] E. W. Weisstein, Natural invariant, from MathWorld—A Wolfram Web Resource, <http://mathworld.wolfram.com/NaturalInvariant.html>.

- [49] E. Ott, *Chaos in Dynamical Systems*, 2nd ed. (Cambridge University Press, Cambridge, 2002).
- [50] N. Marwan, J. F. Donges, Y. Zou, R. V. Donner, and J. Kurths, *Phys. Lett. A* **373**, 4246 (2009).
- [51] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, *EPL* **4**, 973 (1987).
- [52] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, *Phys. Rep.* **438**, 237 (2007).
- [53] S. S. Ray, *Graph Theory with Algorithms and Its Applications* (Springer India, Delhi, 2013).
- [54] M. K. Krotter, I. C. Christov, J. M. Ottino, and R. M. Lueptow, *Int. J. Bifurcation Chaos* **22**, 1230041 (2012).
- [55] M. Keane, *Math. Z.* **141**, 25 (1975).
- [56] W. A. Veech, *J. Anal. Math.* **33**, 222 (1978).
- [57] M. Viana, *Rev. Mat. Comput.* **19**, 7 (2006).
- [58] E. W. Dijkstra, *Numer. Math.* **1**, 269 (1959).