Neighborhoods of periodic orbits and the stationary distribution of a noisy chaotic system

Jeffrey M. Heninger,1 Domenico Lippolis,2,* and Predrag Cvitanović3
1Department of Physics, University of Texas, Austin, Texas 78712, USA
2Institute for Advanced Study, Tsinghua University, Beijing 100084, China
3Center for Nonlinear Science and School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

(Received 17 July 2015; published 23 December 2015)

The finest state-space resolution that can be achieved in a physical dynamical system is limited by the presence of noise. In the weak-noise approximation, the stochastic neighborhoods of deterministic periodic orbits can be computed from distributions stationary under the action of a local Fokker-Planck operator and its adjoint. We derive explicit formulas for widths of these distributions in the case of chaotic dynamics, when the periodic orbits are hyperbolic. The resulting neighborhoods form a basis for functions on the attractor. The global stationary distribution, needed for calculation of long-time expectation values of observables, can be expressed in this basis.

DOI: 10.1103/PhysRevE.92.062922 PACS number(s): 05.45.—a, 45.10.Db, 45.50.Pk, 47.11.—j

I. INTRODUCTION

This paper investigates the interplay of deterministic chaotic dynamics and weak stochastic noise, and proposes a definition of the neighborhood of a noisy hyperbolic state-space point. Such neighborhoods are conjectured to partition the state space in an optimal way and provide a basis function set for the evaluation of the stationary distribution.

A. Width of a noisy trajectory

The basic idea of a stochastic “neighborhood” is that the balance between the noise broadening of a trajectory and the deterministic contraction leads to a probability distribution of finite width, as opposed to one that spreads with time (diffusion only). For an orbit that converges to a linearly stable, attractive equilibrium, this neighborhood was computed in 1810 by Laplace [1,2] and is today known as a solution to the Lyapunov equation [3], or the Ornstein-Uhlenbeck process [4]: for a one-dimensional flow, the deterministic equilibrium point is smeared into a Gaussian probability density centered on it, whose covariance \( Q = -D/\lambda \) is a balance of the expansion rate \( D \) (diffusion constant) against the contraction rate \( \lambda = 0 \). Fokker-Planck equation [5] generalizations to higher-dimensional stable equilibria and limit cycles (stable periodic orbits) are immediate, provided proper care is taken of the diffusion along the periodic orbit [6,7].

What if a periodic orbit is unstable? Both the diffusion rate and the linearized stability rate \( \lambda > 0 \) now expand forward in time and cannot balance each other. This problem was solved in Refs. [8,9] for repelling periodic orbits with no contracting directions, by balancing the stochastic diffusion against the contraction by the adjoint Fokker-Planck operator. The resulting covariance matrix defines the stochastic neighborhood for a repelling orbit, while the Ornstein-Uhlenbeck covariance defines it for a stable orbit. However, neither these stable nor repelling orbits play a role in chaotic dynamics. The long-time attractors of chaotic dynamics are organized by an infinity of hyperbolic periodic orbits [10–12], i.e., orbits which an ergodic trajectory visits by approaching them along their stable eigendirections and leaves along their unstable eigendirections.

The central result of this paper is that techniques developed for solving the Lyapunov equations [13–15] enable us to define the neighborhood of a hyperbolic periodic point by splitting the covariance matrix \( Q \) into two (mutually nonorthogonal) covariance matrices: \( Q_{cc} \) for the contracting directions and \( Q_{ee} \) for the expanding directions.

There are two immediate applications of the notion of the neighborhood of a hyperbolic point: (a) “optimal partition” of the attractor, and (b) construction of a basis set for the stationary distribution of a noisy chaotic flow.

B. An optimal partition from periodic orbits

While in the idealized deterministic dynamics the state space can be resolved arbitrarily finely, in physical systems noise always limits the attainable state-space resolution. This observation had motivated the many limiting resolution estimates for state-space granularity of chaotic systems with background noise. The idea of an optimal partition in this context was first introduced in 1983 by Crutchfield and Packard [16], who formulated a state-space resolution criterion in terms of a globally averaged “attainable information.” The approach was later generalized and applied to time-series analysis, where the underlying dynamics is unknown [17,18].

A different strategy consists of computing a transfer matrix between intervals of a uniform grid and estimating averages of observables from its eigenvalues and eigenfunctions. First introduced by Ulam [19], this technique has been developed over the years [20,21]. All of these approaches (see Ref. [9] for a review) are based on global averages and assume that granularity is uniform across the state space. In contrast, the main, computationally precise lesson of our work is that even when the external noise is white, additive, and globally homogeneous, the interplay of noise and nonlinear dynamics always results in a local stochastic neighborhood, whose covariance depends on both the past and the future noise integrated and nonlinearly convolved with deterministic evolution along the trajectory. The optimal resolution thus varies from neighborhood to neighborhood and has to be computed locally. As was shown in Ref. [8] for a strictly...
expanding 1d chaotic map and a given noise, the maximal set of nonoverlapping neighborhoods of periodic orbits can be used to construct an optimal partition of the state space, and compute dynamical averages from the associated approximate matrix Fokker-Planck operator.

C. The stationary distribution

In this paper, we utilize our construction of optimal partitions to approximate the stationary probability distribution function by a finite sum over Gaussians, one for each neighborhood. When the dynamics is chaotic, the most one can predict accurately for long times are the statistical properties of the system, given by the state-space averages of observables \( a(x) \),

\[
(a) = \int dx \rho(x) a(x),
\]

where the stationary distribution (natural measure [22–24]) \( \rho(x) \) is the probability of finding the system in the state \( x \) on the attractor. For a deterministic system, \( \rho(x) \) is a singular, nowhere differentiable distribution with support on the attractor. For a deterministic map \( f(x) \), a deterministic map \( f(x) \), an initial Dirac function \( \delta(x) \) located at \( x = x_a \) is smeared out into a Gaussian ellipsoid \( \rho_a(x) \) located at \( x = x_a \) with covariance \( \Delta(x) \). This defines the kernel of the Fokker-Planck operator [5], with kernel

\[
\mathcal{L}(y,x) \, dy = e^{-\frac{1}{2} \|y - f(x)\|^2} \delta(y - f(x)) \, dy.
\]

Like in the forward evolution, we may substitute a Gaussian density \( \rho_a \) as the convolution of its image \( \rho_{a+1} \) with the noisy dynamics

\[
\rho_a(z_a, Q_a) = \int dz_{a+1} \mathcal{L}(z_{a+1}, z_a) \rho_{a+1}(z_{a+1}, Q_{a+1}).
\]

We show in what follows that if the Jacobian matrices \( M_a \) have all eigenvalues strictly contracting (expanding), any initial Gaussian converges to an invariant density under the action of the (adjoint) Fokker-Planck operator. Consider first the case of a map \( f(x) \) with a stable fixed point at \( x_a \) (at \( z = z_a = 0 \)). The covariance matrix transforms as

\[
Q = \Delta + M \Delta M^T + M^2 \Delta (M^T)^2 + \cdots = \sum_{m,n=0}^\infty M^m \Delta (M^T)^n \delta_{mn}.
\]

By inserting the Fourier representation of Kronecker \( \delta_{mn} \) into (6), we can recast this expression into the resolvent form

\[
Q = \left( \frac{1}{2\pi} \int_0^{2\pi} d\theta \right) \sum_{m,n=0}^\infty \frac{(e^{-i\theta} M)^n \Delta (e^{i\theta} M^T)^m}{1 - e^{-i\theta} M \Delta (1 - e^{i\theta} M^T)}
\]
We do the same in the expanding case, by using the adjoint evolution
\[
M^TQ = \sum_{m,n=0}^{\infty} \frac{1}{M^n} \frac{1}{(M^r)^n} \delta_{mn} = \int_0^{2\pi} \frac{d\theta}{2\pi} \sum_{m,n=0}^{\infty} \left( \frac{e^{i\theta}}{M} \right)^n \Delta \left( \frac{e^{-i\theta}}{M^r} \right)^m = \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{1}{1 - e^{-i\theta} / M} \Delta \left( 1 - \frac{e^{-i\theta} / M^r}{1 - e^{-i\theta} / M^r} \right),
\]
which is then easily reduced to (7), so that the resolvent form is the same regardless of whether \( M \) is expanding or contracting. This result becomes particularly handy when we deal with a hyperbolic fixed point, that is, when \( M_a \) has both expanding and contracting eigenvalues. The monodromy matrix is not symmetric and it cannot be diagonalized by an orthogonal transformation, but its expanding and contracting parts can be separated with a similarity transformation \( S \) that brings \( M \) to a block-diagonal form,
\[
\Lambda \equiv S^{-1}MS = \begin{pmatrix} \Lambda_e & 0 \\ 0 & \Lambda_c \end{pmatrix}.
\]
Here the blocks \( \Lambda_e \) and \( \Lambda_c \) contain all expanding and contracting eigenvalues of the monodromy matrix, respectively. The covariance matrix \( Q = S^TQ^S \) is not block diagonalized by the above similarity transformation, but consider the four blocks
\[
\hat{Q}_{ij} = \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{1}{1 - e^{-i\theta} \Lambda_i} \hat{\Delta}_{ij} \frac{1}{1 - e^{i\theta} \Lambda_j},
\]
where \( \hat{\Delta} \equiv S^{-1} \Delta(S^{-1})^T \), and \( i, j \in \{c,e\} \), where \( \{c,e\} \) denotes \{ contracting, expanding \}. This expression may be evaluated as a contour integral around the unit circle in the complex plane [13,15],
\[
\hat{Q}_{ij} = \frac{1}{2\pi i} \oint z^{-1 - \Lambda_i} \hat{\Delta}_{ij} \frac{1}{1 - z \Lambda_j}.
\]
The diagonal blocks \( \hat{Q}_{ee}, \hat{Q}_{cc} \) have either all expanding or all contracting eigenvalues, meaning at least one pole inside and one pole outside the unit circle, and the residue theorem yields a nonvanishing result for the integral. Consider next the off-diagonal block \( \hat{Q}_{ec} \) with \( \Lambda_i \) contracting and \( \Lambda_j \) expanding: in this case, the poles all lie outside the unit circle and the integral vanishes. The remaining off-diagonal block having \( \Lambda_i \) expanding and \( \Lambda_j \) contracting must also vanish when integrated, due to the symmetry of \( \hat{Q} \), which is therefore block diagonal,
\[
Q = S \begin{pmatrix} \hat{Q}_{ee} & 0 \\ 0 & \hat{Q}_{cc} \end{pmatrix} S^T.
\]
These results are easily extended to a periodic orbit \( p \) of period \( n_p \), since any point \( x_a \) of the orbit is a fixed point of the \( n_p \)th iteration of the map. The forward and adjoint evolution equations (4) and (5) for the covariance matrix, as well as the resolvent (7), all still hold, with some changes in the notation: each periodic point \( x_a \) has its own neighborhood, with its own covariance matrix \( Q_a \). The monodromy matrix \( M_a \) of \( x_a \) now evolves \( n_p \) steps along the orbit,
\[
M^{n_p}_a = M_{a+n_p-1} \cdots M_{a+2} M_{a+1} M_a,
\]
while the diffusion tensor \( \Delta_a \) now accounts for the total noise accumulated along the periodic orbit,
\[
\Delta_{p,a} \equiv \sum_{i=0}^{n_p-1} M_m^{i,i+1} \Delta_{a+i} M_m^{i+1,1} \cdot \cdot \cdot.
\]

### III. Optimal Partition and Stationary Distribution

At this point, our strategy is to build a partition out of neighborhoods of the periodic points, each defined by means of the stationarity condition (7): solve for the expanding and contracting blocks of (12) separately and draw a parallelogram on the supports of the resulting Gaussians, with axes oriented along the eigenvectors of the covariance matrices \( Q_{ee} \) and \( Q_{cc} \) and their widths given by one standard deviation along each direction. We say that two neighborhoods overlap if they do so by at least 50% of their areas (consistent with the 1σ confidence interval chosen as the overlapping threshold in Ref. [8]). For a typical chaotic map, periodic points are dense in the deterministic attractor [27], which we now aim to cover entirely with the minimum number of neighborhoods possible. We do so via the following algorithm: (i) Find periodic points of period \( n_p = 1, 2, \ldots, \) and their corresponding neighborhoods. (ii) If any neighborhood overlaps with the neighborhood of a shorter periodic point, then it is discarded and the neighborhood of lower period occupying the same area is instead kept in the partition. (iii) Among groups of neighborhoods of the same period, discard those that overlap, while keep the rest in the partition. (iv) The algorithm stops when the attractor is fully covered and no further nonoverlapping neighborhoods can be found. An example is shown in Fig. 1 for the two-dimensional Lozi attractor [25].

The main utility of a good partition is that it provides a basis for an accurate and efficient estimate of long-time averages of observables defined on the dynamical system, of the form (1). As explained in Sec. I C, our goal here is to determine the stationary distribution \( \rho(x) \). For that purpose, we use as our basis the Gaussian ellipsoids that satisfy the local stationarity condition (7) in each neighborhood of the optimal partition. A set of Gaussians centered at every point in the state space \( \mathcal{M} \) forms an overcomplete, nonorthogonal basis for functions in \( L^2(\mathcal{M}) \), as is well known from the study of coherent states of quantum harmonic oscillators [28]. Our (also overcomplete and nonorthogonal) set of Gaussians is centered only on periodic points, which are dense in the deterministic attractor, but not in the entire state space. Therefore, our basis is designed to resolve the structure of any function with support on the hyperbolic “strange set” (an attractor or a repeller). The Gaussians are constructed so that their widths balance the noise spreading and the (time-forward or -backward) contraction of the deterministic dynamics. In the transverse directions, the basis gives the width of the global stationary distribution, locally everywhere determined by the balance between noise and dynamics. Along the attractor, the basis determines the
The deterministic attractor is the fractal structure in the background of each picture. Stochastic neighborhoods of a set of periodic points are indicated by their standard deviation parallelograms. (a) An initial partition: with only periodic points of periods $\leq 5$, much of the attractor remains to be covered. (b) The final, optimal partition covers the whole attractor, with no pair of neighborhoods overlapping by more than 50%. (c) $N$, the number of neighborhoods needed to achieve the optimal partition for a given noise strength $\Delta$.

There are numerical methods (such as refinements of Ulam’s method [29]) that identify the asymptotic attractor by running long noisy trajectories, dropping the transients, and covering the attractor so revealed by a finite number of boxes. These algorithms have no \textit{a priori} information about how the stationary distribution behaves transversely to the deterministic attractor, and they may easily overestimate the number of basis elements needed to resolve this structure.

In contrast, in our approach the transverse structure is identified \textit{a posteriori} by running long noisy trajectories, dropping the transients, and covering the attractor so revealed by a finite number of boxes. These algorithms have no \textit{a priori} information about how the stationary distribution behaves transversely to the deterministic attractor, and they may easily overestimate the number of basis elements needed to resolve this structure.

In discrete time dynamics, the stationary distribution is the ground-state eigenfunction of the Fokker-Planck evolution operator (2) with escape rate $\gamma$,

$$\mathcal{L}\rho(x) = e^{-\gamma} \rho(x). \tag{14}$$

In order to estimate the stationary distribution, we write it as a sum over the neighborhoods of the periodic points:

$$\rho(x) = \sum_{a=1}^{N} h_a \phi_a(x), \tag{15}$$

where $\phi_a = e^{-\gamma} Q_a \rho$ are the Gaussian basis functions, with $Q_a$ given by (12), and the coefficients $(h_a)$ to be determined. The truncation of the expansion (15) to $N$ basis functions follows from our optimal partition. We estimate the coefficients $h_a$ by minimizing the cost function,

$$\int \left[ \sum_{a=1}^{N} h_a (\mathcal{L} - e^{-\gamma}) \phi_a(x) \right]^2 dx, \tag{16}$$

together with the normalization constraint for $\rho(x)$. We can also estimate the escape rate of the system by minimizing the error with respect to $e^{-\gamma}$.

As an example, we apply the procedure to the Lozi map [25],

$$x_{n+1} = 1 - a|x_n| + b y_n,$$

$$y_{n+1} = x_n, \tag{17}$$

with parameters $a = 1.85, b = 0.3$ and isotropic, constant diffusion tensor $\delta_{ij} $, with $\Delta$ ranging in the interval $[0.003, 0.1]$. Figure 1(c), which shows the number $N$ of neighborhoods required by the optimal partition for a given $\Delta$, illustrates the efficiency of our method: $N$ goes from tens to few hundreds in the noise range considered. In order to test our algorithm, we also estimate $\rho(x)$ and $\gamma$ by a direct numerical simulation. The $(x,y)$ state space is divided into uniform mesh $6.4 \times 10^6$ bins; we follow long stochastic trajectories and count how many times they visit each bin. The stationary distribution $\rho_B(x)$ is then the normalized frequency distribution of the whole grid. The deterministic Lozi map has a fixed point at the edge of the attractor, whose stable manifold is the boundary of the deterministic basin of attraction. The noise makes it possible for a stochastic trajectory to cross this boundary and escape. We compute the escape rate as the ratio of the total number of points in the noisy trajectories to the number of escapes.

Figure 2 shows an example of the stationary distribution estimated with both methods, while in Fig. 3 we quantitatively compare the two procedures. In particular, we estimate the relative error between the stationary distribution $\rho$ computed with the optimal partition and $\rho_B$ computed on the uniform grid, by using a normalized $L^2$ distance, as

$$d(\rho, \rho_B) = \frac{\int (\rho(x) - \rho_B(x))^2 dx}{\int \rho_B(x)^2 dx}. \tag{18}$$

FIG. 1. (Color online) Building the partition for the Lozi attractor, for an isotropic constant diffusion tensor $\delta_{ij} \Delta$. (a), (b) $\Delta = 0.01$. The deterministic attractor is the fractal structure in the background of each picture. Stochastic neighborhoods of a set of periodic points are indicated by their standard deviation parallelograms. (a) An initial partition: with only periodic points of periods $\leq 5$, much of the attractor remains to be covered. (b) The final, optimal partition covers the whole attractor, with no pair of neighborhoods overlapping by more than 50%. (c) $N$, the number of neighborhoods needed to achieve the optimal partition for a given noise strength $\Delta$.

FIG. 2. (Color online) The stationary distribution of the Lozi map with $\Delta = 0.01$. (a) A direct numerical calculation obtained by running noisy trajectories for a long time. (b) The stationary distribution calculated with the optimal partition method.
The two distributions are within 5% of each other, whereas the escape rates differ by at most 10%, over a range of \( \Delta \) that spans two orders of magnitude.

IV. DISCUSSION

In conclusion, we have generalized the optimal partition hypothesis first formulated in [8] to hyperbolic maps in arbitrary dimension and tested the method on a two-dimensional system with weak white isotropic noise. As noise induces a finite resolution of the state space of any physical system, finite numbers of neighborhoods suffice to partition the state space explored by chaotic dynamics and to estimate long-time averages of observables. Here we have used the deterministic unstable periodic orbits as the skeleton on which to build an optimal partition for the noisy state space. First we determine a local stationary distribution in the neighborhood of each periodic point by balancing the noise against the deterministic expansion or contraction. From the separation of expanding and contracting blocks in the covariance matrix that characterizes the Gaussian approximation to the local stationary distribution, we carve out a precise definition of neighborhood, i.e., the constituent of our partition, which is then used to approximate the global stationary distribution, estimate the escape rate (for open systems that allow escape), and estimate any long-time averaged observable. Numerical tests confirm that the accuracy of our method is comparable to that of a uniform grid discretization, but the number of neighborhoods required for our optimal partition (\( \sim 10 \) to \( 100 \)) is three to four orders of magnitude smaller than the number of bins used in the uniform grid discretization method (\( \sim 10^6 \)).

The problems that dynamical chaos (or “turbulence”) theory faces nowadays are not two dimensional, but high and even infinite dimensional. Today it is possible to compute numerically exact periodic orbits (“recurrent flows” [30]) in a variety of physically realistic turbulent fluid flows [31,32], but these calculations are at the limit of what current codes can do, and we hope that the methods presented here can provide sharp criteria for when a sufficient number of such solutions has been computed. Furthermore, unlike the uniform grid discretization, our partitions are smart, since they rely on the periodic orbits of the deterministic system as skeleton of the dynamics, as well as efficient, due to the finite (and surprisingly optimal) numbers of neighborhoods and corresponding basis functions. This, we believe, should make our algorithm less costly to implement than direct numerical simulations in higher dimensions, where discretizations would be impractical. With some modifications and application of Poincaré sections, the formalism can be applied to continuous time flows as well [9,33]. Outstanding challenges include dealing with the lack of hyperbolicity in higher dimensions (marginal directions were treated in Ref. [9] for 1d maps) as well as extending the definition of neighborhood to other time-invariant sets, such as relative periodic orbits and partially hyperbolic invariant manifolds. Further technical issues, such as improving the efficiency of the minimization algorithm by modifying the basis of functions used in the computation of the stationary distribution, are also part of our agenda.

ACKNOWLEDGMENTS

We are indebted to P. M. Svetlichnyy for suggesting that we use error minimization to find the global stationary distribution. J.M.H. is thankful for the partial support provided by a Presidential Undergraduate Research Award, and to D.L. for the hospitality at IASTU, Beijing. P.C. thanks the family of the late G. Robinson, Jr. and NSF Grant No. DMS-1211827 for partial support. D.L. acknowledges support from the National Science Foundation of China (NSFC), International Young Scientists (Grant No. 11450110057-04323001).


