Exact decomposition of homoclinic orbit actions in chaotic systems: Information reduction

Jizhou Li and Steven Tomsovic
Department of Physics and Astronomy, Washington State University, Pullman, Washington 99164-2814, USA

(Received 9 July 2018; revised manuscript received 6 February 2019; published 11 March 2019)

Homoclinic and heteroclinic orbits provide a skeleton of the full dynamics of a chaotic dynamical system and are the foundation of semiclassical sums for quantum wave packets, coherent states, and transport quantities. Here, the homoclinic orbits are organized according to the complexity of their phase-space excursions, and exact relations are derived expressing the relative classical actions of complicated orbits as linear combinations of those with simpler excursions plus phase-space cell areas bounded by stable and unstable manifolds. The total number of homoclinic orbits increases exponentially with excursion complexity, and the corresponding cell areas decrease exponentially in size as well. With the specification of a desired precision, the exponentially proliferating set of homoclinic orbit actions is expressible by a slower-than-exponentially increasing set of cell areas, which may present a means for developing greatly simplified semiclassical formulas.

DOI: 10.1103/PhysRevE.99.032212

I. INTRODUCTION

Specific sets of rare classically chaotic orbits are central ingredients for sum rules in classical and quantum systems [1]. Classical sum rules over unstable periodic orbits describe various entropies, Lyapunov exponents, escape rates, and the uniformity principle [2]. Gutzwiller's trace formula [3] for quantum spectra is over unstable periodic orbits, closed orbit theory of atomic spectra [4,5] gives the absorption spectrum close to the ionization threshold of atoms placed in magnetic fields, and heteroclinic (homoclinic) orbits arising from intersections between the stable and unstable manifolds of different (same) hyperbolic trajectories describe quantum transport between initial and final localized wave packets [6].

It is often the case that the nonlinear flows of phase-space densities are completely captured by the stable and unstable manifolds of one or just a few short periodic orbits, hence also by the homoclinic and heteroclinic orbits that arise from intersections between these manifolds. These orbits can thus play the important role of providing a "skeleton" of transport for the system. It is not a unique choice, but each choice provides the same information. For example, an unstable periodic orbit gives rise to an infinity of homoclinic orbits, but it is also true that families of periodic orbits of arbitrary lengths accumulate on some point along every homoclinic orbit [7–9], and the periodic orbit points can be viewed as being topologically forced by the homoclinic point on which a particular sequence accumulates [10,11].

Two problems are immediately apparent. The first is the particular importance of having accurate evaluations of classical actions because these quantities are divided by $\hbar$ and play the role of phase factors for the interferences between terms, and their remainder after taking the modulus with respect to 2π must be $\ll 2\pi$. A straightforward calculation would proceed with the numerical construction of the actions, which would be plagued by the sensitive dependence on initial conditions for long orbits. An alternative method has been developed by the authors [11,12]. That scheme converts the calculation of unstable periodic orbit actions into the evaluation of homoclinic orbit action differences. The homoclinic orbit actions can then be stably obtained as phase-space areas via the MacKay-Meiss-Percival principle [13,14], or directly from the stable constructions of homoclinic orbits [9,15–18]. Beside the action functions, another quantity of the periodic orbits, namely their stability exponents, also plays the crucial role of the prefactor in Gutzwiller's trace formula. In Sec. IV D, a relation Eq. (29) is introduced that determines the stability exponents of periodic orbits from ratios between areas bounded by stable and unstable manifolds, or equivalently, distribution of homoclinic points on the manifolds. Therefore, both the action and the stability exponent of periodic orbits can be calculated from the knowledge of homoclinic orbits, without the numerical construction of periodic orbits themselves.

The second problem is more fundamental. Namely, the total number of periodic orbits increases exponentially with increasing period and for the homoclinic orbits with increasingly complicated excursions. This is a reflection of the nonvanishing rate of information entropy production associated with chaotic dynamics, which in an algorithmic complexity sense has been proven equivalent to the Kolmogorov-Sinai entropy [19–23], and hence the Lyapunov exponents via Pesin's theorem [24,25]. On the other hand, entropies introduced for quantum systems [26–28] vanish due to the nonzero size of $\hbar$, if these systems are isolated, bounded, and not undergoing a measurement process. This gives one the intuitive notion and hope that there must be a means to escape the exponential proliferation problem of semiclassical sum rules.

Therefore, a scheme to replace classical and semiclassical sum rules that from the outset clearly have vanishing information entropy content is highly desirable [29]. The pseudo-orbits of the cycle expansion [1,30,31], the primitive orbits of Bogomolny's surface of section method [32], and multiplicative semiclassical propagator [33] were steps in this direction. Building on the methods of [11], we develop exact relations

for the decomposition of homoclinic orbit relative actions with complicated excursions in terms of multiples of the two primary ones and sets of phase-space areas. Accounting for an error tolerance determined by $\hbar$ reduces the exponentially proliferating set of homoclinic orbit actions to combinations of an input set (i.e., phase-space cell areas) that increase more slowly than exponentially (i.e., algebraically) with time, thus resolving the conflict between the entropies of classical and quantum chaotic systems, and directly linking $\hbar$ to the boundary between surviving and nonsurviving information in quantum mechanics.

This paper is organized as follows. Section II introduces the basic concept of homoclinic tangle. Section III introduces the relative action functions between homoclinic orbit pairs. Section IV reviews the concepts of winding number and transition time of homoclinic orbits, and introduces a hierarchical ordering of homoclinic points in terms of their winding numbers. Organizing the homoclinic points using the winding numbers, we identify an asymptotic scaling relation between families of homoclinic points, which puts strong constraints on the distribution of homoclinic points along the manifolds. Section V gives two central results of this paper. The first one (Sec. V B) is an exact formula for the complete expansion of homoclinic orbit actions in terms of primary homoclinic orbits and phase-space cell areas bounded by the manifolds. The second one (Sec. V C) is the demonstration that a coarse-grained scale, determined by $\hbar$, allows for an approximation that eliminates exponentially small areas from the complete expansion, which gives an approximate action expansion that requires a subset of cell areas growing subexponentially.

II. BASIC CONCEPTS

Consider a two-degree-of-freedom autonomous Hamiltonian system. With energy conservation and applying the standard Poincaré surface of section technique [34], the continuous flow leads to a discrete area-preserving map M on the two-dimensional phase space (q, p). Assume the existence of a hyperbolic fixed point $x = (q_x, p_x)$ under M: $M(x) = x$. Associated with it are the one-dimensional stable $[S(x)]$ and unstable $[U(x)]$ manifolds, which are the collections of phase-space points that approach x under successive forward and inverse iterations of M, respectively. Typically, $S(x)$ and $U(x)$ intersect infinitely many times and form a complicated pattern named a *homoclinic tangle* [34–36], as partially illustrated in Fig. 1. This figure demonstrates the simplest but generic type of homoclinic tangle, a "Smale horseshoe" [37,38], which results from the exponential stretching along $U(x)$, compressing along $S(x)$, and eventually a binary folding to create mixing dynamics. Refer to Appendix B for a detailed introduction of the Smale horseshoe. The area-preserving Hénon map [39] shown by Eq. (B3) with parameter $a = 10$ is used to generate this figure, along with all forthcoming numerical implementations in this article.

The main objects of study in this article, the *homoclinic orbits*, arise from intersections between $S(x)$ and $U(x)$. These are the orbits asymptotic to x under both forward and inverse iterations of M. For instance, the point h_0 in Fig. 1 is a homoclinic intersection between the manifolds, and its orbit $\{\ldots, h_{-1}, h_0, h_1, \ldots\}$ approaches x under both forward and inverse iterations. In spite of the infinity of homoclinic orbits arising from the pattern in Fig. 1, for the most part only two of them, $\{h_0\}$ and $\{g_0\}$, have a fundamental importance. They have the special property that the segments $U[x, h_0]$ and $S[h_0, x]$ only intersect at x and h_0. Consequently, the loop $US[x, h_0] \equiv U[x, h_0] + S[h_0, x]$ is a single loop, so the orbit $\{h_0\}$ "circles" around the loop only once. The same is true for $\{g_0\}$. A "winding number" of 1 can thus be associated with both $\{h_0\}$ and $\{g_0\}$, and they are commonly referred as the *primary homoclinic orbits*. All other orbits have winding numbers greater than 1. To be shown later, their classical actions can be built by the two primary orbit actions and certain sets of phase-space areas bounded by $S(x)$ and $U(x)$. More details about the winding numbers will be introduced in Sec. IV.

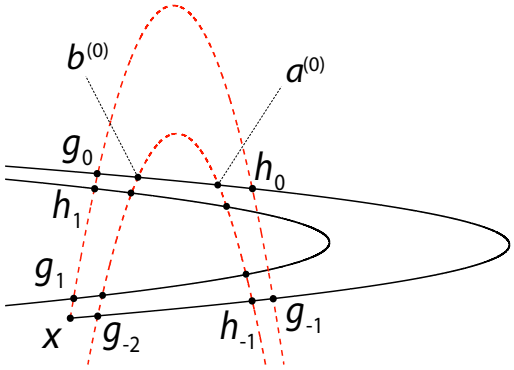

FIG. 1. Horseshoe-shaped homoclinic tangle formed by $S(x)$ (red dashed curve) and $U(x)$ (black solid curve), with two primary homoclinic orbits $\{h_0\}$ and $\{g_0\}$. Notice that the $U(x)$ segments beyond g_0 and h_1 are simply connected and omitted from the figure for clarity, and the same for the $S(x)$ segments beyond g_{-1} and h_{-1}.

The topological structures of homoclinic tangles are well understood nowadays, and they provide a foundation for our analysis on the homoclinic orbit actions in later sections. With the help of certain generating Markov partitions identified from the homoclinic tangle (V_0 and V_1 in Fig. 15 in Appendix B), the nonwandering orbits of the system can be put into a one-to-one correspondence with bi-infinite strings of integers, i.e., the *symbolic dynamics* [40–43] of chaotic systems. For example, the hyperbolic fixed point x in Fig. 1 is labeled by the bi-infinite string $\overline{0}.\overline{0}$, where the overhead bar indicates infinite repetitions of the symbolic string underneath it, and the decimal point indicate the location of the current iteration. This symbolic code reflects the fact that x stays in V_0 under all forward and inverse iterations. The primary homoclinic points h_0 and g_0 are labeled by $h_0 \Rightarrow \overline{0}1.1\overline{0}$ and $g_0 \Rightarrow \overline{0}1.\overline{0}$, respectively. Other than the points on $\{h_0\}$ and $\{g_0\}$, all homoclinic points a of x must have a symbolic string of the form

$$a \Rightarrow \overline{0}1s_{-m}\cdots s_{-1}.s_0s_1\cdots s_n1\overline{0} = \overline{0}1\tilde{s}_-.\tilde{s}_+1\overline{0} \qquad (1)$$

along with all possible shifts of the decimal point, where each digit $s_i \in 0, 1$ $(-m \leqslant i \leqslant n)$. The substrings $\tilde{s}_- = s_{-m}\cdots s_{-1}$ and $\tilde{s}_+ = s_0s_1\cdots s_n$. The $\overline{0}$ on both ends means the orbit approaches the fixed point asymptotically. The orbit $\{a\}$ can then be represented by the same symbolic string:

$$\{a\} \Rightarrow \overline{0}1\tilde{s}_-\tilde{s}_+1\overline{0} \qquad (2)$$

with the decimal point removed, as compared to Eq. (1). The finite symbolic segment "$1\tilde{s}_-\tilde{s}_+1$" is often referred to as the *core* of the symbolic code of a, with its length referred to as the *core length*.

In the horseshoe map, besides the hyperbolic fixed point x, there is another hyperbolic fixed point with reflection, denoted by x'. This fixed point has symbolic code $x' \Rightarrow \overline{1}.\overline{1}$; i.e., it stays in V_1 under all forward and inverse iterations. Denote the stability exponents of x and x' by μ_0 and μ_1, respectively; i.e., the subscripts indicate the symbolic code. These two exponents are of special interest later.

We skip further detailed introduction here and refer the reader to excellent references such as [35,36,44], and to Appendices A and B for the concepts of trellises, symbolic dynamics, and for the definitions of notations adopted throughout this article. The symbolic dynamics will be the main language adapted to identify homoclinic orbits in this study. However, although well resolved [45], the assignment of symbolic codes to homoclinic points is still a nontrivial task in general. The readers are referred to Appendix C for a detailed assignment scheme. In the forthcoming contents, the symbolic codes of all homoclinic points are assumed known.

III. RELATIVE ACTIONS

The classical actions of homoclinic orbits are divergent as they come from the infinite sum over the generating functions associated with each iteration along the orbit. Hence, it is necessary to consider relative actions, which are finite. For any phase-space point $z_n = (q_n, p_n)$ and its image $M(z_n) = z_{n+1} = (q_{n+1}, p_{n+1})$, the mapping M can be viewed as a canonical transformation that maps z_n to z_{n+1} while preserving the symplectic area; therefore a *generating* (*action*) function $F(q_n, q_{n+1})$ can be associated with this transformation such that [13,14]

$$\begin{aligned} p_n &= -\partial F/\partial q_n, \\ p_{n+1} &= \partial F/\partial q_{n+1}. \end{aligned} \tag{3}$$

Despite the fact that F is a function of q_n and q_{n+1}, it is convenient to denote it as $F(z_n, z_{n+1})$. This should cause no confusion as long as it is kept in mind that it is the q variables of z_n and z_{n+1} that go into the expression of F. A special example is the generating function of the fixed point, $F(x, x)$, that maps x into itself under one iteration. For homoclinic orbits $\{h_0\}$, the *classical action* is the sum of generating functions between each step

$$\mathcal{F}_{\{h_0\}} \equiv \lim_{N\to\infty} \sum_{n=-N}^{N-1} F(h_n, h_{n+1}). \tag{4}$$

However, according to the MacKay-Meiss-Percival action principle [13,14], convergent relative actions can be obtained by comparing the classical actions of a homoclinic orbit pair:

$$\begin{aligned} \Delta\mathcal{F}_{\{h_0'\}\{h_0\}} &\equiv \lim_{N\to\infty} \sum_{n=-N}^{N-1} [F(h_n', h_{n+1}') - F(h_n, h_{n+1})] \\ &= \int_{U[h_0,h_0']} p\,dq + \int_{S[h_0',h_0]} p\,dq = \mathcal{A}^{\circ}_{US[h_0,h_0']}, \end{aligned} \tag{5}$$

where the $\circ$ superscript in the last term indicates that the area evaluated is interior to a path that forms a closed loop, and the subscript indicates the path $US[h_0, h_0'] = U[h_0, h_0'] + S[h_0', h_0]$. Such an action difference is referred to as the *relative action* between $\{h_0'\}$ and $\{h_0\}$. A special case of interest is the relative action between a homoclinic orbit $\{h_0\}$ and the fixed point itself $\{x\}$:

$$\begin{aligned} \Delta\mathcal{F}_{\{h_0\}\{x\}} &= \lim_{N\to\infty} \sum_{n=-N}^{N-1} [F(h_n, h_{n+1}) - F(x, x)] \\ &= \mathcal{A}^{\circ}_{US[x,h_0]}, \end{aligned} \tag{6}$$

which gives the action of $\{h_0\}$ relative to the fixed point orbit action, and is simply referred to as the relative action of $\{h_0\}$. An equivalent approach, which makes use of the information about the stable and unstable manifolds of hyperbolic fixed points to obtain convergent expressions of homoclinic and heteroclinic orbit actions as algebraic areas evaluated under these manifolds, was given by Tabacman in Ref. [46]. There, it was shown that the homoclinic and heteroclinic orbits can be calculated as critical values of certain action functions constructed from the generating function of the system and the local stable and unstable manifolds near the fixed points. However, our goal is to identify hidden relations between the homoclinic orbit actions without numerical constructions of the orbits themselves. As shown ahead, this requires information about the global stable and unstable manifolds.

A generalization of Eq. (6) applies to four arbitrary homoclinic orbits of x, namely $\{a_0\}$, $\{b_0\}$, $\{c_0\}$, and $\{d_0\}$. Expressing the relative actions of each of them using Eq. (6) and calculating the action difference between the following two pairs of orbits gives

$$\begin{aligned} &\left(\Delta\mathcal{F}_{\{a_0\}\{x\}} - \Delta\mathcal{F}_{\{b_0\}\{x\}}\right) - \left(\Delta\mathcal{F}_{\{c_0\}\{x\}} - \Delta\mathcal{F}_{\{d_0\}\{x\}}\right) \\ &\quad = \left(\mathcal{A}^{\circ}_{US[x,a_0]} - \mathcal{A}^{\circ}_{US[x,b_0]}\right) - \left(\mathcal{A}^{\circ}_{US[x,c_0]} - \mathcal{A}^{\circ}_{US[x,d_0]}\right) \\ &\quad = \mathcal{A}^{\circ}_{SUSU[a_0,c_0,d_0,b_0]}, \end{aligned} \tag{7}$$

where

$$\begin{aligned} \mathcal{A}^{\circ}_{SUSU[a_0,c_0,d_0,b_0]} \equiv &\int_{S[a_0,c_0]} p\,dq + \int_{U[c_0,d_0]} p\,dq \\ &+ \int_{S[d_0,b_0]} p\,dq + \int_{U[b_0,a_0]} p\,dq \end{aligned} \tag{8}$$

is the curvy parallelogram area bounded by alternating segments of $S(x)$ and $U(x)$ connecting the four homoclinic points.

IV. HIERARCHICAL STRUCTURE OF HOMOCLINIC POINTS

A. Winding numbers and transit times

The infinite set of homoclinic orbits can be put into a hierarchical structure, organized using a winding number [47,48] that characterizes the complexity of phase-space excursion of each individual orbit. The winding number of a homoclinic point h is defined to be the number of single loops (i.e., loops with no self-intersection) that the loop $US[x, h]$ can be decomposed into [48]. The primary homoclinic points h_0 and g_0 points in Fig. 1 are associated with orbits having winding

number 1, since both $US[x, h_0]$ and $US[x, g_0]$ are single loops. They form the complete first hierarchical family.

The nonprimary homoclinic points $a^{(0)}$ and $b^{(0)}$ in Fig. 1 are both associated with winding number 2, i.e., the loop $US[x, a^{(0)}] = US[x, h_0] + US[h_0, a^{(0)}]$, both of which are single loops, and similarly for $b^{(0)}$, $US[x, b^{(0)}] = US[x, g_0] + US[g_0, b^{(0)}]$. All points on a particular orbit are associated with the same winding number. Roughly speaking, a winding-n orbit "circles" the complex region n times from the infinite past to the infinite future, and therefore the winding number characterizes the complexity of its phase-space excursion. Figure 1 of Ref. [48] has a nice illustration.

Within each family, the orbits can be further organized by their *transit times* [35,36], which contains the length of the phase-space excursion of a homoclinic orbit. With the "open system" assumption, there are no homoclinic points on segments U'_n and S_n [for the definition of fundamental segments U_n, U'_n, S_n, and S'_n, see Eq. (A3)]. Therefore, any homoclinic point z_0 must arise from the intersection between some U_n and S'_m segments, with n and m being appropriate integers such that $z_0 \in U_n \cap S'_m$. The transit time of $\{z_0\}$, denoted by t, is defined as the difference in the indices of U_n and S'_m: $t = (n - m)$. Starting from $z_{-n} \in U_0 \cap S'_{m-n}$, and mapping t times, $M^t(z_{-n}) = z_{-m} \in U_{n-m} \cap S'_0$. Thus, t is the number of iterations needed to map the orbit from U_0 to S'_0. Note that, excluding the primary homoclinic orbits, $\{g_0\}$ and $\{h_0\}$, all homoclinic orbits have positive definite t since there are no intersections of U_n with S'_0 with negative integer n or 0; i.e., the first intersection of S'_0 is with U_1.

Since the mapping preserves the topology, $M^k(z_0) = z_k \in U_{n+k} \cap S'_{m+k}$, every orbit $\{z_0\}$ has one and only one point (which is z_0) on S'_m. Therefore, enumerating homoclinic points on S'_m is equivalent to enumerating all distinct homoclinic orbits in the trellis [see Eq. (A5) for the definition of a trellis]. In practice, it is convenient to choose $m = -1$. Equivalently, all homoclinic points on S'_{-1} with a maximum $t = n + 1$ are intersections with the trellis $T_{-1,n}$ ($=T_{-1,t-1}$). The total number of homoclinic orbits increases exponentially rapidly with the transit time. For example, in Fig. 2, U_0 intersects S'_{-1} at two points: $a^{(0)}$ and $b^{(0)}$. U_1 intersects S'_{-1} at four points $a^{(1)}$, $b^{(1)}$, $c^{(1)}$, and $d^{(1)}$. Furthermore, U_2 intersects S'_{-1} at eight points, where the four points $a^{(2)}$, $b^{(2)}$, $c^{(2)}$, and $d^{(2)}$ are winding-2, and the remaining four points, on the upper half of S'_{-1}, are not explicitly labeled and are winding-3. Including g_{-2} and h_{-1}, the total number of homoclinic points on S'_{-1} is exactly $2^{(t+1)}$.

B. Asymptotic accumulation of homoclinic points

Although homoclinic tangles create unimaginably complicated phase-space patterns, their behaviors are highly constrained by a few simple rules of Hamiltonian chaos; namely exponential compression and stretching occurs while preserving phase-space areas, and manifolds cannot intersect themselves or other manifolds of the same type. Therefore, locally near any homoclinic point, unstable (stable) manifolds form fine layers of near-parallel curves, with distances in between the curves scaling down exponentially rapidly as they get closer towards that point. As numerically demonstrated by Eq. (10) in Ref. [48], such asymptotic scaling relations

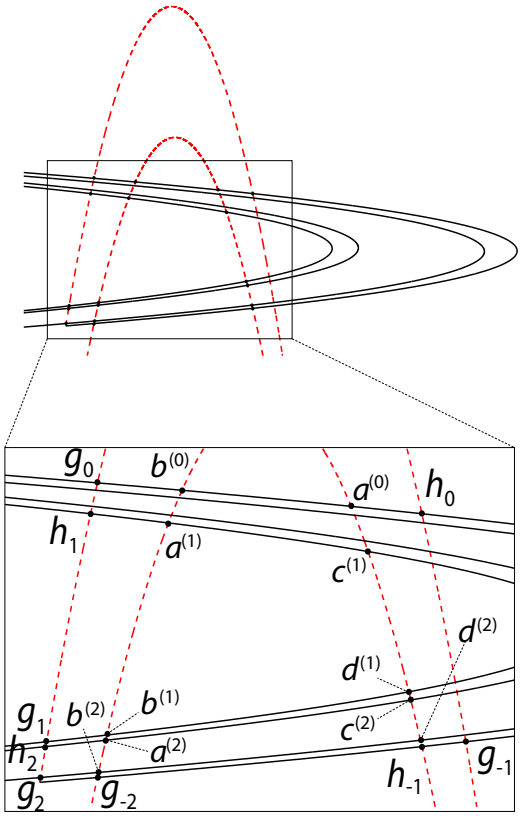

FIG. 2. Upper panel: Trellis $T_{-1,2}$. Lower panel (zoomed-in): The hierarchy of homoclinic points on S'_{-1} can be organized as the following: the winding-1 points g_{-2} and h_{-1} constitute the first-order family. The sequence of winding-2 points $a^{(n)}$ and $b^{(n)}$ ($n \geqslant 0$) form two second-order families that accumulate on g_{-2} asymptotically under Eqs. (9) and (10). Similarly on the right side of S'_{-1}, we have the winding-2 points $c^{(n)}$ and $d^{(n)}$ ($n \geqslant 1$) accumulating on h_{-1}, which form two second-order families as well. Consequently, three families of areas $[\mathcal{A}^{\circ}_{SUSU[g_{-2},b^{(n)},d^{(n)},h_{-1}]}]$, $[\mathcal{A}^{\circ}_{SUSU[b^{(n)},a^{(n)},c^{(n)},d^{(n)}]}]$, and $[\mathcal{A}^{\circ}_{SUSU[a^{(n)},b^{(n-1)},d^{(n-1)},c^{(n)}]}]$ accumulate on the bottom segment $U[g_{-2}, h_{-1}]$ under Eq. (16), with the same asymptotic exponent μ_0.

exist inside every family of homoclinic points. A concrete mathematical description of this phenomenon is given by Lemma 2 in Appendix B 3 of [49], which states "iterates of a curve intersecting the stable manifold approach the unstable manifold." Refer to Appendix D for a brief overview of the lemma.

The asymptotic scaling ratio of the accumulation is determined by the stability exponent of the hyperbolic fixed point, μ_0, as in Eq. (D1). Starting from Eq. (D1), let the z_u base point be g_{-2} in Fig. 2, and the curve $\overline{C}$ that passes through z_u be the stable manifold segment from S'_{-1} that passes through g_{-2}. Furthermore, choose the $\mathcal{C}_0$ curve to be U_0, which intersects S'_{-1} at $a^{(0)}$ and $b^{(0)}$. The pair of points $a^{(0)}$ and $b^{(0)}$ here play the role of the $z^{(0)}$ point in Fig. 18, which are the leading terms of the two families of winding-2 homoclinic points $[a^{(n)}]$ and $[b^{(n)}]$, respectively, that accumulate asymptotically on g_{-2}. The two families of points $[a^{(n)}]$ and $[b^{(n)}]$ are generated from iterating U_0 forward and intersecting the successive images U_n ($n \in \mathbb{Z}^+$) with S'_{-1}, and are located on the upper and lower side of U_n, respectively. The accumulation can be expressed

in the asymptotic relation

$$\lim_{n\to\infty} a^{(n)} = g_{-2},$$
$$\lim_{n\to\infty} |a^{(n)} - g_{-2}|e^{n\mu_0} = C(g_{-2}, a^{(0)}), \tag{9}$$

where $||$ is the standard Euclidean vector norm, and $C(g_{-2}, a^{(0)})$ is a positive constant depending on the base point g_{-2} and the leading term $a^{(0)}$ in the asymptotic family. Similarly for $b^{(n)}$ we have

$$\lim_{n\to\infty} b^{(n)} = g_{-2},$$
$$\lim_{n\to\infty} |b^{(n)} - g_{-2}|e^{n\mu_0} = C(g_{-2}, b^{(0)}). \tag{10}$$

Notice that Eqs. (9) and (10) are obtained directly from Eq. (D1), by the substitutions $z_u \to g_{-2}$ and $z^{(n)} \to a^{(n)}/b^{(n)}$. Therefore, the two families of winding-2 homoclinic points $[a^{(n)}]$ and $[b^{(n)}]$ accumulate asymptotically onto the winding-1 point g_{-2} along the stable manifold, under the scaling relations described by Eqs. (9) and (10). These relations will be denoted symbolically as

$$a^{(n)} \overset{n+1}{\underset{S}{\hookrightarrow}} g_{-2},$$
$$b^{(n)} \overset{n+1}{\underset{S}{\hookrightarrow}} g_{-2}, \tag{11}$$

where the $\overset{n+1}{\underset{S}{\hookrightarrow}}$ symbol indicates that $a^{(n)}$ and $b^{(n)}$ are the $(n+1)$th members of their respective families, $[a^{(0)}, a^{(1)}, \ldots]$ and $[b^{(0)}, b^{(1)}, \ldots]$, that accumulate on g_{-2} along the stable manifold with asymptotic exponent μ_0.

The asymptotic accumulation relations can be used to infer symbolic dynamics of homoclinic points. Given the symbolic codes of the base point, e.g., g_{-2} from Eq. (9), the symbolic codes of the entire families of homoclinic points that accumulate on it can be uniquely determined by suitable additions of $110\cdots$ or $100\cdots$ strings to the left side of the core of g_{-2}. Given $g_{-2} \Rightarrow \bar{0}.01\bar{0}$, it can be inferred that (see Fig. 17)

$$a^{(0)} \Rightarrow \bar{0}1.11\bar{0}$$
$$b^{(0)} \Rightarrow \bar{0}1.01\bar{0} \tag{12}$$

and

$$a^{(n)} \Rightarrow \bar{0}110^{n-1}.01\bar{0}$$
$$b^{(n)} \Rightarrow \bar{0}100^{n-1}.01\bar{0} \qquad (n \geqslant 1), \tag{13}$$

where "0^{n-1}" denotes $(n-1)$ repetitions of 0. The general rule is, the symbolic codes of $a^{(n)}$ and $b^{(n)}$ $(n \geqslant 0)$ are obtained by adding the substrings "110^n" and "100^n", respectively, to the left end of the core of g_{-2}, keeping the position of the decimal point relative to the right end of the core.

Following the same pattern, on the right side of S'_{-1} (see Fig. 17), there are two families of winding-2 homoclinic points $[c^{(n)}]$ and $[d^{(n)}]$ $(n \geqslant 1)$ that accumulate asymptotically along the stable manifold on the winding-1 point h_{-1} under scaling relations similar to Eqs. (9) and (10):

$$c^{(n)} \overset{n}{\underset{S}{\hookrightarrow}} h_{-1},$$
$$d^{(n)} \overset{n}{\underset{S}{\hookrightarrow}} h_{-1}, \tag{14}$$

and their symbolic codes are determined from that of h_{-1}:

$$(c^{(n)} \Rightarrow \bar{0}110^{n-1}.11\bar{0}) \overset{n}{\underset{S}{\hookrightarrow}} (h_{-1} \Rightarrow \bar{0}.11\bar{0}),$$
$$(d^{(n)} \Rightarrow \bar{0}100^{n-1}.11\bar{0}) \overset{n}{\underset{S}{\hookrightarrow}} (h_{-1} \Rightarrow \bar{0}.11\bar{0}), \tag{15}$$

with the same rule of adding the "110^{n-1}" and "100^{n-1}" substrings. This assignment rule for the symbolic code is valid for any homoclinic points in the system. As the construction is rather technical, refer to Appendix C for the detailed systematic assignments of symbolic dynamics.

An important consequence of the above asymptotic relations between homoclinic points is that the phase-space areas spanned by them also scale down at the same rate. Using the present example, three families of areas can be easily identified, which are $[\mathcal{A}^{\circ}_{SUSU[g_{-2},b^{(n)},d^{(n)},h_{-1}]}]$, $[\mathcal{A}^{\circ}_{SUSU[b^{(n)},a^{(n)},c^{(n)},d^{(n)}]}]$, and $[\mathcal{A}^{\circ}_{SUSU[a^{(n)},b^{(n-1)},d^{(n-1)},c^{(n)}]}]$ $(n \geqslant 2)$. Each follows the scaling relation,

$$\lim_{n\to\infty} \frac{\mathcal{A}^{\circ}_{SUSU[g_{-2},b^{(n)},d^{(n)},h_{-1}]}}{\mathcal{A}^{\circ}_{SUSU[g_{-2},b^{(n+1)},d^{(n+1)},h_{-1}]}} = e^{\mu_0}, \tag{16}$$

and similarly for the $[\mathcal{A}^{\circ}_{SUSU[b^{(n)},a^{(n)},c^{(n)},d^{(n)}]}]$ and $[\mathcal{A}^{\circ}_{SUSU[a^{(n)},b^{(n-1)},d^{(n-1)},c^{(n)}]}]$ families as well. These areas are all from the partition of the lobe L'_{-1} using successively propagated lobes L_n. Returning to Fig. 2, where successive intersections between the fundamental segments U_n and S'_{-1} of Eq. (A3) accumulate on g_{-2} and h_{-1}, the following three identifications can be made: $\mathcal{A}^{\circ}_{SUSU[g_{-2},b^{(n)},d^{(n)},h_{-1}]}$ is the area between the lower side of U_n and $U[g_{-2}, h_{-1}]$, $\mathcal{A}^{\circ}_{SUSU[b^{(n)},a^{(n)},c^{(n)},d^{(n)}]}$ is the area between the lower and upper sides of U_n, and $\mathcal{A}^{\circ}_{SUSU[a^{(n)},b^{(n-1)},d^{(n-1)},c^{(n)}]}$ is the area between the upper side of U_n and the lower side of U_{n-1}. As more lobes are added, such areas approach $U[g_{-2}, h_{-1}]$, and the ratio tends to e^{μ_0}. Hence Eq. (16) can be understood as an asymptotic relation between area partitions of L'_{-1} in the neighborhood of $U[g_{-2}, h_{-1}]$.

The above relations are obtained by choosing the z_u base point in Eq. (D1) to be the winding-1 points g_{-2} and h_{-1}, and studying the accumulations of winding-2 homoclinic points on them. Generally speaking, since the choice of the z_u base point is arbitrary, one can just as well choose z_u to be a winding-m homoclinic point on $U(x)$, and there will always be two families of winding-$(m+1)$ homoclinic points that accumulate on z_u along $S(x)$ under similar relations, with the same scaling ratio e^{μ_0}. Therefore, Eq. (16) holds for any winding-m homoclinic point and the winding-$(m+1)$ families of areas that accumulate on it. Such relations are true in the neighborhood of any homoclinic point, and they imply that the computation of a few leading area terms in any $[\mathcal{A}^{\circ}_{SUSU[\cdots]}]$ family suffices to determine the rest of the areas, depending on the desired degree of accuracy.

An important subtlety in the scaling relations concerns the exponent μ_0. Due to the exponential compressing and stretching nature of chaotic dynamics, it is well known that the new cell areas bounded by adjacent stable and unstable segments from a trellis with increasing iteration numbers must become exponentially small. See Appendix A of Ref. [12] for a brief review. In particular, one can anticipate that the new cell areas from $T_{-1,n}$ decrease on average similarly to the horizontal

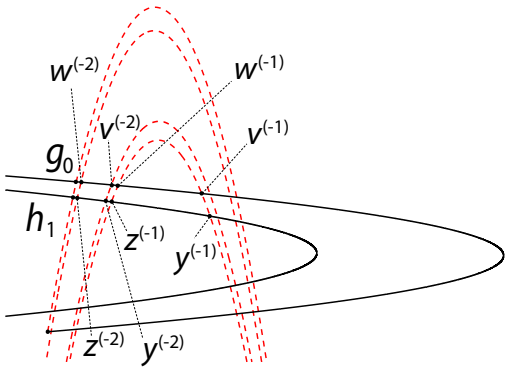

FIG. 3. Accumulation of homoclinic points along $U(x)$ under M^{-1}. Two families of homoclinic points $[v^{(-n)}]$ and $[w^{(-n)}]$ are created from $S'_{-n} \cap U_0$, that accumulate on g_0 along U_0. Notice that only the $n = 1, 2$ cases are plotted here. Similarly, the two families $[y^{(-n)}]$ and $[z^{(-n)}]$ are created from $S'_{-n} \cap U_1$, and accumulate on h_1 along U_1.

strips $H_{s_{-n}\cdots s_{-1}}$ in Figs. 3 and 4 of Ref. [12], which scale at the rate $e^{-n\mu}$, where μ is the system's Lyapunov exponent. However, in general $\mu_0 \neq \mu$, μ_0 measuring the stretching rate of the hyperbolic fixed point, which is expected to be $\geqslant \mu$. This begs the question as to how this could be consistent. In Sec. IV D, it is shown that this presumably larger exponent μ_0 only applies to calculating the ratios between successive areas within the specific families such as those in Eq (16). Between different families, the scaling exponents change to smaller values, which is consistent with the Lyapunov exponent being smaller than μ_0. A shorthand reference to this is to say that Eq. (16) is a fast scaling relation, in the sense that they happen at faster rates than the average instability of the system as a whole, μ.

Identical scaling results hold under the inverse mapping M^{-1} upon switching the roles of the stable and unstable manifolds. Shown in Fig. 3 is a simple example of the inverse case, where families of homoclinic points accumulate along the unstable manifold. For convenience, the $a^{(0)}$ and $b^{(0)}$ points from Fig. 2 are relabeled in this figure as $v^{(-1)}$ and $w^{(-1)}$, respectively. Successive inverse mappings of S'_{-1} intersect with U_0 and create two families of winding-2 points $[v^{(-n)}]$ and $[w^{(-n)}]$ ($n \geqslant 1$), which accumulate on the primary point g_0 along the unstable manifold, under scaling relations similar to Eq. (9). Similarly to Eq. (11), the accumulation along $U(x)$ is denoted by

$$v^{(-n)} \underset{U}{\overset{n}{\hookrightarrow}} g_0,$$
$$w^{(-n)} \underset{U}{\overset{n}{\hookrightarrow}} g_0, \tag{17}$$

where $\underset{U}{\overset{n}{\hookrightarrow}}$ indicates that $v^{(-n)}$ and $w^{(-n)}$ are the nth members of their respective families, $[v^{(-1)}, v^{(-2)}, \ldots]$ and $[w^{(-1)}, w^{(-2)}, \ldots]$, that accumulate on g_0 along the unstable manifold with asymptotic exponent μ_0.

Also shown in Fig. 3 are two other families of winding-2 points $[y^{(-n)}]$ and $[z^{(-n)}]$ generated from $S'_{-n} \cap U_1$, which accumulate on h_1 along the unstable manifold. Notice that points $y^{(-1)}$ and $z^{(-1)}$ are identical to $c^{(1)}$ and $a^{(1)}$ from Fig. 2, respectively. Consequently, three families of areas $[\mathcal{A}^{\circ}_{SUSU[h_1,g_0,w^{(-n)},z^{(-n)}]}]$, $[\mathcal{A}^{\circ}_{SUSU[z^{(-n)},w^{(-n)},v^{(-n)},y^{(-n)}]}]$, and $[\mathcal{A}^{\circ}_{SUSU[y^{(-n)},v^{(-n)},w^{(-n+1)},z^{(-n+1)}]}]$ ($n \geqslant 2$) accumulate on $S[h_1, g_0]$ under the asymptotic ratio e^{μ_0}, similarly to Eq. (16). Therefore, the asymptotic behaviors of the manifolds between M and M^{-1} are identical, upon interchanging the roles of $S(x)$ and $U(x)$. We would like to emphasize that this is a general result that comes from the stability analysis of the system, which holds true whether the system is time-reversal symmetric or not.

There is an interesting special case of the accumulation relations for which z_u is chosen to be the fixed point x itself. For this case, the primary orbits $\{g_i\}$ and $\{h_i\}$ themselves become two families of homoclinic points that accumulate on x with asymptotic ratio e^{μ_0} under both forward and inverse mappings:

$$h_i \underset{S}{\hookrightarrow} x,$$
$$g_i \underset{S}{\hookrightarrow} x, \tag{18}$$

and

$$h_{-i} \underset{U}{\hookrightarrow} x,$$
$$g_{-i} \underset{U}{\hookrightarrow} x, \tag{19}$$

although the meaning of the order number for each point inside these two families now becomes ambiguous, therefore removed from the top of the "$\hookrightarrow$" sign. The hyperbolic fixed point x is now viewed as a "homoclinic point" of winding number 0, on which the winding-1 primaries accumulate.

C. Partitioning of phase-space areas

Of particular relevance to calculating the homoclinic orbit relative actions is the sequence of trellises T_{-1,n_u}, with $n_u = 0, 1, \ldots, N$. New homoclinic points appear on S'_{-1} upon each unit increase of n_u, and their relative actions are closely related to certain phase-space areas called *cells*. Given a trellis T_{n_s,n_u} and four homoclinic points $a, b, c, d \in T_{n_s,n_u}$ that form a simple closed region bounded by the loop $SUSU[a, b, c, d] = S[a, b] + U[b, c] + S[c, d] + U[d, a]$, it is called a *cell of* T_{n_s,n_u} if there are no stable and unstable manifold segments from T_{n_s,n_u} that enter inside the region. Consequently, there are no homoclinic points other than the four vortices on the boundary of the cell. For example, both V_0 and V_1 are cells of $T_{-1,0}$ (Fig. 15). However, in $T_{-1,1}$ (Fig. 4) they get partitioned by U_1 and are not cells anymore since there are unstable segments inside them. Each trellis gives a specific partition to the phase space. By fixing $n_s = -1$ and increasing the n_u value, the resulting sequence of trellises yields a systematic and ever-finer partition of the phase space, which acts as the skeletal-like structure for the study of homoclinic orbits.

In fact, of all the cell areas of T_{-1,n_u}, two subsets are relevant to the action calculations. The first subset, defined as *type-I* cells, are those from the region V_0 partition (Fig. 15). Equivalently, the type-I cells are those with two stable boundary segments located on $S[x, g_0]$ and $S[b^{(0)}, g_{-2}]$, respectively. Similarly, the second subset, or the *type-II* cells, are those from the partition of V' in Fig. 15. Equivalently speaking, the type-II cells are those with two stable segments located on

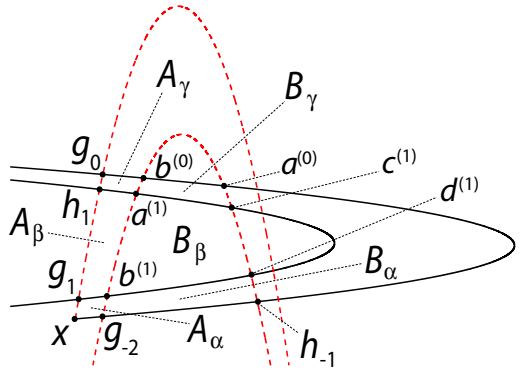

FIG. 4. Partitioning of cell areas in $T_{-1,1}$. The three type-I cells are A_α, A_β, and A_γ. The three type-II cells are B_α, B_β, and B_γ. The A cell from $T_{-1,0}$ is partitioned into three cell areas in $T_{-1,1}$: $A = A_\alpha + A_\beta + A_\gamma$. Similarly for the type-II cell, $B = B_\alpha + B_\beta + B_\gamma$.

$S[b^{(0)}, g_{-2}]$ and $S[h_{-1}, a^{(0)}]$, respectively. Figure 4 shows the examples of $T_{-1,1}$, three type-I cells $A_\alpha, A_\beta, A_\gamma$ and three type-II cells B_α, B_β, B_γ. Section V B shows that the knowledge of these types of cell areas is sufficient for the action calculation of all homoclinic orbits.

In the partitioning of cell areas from increasing trellises, there are families of areas corresponding to fast and slow scaling relations. Since the homoclinic orbit actions are ultimately expressed using these areas, an investigation of this kind is crucial for the understanding of asymptotic clustering of homoclinic orbit actions. The partitioning process is recursive in nature, and the partition of the existing cells of $T_{-1,n}$ by $T_{-1,n+1}$ is the critical step. This process eventually leads to an organization of the cells into treelike structures, and a classification of the scaling rates using the branches of the trees. As introduced in the discussion of Fig. 5, these structures are identical for the type-I and type-II cells, so it suffices to concentrate mostly on the type-I cells.

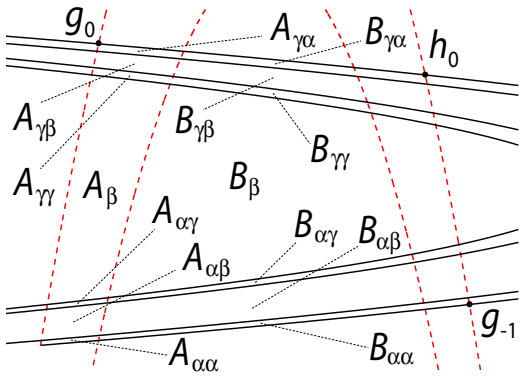

FIG. 5. Zoomed-in graph around the complex region of $T_{-1,2}$ (same as the lower panel of Fig. 2). The A_α and A_γ areas in Fig. 4 are partitioned into three subareas: $A_\alpha = A_{\alpha\alpha} + A_{\alpha\beta} + A_{\alpha\gamma}$ and $A_\gamma = A_{\gamma\alpha} + A_{\gamma\beta} + A_{\gamma\gamma}$. The A_β area does not get partitioned because of the open system assumption, i.e., manifolds outside of the complex region do not revisit the complex region in future iterations. Since the type-I and type-II cells are always partitioned by any lobe L_n simultaneously, the B_α and B_γ cells from Fig. 4 are partitioned in identical ways with A_α and A_γ, respectively.

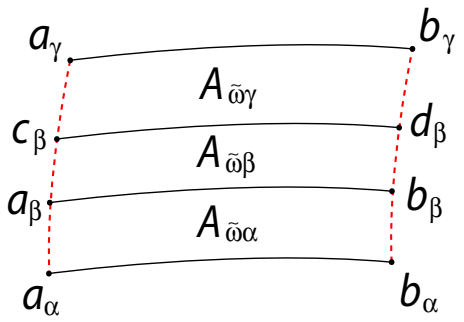

FIG. 6. (Schematic) The cell $A_{\tilde{\omega}}$ in $T_{-1,n}$ is partitioned into three new cells in $T_{-1,n+1}$ by lobe L_{n+1}: $A_{\tilde{\omega}} = A_{\tilde{\omega}\alpha} + A_{\tilde{\omega}\beta} + A_{\tilde{\omega}\gamma}$, where $A_{\tilde{\omega}\beta} \subset L_{n+1}$. The rule of assignment is, $A_{\tilde{\omega}\beta}$ is always assigned to the middle cell, and $A_{\tilde{\omega}\alpha}$ is assigned to the cell with the two corners, namely a_α and b_α, such that a_β, $c_\beta \overset{k}{\underset{S}{\hookrightarrow}} a_\alpha$ and b_β, $d_\beta \overset{k}{\underset{S}{\hookrightarrow}} b_\alpha$, where the order number k is an appropriate integer that depends on $\tilde{\omega}$. Finally, $A_{\tilde{\omega}\gamma}$ is assigned to the last cell. The same pattern applies to all the B cells as well.

The partition starts from $T_{-1,0}$, where the only type-I cell is V_0. In order to introduce a partition subscript, V_0 is denoted A. In the next iteration, A is partitioned by $T_{-1,1}$, in which the lobe L_1 enters A dividing it into three finer cells, namely A_α, A_β, and A_γ, as shown in Fig. 4. Similarly, denote the cell V' of $T_{-1,0}$ by B. B is partitioned by $T_{-1,1}$ in an identical way: $B = B_\alpha + B_\beta + B_\gamma$ since the unstable lobes always enter the type-I and type-II regions simultaneously for the complete horseshoe map, and also for a large class of incomplete horseshoe maps as well.

In the next iteration, $T_{-1,2}$ introduces finer partitions in which L_2 enters A_α and A_γ, dividing both of them into three new cells: $A_\alpha = A_{\alpha\alpha} + A_{\alpha\beta} + A_{\alpha\gamma}$ and $A_\gamma = A_{\gamma\alpha} + A_{\gamma\beta} + A_{\gamma\gamma}$, as labeled in Fig. 5. Therefore, future partitions of a cell correspond to the addition of the α, β, and γ symbols to the end of its existing subscript, except if its subscript ends in β (which terminates that sequence).

In open systems such as the Hénon map, the A_β area does not get partitioned by future iterations because points outside the complex region do not reenter the complex region; therefore no unstable manifolds will extend inside the lobes L_i for all $i \in \mathbb{Z}$. Since A_β belongs to the inside of L_1, it will not be partitioned by any future trellises. The same are true for $A_{\alpha\beta}$, $A_{\gamma\beta}$, and all areas whose subscript ends with β in future trellises, which belong to some future lobes L_i.

The relative position of the new cells is nontrivial. For example, as shown in Fig. 5, $A_{\alpha\alpha}$, $A_{\alpha\beta}$, and $A_{\alpha\gamma}$ are positioned from the bottom to the top, while $A_{\gamma\alpha}, A_{\gamma\beta}, A_{\gamma\gamma}$ are positioned from the top to the bottom, begging the question, how should the order of symbols be assigned for the newly generated cells in a consistent way? The answer is buried in the scaling relations among homoclinic points. As shown in Fig. 6, consider an arbitrary cell area $A_{\tilde{\omega}}$ in $T_{-1,n}$, which is partitioned into three new cells, $A_{\tilde{\omega}\alpha}$, $A_{\tilde{\omega}\beta}$, and $A_{\tilde{\omega}\gamma}$ in $T_{-1,n+1}$ by lobe L_{n+1}. Here $\tilde{\omega}$ denotes a length-n string of symbols composed by arbitrary combinations of α and γ (but not β). The middle cell is always labeled by $\tilde{\omega}\beta$. Let the four homoclinic points on the corners of this cell be a_β, b_β, c_β, and d_β, respectively,

all of which belong to U_{n+1}. The $\tilde{\omega}\alpha$ subscript is then assigned to the cell with the two corners on which a_β, b_β, c_β, and d_β accumulate:

$$a_\beta,\ c_\beta \overset{k}{\underset{S}{\hookrightarrow}} a_\alpha,$$
$$b_\beta,\ d_\beta \overset{k}{\underset{S}{\hookrightarrow}} b_\alpha, \tag{20}$$

where the order number k depends on the detailed forms of $\tilde{\omega}$. The $\tilde{\omega}\gamma$ subscript is assigned to the remaining cell.

If the symbolic codes of a_α and b_α are $a_\alpha \Rightarrow \bar{0}1\tilde{s}_-.\tilde{s}_+1\bar{0}$ and $b_\alpha \Rightarrow \bar{0}1\tilde{s}'_-.\tilde{s}'_+1\bar{0}$, where $\tilde{s}_\pm$ and $\tilde{s}'_\pm$ are substrings composed by 0s and 1s, then it can be inferred using Eq. (C5) that

$$a_\beta \Rightarrow \bar{0}100^{k-1}1\tilde{s}_-.\tilde{s}_+1\bar{0},$$
$$c_\beta \Rightarrow \bar{0}110^{k-1}1\tilde{s}_-.\tilde{s}_+1\bar{0}, \tag{21}$$

and

$$b_\beta \Rightarrow \bar{0}100^{k-1}1\tilde{s}'_-.\tilde{s}'_+1\bar{0},$$
$$d_\beta \Rightarrow \bar{0}110^{k-1}1\tilde{s}'_-.\tilde{s}'_+1\bar{0}. \tag{22}$$

For a concrete example, consider the partition $A = A_\alpha + A_\beta + A_\gamma$ in Fig. 4, where $A_{\tilde{\omega}} = A$ with $\tilde{\omega}$ being an empty string. The A_β is first identified as the one in the middle. Notice that its corners $g_1, h_1 \overset{}{\underset{S}{\hookrightarrow}} x$ and $a^{(1)}, b^{(1)} \overset{2}{\underset{S}{\hookrightarrow}} g_{-2}$; thus A_α is assigned to the cell at the bottom, and A_γ is thus the cell at the top. One can verify that the assignments of cells in Fig. 5 follow the same pattern. In particular, the relative positions of the $A_{\gamma\alpha}$, $A_{\gamma\beta}$, and $A_{\gamma\gamma}$ cells are indeed reversed. This can be seen from the zoomed-in Fig. 7, where the four corners of $A_{\gamma\beta}$, namely v, w, $r^{(1)}$, and $s^{(1)}$, accumulate on g_0 and $b^{(0)}$: $v, w \overset{1}{\underset{S}{\hookrightarrow}} g_0$ and $r^{(1)}, s^{(1)} \overset{1}{\underset{S}{\hookrightarrow}} b^{(0)}$. Thus, $A_{\gamma\alpha}$ is assigned to the cell on the top of $A_{\gamma\beta}$, and $A_{\gamma\gamma}$ the one at the bottom. The partition of the B cells follows an identical scheme.

A complete assignment of the areas' symbols are determined by the accumulation relations between homoclinic points along $S(x)$, which can be carried on with increasing iterations of $T_{-1,n}$ to obtain ever-finer partitions of type-I and type-II cell areas. The progressive partitioning of the type-I cells can be represented by a *partition tree* shown in Fig. 8. Defining the node A to be the 0th level of the tree, which is a cell generated by $T_{-1,0}$, then nodes at the nth level along the tree represent the cells newly generated by $T_{-1,n}$. Notice the β nodes do not get expanded at the next level, because of the open system assumption. A finite truncation of the partition tree to the nth level corresponds to the partition of the type-I areas up to $T_{-1,n}$. Note that the partition tree of type-II cell areas is identical to the type-I tree upon changing the symbols A into B.

D. Scaling relations and periodic orbit exponents

In this section we demonstrate numerically a fundamental relation between the stability exponents of periodic orbits and the scaling ratios in certain families of areas of the partition tree. The relation provides an efficient way to compute the stability exponents of periodic orbits from the areas bounded by stable and unstable manifolds, which does not require the numerical construction of periodic orbits.

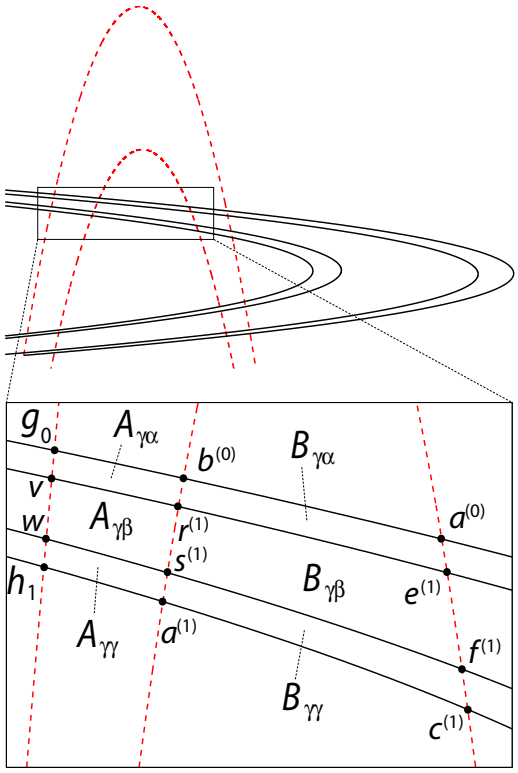

FIG. 7. (Zoomed-in graph) The A_γ and B_γ cells from Fig. 4 are partitioned by L_2 into three new cells each. The $A_{\gamma\beta}$ area is assigned to the middle one. Since $v, w \overset{1}{\underset{S}{\hookrightarrow}} g_0$ and $r^{(1)}, s^{(1)} \overset{1}{\underset{S}{\hookrightarrow}} b^{(0)}$, $A_{\gamma\alpha}$ is assigned to the top one, leaving $A_{\gamma\gamma}$ to be the bottom one. The same rules apply to the B cells as well.

The complete and exact decomposition of the homoclinic orbit actions requires only the areas of the partition trees. On the other hand, their areas scale down asymptotically with the tree level exponentially, with the exponents determined by the specific paths that one moves down the trees. The simplest example is a path of consecutive "α" directions. Start from any α, β, or γ node of the tree, denoted by $A_{\tilde{\omega}\alpha}$, $A_{\tilde{\omega}\beta}$, and

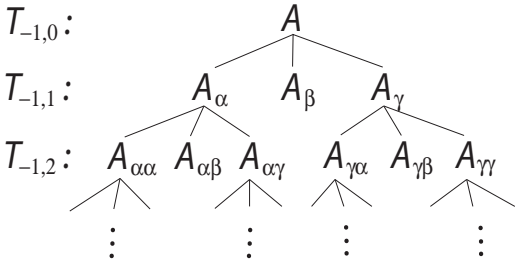

FIG. 8. The partition tree of type-I cell areas. Nodes at the nth level along the tree are areas generated from the partition of $T_{-1,n-1}$ by $T_{-1,n}$. Every α and γ nodes are partitioned into three new nodes at the next level, while the β nodes do not get partitioned any further. The partition tree of type-II areas follows an identical pattern upon changing the symbols A into B. To order the cells as in the trellis, proceeding from the top of the tree, reverse the order for the next level down each time the number of γ symbols is odd.

$A_{\tilde{\omega}\gamma}$, respectively, and move to deeper levels along the left directions. The successive cells areas visited by such paths form three families: $[A_{\tilde{\omega}\alpha^n}]$, $[A_{\tilde{\omega}\alpha^{n-1}\beta}]$, and $[A_{\tilde{\omega}\alpha^{n-1}\gamma}]$, that scale down with the stability exponent of the fixed point:

$$\lim_{n\to\infty}\frac{A_{\tilde{\omega}\alpha^n}}{A_{\tilde{\omega}\alpha^{n+1}}}=\lim_{n\to\infty}\frac{A_{\tilde{\omega}\alpha^{n-1}\beta}}{A_{\tilde{\omega}\alpha^n\beta}}=\lim_{n\to\infty}\frac{A_{\tilde{\omega}\alpha^{n-1}\gamma}}{A_{\tilde{\omega}\alpha^n\gamma}}=e^{\mu_0}, \quad (23)$$

where α^n denotes n consecutive α characters in the string. Identical relations hold for the B cells as well.

The exponents in Eqs. (16) and (23) are identical, and this is not a coincidence. Returning to Sec. IV B, the three families of areas $[\mathcal{A}^{\circ}_{SUSU[g_{-2},b^{(n)},d^{(n)},h_{-1}]}]$, $[\mathcal{A}^{\circ}_{SUSU[b^{(n)},a^{(n)},c^{(n)},d^{(n)}]}]$, and $[\mathcal{A}^{\circ}_{SUSU[a^{(n)},b^{(n-1)},d^{(n-1)},c^{(n)}]}]$ $(n\geqslant 2)$ are just $[B_{\tilde{\omega}\alpha^n}]$, $[B_{\tilde{\omega}\alpha^{n-1}\beta}]$, and $[B_{\tilde{\omega}\alpha^{n-1}\gamma}]$ $(n\geqslant 2)$, respectively, upon letting $\tilde{\omega}=\emptyset$ (null string). Therefore, Eq. (16) is just a special case of Eq. (23). In fact, just as Eq. (16) is a direct consequence of the accumulation relations in Eqs. (11) and (14), the general formula Eq. (23) also comes from the accumulation of corresponding homoclinic points at the vertices of the cells. This can be demonstrated by Fig. 9, where three families of areas $[A_{\tilde{\omega}\alpha^n}]$, $[A_{\tilde{\omega}\alpha^{n-1}\beta}]$, and $[A_{\tilde{\omega}\alpha^{n-1}\gamma}]$ $(n\geqslant 2)$ accumulate on $U[a_\alpha,b_\alpha]$. Starting from the $A_{\tilde{\omega}}$ cell in $T_{-1,n}$ and mapping to higher iterations, the addition of L_{n+m} $(m=1,2,\ldots)$ partitions $A_{\tilde{\omega}\alpha^{m-1}}$ into three new areas: $A_{\tilde{\omega}\alpha^m}$, $A_{\tilde{\omega}\alpha^{m-1}\beta}$, and $A_{\tilde{\omega}\alpha^{m-1}\gamma}$, which approach the $U[a_\alpha,b_\alpha]$ segment asymptotically. The two sequences of points $[a_{\alpha^{m-1}\beta}]$ and $[c_{\alpha^{m-1}\beta}]$ $(m\geqslant 1)$, which are created from successive intersections between U_{n+m} and $S[a_\alpha,a_\gamma]$, give rise to two families of points that accumulate on the base point a_α:

$$a_{\alpha^{m-1}\beta},\, c_{\alpha^{m-1}\beta}\overset{k+m-1}{\underset{S}{\hookrightarrow}} a_\alpha, \quad (24)$$

with exponent μ_0, where k depends on the detailed form of $\tilde{\omega}$.

Similarly, the two sequences of points $[b_{\alpha^{m-1}\beta}]$ and $[d_{\alpha^{m-1}\beta}]$ $(m\geqslant 1)$, generated from successive intersections between U_{n+m} and $S[b_\alpha,b_\gamma]$, give rise to two families of points that accumulate on the base point b_α:

$$b_{\alpha^{m-1}\beta},\, d_{\alpha^{m-1}\beta}\overset{k+m-1}{\underset{S}{\hookrightarrow}} b_\alpha, \quad (25)$$

with the same exponent μ_0 as well.

The scaling relations for the cell areas in Eq. (23) come from the scaling relations of their vertices in Eqs. (24) and (25). In particular, denote the length of the stable manifold segment $S[a,b]$ by $d_s(a,b)$; then the lengths $d_s(a_\alpha,a_{\alpha^{m-1}\beta})$, $d_s(a_{\alpha^{m-1}\beta},c_{\alpha^{m-1}\beta})$, and $d_s(c_{\alpha^{m-1}\beta},a_{\alpha^{m-2}\beta})$ scale as (see Fig. 9)

$$\lim_{m\to\infty}\frac{d_s(a_\alpha,a_{\alpha^{m-1}\beta})}{d_s(a_\alpha,a_{\alpha^m\beta})}=\lim_{m\to\infty}\frac{d_s(a_{\alpha^{m-1}\beta},c_{\alpha^{m-1}\beta})}{d_s(a_{\alpha^m\beta},c_{\alpha^m\beta})}=\lim_{m\to\infty}\frac{d_s(c_{\alpha^{m-1}\beta},a_{\alpha^{m-2}\beta})}{d_s(c_{\alpha^m\beta},a_{\alpha^{m-1}\beta})}=e^{\mu_0}. \quad (26)$$

Considering that the points in Eq. (26) are infinitely close under the $m\to\infty$ limit, so the stable manifold segments connecting them are infinitely close to straight-line segments, the distances between homoclinic points can be replaced by the differences in their p (or q) coordinates (assuming the generic cases in which the local manifolds do not form

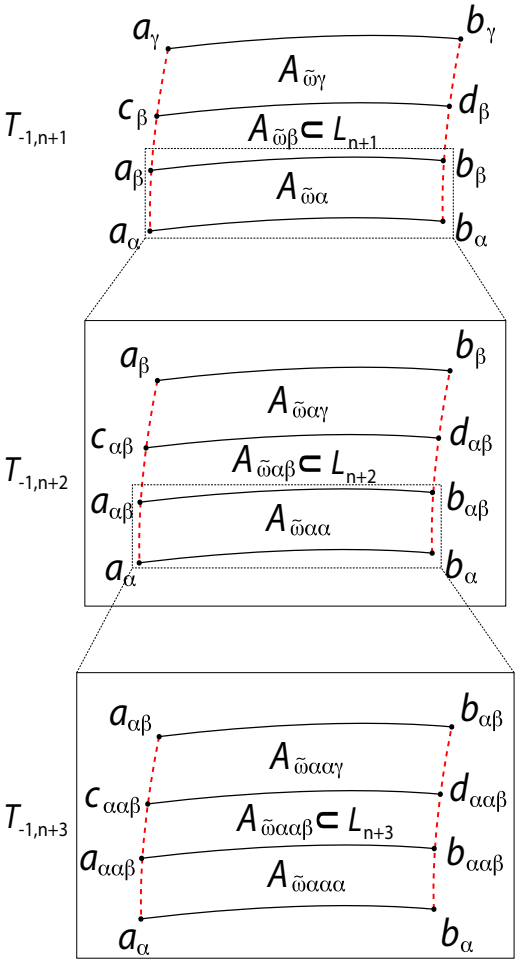

FIG. 9. (Schematic) Successive partitions of $A_{\tilde{\omega}}\subset T_{-1,n}$ in later trellis $T_{-1,n+m}$. Upper panel ($T_{-1,n+1}$): The same as Fig. 6, where $A_{\tilde{\omega}}$ is partitioned into three areas by L_{n+1}. Middle panel ($T_{-1,n+2}$): Zoomed-in graph of $A_{\tilde{\omega}\alpha}$ in $T_{-1,n+2}$, where $A_{\tilde{\omega}\alpha}$ is partitioned by L_{n+2} into three new areas. Lower panel ($T_{-1,n+3}$): Zoomed-in graph of $A_{\tilde{\omega}\alpha\alpha}$ in $T_{-1,n+3}$, where $A_{\tilde{\omega}\alpha\alpha}$ is partitioned by L_{n+3} into three new areas. The addition of successive lobes create four families of homoclinic points, $[a_{\alpha^{m-1}\beta}]$, $[c_{\alpha^{m-1}\beta}]$, $[b_{\alpha^{m-1}\beta}]$, and $[d_{\alpha^{m-1}\beta}]$, that accumulate on a_α and b_α under Eqs. (24) and (25) with exponent μ_0. Therefore, the three families of areas $[A_{\tilde{\omega}\alpha^m}]$, $[A_{\tilde{\omega}\alpha^{m-1}\beta}]$, and $[A_{\tilde{\omega}\alpha^{m-1}\gamma}]$ also converge onto $U[a_\alpha,b_\alpha]$ with exponent μ_0, as described by Eq. (23).

caustics):

$$\lim_{m\to\infty}\frac{p(a_{\alpha^{m-1}\beta})-p(a_\alpha)}{p(a_{\alpha^m\beta})-p(a_\alpha)}=\lim_{m\to\infty}\frac{p(c_{\alpha^{m-1}\beta})-p(a_{\alpha^{m-1}\beta})}{p(c_{\alpha^m\beta})-p(a_{\alpha^m\beta})}=\lim_{m\to\infty}\frac{p(a_{\alpha^{m-2}\beta})-p(c_{\alpha^{m-1}\beta})}{p(a_{\alpha^{m-1}\beta})-p(c_{\alpha^m\beta})}=e^{\mu_0}, \quad (27)$$

where $p(a)$ denotes the p-coordinate value of a. The same relations hold for the q-coordinate values as well. The leading terms of the homoclinic families in Eq. (27) are shown in Fig. 9.

Thus, the asymptotic area scaling relations originate from the asymptotic relations between the positions of homoclinic

TABLE I. e^{μ_0} scaling: The scaling exponents in the "α" direction starting from A_α, A_β, and A_γ are listed in the three columns, respectively. Clearly, they all converge to μ_0 asymptotically.

n	$\ln\left(\frac{A_{\alpha^n}}{A_{\alpha^{n+1}}}\right)$	$\ln\left(\frac{A_{\alpha^{n-1}\beta}}{A_{\alpha^n\beta}}\right)$	$\ln\left(\frac{A_{\alpha^{n-1}\gamma}}{A_{\alpha^n\gamma}}\right)$
1	2.144099	2.103342	2.197343
2	2.142725	2.142323	2.156467
3	2.142084	2.142521	2.144631
4	2.141952	2.142060	2.142364
5	2.141929	2.141949	2.141991
6	2.141927	2.141929	2.141933
μ_0	2.141926	2.141926	2.141926

points on the invariant manifolds. Furthermore, the scaling relations between the phase-space positions of certain homoclinic points give rise to the stability exponent of the fixed point x. In fact, the same relations exist for the stability exponent of any unstable periodic orbit in general [50].

As an example of Eq. (23), the three families of areas $[\mathcal{A}^\circ_{SUSU[g_{-2},b^{(n)},d^{(n)},h_{-1}]}]$, $[\mathcal{A}^\circ_{SUSU[b^{(n)},a^{(n)},c^{(n)},d^{(n)}]}]$, and $[\mathcal{A}^\circ_{SUSU[a^{(n)},b^{(n-1)},d^{(n-1)},c^{(n)}]}]$ $(n \geqslant 2)$ from Eq. (16) can be identified as $[B_{\tilde{\omega}\alpha^n}]$, $[B_{\tilde{\omega}\alpha^{n-1}\beta}]$, and $[B_{\tilde{\omega}\alpha^{n-1}\gamma}]$, respectively, by letting $\tilde{\omega}$ be an empty string. Comparing the areas in Fig. 2 and Fig. 5, the leading terms in the tree families are identified as $\mathcal{A}^\circ_{SUSU[g_{-2},b^{(2)},d^{(2)},h_{-1}]} = B_{\alpha\alpha}$, $\mathcal{A}^\circ_{SUSU[b^{(2)},a^{(2)},c^{(2)},d^{(2)}]} = B_{\alpha\beta}$, and $\mathcal{A}^\circ_{SUSU[a^{(2)},b^{(1)},d^{(1)},c^{(2)}]} = B_{\alpha\gamma}$. Although not plotted in the figure, future lobes partition $B_{\alpha\alpha}$ into ever-finer areas and create the three infinite families of areas that converge to the bottom segment $U[g_{-2}, h_{-1}]$.

To check the accuracy of Eq. (23), the first seven areas of the three families $[A_{\alpha^n}]$, $[A_{\alpha^{n-1}\beta}]$, and $[A_{\alpha^{n-1}\gamma}]$ are given in Table I. The three columns give the scaling exponents obtained from $[A_\alpha, A_{\alpha\alpha}, A_{\alpha\alpha\alpha}, \ldots]$, $[A_\beta, A_{\alpha\beta}, A_{\alpha\alpha\beta}, \ldots]$, and $[A_\gamma, A_{\alpha\gamma}, A_{\alpha\alpha\gamma}, \ldots]$, respectively. Even for the first ratio (worst case), the predicted exponent is good to better than two decimal places. By the bottom of each column, the distinction first appears only in the sixth digit.

The opposite direction down the tree follows increasing repetitions of γ leading to the families $[A_{\tilde{\omega}\gamma^{n-1}\alpha}]$, $[A_{\tilde{\omega}\gamma^{n-1}\beta}]$, and $[A_{\tilde{\omega}\gamma^n}]$ $(n \geqslant 1)$, respectively. The exponential shrinking rate is much slower, and numerical evidence with specific families of cells shown in Tables II and III indicate that the scaling along such "γ" directions converge to the stability exponent μ_1 of x', i.e., the hyperbolic fixed point with reflection:

$$\lim_{n\to\infty}\frac{A_{\tilde{\omega}\gamma^{n-1}\alpha}}{A_{\tilde{\omega}\gamma^n\alpha}} = \lim_{n\to\infty}\frac{A_{\tilde{\omega}\gamma^{n-1}\beta}}{A_{\tilde{\omega}\gamma^n\beta}} = \lim_{n\to\infty}\frac{A_{\tilde{\omega}\gamma^n}}{A_{\tilde{\omega}\gamma^{n+1}}} = e^{\mu_1}, \quad (28)$$

which is in complete analogy to Eq. (23), except for a different direction along the tree, and with a different scaling exponent.

The above-mentioned tables indicate that the scaling of cells along consecutive "α" directions yield the exponent μ_0, and cells along consecutive "γ" directions yield the exponent μ_1. Such phenomena are still just special cases of a general relation that links the scaling exponents along different directions to the symbolic codes of periodic orbits. The association is simple: a scaling step in the "α" direction contributes a symbolic digit "0", and a scaling step in the "γ" direction contributes a digit "1". To formulate this process, define a mapping Ψ that maps a string of Greek letters "α" and "γ" to a string of symbolic codes of "0" and "1", with the grammar $\alpha \mapsto 0$ and $\gamma \mapsto 1$. For example, $\Psi(\gamma\alpha\gamma) = 101$, and the asymptotic scaling exponent in successive "$\gamma\alpha\gamma$" directions is the stability exponent of the $\overline{101}$ periodic orbit, μ_{101}.

TABLE II. e^{μ_1} scaling: The scaling exponents in the "γ" direction starting from A_α, A_β, and A_γ are listed in the three columns, respectively. Clearly, they all converge to μ_1, the stability exponent of the periodic orbit $\overline{1}$.

n	$\ln\left(\frac{A_{\gamma^{n-1}\alpha}}{A_{\gamma^n\alpha}}\right)$	$\ln\left(\frac{A_{\gamma^{n-1}\beta}}{A_{\gamma^n\beta}}\right)$	$\ln\left(\frac{A_{\gamma^n}}{A_{\gamma^{n+1}}}\right)$
1	1.320085	2.365152	1.468471
2	1.707766	1.384612	1.446403
3	1.343392	1.460855	1.500372
4	1.535619	1.496668	1.477362
5	1.467206	1.478053	1.484760
6	1.487618	1.484611	1.482549
7	1.481780	1.482579	1.483168
8	1.483367	1.483164	1.482999
μ_1	1.483036	1.483036	1.483036

In the most general case, consider beginning with an arbitrary node (denoted by either $A_{\tilde{\omega}\alpha}$, $A_{\tilde{\omega}\beta}$, or $A_{\tilde{\omega}\gamma}$, depending on its location) in the type-I partition tree, and study the scaling exponent in an arbitrary direction $\tilde{\eta}$ deepening along the tree. Here $\tilde{\eta}$ is a Greek letter string composed by "α"s and "γ"s that specifies the scaling path. The scaling exponent along $\tilde{\eta}$ is determined by the stability exponent of the periodic orbit $\overline{\Psi(\tilde{\eta})}$, $\mu_{\Psi(\tilde{\eta})}$:

$$\lim_{n\to\infty}\frac{A_{\tilde{\omega}\tilde{\eta}^{n-1}\alpha}}{A_{\tilde{\omega}\tilde{\eta}^n\alpha}} = \lim_{n\to\infty}\frac{A_{\tilde{\omega}\tilde{\eta}^{n-1}\beta}}{A_{\tilde{\omega}\tilde{\eta}^n\beta}} = \lim_{n\to\infty}\frac{A_{\tilde{\omega}\tilde{\eta}^{n-1}\gamma}}{A_{\tilde{\omega}\tilde{\eta}^n\gamma}} = e^{\mu_{\Psi(\tilde{\eta})}}, \quad (29)$$

which is in complete analogy to Eqs. (23) and (28). Notice the relations are independent of $\tilde{\omega}$; i.e., any node of the tree can be used as a starting node (the $n = 1$ terms) of the scaling. Identical relations hold for B cells in the type-II partition tree as well. See Table IV for a simple example of Eq. (29).

TABLE III. e^{μ_1} scaling: The scaling exponents in the "γ" direction starting from $A_{\alpha\alpha}$, $A_{\alpha\beta}$, and $A_{\alpha\gamma}$ are listed in the three columns, respectively. They all converge to μ_1.

n	$\ln\left(\frac{A_{\alpha\gamma^{n-1}\alpha}}{A_{\alpha\gamma^n\alpha}}\right)$	$\ln\left(\frac{A_{\alpha\gamma^{n-1}\beta}}{A_{\alpha\gamma^n\beta}}\right)$	$\ln\left(\frac{A_{\alpha\gamma^n}}{A_{\alpha\gamma^{n+1}}}\right)$
1	1.364533	2.471588	1.449553
2	1.703491	1.352048	1.444654
3	1.332763	1.460781	1.502057
4	1.541193	1.497815	1.476780
5	1.465512	1.477561	1.484950
6	1.488134	1.484780	1.482495
7	1.481634	1.482527	1.483189
μ_1	1.483036	1.483036	1.483036

TABLE IV. $e^{\mu_{01}}$ scaling: The scaling exponents in the "$\alpha\gamma$" direction starting from $A_{\alpha\alpha}$, $A_{\alpha\beta}$, and $A_{\alpha\gamma}$ are listed in the three columns, respectively. Clearly, they all converge to μ_{01}, the stability exponent of the periodic orbit $\overline{01}$.

n	$\ln\left(\frac{A_{\alpha(\alpha\gamma)^{n-1}\alpha}}{A_{\alpha(\alpha\gamma)^n\alpha}}\right)$	$\ln\left(\frac{A_{\alpha(\alpha\gamma)^{n-1}\beta}}{A_{\alpha(\alpha\gamma)^n\beta}}\right)$	$\ln\left(\frac{A_{\alpha(\alpha\gamma)^{n-1}\gamma}}{A_{\alpha(\alpha\gamma)^n\gamma}}\right)$
1	3.520098	4.629501	3.603747
2	3.226675	3.202485	3.292394
3	3.259026	3.255664	3.248603
4	3.256531	3.256733	3.257234
μ_{01}	3.256614	3.256614	3.256614

Similarly to Eq. (23), the origin of Eq. (29) comes from a fundamental relation linking the stability exponents of unstable periodic orbits to the distribution of certain families of homoclinic points (which can be identified as the vertices of the cell areas) on the invariant manifolds.

V. HOMOCLINIC ACTION FORMULAS

All the tools are now in place to develop exact relations expressing the classical actions of any homoclinic orbit in $T_{-1,N}$ (therefore up to transition time $N+1$), in terms of the type-I and type-II cell areas of $T_{-1,N}$. In this method, the calculation of numerical orbits, which suffers from sensitive dependence on initial errors and is unstable in nature, is converted into the calculation of areas bounded by $S(x)$ and $U(x)$, which can be evaluated in stable ways. The exact relations of Sec. V B are perfectly adapted for the development of approximations in Sec. V C that make use of the asymptotic scaling relations among the areas, and that lead to approximate expressions for the homoclinic orbit actions in $T_{-1,N}$ using only the type-I and type-II cell areas from $T_{-1,d(N)}$, where $d(N)$ is an integer much smaller than N. Consequently, it is possible to express the exponentially increasing set of homoclinic orbit actions using a set of areas that is increasing at a much slower rate (e.g., algebraic or linear).

A. Projection operations

The main process leading to the homoclinic action formulas in this section is to express the actions of the homoclinic orbits with large winding numbers in terms of those with small winding numbers, i.e., the decomposition of orbits according to their hierarchical structure. To accomplish this, there are some projection operations to be defined which establish mappings between orbits with different winding numbers.

Given a winding-n ($n \geqslant 1$) homoclinic point y and two winding-$(n+1)$ points z and w such that $z \overset{k}{\underset{S}{\hookrightarrow}} y$ and $w \overset{k}{\underset{S}{\hookrightarrow}} y$ ($\forall k \geqslant 1$) and $S[y,w] \subset S[y,z]$, define the *projection operation along the stable manifold*, denoted by P_S, to be the mapping that maps z and w into the base point y:

$$P_S(z) = P_S(w) = y. \tag{30}$$

The corresponding operation on the symbolic strings, denoted by π_S, can be readily obtained by working backward from Eq. (C5). Namely, given the symbolic codes of z and w, the π_S operation deletes the substrings "110^{k-1}" and "100^{k-1}", respectively, from the left ends of the cores of z and w, while maintaining the position of the decimal point relative to the right end of the core. The resulting symbolic code is then y. Take the points $a^{(0)} \Rightarrow \overline{0}1.11\overline{0}$, $b^{(0)} \Rightarrow \overline{0}1.01\overline{0}$, and $g_{-2} \Rightarrow \overline{0}.01\overline{0}$ in Fig. 1 as examples; we know $a^{(0)}, b^{(0)} \overset{1}{\underset{S}{\hookrightarrow}} g_{-2}$, thus $P_S(a^{(0)}) = P_S(b^{(0)}) = g_{-2}$. Correspondingly for the symbolic codes

$$\begin{aligned} \pi_S(\overline{0}1.11\overline{0}) &= \overline{0}.01\overline{0}, \\ \pi_S(\overline{0}1.01\overline{0}) &= \overline{0}.01\overline{0}, \end{aligned} \tag{31}$$

where the π_S operation deletes either the "11" (for $a^{(0)}$) or "10" (for $b^{(0)}$) substring from the left of the cores while keeping the position of the decimal points relative to the right end of the core unchanged.

Similar operations can be defined for the accumulating homoclinic families along the unstable manifold under the inverse mappings as well. Given a winding-n homoclinic point y', and the winding-$(n+1)$ points z' and w' such that $z' \overset{k}{\underset{U}{\hookrightarrow}} y'$ and $w' \overset{k}{\underset{U}{\hookrightarrow}} y'$ and $U[y',w'] \subset U[y',z']$, define the *projection operation along the unstable manifold*, denoted by P_U, to be the mapping

$$P_U(z') = P_U(w') = y'. \tag{32}$$

The corresponding operation π_U on the symbolic codes is then defined by working backward from Eq. (C6). Namely, given the symbolic codes of z' and w', the π_U operation deletes the substrings "$0^{k-1}11$" and "$0^{k-1}01$", respectively, from the right ends of the cores of z' and w', while maintaining the position of the decimal point relative to the left end of the core. The resulting symbolic code then gives y'.

In the preceding definitions, the projection operations must be applied to homoclinic points with winding numbers $\geqslant 2$. However, they can be naturally extended to apply to the primary (winding-1) points as well. The extension is straightforward: for any primary homoclinic point g_i or h_i, define

$$P_S(g_i) = P_U(g_i) = P_S(h_i) = P_U(h_i) = x \tag{33}$$

with corresponding π_S and π_U operations mapping the symbolic codes of h_i and g_i into $\overline{0}.\overline{0}$, i.e., that of the hyperbolic fixed point x. This is consistent with the scaling relations of Eqs. (18) and (19) as well.

Since π_S and π_U operate on different sides of the cores, it is easy to see that they commute: $\pi_S\pi_U = \pi_U\pi_S$. Since the symbolic codes are in one-to-one correspondences with the phase space points, the projection operations P_S and P_U also commute: $P_SP_U = P_UP_S$. Therefore, a mixed string of operations consisting of n applications of P_S and m applications of P_U, disregarding their relative orders, can always be written as $P_S^nP_U^m$, and similarly for the mixed string of operations of π_S and π_U as well. Such operations are extensively used in the decomposition scheme in Sec. V B.

As an example, consider the $c^{(1)} \Rightarrow \overline{0}11.11\overline{0}$, $h_{-1} \Rightarrow \overline{0}.11\overline{0}$, and $h_1 \Rightarrow \overline{0}11.\overline{0}$ points from Fig. 10. The accumulation relations are $c^{(1)} \overset{1}{\underset{S}{\hookrightarrow}} h_{-1} \underset{U}{\hookrightarrow} x$ and $c^{(1)} \overset{1}{\underset{U}{\hookrightarrow}} h_1 \underset{S}{\hookrightarrow} x$, thus $P_UP_S(c^{(1)}) = P_U(h_{-1}) = x$ and $P_SP_U(c^{(1)}) = P_S(h_1) = x$. On the other hand, using the symbolic dynamics we have

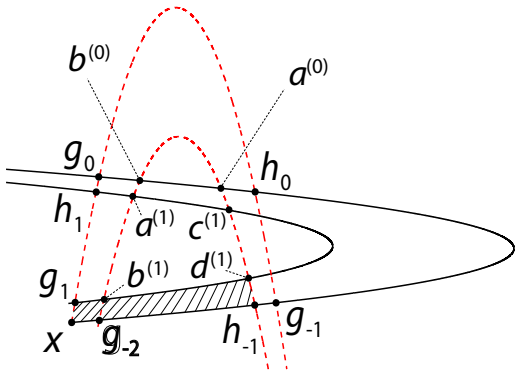

FIG. 10. An example of the homoclinic orbit action decomposition. As shown by Eq. (37), the relative action of the winding-2 orbit $\{d^{(1)}\}$ is decomposed into the sum of the relative actions of the winding-1 orbits $\{h_{-1}\}$ and $\{g_1\}$, and a phase-space area $\mathcal{A}^\circ(d^{(1)}) = \mathcal{A}^\circ_{SUSU[d^{(1)},h_{-1},x,g_1]}$ marked by the hatched region in the figure. Similar decomposition can be done for any homoclinic point on S'_{-1}.

$\pi_U\pi_S(\bar{0}11.11\bar{0}) = \pi_U(\bar{0}.11\bar{0}) = \bar{0}.\bar{0}$ and $\pi_S\pi_U(\bar{0}11.11\bar{0}) = \pi_S(\bar{0}11.\bar{0}) = \bar{0}.\bar{0}$, consistent with the results from the accumulation relations.

B. Exact decomposition

The derivation of the exact formula makes repeated use of the MacKay-Meiss-Percival action principle described by Eqs. (5) and (6), and expresses the relative classical actions of homoclinic orbits as sums of phase-space areas bounded by $S(x)$ and $U(x)$. The fixed-point orbit $\{x\}$ becomes a natural candidate for a reference orbit, and the actions of all homoclinic orbits $\{h\}$ can be expressed relative to $\{x\}$ in the form of $\Delta\mathcal{F}_{\{h\}\{x\}}$, as shown by Eq. (6).

Start by calculating the actions of the two primary orbits $\{g_0\}$ and $\{h_0\}$, which readily follow from Eq. (6). The two areas $\mathcal{A}^\circ_{US[x,h_0]}$ and $\mathcal{A}^\circ_{US[x,g_0]}$ are straightforward to evaluate since only short segments of $S(x)$ and $U(x)$ are required. Having the primary relative orbit actions available, the actions of all winding-n orbits ($n \geqslant 2$) can be determined recursively from the actions of the winding-$(n-1)$ and winding-$(n-2)$ orbits. In particular, given any winding-n ($n \geqslant 2$) homoclinic point $y \in S'_{-1} \cap U_m$, the action of $\{y\}$ can be expressed using three auxiliary orbits: $\{P_S(y)\}$, $\{P_U(y)\}$, and $\{P_SP_U(y)\}$. Substituting $\{y\}$, $\{P_U(y)\}$, $\{P_S(y)\}$, and $\{P_SP_U(y)\}$ into Eq. (7) gives

$$\left(\Delta\mathcal{F}_{\{y\}\{x\}} - \Delta\mathcal{F}_{\{P_U(y)\}\{x\}}\right) - \left(\Delta\mathcal{F}_{\{P_S(y)\}\{x\}} - \Delta\mathcal{F}_{\{P_SP_U(y)\}\{x\}}\right) = \mathcal{A}^\circ_{SUSU[y,P_S(y),P_SP_U(y),P_U(y)]} \tag{34}$$

and therefore

$$\Delta\mathcal{F}_{\{y\}\{x\}} = \Delta\mathcal{F}_{\{P_S(y)\}\{x\}} + \Delta\mathcal{F}_{\{P_U(y)\}\{x\}} - \Delta\mathcal{F}_{\{P_SP_U(y)\}\{x\}} + \mathcal{A}^\circ_{SUSU[y,P_S(y),P_SP_U(y),P_U(y)]}. \tag{35}$$

Notice that the P_S and P_U operations reduce the winding number of y by 1. Similarly, from Eqs. (C5) and (C6) the core length is reduced by at least 2, since their effect is to delete substrings of a minimum of two digits from the original core ("110^{k-1}" or "100^{k-1}" for P_S, "$0^{k-1}11$" or "$0^{k-1}01$" for P_U). Therefore, the three auxiliary orbits are guaranteed to have simpler and shorter phase-space excursions than $\{y\}$. In this sense, Eq. (35) provides a decomposition of the relative action of any arbitrary homoclinic orbit into the relative actions of three simpler auxiliary homoclinic orbits, plus a phase-space area bounded by the manifolds. By repeated contractions, the decomposition could be pushed to involving only the primary homoclinic orbits, the fixed point, and a set of $\mathcal{A}^\circ_{SUSU[\cdots]}$ areas. Implied by this process is that the inverse sequences could be used beginning with the two primary homoclinic orbits, fixed point, and a set of areas to construct the relative actions of all the homoclinic orbits.

The particular form of $\mathcal{A}^\circ_{SUSU[y,P_S(y),P_SP_U(y),P_U(y)]}$ indicates that the area depends only on the homoclinic point y. Once y is chosen, the uniqueness of $P_S(y)$, $P_U(y)$, and $P_SP_U(y)$ means that the area is uniquely calculated. Thus, in the forthcoming contents the shorthand notation

$$\mathcal{A}^\circ(y) \equiv \mathcal{A}^\circ_{SUSU[y,P_S(y),P_SP_U(y),P_U(y)]} \tag{36}$$

will be used frequently to simplify the notation.

An important outcome, buried in Eq. (35), relates to the particular form of $\mathcal{A}^\circ(y)$. For any $y \in S'_{-1} \cap U_m$, the locations of its projections are highly constrained: $P_S(y) \in S'_{-1}$, $P_U(y) \in S[x, g_0]$, and $P_SP_U(y) \in S[x, g_0]$. As a consequence, $\mathcal{A}^\circ(y)$ is always expressible by the type-I and -II cell areas of $T_{-1,m}$. Consider $d^{(1)} \in (S'_{-1} \cap U_1)$ from Fig. 10 for example; the use of Eq. (35) yields

$$\Delta\mathcal{F}_{\{d^{(1)}\}\{x\}} = \Delta\mathcal{F}_{\{h_{-1}\}\{x\}} + \Delta\mathcal{F}_{\{g_1\}\{x\}} - \Delta\mathcal{F}_{\{x\}\{x\}} + \mathcal{A}^\circ(d^{(1)}), \tag{37}$$

where $\Delta\mathcal{F}_{\{x\}\{x\}} = 0$ gives zero contributions. Comparing Fig. 10 with Fig. 4, the $\mathcal{A}^\circ(d^{(1)})$ term (hatched region in Fig. 10) is expressible by two cell areas from the type-I and type-II partition trees of $T_{-1,1}$:

$$\mathcal{A}^\circ(d^{(1)}) = A_\alpha + B_\alpha, \tag{38}$$

both of which are finite curvy trapezoids bounded by the manifolds that can be evaluated simply. The same results hold for all homoclinic points on S'_{-1} with a single exception—$a^{(0)}$. The use of Eq. (35) on $a^{(0)}$ gives

$$\Delta\mathcal{F}_{\{a^{(0)}\}\{x\}} = \Delta\mathcal{F}_{\{g_{-2}\}\{x\}} + \Delta\mathcal{F}_{\{g_0\}\{x\}} + \mathcal{A}^\circ(a^{(0)}),$$

where the evaluation of $\mathcal{A}^\circ(a^{(0)}) = \mathcal{A}^\circ_{SUSU[a^{(0)},g_{-2},x,g_0]}$ requires the additional area $\mathcal{A}^\circ_{SU[a^{(0)},b^{(0)}]}$ that is not part of the partition tree areas. Although the calculation of $\mathcal{A}^\circ_{SU[a^{(0)},b^{(0)}]}$ is not difficult, to make the scheme consistent for all homoclinic points, an alternate form of Eq. (35) is used for $a^{(0)}$ only:

$$\Delta\mathcal{F}_{\{a^{(0)}\}\{x\}} = \Delta\mathcal{F}_{\{h_{-1}\}\{x\}} + \Delta\mathcal{F}_{\{g_0\}\{x\}} + \mathcal{A}^\circ_{SUSU[a^{(0)},h_{-1},x,g_0]} \tag{39}$$

so $\mathcal{A}^\circ_{SUSU[a^{(0)},h_{-1},x,g_0]}$ is expressible by cell areas $A + B$.

Although $\mathcal{A}^\circ(y)$ is expressible by linear combinations of type-I and type-II partition tree areas, $A_{\tilde{\omega}}$ and $B_{\tilde{\omega}}$, the precise mapping between this area and the tree area symbols must be determined. Given the symbolic code of any homoclinic point $y \in (S'_{-1} \cap U_m)$, the explicit mapping links $\mathcal{A}^\circ(y)$ with specific linear combinations of cell areas from the type-I and type-II partition trees of $T_{-1,m}$. Since the transition time of y is $m+1$, according to Eq. (C2), its core length is $m+3$. Let $\tilde{s} = s_1s_2\cdots s_{m+2}s_{m+3}$ ($s_i \in \{0, 1\}$, $s_1 = s_{m+3} = 1$) be the core of the symbolic code of y; then the linear combination of cell

areas depends solely on $\tilde{s}$. As the association is rather technical, the details are given in Appendix E. The correspondence is given by Eq. (E4) using the notation and other relations also defined in the Appendix.

Even though the actions of individual homoclinic orbits can always be calculated directly with the MacKay-Meiss-Percival action principle, $\Delta\mathcal{F}_{\{y\}\{x\}} = \mathcal{A}^{\circ}_{US[x,y]}$, for those orbits with large transit times, the integration path $US[x, y]$ will be stretched exponentially long and extend far from the fixed point. Accurate interpolation of the path will require an exponentially growing set of points on the manifolds to maintain a reasonable density, an impractical task given the formidable computation time and memory space. On the other hand, using Eqs. (35) and (E4), the entire set of the homoclinic orbit actions arising from any trellis $T_{-1,N}$ can be calculated with the two primary orbit actions, $\Delta\mathcal{F}_{\{h_0\}\{x\}}$ and $\Delta\mathcal{F}_{\{g_0\}\{x\}}$, and the areas of the cells of the type-I and type-II partition trees of $T_{-1,N}$. These areas are confined to a finite region of the phase space, and bounded by stable and unstable manifolds with small curvatures, which are far easier to compute. Notice that both the symbolic codes of homoclinic points and the numerical areas in the partition trees can be generated with straightforward computer algorithms, so the recursive use of Eqs. (35) and (E4) give rise to an automated computational scheme for the exact calculation of homoclinic orbit actions.

Equivalently, one may carry out the recursive process explicitly, which leads to an expression of the homoclinic orbit action as a cell-area expansion. This is done by expanding the three auxiliary homoclinic orbit actions in Eq. (35) using the equation itself, repeatedly, until all auxiliary orbits reduce to the primary homoclinic orbits. However, there is a technical difficulty of Eq. (35) to take into account: the point $P_U(y)$ is no longer on S'_{-1}, so the area term in its own expansion, $\mathcal{A}^{\circ}(P_U(y))$, is no longer being expressed by the type-I and type-II cell areas. Consequently, Eq. (E4) breaks down for $P_U(y)$. The same is true for point $P_S P_U(y)$ as well. To adjust for this problem, all that is needed is to identify the representative point of the orbit $\{P_U(y)\}$ on S'_{-1}, denoted by $P'_U(y)$. In fact, $P'_U(y)$ is just an image of $P_U(y)$ under several inverse mappings. The number of inverse mappings is straightforwardly identified. All homoclinic points on S'_{-1} have symbolic codes of the form $\overline{0}\tilde{\zeta}.01\overline{0}$ (if they are located on $S[b^{(0)}, g_{-2}]$) or $\overline{0}\tilde{\zeta}.11\overline{0}$ (if they are located on $S[h_{-1}, a^{(0)}]$), where $\tilde{\zeta}$ denotes an arbitrary symbolic string of binary digits. Equivalently stated, the decimal point in the symbolic code of any homoclinic point on S'_{-1} is always two digits left of the right end of its core. Hence, the resultant shift of the decimal point of $P_U(y)$ yields $P'_U(y)$. Suppose the decimal point of $P_U(y)$ is n' digits to the right side of the right end of its core; then the P'_U operation can be defined as

$$P'_U(y) \equiv M^{-(n'+2)} P_U(y). \tag{40}$$

The corresponding symbolic operation π'_U can be defined as a shift of the decimal point for $n' + 2$ digits towards the left, after the operation π_U.

For the special cases of $y = h_i$ or $y = g_i$, i.e., a primary homoclinic point, $P_U(y)$ reduces to x, and n' loses its meaning. For those cases, define

$$P'_U(h_i) = P'_U(g_i) = x \tag{41}$$

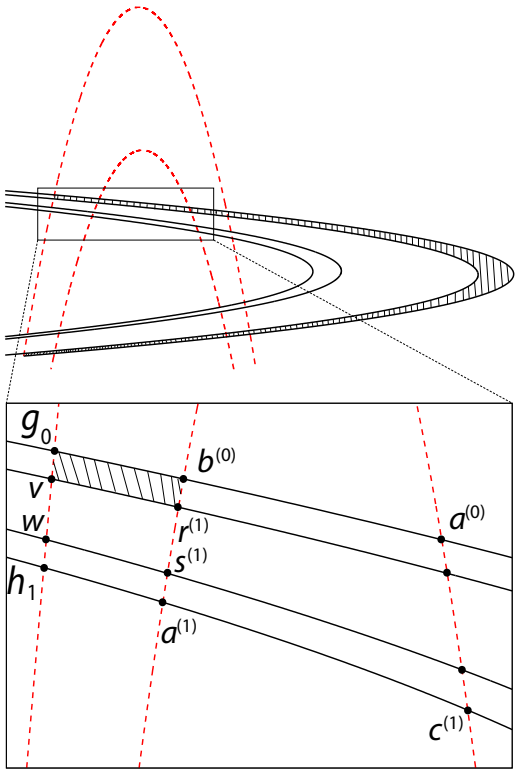

FIG. 11. Relative areas for the decomposition of the winding-3 orbit $\{r^{(1)}\}$. The $\mathcal{A}^{\circ}(r^{(1)})$ term in Eq. (43) is marked as the hatched region in the lower panel, which is just $-A_{\gamma\alpha}$. The long and curvy, hatched region in the upper panel is the $\mathcal{A}^{\circ}(v)$ term in Eq. (45). Areas like this may not be expressible by the type-I and type-II cell areas.

and the corresponding π'_U operation maps the symbolic codes of the primary homoclinic points into $\overline{0}.\overline{0}$, i.e., that of the hyperbolic fixed point.

The commutative relations hold for both the projection operations and their symbolic counterparts: $P_S P'_U = P'_U P_S$ and $\pi_S \pi'_U = \pi'_U \pi_S$. Using the P'_U operation, Eq. (35) can be written alternatively as

$$\begin{aligned}\Delta\mathcal{F}_{\{y\}\{x\}} = {} & \Delta\mathcal{F}_{\{P_S(y)\}\{x\}} + \Delta\mathcal{F}_{\{P'_U(y)\}\{x\}} \\ & - \Delta\mathcal{F}_{\{P_S P'_U(y)\}\{x\}} + \mathcal{A}^{\circ}(y),\end{aligned} \tag{42}$$

in which the representative points $P'_U(y)$ and $P_S P'_U(y)$ of the auxiliary homoclinic orbits $\{P'_U(y)\}$ and $\{P_S P'_U(y)\}$ both locate on S'_{-1} now. Therefore, the recursive expansion of Eq. (42) can be continued until all auxiliary orbits involved are primary homoclinic orbits.

The above motivation for introducing this extra P'_U operation is better demonstrated with the example in Fig. 11. For the winding-3 homoclinic point $r^{(1)}$, we have $r^{(1)} \xrightarrow[S]{1} b^{(0)}$ and $r^{(1)} \xrightarrow[U]{1} v$; therefore the projection operations on it give $P_S(r^{(1)}) = b^{(0)}$, $P_U(r^{(1)}) = v$, and $P_S P_U(r^{(1)}) = g_0$. Thus, Eq. (35), when applied to $r^{(1)}$, reads

$$\begin{aligned}\Delta\mathcal{F}_{\{r^{(1)}\}\{x\}} = {} & \Delta\mathcal{F}_{\{b^{(0)}\}\{x\}} + \Delta\mathcal{F}_{\{v\}\{x\}} \\ & - \Delta\mathcal{F}_{\{g_0\}\{x\}} + \mathcal{A}^{\circ}(r^{(1)}),\end{aligned} \tag{43}$$

where $\mathcal{A}^\circ(r^{(1)}) = \mathcal{A}^\circ_{SUSU[r^{(1)},b^{(0)},g_0,v]}$ is the negative area of the hatched region ($-A_{\gamma\alpha}$) from the lower panel of Fig. 11. Among the three auxiliary orbit actions in the above expression, $\Delta\mathcal{F}_{\{g_0\}\{x\}}$ is already a primary orbit action; therefore no further decomposition is needed for it. The other two, $\Delta\mathcal{F}_{\{b^{(0)}\}\{x\}}$ and $\Delta\mathcal{F}_{\{v\}\{x\}}$, are both winding-2 orbits, and thus need to be further decomposed via Eq. (35) again. This is fine for $\Delta\mathcal{F}_{\{b^{(0)}\}\{x\}}$, since $b^{(0)}$ is already on S'_{-1}, and thus,

$$\Delta\mathcal{F}_{\{b^{(0)}\}\{x\}} = \Delta\mathcal{F}_{\{g_{-2}\}\{x\}} + \Delta\mathcal{F}_{\{g_0\}\{x\}} - \Delta\mathcal{F}_{\{x\}\{x\}} + \mathcal{A}^\circ(b^{(0)}), \tag{44}$$

where $\mathcal{A}^\circ(b^{(0)}) = \mathcal{A}^\circ_{SUSU[b^{(0)},g_{-2},x,g_0]} = A_\alpha + A_\beta + A_\gamma$. However, the same procedure, when applied to $\Delta\mathcal{F}_{\{v\}\{x\}}$, gives rise to undesired subtleties. Notice that $v \notin S'_{-1}$, $P_S(v) = g_0$, and $P_U(v) = g_2$, which lead to the expansion

$$\Delta\mathcal{F}_{\{v\}\{x\}} = \Delta\mathcal{F}_{\{g_0\}\{x\}} + \Delta\mathcal{F}_{\{g_2\}\{x\}} - \Delta\mathcal{F}_{\{x\}\{x\}} + \mathcal{A}^\circ(v), \tag{45}$$

where $\mathcal{A}^\circ(v) = \mathcal{A}^\circ_{SUSU[v,g_0,x,g_2]}$ is a long, thin, and folded area indicated by the hatched region in the upper panel of Fig. 11. The expressions of such areas in terms of the type-I and -II cells are not immediately apparent, and the correspondence relation Eq. (E4) will fail. The fix, however, is simple and straightforward: use the representative point of $\{v\}$ on S'_{-1}. This point can be easily identified from the symbolic dynamics. Given that $b^{(0)} \Rightarrow \bar{0}1.01\bar{0}$ and $r^{(1)} \overset{1}{\underset{S}{\hookrightarrow}} b^{(0)}$, we know from Eq. (C5) that $r^{(1)} \Rightarrow \bar{0}101.01\bar{0}$. Since $v = P_U(r^{(1)})$, its symbolic code is then $v \Rightarrow \pi_U(\bar{0}101.01\bar{0}) = \bar{0}101.\bar{0}$, which indicates that $v = M^2(b^{(0)})$. Therefore, the representative point of $\{v\}$ on S'_{-1} is identified to be $b^{(0)}$. Correspondingly, one can verify the validity of Eq. (40) since $P'_U(r^{(1)}) = b^{(0)}$; i.e., $P'_U(y)$ indeed yields the correct representative point of $\{P_U(y)\}$ on S'_{-1}. Therefore, $\Delta\mathcal{F}_{\{v\}\{x\}} = \Delta\mathcal{F}_{\{b^{(0)}\}\{x\}}$, which is expressible via Eq. (44) again. The final expression for $\{r^{(1)}\}$ is then

$$\begin{aligned}\Delta\mathcal{F}_{\{r^{(1)}\}\{x\}} &= 3\Delta\mathcal{F}_{\{g_0\}\{x\}} + 2\mathcal{A}^\circ(b^{(0)}) + \mathcal{A}^\circ(r^{(1)}) \\ &= 3\Delta\mathcal{F}_{\{g_0\}\{x\}} + 2A - A_{\gamma\alpha},\end{aligned} \tag{46}$$

which only involves $\Delta\mathcal{F}_{\{g_0\}\{x\}}$ and several type-I cell areas. As shown by Eq. (49) later, similar decomposition can be written for any homoclinic orbit, and the resulting expansions will only involve the two primary orbit actions, $\Delta\mathcal{F}_{\{g_0\}\{x\}}$ and $\Delta\mathcal{F}_{\{h_0\}\{x\}}$, plus a linear combination of some type-I and type-II cell areas.

The general process proceeds as follows. Consider the case of $\{y\}$ with winding-2. Then $P_S P_U(y) = x$, thus $\Delta\mathcal{F}_{\{P_S P_U(y)\}\{x\}} = 0$. The two nonvanishing auxiliary orbits are $\{P_S(y)\}$ and $\{P'_U(y)\}$, both of which are primary orbits, so Eq. (42) is already a complete expansion. For all higher winding cases, $n \geqslant 3$, it is possible to expand the $\Delta\mathcal{F}_{\{P_S(y)\}\{x\}}$ and $\Delta\mathcal{F}_{\{P'_U(y)\}\{x\}}$ terms in Eq. (42) using the equation itself to obtain a twice-iterated formula

$$\begin{aligned}\Delta\mathcal{F}_{\{y\}\{x\}} = {} & \Delta\mathcal{F}_{\{P_S^2(y)\}\{x\}} + \Delta\mathcal{F}_{\{P_S P'_U(y)\}\{x\}} \\ & + \Delta\mathcal{F}_{\{P_U'^2(y)\}\{x\}} - \Delta\mathcal{F}_{\{P_S^2 P'_U(y)\}\{x\}} - \Delta\mathcal{F}_{\{P_S P_U'^2(y)\}\{x\}} \\ & + \mathcal{A}^\circ(y) + \mathcal{A}^\circ(P_S(y)) + \mathcal{A}^\circ(P'_U(y)).\end{aligned} \tag{47}$$

Since y, $P_S(y)$, and $P'_U(y)$ are all located on S'_{-1}, with the help of Eq. (E4), the three $\mathcal{A}^\circ$ areas in the above formula are all expressible using type-I and type-II areas. For the orbits with $n = 3$, both $\Delta\mathcal{F}_{\{P_S^2 P'_U(y)\}\{x\}}$ and $\Delta\mathcal{F}_{\{P_S P_U'^2(y)\}\{x\}}$ vanish, so no more expansions are needed. An example of this is already provided by Eq. (46) previously. For the $n \geqslant 4$ cases, the above procedure can be carried on repeatedly, until the $P_S^{n-i} P_U'^i(y)$ $(1 \leqslant i \leqslant n-1)$ action terms are present, which reduce y into x. To further simplify the notations, define the mixed projections of P_S and P'_U on y as

$$P(y;i;j) \equiv P_S^{i-j} P_U'^j(y), \quad (i \geqslant j). \tag{48}$$

Then, a general formula for the complete action decomposition of any winding-n homoclinic orbit $\{y\}$ (where $y \in S'_{-1} \cap U_m$) can be written as

$$\Delta\mathcal{F}_{\{y\}\{x\}} = \sum_{i=0}^{n-1} \Delta\mathcal{F}_{\{P(y;n-1;i)\}\{x\}} + \sum_{i=0}^{n-2}\sum_{j=0}^{i} \mathcal{A}^\circ(P(y;i;j)), \tag{49}$$

where $\Delta\mathcal{F}_{\{P(y;n-1;i)\}\{x\}} = \Delta\mathcal{F}_{\{P_S^{n-1-i} P_U'^i(y)\}\{x\}}$ are relative actions of the primary homoclinic orbits, therefore either $\Delta\mathcal{F}_{\{h_0\}\{x\}}$ or $\Delta\mathcal{F}_{\{g_0\}\{x\}}$. The $\mathcal{A}^\circ(P(y;i;j))$ terms in the double sum are areas of the curvy parallelograms spanned by four homoclinic points of various winding numbers, generated from mixed projections of y. By design, all $P(y;i;j)$ points in these areas are located on S'_{-1}; thus the $\mathcal{A}^\circ(P(y;i;j))$ terms are expressible using the type-I and type-II cells via Eq. (E4).

Equation (49) gives a complete expansion of the homoclinic orbit actions in terms of the primary homoclinic orbit actions plus the cell areas of type-I and type-II partition trees. It converts the determinations of numerical orbits into area calculations in a finite region of the phase space, and avoids exponentially extending integration paths associated with complicated orbits. Furthermore, the two types of cells come from a nearly parallel and linear foliated phase-space region with relatively small curvature along the manifolds, so the numerical interpolation of the manifolds does not require a very dense set of points, and therefore renders the calculations practical.

Nevertheless, the total number of the cell areas proliferates with the same rate as the homoclinic points on S'_{-1}. This is because the cells can be put into a one-to-one correspondence with the nonprimary homoclinic points on S'_{-1}, such that each cell corresponds to the homoclinic point at its upper right corner. For example, in Fig. 4, the cells A_α, A_β, A_γ, B_α, B_β, and B_γ correspond to points $b^{(1)}$, $a^{(1)}$, $b^{(0)}$, $d^{(1)}$, $c^{(1)}$, and $a^{(0)}$, respectively. As we increase the integer N of the trellis $T_{-1,N}$, new cells emerge at an identical rate with new homoclinic points on S'_{-1}, both of which proliferate as $2^{N+2} = e^{\alpha(N+2)}$, where $\alpha = \ln 2$ is the topological entropy of the system. Therefore, the exact evaluation of homoclinic orbit actions, Eq. (49), requires an exponentially increasing set of areas for its input, as must happen.

A few words are in order for the symbolic dynamics. In all the derivations up till now, we have assumed the homoclinic tangle forms a complete horseshoe structure, which allows all possible sequences of binary digits. Although this is often true for highly chaotic systems, for other types of systems with

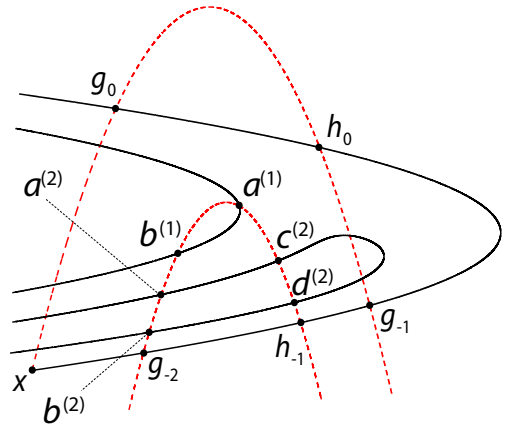

FIG. 12. (Schematic) Homoclinic tangle forming an incomplete horseshoe. Comparing to the complete horseshoe case (Fig. 2), the points $a^{(0)}$, $b^{(0)}$, $c^{(1)}$, and $d^{(1)}$ are pruned. However, the accumulation relations (thus the projection operations) for the unpruned homoclinic points remain the same. Therefore, for the unpruned homoclinic points, Eqs. (35) and (49) remain valid.

mixed dynamics, the homoclinic tangles will in general form incomplete horseshoe structures that coexist with stability islands in phase space. A simple kind of incomplete horseshoe is shown by Fig. 12. The symbolic dynamics of such systems are more complicated as certain substrings are not admissible by the dynamics and therefore "pruned" from the symbol plane [51,52]. Therefore, not all symbolic strings may exist, and their very existence is determined by a "pruning front" [51] which separates the allowed and disallowed orbits in the symbol plane. In spite of this apparent complication, the foundations of our final result Eq. (49) hold true in general, even for incomplete horseshoes. Namely, the accumulation of homoclinic points along the manifolds, and the projection operations defined accordingly, remain valid for all types of horseshoe structures. For instance, in Fig. 12, although $a^{(0)}$ and $b^{(0)}$ are pruned, we still have $a^{(n)} \overset{n}{\underset{S}{\hookrightarrow}} g_{-2}$ and $b^{(n)} \overset{n}{\underset{S}{\hookrightarrow}} g_{-2}$ (where $n \geqslant 1$). As compared to Eq. (11), the pruning removes the first members ($a^{(0)}$ and $b^{(0)}$) of the two accumulating families, but leaves the rest unchanged. Therefore, as long as the pruning front (or a finite approximation of it) has been established by methods such as [53], Eqs. (35) and (49) will be applicable to any admissible homoclinic orbit $\{y\}$ since all the projections involved are admissible as well. Therefore, their range of applicability is not limited to the complete horseshoes.

A complication that does arise in the incomplete horseshoe cases is the pruning of the area partition trees. Depending on the complexities of the horseshoes, more types of trees might be needed, and their structures will not be as simple as the one in Fig. 8. Certain nodes will be pruned away, and there may not exist a finite grammar rule. Just like the pruning fronts, the partition trees are also system-specific, and we anticipate that the numerical algorithms for generating the pruning front should already contain adequate information for generating the partitions trees as well, although more sophisticated investigations along this direction are needed.

C. Information reduction

In semiclassical approximations, the classical actions divided by $\hbar$ determine phase angles, and as it is already an approximation to begin with, it is possible to tolerate small errors, say $\epsilon = \delta\mathcal{F}/\hbar$, measured in radians. As a practical matter, once this ratio is $\lesssim 0.1$ or some similar scale, constructive and destructive interferences are properly predicted, and much greater precision becomes increasingly irrelevant. Given that the areas in Eq. (49), or similarly of the partition tree cells, shrink exponentially rapidly, most of these corrections can be dropped or ignored.

Identifying the necessary information begins with an estimate of orders of magnitudes of the areas terms in Eq. (49). Given any trellis $T_{-1,N}$, the maximum winding number of a homoclinic orbit is $n_{\max} = N/2 + 2$. Due to the slow scaling direction of the tree structure, the orbit $y \Rightarrow \bar{0}1^{N+1}.11\bar{0}$ yields an expansion with the largest possible number of significant $\mathcal{A}^{\circ}(P(y;i;j))$ terms, and hence an upper bound on the number of necessary areas.

It is reasonable to assume the cell areas A and B of $T_{-1,0}$ are of the same magnitude, and it is sufficient to consider the ratios $\mathcal{R} = \mathcal{A}^{\circ}(P(y;i;j))/A$. Via Eq. (E4), $\mathcal{A}^{\circ}(P(y;i;j))$ is expressible as a linear combination of cell areas of partition trees of $T_{-1,N-2i}$. These cell areas are at the $(N-2i)$th level of the partition trees; hence the scaling relation, Eq. (28), gives ratio estimates $\sim e^{-\mu_1(N-2i)}$. As a result, the inner area sum of Eq. (49) gives

$$\sum_{j=0}^{i} \mathcal{A}^{\circ}(P(y;i;j)) \sim A \cdot O((i+1)e^{-\mu_1(N-2i)}). \tag{50}$$

Comparing this estimate with the threshold $\delta\mathcal{F}$ yields a maximum value of the depth $d \equiv N - 2i$ of the tree needed:

$$A(i+1)e^{-\mu_1 d} \geqslant \delta\mathcal{F}, \tag{51}$$

therefore

$$e^{-\mu_1 d} \geqslant \frac{\epsilon\hbar}{A(i+1)}. \tag{52}$$

A slightly more conservative bound replaces $i+1$ with $n_{\max} \approx N/2$ and gives after some algebra

$$d \leqslant \frac{1}{\mu_1} \ln \frac{NA}{2\epsilon\hbar}. \tag{53}$$

Therefore, in order to calculate all homoclinic orbit actions arising from $T_{-1,N}$ within the error tolerance $\epsilon\hbar$, we only need to determine numerically the type-I and type-II cell areas of the partition trees of $T_{-1,d}$. Recall that the number of cell areas in $T_{-1,d}$ is estimated by

$$e^{\alpha d} \sim \left(\frac{NA}{2\epsilon\hbar}\right)^{\frac{\alpha}{\mu_1}}, \tag{54}$$

whereas the number of homoclinic orbits in $T_{-1,N}$ is $\propto e^{\alpha N}$, where $\alpha = \ln 2$ is the topological entropy of the system. Thus, the exponentially proliferating homoclinic orbit actions in $T_{-1,N}$ are expressible by the algebraically proliferating cell areas from $T_{-1,d}$, a significant information reduction.

In practice, the use of $T_{-1,d}$ to construct the relative actions of $T_{-1,N}$ alters the area sum in Eq. (49), such that any

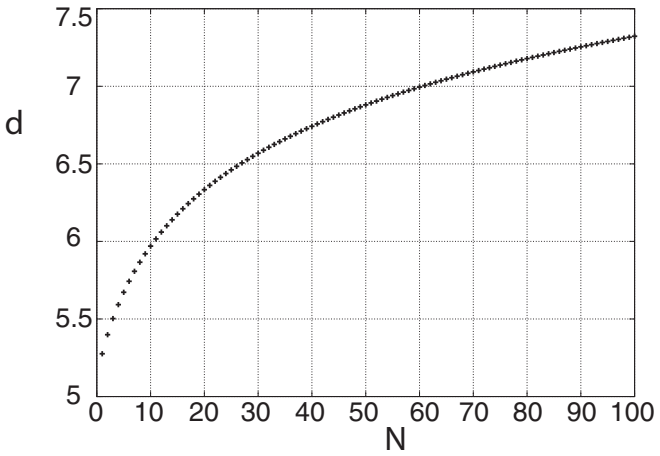

FIG. 13. The logarithmic dependence of d with respect to N using the slow scaling exponent $\mu_1 = 1.483$, error tolerance $\delta\mathcal{F}/A = 0.001$, and $a = 10$ for the Hénon map. For the computation of large trellises such as $N = 100$, the approximate formula of Eq. (55) only requires the computation of cell areas up to trellis number $d = 8$.

$\mathcal{A}^{\circ}(P(y;i;j))$ terms with $P(y;i;j) \notin T_{-1,d}$ will be excluded from the double sum, leading to the reduced action formula

$$\Delta\mathcal{F}_{\{y\}\{x\}} = \sum_{i=0}^{n-1} \Delta\mathcal{F}_{\{P(y;n-1;i)\}\{x\}} + \sum_{i=0}^{n-2} \sum_{\substack{j=0 \\ P(y;i;j)\in T_{-1,d}}}^{i} \mathcal{A}^{\circ}(P(y;i;j)) + O(\delta\mathcal{F}), \tag{55}$$

where the constraint $P(y;i;j) \in T_{-1,d}$ is imposed, therefore eliminating the (exponentially many) cell areas smaller than the error threshold $\delta\mathcal{F}$.

D. Numerical example

For the Hénon map in Eq. (B3) with $a = 10$, and an error tolerance $\delta\mathcal{F}/A = 0.001$, the natural logarithmic dependence of d on N is shown in Fig. 13. The information reduction is significant: even for the calculation of homoclinic orbit actions of $T_{-1,100}$, which is obviously impossible via traditional methods, our scheme only requires the numerical computation of cell areas up to $T_{-1,8}$, an effortless task for personal computers.

For the numerical verification of Eqs. (49) and (55), we calculate the relative actions of the homoclinic orbits of $T_{-1,N}$ in three different ways. The first method is to implement the orbit finder method introduced in our previous work [18], which determines the numerical orbits $\{y\}$ and thus their relative actions, $\Delta\mathcal{F}^{(\mathrm{ref.})}_{\{y\}\{x\}}$. These actions are the standard reference actions for comparison. The second method is to calculate the cell areas in the partition trees of $T_{-1,N}$, and evaluate the actions $\Delta\mathcal{F}^{(\mathrm{exact})}_{\{y\}\{x\}}$ using Eqs. (49) and (E4). These should only differ from $\Delta\mathcal{F}^{(\mathrm{ref.})}_{\{y\}\{x\}}$ due to relying on double precision computation since both are exact evaluations with no approximations involved. On the contrary, in the third method the tolerance is $\delta\mathcal{F}/A = 0.001$ (where $A \approx 10.973$ for the current case of $a = 10$), and only cell areas of the partition trees up to the reduced trellis $T_{-1,d}$ are used with Eqs. (55) and (E4) to obtain the approximate actions, $\Delta\mathcal{F}^{(\mathrm{approx.})}_{\{y\}\{x\}}$.

Every homoclinic orbit up to iteration number $N = 10$ is constructed, which corresponds to trellis $T_{-1,10}$. The total number of orbits is $2^{12} = 4096$. The reduced iteration number for this case is $d = 6$; i.e., the relative homoclinic orbit actions in $T_{-1,10}$ should be given to an accuracy $A \times O(10^{-3}) \sim 1 \times 10^{-2}$ or better using only the cell areas from $T_{-1,6}$.

Due to the large number of orbits, it is impractical to list the results for $\Delta\mathcal{F}^{(\mathrm{exact})}_{\{y\}\{x\}}$ and $\Delta\mathcal{F}^{(\mathrm{approx.})}_{\{y\}\{x\}}$ for every orbit. Instead, we show the two orbits that yield the maximum errors. The homoclinic orbit that leads to the maximum error in $\Delta\mathcal{F}^{(\mathrm{exact})}_{\{y\}\{x\}}$ out of all 4096 orbits is $\{y\} \Rightarrow \bar{0}10100011000110\bar{0}$, for which

$$\Delta\mathcal{F}^{(\mathrm{exact})}_{\{y\}\{x\}} - \Delta\mathcal{F}^{(\mathrm{ref.})}_{\{y\}\{x\}} = 8.08 \times 10^{-8}. \tag{56}$$

Compared to the orbit action itself, $\Delta\mathcal{F}^{(\mathrm{ref.})}_{\{y\}\{x\}} = -466.602\,850\,894\,90$, the relative error is around 1.7×10^{-10}, almost as good as possible due to the presence of interpolation error. This demonstrates the accuracy of Eq. (49).

As for $\Delta\mathcal{F}^{(\mathrm{approx.})}_{\{y\}\{x\}}$, the maximum error emerges for the orbit $\{y\} \Rightarrow \bar{0}111111111111\bar{0}$, for which

$$\Delta\mathcal{F}^{(\mathrm{approx.})}_{\{y\}\{x\}} - \Delta\mathcal{F}^{(\mathrm{ref.})}_{\{y\}\{x\}} = -5.453 \times 10^{-3}, \tag{57}$$

which is well below the error tolerance 1×10^{-2}. Compared to the orbit action itself, $\Delta\mathcal{F}^{(\mathrm{ref.})}_{\{y\}\{x\}} = -628.514\,708\,240\,16$, the relative error is around 8.7×10^{-6}.

VI. CONCLUSIONS

It is possible to construct the complete set of homoclinic orbit relative actions arising from horseshoe-shaped homoclinic tangles in terms of the primitive orbits' relative actions and an exponentially decreasing set of parallelogram-like areas bounded by stable and unstable manifolds. Important constraints exist on the distribution of homoclinic points [48,49], which are imposed by the topology of the homoclinic tangle. This enables an organizational scheme for the orbits by their winding numbers and assigns binary symbolic codes to each of them. The projection operations, P_S and P_U, together with the corresponding symbolic operations, π_S and π_U, link homoclinic points of different winding numbers. Based on a judicious use of the MacKay-Meiss-Percival action principle and mixed projections of all degrees, an exact geometric formula [Eq. (49)] emerges that determines their relative actions in terms of cell areas from a finite region of phase space, which are bounded by manifolds with low curvatures. However, these areas still proliferate at the same rate as the homoclinic points, which become exponentially hard to compute for large iteration numbers N. To overcome this, we made use of the exponential decay of cell areas in the partition trees, and eliminated all small areas that are asymptotically negligible. The exponentially shrinking areas have their origins in the asymptotic foliations of stable and unstable manifolds, and are thus generic to all chaotic systems. The resulting approximate expression [Eq. (55)] relies on a logarithmically reduced amount of information relative to the exact Eq. (49). It gives

the relative actions or orbits in $T_{-1,N}$ using only the areas from $T_{-1,d}$, in exchange for comprising the accuracy by a designated order of magnitude $O(\delta\mathcal{F} = \epsilon\hbar)$.

For semiclassical trace formulas, once the actions are determined to within an appropriate tolerance level such as mentioned above, additional accuracy becomes irrelevant and of no consequence. Straightforward computations of the actions rely on the numerical constructions of orbits, for which the difficulties are twofold. First, in highly chaotic systems, numerical determination of individual long orbits suffer from sensitive dependence on initial errors. Second, the total number of orbits proliferates exponentially rapidly with relevant timescales (the trellis number N in our case). For homoclinic, heteroclinic, and periodic orbits in Hamiltonian chaos with two degrees of freedom, the first difficulty is not fundamental, and solvable in many ways. The second difficulty, addressed in the present article, illustrates in great detail how information entropy vanishes for quantum systems (isolated, bounded, nonmeasured) from the perspective of semiclassical theory. The reduction of information implied by $\hbar$ or any error tolerance criterion produces an exponentially increasing set of output calculations using a slower-than-exponentially (i.e., algebraically) increasing set of input information.

This method has the potential to serve as a generic paradigm for the information reduction of semiclassical calculations of chaotic systems. Although the present work is focused on homoclinic orbit actions, the results can be immediately generalized into broader contexts, such as the evaluation of unstable periodic orbit actions. Such connections are given by Eqs. (27), (38), and (45) in Ref. [12]. These equations convert the evaluation of periodic orbit actions into the calculation of action differences between certain auxiliary homoclinic orbits constructed from the symbolic codes of the periodic orbit. Therefore, upon the determination of homoclinic orbit actions, the determination of periodic orbit actions becomes a simple manipulation of symbolic strings and subtractions within the homoclinic action set, a trivial task that poses no serious difficulties. Therefore, just like the homoclinic orbit actions, the exponentially increasing set of periodic orbit actions is expressible with the same reduced set of cell areas as well. Further extension of the current method concerns the stability exponents of unstable periodic orbits, which is a topic under current investigation.

ACKNOWLEDGMENT

J.L. gratefully acknowledges many inspiring discussions with Akira Shudo during several productive visits to Tokyo Metropolitan University.

APPENDIX A: HOMOCLINIC TANGLE

In this Appendix we illustrate the fundamental concepts and definitions related to homoclinic tangles that are used throughout this article. Consider a two-degree-of-freedom autonomous Hamiltonian system. With energy conservation and applying the standard Poincaré surface of section technique [34], the continuous flow leads to a discrete area-preserving map M on the two-dimensional phase space (q, p). Assume the existence of a hyperbolic fixed point $x = (q_x, p_x)$ under M:

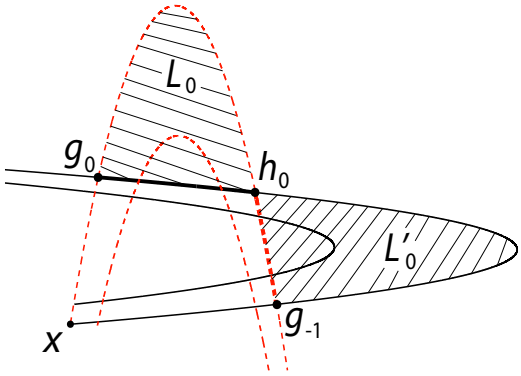

FIG. 14. Trellis $T_{-1,1}$. The fundamental segments U_0 and S'_0 are indicated by thick solid and thick dashed curve segments, respectively. The lobes L_0 and L'_0 (hatched regions) form the turnstile which governs the phase-space transport.

$M(x) = x$. Associated with it are the one-dimensional stable $[S(x)]$ and unstable $[U(x)]$ manifolds, which are the collections of phase-space points that approach x under successive forward and inverse iterations of M, respectively. Typically, $S(x)$ and $U(x)$ intersect infinitely many times and form a complicated pattern named a *homoclinic tangle* [34–36], as partially illustrated in Fig. 1.

Homoclinic tangles have been extensively studied as the organizing structures for classical transport and escape problems [13,35,36,44,49,54–57]. Of particular interest are the homoclinic orbits, which lie along intersections between $S(x)$ and $U(x)$,

$$h_0 = S(x) \bigcap U(x), \tag{A1}$$

whose images under both M and M^{-1} approach x asymptotically: $M^{\pm\infty}(h_0) = h_{\pm\infty} = x$. The bi-infinite collection of images $M^n(h_0) = h_n$ is often referred to as a *homoclinic orbit*

$$\{h_0\} = \{\ldots, h_{-1}, h_0, h_1, \ldots\}. \tag{A2}$$

A *primary homoclinic point*, h_0, arises if the stable and unstable segments, $S[h_0, x]$ and $U[x, h_0]$, intersect only at x and h_0. The resulting closed loop $US[x, h_0] = U[x, h_0] + S[h_0, x]$ is topologically equivalent to a circle. As a result, the phase-space excursion of the *primary homoclinic orbit* $\{h_0\}$ takes the simplest possible form. It "circles" around the loop once from infinite past to infinite future. Figure 1 shows the simplest kind of homoclinic tangle having only two primary homoclinic orbits, $\{h_0\}$ and $\{g_0\}$. In practice, more complicated homoclinic tangles are possible. However, generalizations are straightforward and not considered here.

The entire homoclinic tangle, as an infinite entity, can be constructed from iterations of finite segments on $S(x)$ and $U(x)$. Identifying the *fundamental segments* as

$$\begin{aligned} U_n &\equiv U[h_n, g_n], \quad U'_n \equiv U(g_{n-1}, h_n), \\ S_n &\equiv S(g_n, h_n), \quad S'_n \equiv S[h_n, g_{n-1}], \end{aligned} \tag{A3}$$

$U_{n+k} = M^k(U_n)$, and similarly for U'_n, S_n, and S'_n. Shown in Fig. 14 are examples of U_0 (thick solid segment) and S'_0 (thick dashed segment). The manifolds can be built as

nonoverlapping unions of the respective fundamental segments:

$$U(x) = \bigcup_{n=-\infty}^{\infty} (U_n \cup U_n'),$$
$$S(x) = \bigcup_{n=-\infty}^{\infty} (S_n \cup S_n'), \tag{A4}$$

and likewise for the homoclinic tangle. The topology of a homoclinic tangle contains important dynamical information, and is often studied over its truncations, namely a *trellis* [35,36] defined as

$$T_{n_s,n_u} \equiv \left(\bigcup_{i=n_s}^{\infty}(S_i \cup S_i')\right) \bigcup \left(\bigcup_{i=-\infty}^{n_u} (U_i \cup U_i')\right), \tag{A5}$$

where the integers n_s and n_u give the lower and upper bounds for the indices of the stable and unstable fundamental segments, respectively. For example, the pattern shown in Fig. 14 is $T_{-1,1}$.

For the study of chaotic transport, it is customary to define some special regions inside the homoclinic tangle, which govern the flux in and out of the tangle. Following the conventions [13,36], the phase-space region bounded by loop $US[x, g_0]$ is the *complex* (also referred to as the *resonance zone* by Easton [35]), and the regions bounded by the loops $US[h_n, g_n]$ and $US[g_{n-1}, h_n]$ are *lobes* denoted by L_n and L_n', respectively. The union of lobes L_0 and L_0' is often called a *turnstile* [13], as demonstrated by the hatched regions in Fig. 14.

A simplifying assumption adopted here is the "open system" condition [36,49], which assumes that the lobes L_n' and L_{-n} with $n \geqslant 1$ extend out to infinity as n increases and never enter the complex region. Consequently, there are no homoclinic points distributed on the segments, $S(g_n, h_n)$ and $U(g_{n-1}, h_n)$, which simplifies addressing the homoclinic orbits. However, this restriction is not essential and can be removed to accommodate closed systems as well.

APPENDIX B: SYMBOLIC DYNAMICS

Symbolic dynamics [40–43] is a powerful construct that characterizes the topology of orbits in chaotic systems. In essence, it encodes the trajectories of various initial conditions under the mapping into infinite strings of alphabets, assigned using their phase-space itineraries with respect to a generating Markov partition [58,59]. Constructions of exact generating partitions for general mixed systems, if possible, still remain challenging. However, finite approximations can be obtained via efficient techniques introduced in [60–64].

Assume that the system is highly chaotic and the homoclinic tangle forms a complete Smale horseshoe [37,38], as the one depicted in Fig. 15. The generating partition is then the collection of two regions $[V_0, V_1]$ (marked as hatched regions in the upper panel of the figure), where V_0 is the closed region bounded by $USUS[x, g_{-2}, b^{(0)}, g_0] = U[x, g_{-2}] + S[g_{-2}, b^{(0)}] + U[b^{(0)}, g_0] + S[g_0, x]$, and V_1 is the closed region bounded by $USUS[h_{-1}, g_{-1}, h_0, a^{(0)}]$. Note that the curvy-trapezoid region between V_0 and V_1 is also labeled in the figure as V', which is bounded by loop $USUS[g_{-2}, h_{-1}, a^{(0)}, b^{(0)}]$. The deformation of these regions

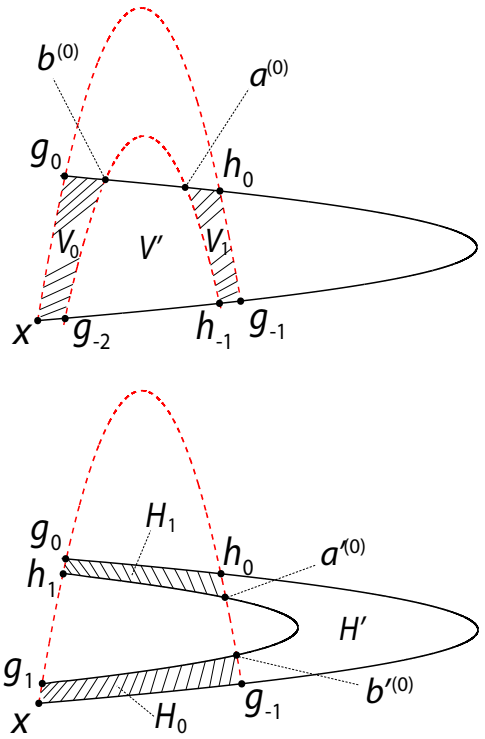

FIG. 15. Smale horseshoe formed by $S(x)$ (red dashed curve) and $U(x)$ (black solid curve). The vertical strips V_0 and V_1 in the upper panel (hatched regions) are the generating partitions of the symbolic dynamics, and are mapped into the horizontal strips H_0 and H_1 in the lower panel (hatched regions), respectively, under one iteration. The fixed point x has symbolic string $\overline{0}.\overline{0}$, and the primary homoclinic points h_0 and g_0 have symbolic strings $\overline{0}1.1\overline{0}$ and $\overline{0}1.\overline{0}$, respectively.

under the dynamics can be visualized in a simple way: under one iteration of M, the curvy-trapezoid region bounded by $USUS[x, g_{-1}, h_0, g_0]$ (the union of V_0, V', and V_1) from the upper panel of Fig. 15 is compressed along its stable boundary and stretched along its unstable boundary while preserving the total area, folded into a U-shaped region bounded by $USUS[x, g_0, h_1, g_1]$ in the lower panel, which is the union of H_0, H_1, and H'. During this process, the vertical strips V_0 and V_1 are mapped into the horizontal strips H_0 and H_1, respectively, marked by the hatched regions in the lower panel of Fig. 15. In the meantime, V' is mapped into the U-shaped region H' bounded by $USUS[g_{-1}, h_0, a'^{(0)}, b'^{(0)}]$ and will escape the complex region under further iterations. The inverse mapping of M^{-1} has similar but reversed effects, with $M^{-1}(H_i) = V_i$ $(i = 0, 1)$.

Under the symbolic dynamics, each point z_0 inside the complex that never escapes under forward and inverse mappings can be put into a one-to-one correspondence with a bi-infinite symbolic string

$$z_0 \Rightarrow \cdots s_{-2}s_{-1}.s_0s_1s_2\cdots, \tag{B1}$$

where each digit s_n indicates the region that $M^n(z_0)$ lies in: $M^n(z_0) = z_n \in V_{s_n}$, where $s_n \in \{0, 1\}$. The position of the decimal point indicates the present location of z_0 since $z_0 \in V_{s_0}$. The symbolic string gives an "itinerary" of z_0 under successive forward and inverse iterations, in terms of the regions V_0

and V_1 in which each iteration lies. The mapping M then corresponds to a Bernoulli shift on symbolic strings composed by "0"s and "1"s,

$$M^n(z_0) \Rightarrow \cdots s_{n-2}s_{n-1}.s_n s_{n+1}s_{n+2}\cdots , \tag{B2}$$

therefore encoding the dynamics with simple strings of integers. Assume a complete horseshoe structure here in which all possible combinations of substrings exist; i.e., no "pruning" [51,52] is needed.

The area-preserving Hénon map [39] is used as a confirmation of the theory and its approximations:

$$p_{n+1} = q_n, \quad q_{n+1} = a - q_n^2 - p_n. \tag{B3}$$

With parameter $a = 10$, it gives rise to a complete horseshoe-shaped homoclinic tangle; see Fig. 15. As it satisfies both the complete horseshoe and open system assumptions, the theory is directly applicable. Nevertheless, the results derived mostly carry over into more complicated systems possessing incomplete horseshoes [53], or systems with more than binary symbolic codes, though more work is needed to address such complications.

The fixed point x has the symbolic string $x \Rightarrow \cdots 0.0\cdots = \overline{0}.\overline{0}$, where the overhead bar denotes infinite repetitions of "0"s since it stays in (on the boundary of) V_0 forever. Consequently, other than the orbit containing the point $\overline{0}1.\overline{0}$, any homoclinic point h of x must have a symbolic string of the form

$$h \Rightarrow \overline{0}1s_{-m}\cdots s_{-1}.s_0 s_1\cdots s_n 1\overline{0}, \tag{B4}$$

along with all possible shifts of the decimal point. The $\overline{0}$ on both ends means the orbit approaches the fixed point asymptotically. The orbit $\{h\}$ can then be represented by the same symbolic string:

$$\{h\} \Rightarrow \overline{0}1s_{-m}\cdots s_{-1}s_0 s_1\cdots s_n 1\overline{0}, \tag{B5}$$

with the decimal point removed, as compared to Eq. (B4). The finite symbolic segment "$1s_{-m}\cdots s_{-1}s_0 s_1\cdots s_n 1$" is often referred to as the *core* of the symbolic code of h, with its length referred to as the *core length*. To be discussed in Appendix C, the core length is a measure of the length of the phase-space excursion of $\{h\}$.

The identification of symbolic strings associated with arbitrary homoclinic points, as well as the ordering of homoclinic points on the fundamental segments S_n' or U_n, are nontrivial tasks in general. Pioneering works along this line can be found in Ref. [45], where the symbolic assignment and relative ordering of homoclinic points on S_0' were explicitly given for the Hénon map. Refer to Fig. 3 of [45] for a nice pictorial demonstration. However, [45] starts from the anti-integrable limit [65,66] and derives the results as continuations of the limit. In Appendix C, we introduce a different analytic scheme, which makes use of the hierarchical structure of the homoclinic tangle (see Sec. IV) to provide the ordering of homoclinic points on S_{-1}' in terms of their symbolic codes. Based on the symbolic codes of the two primary homoclinic points on S_{-1}', which are $h_{-1} \Rightarrow \overline{0}.11\overline{0}$ and $g_{-2} \Rightarrow \overline{0}.01\overline{0}$, it recursively builds up the codes of the more complicated homoclinic orbits by adding certain symbolic strings of finite lengths to the primaries, according to their positions in the hierarchic structure. The results are equivalent to those of [45] upon changing the alphabets "0" → "+" and "1" → "−". This approach naturally facilitates an important accumulation relation (introduced in Sec. IV B) and thus better integrates into the scheme of the present work.

APPENDIX C: SYSTEMATIC ASSIGNMENTS OF SYMBOLIC CODES

Although the symbolic codes of some simple homoclinic orbits, such as the primary ones, can be easily determined by following the numerical orbits, such tasks become prohibitive for the exponentially proliferating ensemble of more complicated, nonprimary orbits. In addition, a computational method does not reveal the patterns and structural relations buried in substrings of the symbolic codes. In fact, as shown by [45], symbolic codes provide a natural ordering of homoclinic points along the fundamental segments, which is otherwise unattainable from numerical methods. Although this problem is essentially solved by [45] for the Hénon maps in the complete horseshoe region, their approach starts from the anti-integrable limit [65,66], and identifies each homoclinic orbit near the limit as continuations from the anti-integrable limit. Although exact and efficient, it does not make use of the accumulation relations (Sec. IV B), which are the theoretical foundations of the present paper. This Appendix introduces a different approach. Taking advantage of the hierarchical structure of the homoclinic orbits (see Sec. IV), a recursive scheme is introduced that systematically determines the symbolic codes of the families of winding-$(n+1)$ homoclinic orbits based on the symbolic code of the winding-n orbit on which they accumulate. It results in an ordering of homoclinic points on the fundamental segment S_{-1}' in terms of their symbolic codes, which is equivalent to Lemma 7 of [45] upon switching the alphabets "0" → "+" and "1" → "−". This provides a foundation for the exact relations and approximations of Sec. V.

Every homoclinic orbit has one and only one representative point on S_{-1}' and labeling the entire set of orbits can be reduced to labeling the homoclinic points on S_{-1}'. Starting from $T_{-1,-1}$, in which S_{-1}' is not intersected by any unstable fundamental segment, the only homoclinic points are the primaries $h_{-1} \Rightarrow \overline{0}.11\overline{0}$ and $g_{-2} \Rightarrow \overline{0}.01\overline{0}$, both of which are winding-1. Proceeding to the intersections of S_{-1}' with $T_{-1,0}$, there are two winding-2 points, $a^{(0)}$ and $b^{(0)}$, as shown by Fig. 16, which are the leading terms of the two winding-2 families $[a^{(n)}]$ and $[b^{(n)}]$ from the future $T_{-1,n}$ that accumulate on g_{-2}. Their symbolic codes are $a^{(0)} \Rightarrow \overline{0}1.11\overline{0}$ and $b^{(0)} \Rightarrow \overline{0}1.01\overline{0}$, which emerge quickly by following their excursions. The hierarchical relationship at this stage can be denoted alternatively as

$$\begin{aligned}(a^{(0)} \Rightarrow \overline{0}1.11\overline{0}) &\overset{1}{\underset{S}{\hookrightarrow}} (g_{-2} \Rightarrow \overline{0}.01\overline{0}),\\ (b^{(0)} \Rightarrow \overline{0}1.01\overline{0}) &\overset{1}{\underset{S}{\hookrightarrow}} (g_{-2} \Rightarrow \overline{0}.01\overline{0}),\end{aligned} \tag{C1}$$

where the notations "$\overset{1}{\underset{S}{\hookrightarrow}}$" are defined in Eq. (11). Notice that the hierarchical relations imply the symbolic code assignments: the codes of $a^{(0)}$ and $b^{(0)}$ can be obtained by adding the substrings "11" and "10", respectively, to the left end of the core of g_{-2}, while maintaining the position of the decimal

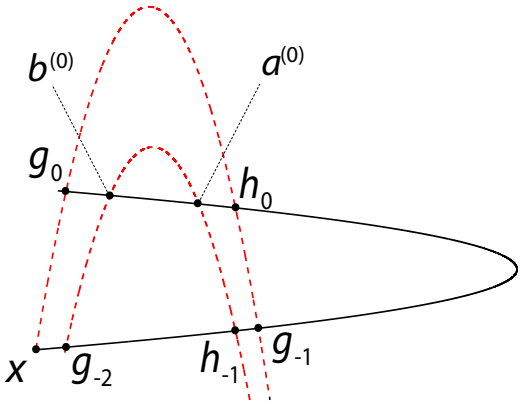

FIG. 16. Homoclinic points in $T_{-1,0}$. The symbolic codes are $h_{-1} \Rightarrow \bar{0}.11\bar{0}$, $g_{-2} \Rightarrow \bar{0}.01\bar{0}$, $a^{(0)} \Rightarrow \bar{0}1.11\bar{0}$, and $b^{(0)} \Rightarrow \bar{0}1.01\bar{0}$. The hierarchical relations are $a^{(0)}, b^{(0)} \overset{1}{\underset{S}{\hookrightarrow}} g_{-2}$. Notice that the hierarchical relations are indicative for the assignments of symbolic codes: the codes of $a^{(0)}$ and $b^{(0)}$ can be obtained by adding the substrings "11" and "10", respectively, to the left end of the core of g_{-2}, while maintaining the position of the decimal point relative to the right end of the core.

point relative to the right end of the core. Also, the transit times of $a^{(0)}$ and $b^{(0)}$ are both unity, and their core lengths are both 3. It turns out in general that

$$\textit{core length} = \textit{transit time} + 2, \tag{C2}$$

which holds true for all nonprimary homoclinic points. Another important observation is that $a^{(0)}$ and $b^{(0)}$ with core lengths 3 emerged from $S'_{-1} \cap U_0$ in trellis $T_{-1,0}$. This leads to the simple fact that any nonprimary homoclinic point that emerges from $S'_{-1} \cap U_n$ in $T_{-1,n}$ must have core length $n + 3$.

There are four new intersections generated by $T_{-1,1}$, i.e., $S'_{-1} \cap U_1$. Figure 17 shows the four new winding-2 points labeled $a^{(1)}$, $b^{(1)}$, $c^{(1)}$, and $d^{(1)}$. An important distinction between them is that $a^{(1)}$ and $b^{(1)}$ are the second realizations of their respective families $[a^{(n)}]$ and $[b^{(n)}]$ $(n \geqslant 0)$ that accumulate on g_{-2}, whereas $c^{(1)}$ and $d^{(1)}$ are the first terms of their

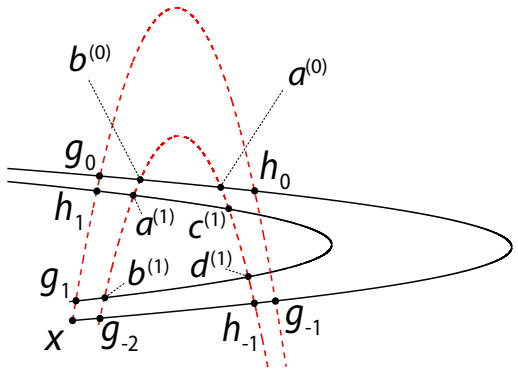

FIG. 17. Homoclinic points in $T_{-1,1}$. The symbolic codes are $a^{(1)} \Rightarrow \bar{0}11.01\bar{0}$, $b^{(1)} \Rightarrow \bar{0}10.01\bar{0}$, $c^{(1)} \Rightarrow \bar{0}11.11\bar{0}$, and $d^{(1)} \Rightarrow \bar{0}10.11\bar{0}$. The hierarchical relations are $a^{(1)}, b^{(1)} \overset{2}{\underset{S}{\hookrightarrow}} g_{-2}$ and $c^{(1)}, d^{(1)} \overset{1}{\underset{S}{\hookrightarrow}} h_{-1}$.

respective families, $[c^{(n)}]$ and $d^{(n)}$ $(n \geqslant 1)$, that accumulate on h_{-1}. Therefore, following the pattern of Eq. (C1), the symbolic codes of $c^{(1)}$ and $d^{(1)}$ should be obtained by adding the substrings "11" and "10", respectively, to the left end of the core of h_{-1} (which is "11"), while keeping the position of the decimal point relative to the right end of the core. This leads to the assignments $c^{(1)} \Rightarrow \bar{0}11.11\bar{0}$ and $d^{(1)} \Rightarrow \bar{0}10.11\bar{0}$ according to the hierarchical relations

$$\begin{aligned} (c^{(1)} \Rightarrow \bar{0}11.11\bar{0}) &\overset{1}{\underset{S}{\hookrightarrow}} (h_{-1} \Rightarrow \bar{0}.11\bar{0}), \\ (d^{(1)} \Rightarrow \bar{0}10.11\bar{0}) &\overset{1}{\underset{S}{\hookrightarrow}} (h_{-1} \Rightarrow \bar{0}.11\bar{0}). \end{aligned} \tag{C3}$$

As for the symbolic codes of $a^{(1)}$ and $b^{(1)}$, since they are the second terms in their respective accumulating families, the substrings "110" and "100", instead of "11" and "10", should be added to the left end of the core of g_{-2}, respectively, while keeping the position of the decimal point relative to the right end of the core unchanged:

$$\begin{aligned} (a^{(1)} \Rightarrow \bar{0}11.01\bar{0}) &\overset{2}{\underset{S}{\hookrightarrow}} (g_{-2} \Rightarrow \bar{0}.01\bar{0}), \\ (b^{(1)} \Rightarrow \bar{0}10.01\bar{0}) &\overset{2}{\underset{S}{\hookrightarrow}} (g_{-2} \Rightarrow \bar{0}.01\bar{0}). \end{aligned} \tag{C4}$$

Calculating the orbits numerically, one readily verifies that Eqs. (C3) and (C4) indeed give the correct desired symbolic codes for the orbits.

Generalization of the above relations gives the general rule for the assignment of symbolic codes. Given an arbitrary winding-m homoclinic point y, and two winding-$(m+1)$ homoclinic points z and w from the two winding-$(m+1)$ families accumulating on y, such that $z \overset{k}{\underset{S}{\hookrightarrow}} y$ and $w \overset{k}{\underset{S}{\hookrightarrow}} y$ $(k \geqslant 1)$ and $S[y, w] \subset S[y, z]$, then the symbolic codes of z and w can be obtained by adding the substrings "110^{k-1}" and "100^{k-1}", respectively, to the left end of the core of y, keeping the position of the decimal point relative to the right end of the core. The notation "0^{k-1}" denotes a string composed of $(k-1)$ consecutive "0"s. Or equivalently, let the symbolic code of the orbit $\{y\}$ be $\{y\} \Rightarrow \bar{0}\tilde{s}\bar{0}$, where the string $\tilde{s}$ denotes the core; then the symbolic codes of orbits $\{z\}$ and $\{w\}$ are determined as

$$\begin{aligned} \{z\} &\Rightarrow \bar{0}110^{k-1}\tilde{s}\bar{0}, \\ \{w\} &\Rightarrow \bar{0}100^{k-1}\tilde{s}\bar{0}, \end{aligned} \tag{C5}$$

and the position of the decimal points in the symbolic codes of z and w are identical to that of y, when counted from the right ends of their cores.

Concrete examples of the preceding assignment rules are labeled in Fig. 7. Choose the winding-2 point $a^{(0)} \Rightarrow \bar{0}1.11\bar{0}$ as the base, and notice the accumulating points $e^{(k)}, f^{(k)} \overset{k}{\underset{S}{\hookrightarrow}} a^{(0)}$, where the $k = 1$ case is explicitly shown in the figure. According to the preceding assignment rules, the symbolic codes of $e^{(k)}$ and $f^{(k)}$ are constructed as $f^{(k)} \Rightarrow \bar{0}110^{k-1}1.11\bar{0}$ and $e^{(k)} \Rightarrow \bar{0}100^{k-1}1.11\bar{0}$, which was verified numerically.

The proof of Eq. (C5) involves mapping the base point y simultaneously with z and w forward and inversely, to study the deformation of $S[y, z/w]$ under forward iterations, and the

deformation of $U[y, z/w]$ under inverse iterations. Notice that the stable segments $S[y, z/w]$ belong to either $S[g_{-2}, b^{(0)}]$ or $S[h_{-1}, a^{(0)}]$, which will become even shorter under forward iterations. Therefore, forward iterations of y and z/w are guaranteed to locate on the same side of S'_{-1}, thus in the same generating partition (V_0 or V_1). For the inverse mappings, the unstable segments $U[y, z/w]$ are constrained to deform in a specific way such that the images of y and z/w must locate in the same partition along the code segment "$0^{k-1}\tilde{s}$" first. After that, the backward images of z immediately visit V_1 twice, then stay in V_0 as they approach x; on the contrary, the backward images of w visit V_0 and V_1 consecutively, and then stay in V_0 as they approach x. The slight difference in their behaviors gives rise to the "$\bar{0}11$" and "$\bar{0}10$" in their respective symbolic codes in Eq. (C5). The detailed derivation is quite lengthy and skipped here for brevity.

With Eq. (C5), the complete set of symbolic codes is generated based on just the symbolic codes of the two primary orbits. For a finite trellis $T_{-1,N}$ (presumably with large N), the maximum transition time of homoclinic orbits is $N+1$; i.e., those arise from $S'_{-1} \cap U_N$. According to Eq. (C2), the corresponding maximum core length is $N+3$. Therefore, starting from $\{h_{-1}\} \Rightarrow \bar{0}11\bar{0}$ and $\{g_{-2}\} \Rightarrow \bar{0}1\bar{0}$, by intersecting S'_{-1} with successive U_i where $0 \leqslant i \leqslant N$ and recursive use of Eq. (C5) up to core length $N+3$, the symbolic codes of all homoclinic orbits present in $T_{-1,N}$ are generated according to the relative positions of their representative points on S'_{-1}. This process is equivalent to the $>_s$ ordering in Lemma 7 of [45].

A similar prescription could have been generated for the accumulating homoclinic families along the unstable manifold under inverse mappings. Given any winding-n homoclinic point y', and two winding-$(n+1)$ homoclinic points z' and w' such that $z' \underset{U}{\overset{k}{\hookrightarrow}} y'$, $w' \underset{U}{\overset{k}{\hookrightarrow}} y'$ $(k \geqslant 1)$ and $U[y', w'] \subset U[y', z']$, the symbolic codes of z' and w' can be constructed by adding the substrings "$0^{k-1}11$" and "$0^{k-1}01$", respectively, to the right end of the core of the symbolic code of y', while keeping the position of the decimal point relative to the left end of the core unchanged. Or equivalently, if we let the symbolic code of the orbit be $\{y'\} \Rightarrow \bar{0}\tilde{s}'\bar{0}$ where $\tilde{s}'$ denotes the core, then the symbolic codes of orbits $\{z'\}$ and $\{w'\}$ are constructed as

$$\begin{aligned} \{z'\} &\Rightarrow \bar{0}\tilde{s}'0^{k-1}11\bar{0}, \\ \{w'\} &\Rightarrow \bar{0}\tilde{s}'0^{k-1}01\bar{0}, \end{aligned} \tag{C6}$$

which is in complete analogy to Eq. (C5), and equivalent to the $>_u$ ordering in Lemma 7 of [45]. For example, in Fig. 3 we have $g_0 \Rightarrow \bar{0}1.\bar{0}$, and $v^{(-k)}, w^{(-k)} \underset{U}{\overset{k}{\hookrightarrow}} g_0$, where the $k = 1, 2$ cases are explicitly shown in the figure. Then according to the preceding rules, the symbolic codes of $v^{(-k)}$ and $w^{(-k)}$ are constructed from the symbolic code of g_0 as $v^{(-k)} \Rightarrow \bar{0}1.0^{k-1}11\bar{0}$ and $w^{(-k)} \Rightarrow \bar{0}1.0^{k-1}01\bar{0}$, respectively.

APPENDIX D: ASYMPTOTIC ACCUMULATION EXPONENT

The foundation of Sec. IV B is established by Lemma 2 in Appendix B 3 of [49], and a brief overview of their results is given here. The setting of the lemma is demonstrated schematically by Fig. 18. Let z_u be an arbitrary point on $U(x)$, and $\bar{\mathcal{C}}$ an arbitrary differentiable curve passing transversely through $U(x)$ at z_u. Consider another arbitrary differentiable curve, $\mathcal{C}_0$, which passes through $S(x)$ transversely at $r^{(0)}$, and intersects $\bar{\mathcal{C}}$ at $z^{(0)}$. Then, its future iterations $\mathcal{C}_n = M^n(\mathcal{C}_0)$ $(n \geqslant 1)$ pass through $S(x)$ transversely at $r^{(n)}$, and intersect $\bar{\mathcal{C}}$ at $z^{(n)}$, which form a family of points $[z^{(n)}]$ that accumulate asymptotically on the base point z_u:

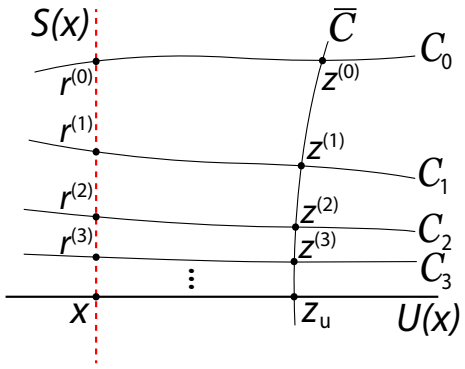

FIG. 18. (Schematic) Iterates of a curve intersecting the stable manifold approach the unstable manifold. Future iterations of the curve $\mathcal{C}_0$ create a family of curves $[\mathcal{C}_n]$, which intersect $\bar{\mathcal{C}}$ at a family of points $[z^{(n)}]$. $[z^{(n)}]$ accumulates on z_u under the exponent μ_x, as given by Eq. (D1).

$$\begin{aligned} &\lim_{n\to\infty} z^{(n)} = z_u, \\ &\lim_{n\to\infty} |z^{(n)} - z_u| e^{n\mu_x} = C(z_u, z^{(0)}), \end{aligned} \tag{D1}$$

where $||$ is the standard Euclidean vector norm, μ_x is the stability exponent of x, and $C(z_u, z^{(0)})$ is a positive constant depending on the base point z_u and the leading term $z^{(0)}$ in the asymptotic family. Notice that Eq. (D1) is just a reexpression of Eqs. (B5) and (B6) of [49]. What is surprising here is that even though the manifolds explore the vast majority of phase space with clearly nonuniform expansion rates, the asymptotic exponent in the above equation is still that of the hyperbolic fixed point.

APPENDIX E: AREA CORRESPONDENCE RELATIONS

Given any homoclinic point $y \in (S'_{-1} \cap U_m)$, there is an explicit relation that links $\mathcal{A}^\circ(y) = \mathcal{A}^\circ_{SUSU[y,P_S(y),P_SP_U(y),P_U(y)]}$ with specific linear combinations of cell areas from the type-I and type-II partition trees of $T_{-1,m}$. The transition time of y is $m+1$, so its core length is $m+3$. Let $\tilde{s} = s_1 s_2 \cdots s_{m+2} s_{m+3}$ $(s_i \in \{0, 1\},\ s_1 = s_{m+3} = 1)$ be the core of the symbolic code of y; then the linear combination of cell areas depends solely on $\tilde{s}$. The correspondence relation is established in the following step.

(1) Define $\Phi_{B\mapsto A}$ to be a mapping from the cells of the type-II partition trees to the cells of the type-I partition trees, such that for any finite Greek alphabet string $\tilde{\omega}$ composed of α, γ, and β ($\tilde{\omega}$ could also be an empty string) we have

$$\Phi_{B\mapsto A}(B_{\tilde{\omega}}) = A_{\tilde{\omega}}, \quad \Phi_{B\mapsto A}(A_{\tilde{\omega}}) = \varnothing, \quad \Phi_{B\mapsto A}(\varnothing) = \varnothing, \tag{E1}$$

where $\varnothing$ denotes a null cell that gives zero contribution to the action calculations.

(2) Define $\Phi_{\beta\mapsto\alpha}$ to be a mapping between the cells of the partition trees, such that for any finite Greek alphabet string $\tilde{\omega}$ composed of α and γ (but not β; note also that $\tilde{\omega}$ could be an empty string), we have

$$\begin{aligned}
&\Phi_{\beta\mapsto\alpha}(A_{\tilde{\omega}\beta}) = A_{\tilde{\omega}\alpha},\\
&\Phi_{\beta\mapsto\alpha}(B_{\tilde{\omega}\beta}) = B_{\tilde{\omega}\alpha},\\
&\Phi_{\beta\mapsto\alpha}(A_{\tilde{\omega}\alpha}) = \Phi_{\beta\mapsto\alpha}(B_{\tilde{\omega}\alpha}) = \varnothing,
\end{aligned} \tag{E2}$$

$$\begin{aligned}
&\Phi_{\beta\mapsto\alpha}(A) = \Phi_{\beta\mapsto\alpha}(B) = \varnothing,\\
&\Phi_{\beta\mapsto\alpha}(\varnothing) = \varnothing.
\end{aligned}$$

(3) Define Γ to be a mapping from the core $\tilde{s} = s_1 s_2 \cdots s_{m+2} s_{m+3}$ ($s_1 = s_{m+3} = 1$) of the symbolic code of any nonprimary homoclinic point $y \in (S'_{-1} \cap U_m)$ to the cells of the partition trees, such that depending on the detailed forms of $\tilde{s}$, the mapping Γ takes the forms

$$\Gamma(\tilde{s}) = \begin{cases}
\Gamma(101) = A,\\
\Gamma(111) = B,\\
\Gamma\ \underbrace{(10}\ \cdots\ \underbrace{0}\ \cdots\ \underbrace{1}\ \cdots\ \underbrace{01)} = A_{\cdots\gamma\cdots\alpha\cdots\alpha},\\
\quad\ \downarrow\quad\ \downarrow\quad\ \downarrow\quad\ \downarrow\\
\quad\ \alpha\ \cdots\ \alpha\ \cdots\ \gamma\ \cdots\ A\\
\Gamma\ \underbrace{(10}\ \cdots\ \underbrace{0}\ \cdots\ \underbrace{1}\ \cdots\ \underbrace{11)} = B_{\cdots\gamma\cdots\alpha\cdots\alpha},\\
\quad\ \downarrow\quad\ \downarrow\quad\ \downarrow\quad\ \downarrow\\
\quad\ \alpha\ \cdots\ \alpha\ \cdots\ \gamma\ \cdots\ B\\
\Gamma\ \underbrace{(11}\ \cdots\ \underbrace{0}\ \cdots\ \underbrace{1}\ \cdots\ \underbrace{01)} = A_{\cdots\gamma\cdots\alpha\cdots\beta},\\
\quad\ \downarrow\quad\ \downarrow\quad\ \downarrow\quad\ \downarrow\\
\quad\ \beta\ \cdots\ \alpha\ \cdots\ \gamma\ \cdots\ A\\
\Gamma\ \underbrace{(11}\ \cdots\ \underbrace{0}\ \cdots\ \underbrace{1}\ \cdots\ \underbrace{11)} = B_{\cdots\gamma\cdots\alpha\cdots\beta},\\
\quad\ \downarrow\quad\ \downarrow\quad\ \downarrow\quad\ \downarrow\\
\quad\ \beta\ \cdots\ \alpha\ \cdots\ \gamma\ \cdots\ B
\end{cases} \tag{E3}$$

in which the $\tilde{s} = 101$ and $\tilde{s} = 111$ cases yield cells A and B, respectively, and all the rest of the cases with core lengths $\geqslant 4$ (or equivalently $m \geqslant 1$) are categorized into four cases, $\{s_2 = 0, s_{m+2} = 0\}$, $\{s_2 = 0, s_{m+2} = 1\}$, $\{s_2 = 1, s_{m+2} = 0\}$, and $\{s_2 = 1, s_{m+2} = 1\}$, which correspond to the third, fourth, fifth, and sixth lines of Eq. (E3), respectively. Notice in those four cases, the letters A and B of the cell names are given by the last two digits $s_{m+2}s_{m+3}$ of $\tilde{s}$ with grammar "$01 \mapsto A$" and "$11 \mapsto B$". The Greek alphabet string of the cell names are given by the first $m+1$ digits of $\tilde{s}$ in a reversed order: $s_1 s_2$ gives the last alphabet in the Greek string, with grammar "$10 \mapsto \alpha$" and "$11 \mapsto \beta$", and $s_{m+1}s_m \cdots s_4 s_3$ (reversed string of $s_3 s_4 \cdots s_m s_{m+1}$) gives the first $m-1$ alphabets in the Greek string, with grammar "$0 \mapsto \alpha$" and "$1 \mapsto \gamma$".

(4) Finally, $\mathcal{A}^\circ(y)$ can be calculated as

$$\begin{aligned}
\mathcal{A}^\circ(y) = (-1)^{n_\gamma(\Gamma(\tilde{s}))} \cdot [&\Gamma(\tilde{s}) + \Phi_{\beta\mapsto\alpha}(\Gamma(\tilde{s}))\\
&+ \Phi_{B\mapsto A}(\Gamma(\tilde{s})) + \Phi_{B\mapsto A}(\Phi_{\beta\mapsto\alpha}(\Gamma(\tilde{s})))],
\end{aligned} \tag{E4}$$

where $n_\gamma(\Gamma(\tilde{s}))$ is a function that returns the total number of γ in the Greek alphabet string of the cell $\Gamma(\tilde{s})$. For example, $n_\gamma(A_{\alpha\beta}) = 0$ and $n_\gamma(B_{\gamma\beta}) = 1$. Again, we emphasize that Eq. (E4) only applies to nonprimary homoclinic points y located on S'_{-1}.

Equation (E4) gives a systematic way of identifying the $\mathcal{A}^\circ(y)$ term in the homoclinic action decomposition [Eq. (35)] in terms of a linear combination of cell areas from the type-I and type-II partition trees. In practice, some of the terms in Eq. (E4) will vanish due to the presence of null areas ($\varnothing$) in Eqs. (E1) and (E2). Depending on y, Eq. (E4) may take four possible forms, as listed here:

(1) A single type-I cell area: A (for $y = b^{(0)}$ only) or $A_{\tilde{\omega}\alpha}$, where $\tilde{\omega}$ denotes some Greek alphabet string composed by α and γ. Examples are

(i) In Fig. 16, let $y = b^{(0)} \in (S'_{-1} \cap U_0)$, then

$$\mathcal{A}^\circ(y) = A,$$

which is a type-I cell of $T_{-1,0}$.

(ii) In Fig. 4, let $y = b^{(1)} \in (S'_{-1} \cap U_1)$, then

$$\mathcal{A}^\circ(y) = A_\alpha,$$

which is a type-I cell of $T_{-1,1}$.

(iii) In Fig. 7, let $y = r^{(1)} \in (S'_{-1} \cap U_2)$, then

$$\mathcal{A}^\circ(y) = -A_{\gamma\alpha},$$

which is a type-I cell area of $T_{-1,2}$.

(2) Two type-I areas: $A_{\tilde{\omega}\beta} + A_{\tilde{\omega}\alpha}$. Examples are

(i) In Fig. 4, let $y = a^{(1)} \in (S'_{-1} \cap U_1)$, then

$$\mathcal{A}^\circ(y) = A_\beta + A_\alpha,$$

which is the sum of two type-I areas of $T_{-1,1}$.

(ii) In Fig. 7, let $y = s^{(1)} \in (S'_{-1} \cap U_2)$, then

$$\mathcal{A}^\circ(y) = -(A_{\gamma\beta} + A_{\gamma\alpha}),$$

which is the sum of two type-I areas of $T_{-1,2}$.

(3) A type-I area and a type-II area: $A + B$ (for $y = a^{(0)}$ only) or $A_{\tilde{\omega}\alpha} + B_{\tilde{\omega}\alpha}$. Examples are

(i) In Fig. 16, let $y = a^{(0)} \in (S'_{-1} \cap U_0)$. Recall that only for the special case of $y = a^{(0)}$, we alter Eq. (35) into Eq. (39), whose area term gives

$$\mathcal{A}^{\circ}_{SUSU[a^{(0)},h_{-1},x,g_0]} = A + B,$$

which is the sum of a type-I and a type-II area of $T_{-1,0}$.

(ii) In Fig. 4, let $y = d^{(1)} \in (S'_{-1} \cap U_1)$, then

$$\mathcal{A}^{\circ}(y) = A_\alpha + B_\alpha,$$

which is the sum of a type-I and a type-II area.

(iii) In Fig. 7, let $y = e^{(1)} \in (S'_{-1} \cap U_2)$, then

$$\mathcal{A}^{\circ}(y) = -(A_{\gamma\alpha} + B_{\gamma\alpha}),$$

which is the sum of a type-I and a type-II area of $T_{-1,2}$.

(4) Two type-I areas plus two type-II areas: $A_{\tilde{\omega}\alpha} + A_{\tilde{\omega}\beta} + B_{\tilde{\omega}\alpha} + B_{\tilde{\omega}\beta}$. Examples are

(i) In Fig. 4, let $y = c^{(1)} \in (S'_{-1} \cap U_1)$, then

$$\mathcal{A}^{\circ}(y) = A_\alpha + A_\beta + B_\alpha + B_\beta,$$

which is the sum of two type-I and two type-II areas of $T_{-1,1}$.

(ii) In Fig. 7, let $y = f^{(1)} \in (S'_{-1} \cap U_2)$, then

$$\mathcal{A}^{\circ}(y) = -(A_{\gamma\alpha} + A_{\gamma\beta} + B_{\gamma\alpha} + B_{\gamma\beta}),$$

which is the sum of two type-I and two type-II areas of $T_{-1,2}$.

[1] P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay, *Chaos: Classical and Quantum* (Niels Bohr Institute, Copenhagen, 2016).
[2] P. So, Scholarpedia **2**, 1353 (2007).
[3] M. C. Gutzwiller, J. Math. Phys. **12**, 343 (1971), and references therein.
[4] M. L. Du and J. B. Delos, Phys. Rev. A **38**, 1896 (1988).
[5] M. L. Du and J. B. Delos, Phys. Rev. A **38**, 1913 (1988).
[6] S. Tomsovic and E. J. Heller, Phys. Rev. E **47**, 282 (1993).
[7] G. D. Birkhoff, Acta Math. **50**, 359 (1927).
[8] J. Moser, Commun. Pure Appl. Math. **9**, 673 (1956).
[9] G. L. da Silva Ritter, A. M. Ozorio de Almeida, and R. Douady, Physica D **29**, 181 (1987).
[10] A. M. Ozorio de Almeida, Nonlinearity **2**, 519 (1989).
[11] J. Li and S. Tomsovic, Phys. Rev. E **95**, 062224 (2017).
[12] J. Li and S. Tomsovic, Phys. Rev. E **97**, 022216 (2018).
[13] R. S. MacKay, J. D. Meiss, and I. C. Percival, Physica D **13**, 55 (1984).
[14] J. D. Meiss, Rev. Mod. Phys. **64**, 795 (1992).
[15] E. J. Doedel and M. J. Friedman, J. Comput. Appl. Math. **26**, 155 (1989).
[16] W. J. Beyn, IMA J. Numer. Anal. **9**, 379 (1990).
[17] G. Moore, IMA J. Numer. Anal. **15**, 245 (1995).
[18] J. Li and S. Tomsovic, J. Phys. A: Math. Theor. **50**, 135101 (2017).
[19] A. A. Brudno, Russ. Math. Surv. **33**, 197 (1978).
[20] V. M. Alekseev and M. V. Yakobson, Phys. Rep. **75**, 287 (1981).
[21] A. N. Kolmogorov, Dokl. Russ. Acad. Sci. **119**, 861 (1958).
[22] A. N. Kolmogorov, Dokl. Russ. Acad. Sci. **124**, 754 (1959).
[23] Y. G. Sinai, Dokl. Russ. Acad. Sci. **124**, 768 (1959).
[24] Y. B. Pesin, Russ. Math. Surv. **32**, 55 (1977).
[25] P. Gaspard and G. Nicolis, Phys. Rev. Lett. **65**, 1693 (1990).
[26] A. Connes, H. Narnhofer, and W. Thirring, Commun. Math. Phys. **112**, 691 (1987).
[27] R. Alicki and M. Fannes, Lett. Math. Phys. **32**, 75 (1994).
[28] G. Lindblad, in *Quantum Probability and Applications*, edited by L. Accardi and W. von Waldenfels (Springer, Berlin, 1988), pp. 183–191, Vol. III.
[29] P. Cvitanović, Chaos **2**, 1 (1992).
[30] P. Cvitanović, Phys. Rev. Lett. **61**, 2729 (1988).
[31] P. Cvitanović and B. Eckhardt, Phys. Rev. Lett. **63**, 823 (1989).
[32] E. B. Bogomolny, Chaos **2**, 5 (1992).
[33] L. Kaplan, Phys. Rev. Lett. **81**, 3371 (1998).
[34] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, Vol. 3 (Gauthier-Villars et fils, Paris, 1899).
[35] R. W. Easton, Trans. Am. Math. Soc. **294**, 719 (1986).
[36] V. Rom-Kedar, Physica D **43**, 229 (1990).
[37] S. Smale, *Differential and Combinatorial Topology*, edited by S. S. Cairns (Princeton University Press, Princeton, 1963).
[38] S. Smale, *The Mathematics of Time: Essays on Dynamical Systems, Economic Processes and Related Topics* (Springer-Verlag, New York, 1980).
[39] M. Hénon, Commun. Math. Phys. **50**, 69 (1976).
[40] J. Hadamard, J. Math. Pures Appl. Series 5 **4**, 27 (1898).
[41] G. D. Birkhoff, *A.M.S. Coll. Publications*, Vol. 9 (American Mathematical Society, Providence, 1927).
[42] G. D. Birkhoff, Mem. Pont. Acad. Sci. Novi Lyncaei **1**, 85 (1935).
[43] M. Morse and G. A. Hedlund, Amer. J. Math. **60**, 815 (1938).
[44] S. Wiggins, *Chaotic Transport in Dynamical Systems* (Springer, New York, 1992).
[45] D. Sterling, H. R. Dullin, and J. D. Meiss, Physica D **134**, 153 (1999).
[46] E. Tabacman, Physica D **85**, 548 (1995).
[47] K. Hockett and P. Holmes, Ergodic Theory Dyn. Syst. **6**, 205 (1986).
[48] D. Bevilaqua and M. Basílio de Matos, Physica D **145**, 13 (2000).
[49] K. A. Mitchell, J. P. Handley, B. Tighe, J. B. Delos, and S. K. Knudson, Chaos **13**, 880 (2003).
[50] J. Li and S. Tomsovic (unpublished).
[51] P. Cvitanović, G. H. Gunaratne, and I. Procaccia, Phys. Rev. A **38**, 1503 (1988).
[52] P. Cvitanović, Physica D **51**, 138 (1991).
[53] R. Hagiwara and A. Shudo, J. Phys. A **37**, 10521 (2004).
[54] K. A. Mitchell, J. P. Handley, J. B. Delos, and S. K. Knudson, Chaos **13**, 892 (2003).
[55] K. A. Mitchell and J. B. Delos, Physica D **221**, 170 (2006).
[56] J. Novick, M. L. Keeler, J. Giefer, and J. B. Delos, Phys. Rev. E **85**, 016205 (2012).
[57] J. Novick and J. B. Delos, Phys. Rev. E **85**, 016206 (2012).

[58] R. Bowen, *Lect. Notes in Math.*, Vol. 470 (Springer-Verlag, Berlin, 1975).
[59] P. Gaspard, *Chaos, Scattering and Statistical Mechanics* (Cambridge University Press, Cambridge, 1998).
[60] P. Grassberger and H. Kantz, Phys. Lett. A **113**, 235 (1985).
[61] F. Christiansen and A. Politi, Phys. Rev. E **51**, R3811 (1995).
[62] F. Christiansen and A. Politi, Nonlinearity **9**, 1623 (1996).
[63] F. Christiansen and A. Politi, Physica D **109**, 32 (1997).
[64] N. Rubido, C. Grebogi, and M. S. Baptista, Chaos **28**, 033611 (2018).
[65] S. Aubry and G. Abramovici, Physica D **43**, 199 (1990).
[66] S. Aubry, Physica D **86**, 284 (1995).