

# A TRAJECTORY-DRIVEN ALGORITHM FOR DIFFERENTIATING SRB MEASURES ON UNSTABLE MANIFOLDS\*

ADAM A. ŚLIWIAK<sup>†</sup> AND QIQI WANG<sup>†</sup>

**Abstract.** SRB measures are limiting stationary distributions describing the statistical behavior of chaotic dynamical systems. Directional derivatives of SRB measure densities conditioned on unstable manifolds are critical in the sensitivity analysis of hyperbolic chaos. These derivatives, known as the *SRB density gradients*, are by-products of the regularization of Lebesgue integrals appearing in the original *linear response* expression. In this paper, we propose a novel trajectory-driven algorithm for computing the SRB density gradient defined for systems with high-dimensional unstable manifolds. We apply the concept of measure preservation together with the chain rule on smooth manifolds. Due to the recursive one-step nature of our derivations, the proposed procedure is memory-efficient and can be naturally integrated with existing Monte Carlo schemes widely used in computational chaotic dynamics. We numerically show the exponential convergence of our scheme, analyze the computational cost, and present its use in the context of Monte Carlo integration.

**Key words.** SRB measure, SRB density gradient, Measure preservation, Chaotic dynamical systems, Sensitivity analysis, Monte Carlo integration

**AMS subject classifications.** 65P20, 65C05, 37A05, 28A25

**1. Introduction.** Due to their seemingly irregular and quasi-random behavior, a mathematical description of chaotic dynamical systems might be challenging. A major breakthrough in the analysis of chaos was the introduction of the SRB (Sinai-Ruelle-Bowen) measure  $\mu$  [26]. This scalar quantity, defined on a compact Riemannian manifold, contains a coherent statistical description of the dynamics. Intuitively, the SRB measure represents the likelihood of the trajectory passing through a non-zero-volume region of a strange attractor. Although the concept of SRB measures was originally applied to Axiom A systems, several rigorous studies extended this idea even beyond the universe of uniformly hyperbolic systems [34, 11, 12].

Lebesgue integrals with respect to  $\mu$ , which represent expected values of certain smooth observables, are fundamental in the analysis of chaos. Under the assumption of *ergodicity*, they equal the time-average of an infinitely-long sequence generated along a trajectory. Integrals of this type can thus be approximated using a Monte Carlo method. If the integrand involves highly-oscillatory derivatives, then the Monte Carlo integration might be prohibitively expensive due to a large variance of the sample [31]. In case of derivatives of functions evaluated at a future time (see examples of such integrands in [10, 29, 15, 2]), the direct use of any integration scheme might be impossible due to the *butterfly effect*. Indeed, the application of the chain rule results in a product of the system's Jacobian matrices whose norms increase exponentially in time. A remedy for this computational difficulty is integration by parts, which moves the differentiation operator away from the problematic function to the SRB measure. This is in fact a consequence of the generalized fundamental theorem of calculus. In addition to the boundary term, we effectively obtain a new Lebesgue integral involving a product of the antiderivative of the original integrand and the *SRB density gradient*  $g = \partial \log \rho = \partial \rho / \rho$ , where  $\rho$  denotes the density of  $\mu$  (i.e., the

\*Submitted to SIAM.

**Funding:** This work was funded by Air Force Office of Scientific Research Grant No. FA8650-19-C-2207 and U.S. Department of Energy Grant No. DE-FOA-0002068-0018.

<sup>†</sup>Center for Computational Science and Engineering, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, United States of America (asliwiak@mit.edu, qiqi@mit.edu).

Radon-Nikodym derivative [23]).

The SRB density gradient is critical in the sensitivity analysis of chaos. The major implication of Ruelle’s *linear response* theory is a closed-form expression for the parametric derivative of long-time averages (a.k.a. the system’s sensitivity) [27, 28]. The space-split sensitivity (S3) method [8, 10] reformulates Ruelle’s formula to a computable form by splitting the perturbation vector and performing integration by parts on unstable manifolds. Using the S3 formula, one can construct an efficient and provably convergent Monte Carlo algorithm for sensitivities in uniformly hyperbolic systems. This algorithm requires computing the SRB density gradient defined as a directional derivative of  $\rho$  conditioned on the unstable manifold. Indeed, the SRB measure is generally singular with respect to Lebesgue measure in the stable direction [34]. This approach of the regularization of Ruelle’s integrals on unstable manifolds has also been applied in [24] to derive a fast linear response algorithm for differentiating SRB states. Several algorithms for sensitivity analysis that stem from the Fluctuation-Dissipation Theorem (FDT) [17] also require  $g$  [2, 1]. Motivated by empirical data of certain chaotic models, some FDT-based methods assume Gaussian distribution of measure [15]. Such an assumption reduces the FDT linear response operator to a simple time autocorrelation function, which dramatically facilitates the sensitivity computation for the cost of limited applicability. The density gradient can also be used as a reliable indicator of the differentiability of statistical quantities [30] in chaotic systems. In particular, the slope of the distribution tail of  $g$  have been shown to be strictly associated with the existence of parametric derivatives of statistics. Therefore, we seek a numerical procedure for  $g$  that does not make any assumptions about the statistical behavior of the system and is thus generalizable to any chaotic dynamical system that admits SRB measures.

There already exist algorithms for the SRB density gradient derived for systems with one-dimensional unstable manifolds. In case of simple one-dimensional maps, one can derive an exponentially convergent recursion for  $g$  using the measure preservation property [29]. The same formula can be inferred using the fact the SRB density is an eigenfunction of the Frobenius-Perron operator with eigenvalue 1 [30]. The authors of [9] propose an ergodic-averaging algorithm for self-derivatives (i.e., directional derivatives along one-dimensional expanding directions) of covariant Lyapunov vectors (CLVs) corresponding to the only positive Lyapunov exponent, which are tangent to unstable manifolds at any point on the attractor. Using the chain rule on smooth manifolds, one can show  $g$  depends on the self-derivative of CLV at the previous time step, and this relation is governed by a second-order tangent equation [9, 30]. In a recent work, Ni [24] proposes an algorithm for divergence on the unstable manifold using approximate *shadowing* coordinates instead of the full basis of the expanding subspace, as opposed to the S3 method.

In this paper, we systematically derive a trajectory-driven algorithm for the SRB density gradient by extending the measure preservation property to high-dimensional smooth manifolds. Using the density-based parameterization of unstable manifolds and the chain rule, it is possible to establish a recursive relation for the evolution of first- and second-order parametric derivatives of the coordinate chart. By definition, this chart is strictly associated with  $g$  and can be interpreted as an SRB inverse cumulative distribution (quantile function). This type of parameterization, motivated by popular methods of statistical inference [13], has been thoroughly explained by the authors in [31] in the context of simple Lebesgue measures. Through the relation of  $g$ , the coordinate map and its parametric derivatives, we show the density gradient can be computed by solving a collection of first- and second-order tangent equations. We

also show that the recurring problem of the butterfly effect, which leads to exploding norms of tangent solutions, can be eliminated by iterative orthonormalization of the chart gradient. The major benefit of our derivation is that it is naturally translatable to a practicable algorithm that can be easily integrated with existing methods for sensitivity analysis of chaos.

This paper is structured as follows. In Section 2, we introduce the SRB measure, its gradient, and highlight their importance in the field of chaotic dynamics. Subsequently, in Section 3, we apply the density-based parameterization for the description of unstable manifolds to derive recursive relations for the SRB density gradient. This derivation is followed by a numerical example involving a chaotic map with straight one-dimensional expanding subspaces. Section 4 generalizes all the concepts to high-dimensional chaotic maps with an arbitrary number of positive Lyapunov exponents (LEs). Based on the systematically derived iterative relations, a practicable algorithm for  $g$  is thoroughly described. We analyze its cost, memory requirements and convergence. In the same section, we also demonstrate a numerical example of Monte Carlo integration, which requires the computation of  $g$ . Section 5 summarizes this work.

**2. SRB measure and its gradient: definitions and significance.** Consider a diffeomorphic map  $\varphi : M \rightarrow M$ ,  $M \in \mathbb{R}^n$ ,  $n \in \mathbb{Z}^+$  with an Axiom A attractor. Theorem 1 of [34] asserts that there exists an *invariant* and *physical* probability measure  $\mu$  (and its density  $\rho$ ), which satisfies:

1. Invariance/conservation of measure condition:

$$(2.1) \quad \mu(A) = \mu(\varphi^{-1}(A))$$

for any Borel subset  $A \subset M$ .

2. Physicality condition: for any smooth  $f : M \rightarrow \mathbb{R}$ ,

$$(2.2) \quad \int_M f(x) d\mu(x) = \int_M f(x) \rho(x) d\omega(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} f \circ \varphi^k(x_0),$$

where  $d\omega$  denotes the Riemmanian volume element. The initial state  $x_0$  is assumed to be  $\mu$ -typical and  $\varphi^k(\cdot) = \varphi(\varphi^{k-1}(\cdot))$ ,  $\varphi^1 = \varphi$ ,  $\varphi^0 = \text{Id}$ .

3. Absolute continuity: Conditional measure of  $\mu$  denoted by  $\tilde{\mu}_x$  and defined on the unstable manifold  $U_x$  at point  $x \in U_x$  is absolutely continuous (analogous property applies to the conditional density  $\tilde{\rho}_x$ ).
4. Singularity with respect to Lebesgue measure:  $\mu$  is generally sharp in the stable direction (across unstable manifolds).
5. Unit measure axiom (probability universe):

$$\int_M d\mu(x) = \int_M \rho(x) d\omega(x) = \mu(M) = 1.$$

The measure  $\mu$  and its density  $\rho$  are respectively known as the SRB measure and SRB density distribution. We listed their properties most important in the context of this paper; however, the reader is referred to [34] for a detailed description of other significant features. One can think about Property (1) as the mass conservation law. For example, consider a Borel subset  $B \subset M$  with a uniform measure that is mapped to  $\varphi(B) \subset M$ . If we divide  $\varphi(B)$  into a finite number of subsets occupying the same volume, each of them generally has a different measure. In other words, each subset generally has its unique weight unless  $\varphi$  represents a simple translation and/or

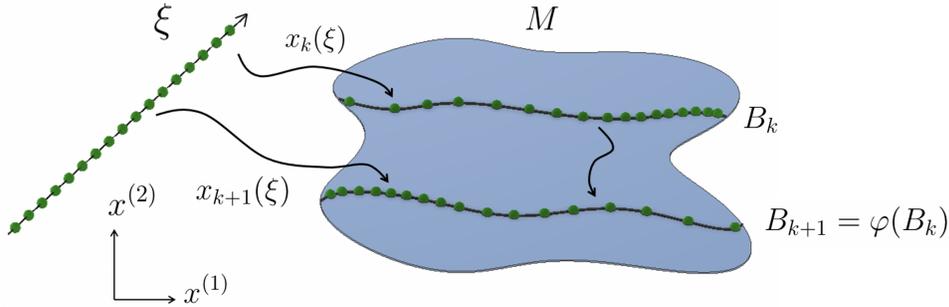


FIG. 1. This figure graphically represents the measure preservation property. The localization of green bullets represents the SRB density on some 1D subspace of a 2D manifold  $M$ . All green bullets are equally weighted. In this sketch, we observe  $\mu(B_k) = \mu(B_{k+1})$  and  $B_{k+1} = \varphi(B_k)$ , where  $B_k \subset M$  and  $B_{k+1} \subset M$  are parameterized by smooth charts,  $x_k(\xi) : [0, 1] \rightarrow B_k$  and  $x_{k+1}(\xi) : [0, 1] \rightarrow B_{k+1}$ , respectively.

rotation. Property (2) states that the SRB measure is physical, which means that by observing the system's evolution for an infinitely long period of time we can assign a weight (density) to each non-zero-volume region of the attractor. The expected value of any smooth function defined on  $M$  can be computed as a simple volume integral over  $M$  of that function multiplied by the density function. Figure 1 graphically explains Property (1), while the remaining four properties and their consequences are further explained and illustrated in the following sections.

As mentioned above, SRB measures are guaranteed to exist in Axiom A (or, uniformly hyperbolic) systems. Different rigorous studies indicate that uniform hyperbolicity is in fact not required for the existence of  $\mu$ . For example, partially hyperbolic systems that have a mostly expanding [3] or contracting [7] central direction also admit SRB measures. In addition, many high-dimensional systems arising from discretization of real-world PDE models behave as uniformly hyperbolic systems, per the *hyperbolicity hypothesis* [14].

In many engineering applications, the expected value of some physically relevant quantity  $f \in L^1(\rho)$ , i.e.,  $\int_M f d\mu$ , is usually of interest. The major challenge in the field of sensitivity analysis of chaos is to find a parametric derivative of the expected value, which is critical in grid adaptation [20], optimization design [16] and uncertainty quantification [32]. Ruelle rigorously derived a closed-form expression for that derivative [27, 28],

$$(2.3) \quad \frac{d}{ds} \int_M f(x) d\mu(x) = \sum_{t=0}^{\infty} \int_M D(f \circ \varphi^t(x)) \cdot \chi(x) d\mu(x),$$

where  $\chi$  denotes the derivative of  $\varphi$  with respect to the map parameter  $s$ , while  $D$  is a phase space differentiation operator. One could potentially apply a Monte Carlo algorithm to the integrals on the RHS of Eq. 2.3. However, owing to the butterfly effect, the direct evaluation of the integrand for a higher  $t$  is computationally infeasible. To illustrate this problem, let us consider the 2D Arnold's cat map  $\varphi : [0, 1]^2 \rightarrow [0, 1]^2$  defined as

$$(2.4) \quad x_{k+1} = Ax_k \bmod 1, \quad A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

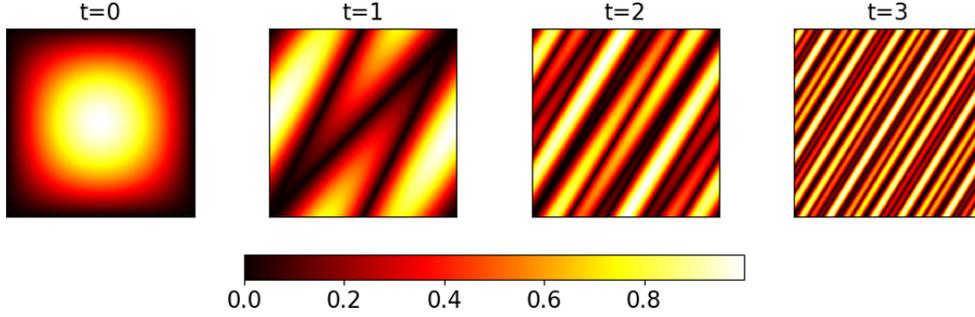


FIG. 2. Evaluation of the composite function  $f \circ \varphi^t(x)$  on the manifold  $M = [0, 1]^2$  at four consecutive steps  $t$ . In this case, the map  $\varphi$  is the Arnold's cat map (Eq. 2.4), while  $f(x^{(1)}, x^{(2)}) = \sin(\pi x^{(1)}) \sin(\pi x^{(2)})$ . This particular  $\varphi$  is a classical representative of an Anosov diffeomorphism.

and some smooth function  $f(x)$ . In Figure 2, we observe that even for a low  $t$ ,  $f \circ \varphi^t$  becomes highly-oscillatory, which implies that  $\|D(f \circ \varphi^t)\|$  grows very fast ( $\|\cdot\|$  denotes the Euclidean norm in  $\mathbb{R}^n$ ). Due to the presence of positive Lyapunov exponents in chaotic systems, the rate of growth is in fact exponential. It means that Ruelle's formula is impractical for a direct Monte Carlo computation.

To circumvent this problem, one can apply integration by parts to move the differentiation operator away from the composite function. This idea was originally applied in the novel S3 method [8, 10], and later also in Ni's approximate method [24] for sensitivity analysis. In case of integrals with respect to a non-uniform measure, integration by parts requires differentiating the measure itself. However, according to Property (3) and Property (4),  $D\rho$  generally does not exist. In this section, let us assume  $\chi$  equals a unit vector  $q$  that is tangent to the one-dimensional unstable manifold at every point on the manifold  $M^1$ . Thus, every integral from the RHS of Eq. 2.3 can be regularized through partial integration as follows,

$$(2.5) \quad I = \int_M Df_t(x) \cdot q(x) d\mu(x)$$

$$(2.6) \quad = \int_{M/U} \int_{U_x} Df_t(s) \cdot q(s) d\tilde{\mu}_x(s) d\hat{\mu}(x)$$

$$(2.7) \quad = \int_{M/U} \int_{U_x} \partial_q f_t(s) \tilde{\rho}_x(s) ds d\hat{\mu}(x)$$

$$(2.8) \quad = \int_{M/U} \int_0^1 \partial_\xi f_t(s(\xi)) \tilde{\rho}_x(s(\xi)) d\xi d\hat{\mu}(x)$$

$$(2.9) \quad = - \int_{M/U} \int_{U_x} f_t(s) \partial_q \tilde{\rho}_x(s) ds d\hat{\mu}(x) + (\text{boundary term})$$

$$(2.10) \quad = - \int_{M/U} \int_{U_x} f_t(s) \frac{\partial_q \tilde{\rho}_x}{\tilde{\rho}_x}(s) d\tilde{\mu}_x(s) d\hat{\mu}(x) + (\text{boundary term})$$

$$(2.11) \quad = - \int_M f_t(x) g(x) d\mu(x) + (\text{boundary term}),$$

<sup>1</sup>In a general case,  $\chi \neq q$  and thus an extra step is required to regularize Ruelle's formula. This step involves a splitting of  $\chi$  into two terms, such that one term belongs to unstable manifolds everywhere on  $M$ . The reader is referred to [8, 10] for a detailed description of this process.

where  $f_t(x) := f \circ (\varphi^t(x))$ . To derive the final form of  $I$ , we perform the following steps. First (Step 2.6), we disintegrate  $\mu$  on a measurable partition  $U$  determined by the geometry of unstable manifolds. The quotient measure  $\hat{\mu}$  is defined such that for all Borel sets  $B \subset M$ ,

$$\mu(B) = \int_{M/U} \tilde{\mu}_x(B \cap U_x) d\hat{\mu}(x),$$

where  $\tilde{\mu}_x$  is a conditional SRB measure with density  $\tilde{\rho}_x$ . Subsequently, in Step 2.7, we use the measure-density relation,  $d\tilde{\mu}_x = \tilde{\rho}_x ds$ , where  $s$  denotes the path length as we move along  $U_x$ . In Step 2.8, we parameterize  $U_x$ , which gives rise to  $ds = \|x'(\xi)\| d\xi$ . Note the multiplicative factor is absorbed by the parametric derivative of  $f$ , because  $\partial_\xi f = \|x'(\xi)\| \partial_s f$ . Integration by parts is applied in Step 2.9, where the differentiation operator is moved from  $f$  to  $\tilde{\rho}$ . In Steps 2.10-2.11, we reshuffle terms and use the above identities again to simplify the final expression. Integration by parts also gives rise to a boundary term, which involves two integrals with respect to the quotient measure of  $f \tilde{\rho}_x$  evaluated at  $\xi = 0$  and  $\xi = 1$ , respectively. From now on, we shall drop the subscript notation for conditional distributions; the tilde ( $\tilde{\cdot}$ ) notation shall imply the given distribution is restricted to a local unstable manifold. Note the boundary term,

$$(2.12) \quad (\text{boundary term}) = \int_{M/U} [\tilde{\rho}(\xi) f_k(\xi)]_{\xi=0}^{\xi=1} d\hat{\mu}(x),$$

can be expressed in terms of a regular volume integral over  $M$  of the divergence on unstable manifolds, which vanishes according to Theorem 3.1(b) of [27]. This is indeed a direct consequence of the fact the boundary terms across two neighboring rectangles of the Markov partition of  $M$  cancel out. To visualize this property, let us consider the Arnold's cat map (Eq. 2.4), for example. Despite its "artificial" discontinuities due to the modulo operator, this nonlinear transformation in fact maps a smooth torus to itself. One could arbitrarily change the boundaries of the square  $M$  in both phase space directions without modifying the map itself, and still describe the same torus.

Since  $\int_M Df_k \cdot q d\mu = -\int_M f_k g d\mu$ , we can thus alternatively apply Monte Carlo to the RHS that involves the SRB density gradient  $g$  [10, 30, 29],

$$(2.13) \quad g(x) = \frac{\partial_q \rho(x)}{\rho(x)} = \frac{\partial_q \tilde{\rho}(x)}{\tilde{\rho}(x)} = \partial_q \log \tilde{\rho}(x).$$

Note that the integrand appearing in the regularized version of  $I$  does not grow exponentially with  $t$  if  $f$  is bounded, which makes the sensitivity formula computable (immune to the butterfly effect). The integration by parts, as presented above, is generally useful if the integrand involves highly-oscillatory functions. The Monte Carlo integral example presented in [31] shows that the partial integration may reduce the number of samples a few orders of magnitude to achieve the desired approximation error. Therefore, the computation of  $g$  might be beneficial not only in the context of the Ruelle/S3/FDT-based method for sensitivity approximation, but also in a general setting when the expected value of an ill-behaved quantity of interest in a chaotic system is needed. The following two sections focus on the computation of  $g$  for systems with an arbitrary number of positive LEs. The primary goal is to derive a recursive procedure compatible with Monte Carlo algorithms, which are widely used in the field.

**3. Computing SRB density gradient for systems with one-dimensional unstable manifolds.** In this section, we consider a generic  $n$ -dimensional,  $n \in \mathbb{Z}^+$ , uniformly hyperbolic dynamical system with one-dimensional unstable manifold governed by the  $C^2$  diffeomorphic map  $\varphi : M \rightarrow M$ .  $M$  is thus a Riemannian manifold immersed in  $\mathbb{R}^n$ . There exists a measurable partition  $U$  of  $M$  such that each member of that partition,  $U_x$ , coincides with the unstable manifold that contains  $x \in M$ . In this particular case, each  $U_x \subseteq M$  is geometrically represented by a curve embedded in  $\mathbb{R}^n$ . We strive to compute the directional derivative of the logarithmic SRB density  $g$  defined by Eq. 2.13.

**3.1. Derivation of the iterative formula.** The following notation is used throughout this section. Let  $x_k(\xi) : [0, 1] \rightarrow U_k \subset M$  denote a  $C^2$  chart (diffeomorphic map) that describes the unstable manifold  $U_k$ ,  $k \in \mathbb{Z}$ . For any  $k$ , two charts  $x_k(\xi)$  and  $x_{k+1}(\xi)$ , defined respectively on  $U_k$  and  $U_{k+1}$ , are related as follows,

$$(3.1) \quad x_{k+1}(\xi) = \varphi(x_k(\xi))$$

for all  $\xi \in [0, 1]$  (see Figure 1 for an illustration of an  $n = 2$  case). We use  $D\varphi$  and  $D^2\varphi$  to respectively denote the Jacobian ( $n \times n$  matrix) and Hessian ( $n \times n \times n$  third-order tensor) of  $\varphi$ . Since  $\varphi$  is invertible, Eq. 3.1 can be viewed as a mathematical description of the evolution of SRB measure. For any observable  $f$  defined on  $M$ , evaluated along a certain trajectory, we use the following short-hand notation,  $f \circ x_k(\xi) := f_k$ . Derivatives of the chart with respect to the parameter  $\xi$  are denoted using the prime ( $'$ ) symbol. A reference to the  $i$ -th component of a vector/matrix/tensor is indicated inside round brackets located in the superscript; for example,  $q^{(i)}$  denotes the  $i$ -th component of  $q$ . Finally, we use  $\partial_i$  to denote differentiation with respect to the  $i$ -th coordinate of phase space.

Let us parameterize  $U_k$  such that

$$(3.2) \quad \xi = \int_{\mathcal{C}_k(\xi)} \tilde{\rho}(x_k(\xi)) ds,$$

where  $\mathcal{C}_k(\xi)$  represents the segment of  $U_k$  between  $x_k(0)$  and  $x_k(\xi)$ , which implies that  $\mathcal{C}_k(1) \equiv U_k$ . Consequently,  $\tilde{\rho}_k$  is the conditional SRB density restricted to  $U_k$  satisfying  $\tilde{\rho}_k = \rho_k / \int_{U_k} \rho_k ds$ . We call it a *measure-based* parameterization, as the value of the parameter  $\xi$  coincides with the value of SRB measure at  $x_k(\xi) \in U_k$ . The variable transformation between  $\xi$  and the arc length  $s$  implies that

$$(3.3) \quad \tilde{\rho}(x_k(\xi)) \|x'_k(\xi)\| = 1.$$

Note Eq. 3.3 is in fact a formula for the density change from a uniform to nonuniform distribution due to the nonlinear variable transformation  $x_k(\xi)$ . Since  $\xi \in [0, 1]$ ,  $\tilde{\rho}_k$ ,  $\xi$ ,  $x_k$  can be respectively viewed as a probability density function (PDF), cumulative distribution (CDF), and inverse cumulative distribution (quantile function). Using the measure-based parametrization as described above, the SRB density gradient, defined in Eq. 2.13, can be expressed in terms of parametric derivatives of the chart, i.e.,

$$(3.4) \quad g(x_k(\xi)) = \frac{\partial_q \tilde{\rho}_k}{\tilde{\rho}_k}(x_k(\xi)) = -\frac{x'_k(\xi) \cdot x''_k(\xi)}{\|x'_k(\xi)\|^3},$$

for any  $\xi \in [0, 1]$ . Here, the derivative  $\partial_q$  is computed in the direction of increasing value of  $\xi$ . The reader is referred to the authors' previous work in [31], where Eq.

3.4 is derived by differentiating Eq. 3.3, and comprehensively described using various numerical examples.

We notice  $x'_k(\xi) = \|x'_k(\xi)\| q(x_k(\xi))$ , and rewrite Eq. 3.4 to

$$(3.5) \quad g(x_k(\xi)) = -q(x_k(\xi)) \cdot \frac{x''_k(\xi)}{\|x'_k(\xi)\|^2} := -q(x_k(\xi)) \cdot a(x_k(\xi)) = -q_k \cdot a_k.$$

Eq. 3.5 indicates that the magnitude of the SRB density gradient equals the length of the projection of the (re-scaled) curve acceleration vector on the line tangent to the curve. We now use Eq. 3.1, differentiate it twice with respect to  $\xi$ , and apply the chain rule to obtain the following expression,

$$(3.6) \quad x''_{k+1}(\xi) = D^2\varphi(x_k(\xi))(x'_k(\xi), x'_k(\xi)) + D\varphi(x_k(\xi)) x''_k(\xi),$$

which means that

$$(3.7) \quad a(x_{k+1}(\xi)) = \frac{x''_{k+1}(\xi)}{\|x'_{k+1}(\xi)\|^2} = \frac{\|x'_k(\xi)\|^2 D^2\varphi(x_k(\xi))(q(x_k(\xi)), q(x_k(\xi)))}{\|x'_{k+1}(\xi)\|^2} + \frac{D\varphi(x_k(\xi))x''_k(\xi)}{\|x'_{k+1}(\xi)\|^2}.$$

The bilinear form that appears in the first term on the RHS of Eq. 3.7 can be expressed using Einstein's summation convention, i.e.,  $[D^2\varphi(q, q)]^{(ijk)} = \partial_i \partial_j \varphi^{(k)} q^{(i)} q^{(j)}$ .

Given  $\|x'_{k+1}(\xi)\| = \alpha(x_k(\xi))\|x'_k(\xi)\|$ , where  $\alpha(x_k(\xi)) = \|D\varphi(x_k) q(x_k(\xi))\|$ , we have

$$(3.8) \quad a_{k+1} = \frac{(D^2\varphi)_k(q_k, q_k) + (D\varphi)_k a_k}{\alpha_k^2}.$$

From the parametric derivative of Eq. 3.1 and the definition of  $\alpha(x_k(\xi))$ , the recursion

$$(3.9) \quad q_{k+1} = \frac{D\varphi_k q_k}{\alpha_k}$$

automatically follows. We emphasize the fact the above procedure for  $g$  (involving Eq. 3.5, 3.8, 3.9) is completely analogous to the algorithm proposed in Section 4.2 of [31], which was meant for simple Lebesgue measures evolving due to a generic non-chaotic diffeomorphisms. Here, however, we consider the evolution of the SRB measure in a chaotic system. Due to the butterfly effect, the tangent solution exponentially increases in norm. Therefore, we need the normalizing factor  $\alpha$  in the iterative formula for  $a$  and  $q$  along the trajectory. Since  $\varphi$  is uniformly hyperbolic, the solution to the tangent equation in Eq. 3.9 converges exponentially in  $k$  to the backward Lyapunov vector that is tangent to the unstable manifold regardless of the choice of the initial condition  $q_0$ . Under the same assumption, the recursion in Eq. 3.8 for the acceleration vector  $a$  also converges uniformly in  $k$  at an exponential rate to the true solution for any initial condition  $a_0$  bounded in norm. The reader is referred to Lemma 7.7 in [10] for the proof of the preceding statement.

To summarize, using the measure-based manifold parameterization, we derived a simple recursive procedure for the SRB density gradient that exponentially converges in case of uniformly hyperbolic systems and does not depend on initial conditions. As for now, we restrict ourselves to systems with one-dimensional unstable manifolds. Our main intention here is to introduce basic concepts before we move to general cases in Section 4.

**3.2. Numerical example: computing SRB density gradient on straight unstable manifolds.** As a pedagogical example, let us consider a family of  $n$ -dimensional maps,  $n \in \mathbb{Z}^+$ , whose unstable manifolds are straight and, without loss of generality, aligned with the first coordinate of the phase space. Certainly, this family includes, but is not limited to, all one-dimensional chaotic maps. In this particular case,  $q^{(i)} = \delta^{(i1)}$ , where  $\delta$  denotes the Kronecker delta. Consequently, the parametric derivative of the chart  $x_k(\xi)$ , for any  $k$ , has all zero entries except the first one and, therefore,  $\alpha(x_k(\xi)) = |\partial_1 \varphi^{(1)}(x_k(\xi))|$ . Thus, our recursive algorithm for  $g$ , which involves Eq. 3.5, 3.8, and 3.9, reduces to a single scalar iterative formula,

$$(3.10) \quad g(x_{k+1}(\xi)) = \frac{g(x_k(\xi))}{\partial_1 \varphi^{(1)}(x_k(\xi))} - \frac{\partial_1^2 \varphi^{(1)}(x_k(\xi))}{(\partial_1 \varphi^{(1)}(x_k(\xi)))^2}$$

for all  $\xi \in [0, 1]$ . We were allowed to drop the absolute values, because  $x'_k(\xi) > 0$ , which is a consequence of our choice of the manifold parameterization. In this simple event of a straight unstable manifold, only two scalars are required to advance the iteration, i.e., first- and second-order derivative (in phase space) of the first component of  $\varphi$ , since the map is expanding only in one direction. This result is fully consistent with early non-systematic attempts to construct such a procedure for  $g$  in [29, 30]. The previous studies used the measure preservation property to derive an analogous version of Eq. 3.10.

To verify the correctness of our procedure, we consider the 2D perturbed Baker's map  $\varphi : M \rightarrow M$ , with  $M = [0, 2\pi]^2$ , defined as follows [10],

$$(3.11) \quad x_{k+1} = \varphi(x_k) = \left( \left[ \begin{array}{c} 2x_k^{(1)} \\ x_k^{(2)}/2 + \pi \lfloor x_k/\pi \rfloor \end{array} \right] + \left[ \begin{array}{c} s_1/2 \sin(x_k^{(1)}/2) + s_2/2 \sin(2x_k^{(1)}) \sin(x_k^{(2)}) \\ s_3 \sin(x_k^{(2)}) + s_4/2 \sin(2x_k^{(1)}) \sin(x_k^{(2)}) \end{array} \right] \right) \bmod 2\pi,$$

where  $s_1, s_2, s_3, s_4$  are real-valued map parameters. If all of them are zero, we obtain the classical Baker's map (first term of the RHS of Eq. 3.11), which is named after the kneading operation that bakers apply to a two-dimensional square dough. In particular, the dough is first stretched horizontally (in the unstable direction) by a constant factor, then compressed vertically (in the stable direction) by the same factor, and so forth. The square-shaped domain is stretched to a  $2 \times 1$  rectangle, cut into two squares, which are subsequently stacked horizontally. The Baker's map is an invertible chaotic map with one positive and one negative Lyapunov exponent.

By introducing an extra term proportional to the four parameters, we perturb the kneading operation in the direction not necessarily aligned with the phase space directions. Indeed, by manipulating these parameters' values, we can control the shape of the unstable manifold, which gives us an excellent study case in the context of the SRB gradient computation. Notice, for example, if  $s_4 = 0$  and  $s_3$  is sufficiently small, the iteration in Eq. 3.9 produces  $q_k$  whose second coordinate,  $q_k^{(2)}$ , converges exponentially to zero with  $k$ . In this case, therefore, unstable manifolds are straight and aligned with the  $x^{(1)}$ -axis. We use this observation to design our first numerical test.

In the first experiment, we consider the Baker's map defined by Eq. 3.11 with  $s_1 = s_3 = s_4 = 0$  and  $s_2 = 0.4$ . The left-hand side plot in Figure 3 illustrates the normalized SRB distribution corresponding to this parameter choice, which represents

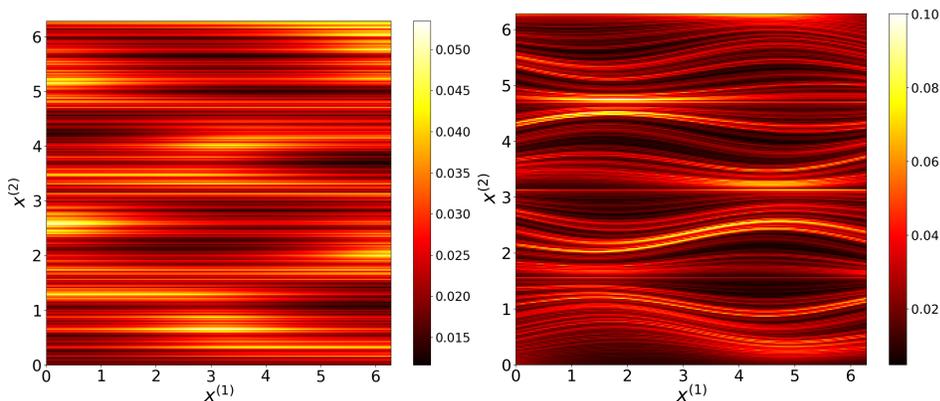


FIG. 3. SRB distribution of the Baker's map with  $s_1 = s_3 = s_4 = 0$ ,  $s_2 = 0.4$  (left plot) and  $s_1 = s_2 = s_3 = 0$ ,  $s_4 = 0.4$  (right plot). We divided  $M$  into  $256^2$  rectangular bins of equal width and counted the number of times the trajectory passed through each of these bins. In this experiment, we generated 8000 trajectories of length 209, 715, 200, which gives us a total of approximately  $1.68 \cdot 10^{12}$  samples.

the probability of the trajectory passing through each square bin everywhere on  $M$  (see the caption of Figure 3 for more details; for completeness, we also included a case with  $s_4 \neq 0$ ). We observe a smooth behavior of the SRB distribution with respect to  $x^{(1)}$  at any vertical level  $x^{(2)}$ . However, as we travel vertically, in the stable direction, the SRB distribution varies sharply. These radically different behaviors are typical symptoms of Property (3) and Property (4) of  $\mu$  described in Section 2, and they can also be observed in Figure 4, where the conditional and marginal SRB distributions are plotted, using data from Figure 3.

In Figure 4, we also plot the SRB density gradients defined on five different unstable manifolds. To compute  $g$ , the simplified recursion from Eq. 3.10 was directly applied. To validate our computation, we approximated  $g$  by applying the central finite-difference method to SRB densities plotted above. We observe a good agreement between the results obtained with these two different approaches, which confirms the correctness of our algorithm.

To conclude, in case of straight unstable manifolds, the SRB density gradient can be computed using the simplified recursive relation along trajectory (Eq. 3.10), which we verify through finite-differencing. This iteration is computationally cheap, as it involves solving a scalar tangent equation featuring both the first and second derivative of the first component of  $\varphi$ . In Appendix A, we show Eq. 3.10 can also be applied to popular one-dimensional maps that are non-injective. We argue that certain non-measure-preserving transformations have their higher-dimensional analogs similar to the classical Baker's map. Appendix B presents a numerical study confirming the hyperbolicity of the Baker's map.

**4. Computing SRB density gradient for systems with general unstable manifolds.** We shall generalize the concepts introduced in Section 3 to systems with  $m$ -dimensional unstable manifolds,  $m \in \mathbb{Z}^+$ . In other words, we consider general  $n$ -dimensional chaotic systems that have  $m$  positive LEs,  $1 \leq m \leq n$ . In this setting, the chart  $x_k(\xi)$ ,  $k \in \mathbb{Z}^+$ , is a diffeomorphism that maps an  $m$ -dimensional hypercube,  $[0, 1]^m$ , to the local unstable manifold  $U_k \subset M$ . For example, if  $m = 2$  and  $n = 3$ , then the system has two positive LEs and its unstable manifolds are planes immersed

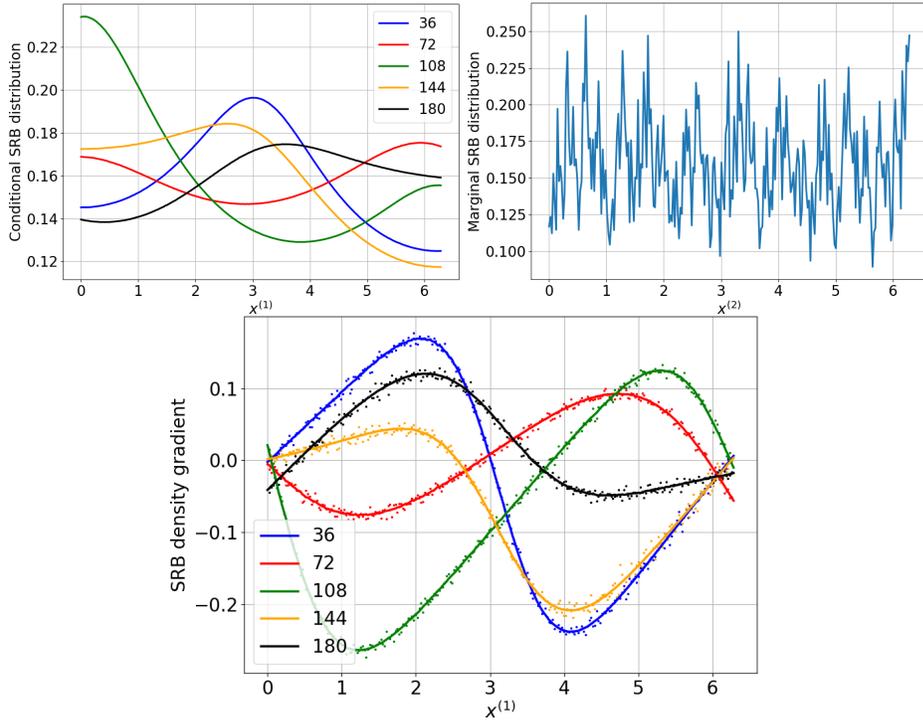


FIG. 4. Upper left plot: conditional SRB distributions (SRB densities) corresponding to five different unstable manifolds. The numbers 36, 72, 108, 144, 180 appearing in the legend represent the index of the horizontal bin row. For example, the red line corresponds to the SRB density defined on the unstable manifold at  $x^{(2)} \approx 72/256 \cdot 2\pi \approx 1.76$ . Upper right plot: marginal SRB distribution obtained through integrating the first coordinate out. Lower plot: SRB density gradient  $g$  corresponding to SRB densities plotted in Figure 4. The  $g$  function was computed using two distinct approaches: through the simplified trajectory-based recursion (Eq. 3.10) (solid lines), and the central finite-difference method (dots). The oscillation of the finite-difference approximation is a manifestation of the statistical noise.

in  $\mathbb{R}^3 \supset M$ .

**4.1. Derivation of the iterative formula.** As introduced above, let us consider an  $m$ -dimensional smooth unstable manifold  $U_k$  described by the chart  $x_k(\xi) : [0, 1]^m \rightarrow U_k \subset M$ . The vectors  $x_k = [x_k^{(1)}, \dots, x_k^{(n)}]^T$  and  $\xi = [\xi^{(1)}, \dots, \xi^{(m)}]^T$  have  $n$  and  $m$  components, respectively, and  $0 \leq \xi^{(i)} \leq 1$ ,  $i = 1, \dots, m$ . We use  $\nabla_\xi x_k(\xi)$  to denote the parametric gradient tensor of the chart. The  $i$ -th column of  $\nabla_\xi x_k(\xi)$  contains the derivative of  $x_k(\xi)$  with respect to  $\xi^{(i)}$ , i.e.,  $\partial_{\xi^{(i)}} x_k(\xi)$ . For any Borel subset  $V \subset [0, 1]^m$  such that  $x_k(V) = B_k \subset U_k$ , the SRB measure-density relation can be expressed as follows,

$$(4.1) \quad \mu(V) = \int_{B_k} \tilde{\rho}_k(x) d\omega(x),$$

where  $d\omega(x)$  denotes the natural volume element defined everywhere on  $U_k$ . Analogously to the 1D case described in Section 3,  $\tilde{\rho}_k$  represents the conditional SRB density defined on  $U_k$ . If we QR-factorize the parametric gradient of  $x_k(\xi)$ ,

$$(4.2) \quad \nabla_\xi x_k(\xi) = Q(x_k(\xi)) R(x_k(\xi))$$

at any  $\xi \in [0, 1]^m$ , the density conservation property can be expressed as

$$(4.3) \quad \tilde{\rho}(x_k(\xi)) |\det R(x_k(\xi))| = 1,$$

which is a generalization of Eq. 3.3. By differentiating Eq. 4.3 with respect to  $\xi$  and applying a non-trivial chain rule, we obtain

$$(4.4) \quad g^{(i)}(x_k(\xi)) = \partial_{Q^{(i)}(x_k(\xi))} \log \tilde{\rho}(x_k(\xi)) = - \frac{\text{tr} (Q^T(x_k(\xi)) \partial_{\xi^{(i)}} \nabla_{\xi} x_k(\xi) R^{-1}(x_k(\xi)))}{\|\partial_{\xi^{(i)}} x_k(\xi)\|},$$

or, equivalently,

$$(4.5) \quad g^{(i)}(x_k(\xi)) = - \frac{Q^{(j)}(x_k(\xi)) \cdot \partial_{\xi^{(i)}} \partial_{\xi^{(k)}} x_k(\xi) (R^{-1})^{(kj)}(x_k(\xi))}{\|\partial_{\xi^{(i)}} x_k(\xi)\|}$$

for all  $\xi \in [0, 1]^m$ , where the repeated indices imply summation (Einstein's convention), while the superscript  $i$  denotes the  $i$ -th column of a matrix. This expression was obtained by employing the orthonogonality of  $Q$  and upper-triangular structure of  $R$ . It is computationally convenient as it does not involve parametric derivatives of the determinant of  $R$ . The reader is referred to [31] for a step-by-step derivation of Eq. 4.3-4.5.

The purpose of this section is to derive an iterative (trajectory-driven) procedure for  $g$ . Analogously to the derivation in Section 3, we combine Eq. 4.5, the evolution equation

$$(4.6) \quad x_{k+1}(\xi) = \varphi(x_k(\xi)),$$

and apply the chain rule. The 1D case, however, was computationally simpler because the tangent equations for  $a$  and  $q$  were regularized by the scalar  $\alpha$  every time step preventing the tangent solutions from blow-ups due to the positive LE (i.e., the butterfly effect). Note that here we need to compute all first- and second-order parametric derivatives of the chart to compute  $g$ . Since we strive to derive a recursive relation, we regularize tangent equations in a fashion analogous to the approach in Section 3. To achieve this goal, one can recursively orthonormalize the parametric gradient through an iterative linear transformation of the parameterization and fixing  $\xi = 0$ . In particular, we change variables from step  $k$  to  $k + 1$  such that

$$\xi_{k+1} = R_{k+1} \xi_k.$$

Note that at  $\xi = 0$  we stay on the same trajectory despite the transformation. This particular choice of  $\xi$  does not restrict our algorithm to concrete trajectories. Indeed, we want to “visit” all infinitesimally small  $\mu$ -typical regions of the attractor after an infinite number of time steps, regardless of the choice of the initial condition. Therefore, we can always linearly re-scale the feasible space of  $\xi$  such that  $\xi = 0$  for our arbitrary choice of the initial condition. To simplify the notation, we skip the argument in our notation whenever  $\xi = 0$ ; for example, we use the short-hand notation  $x_k(0) := x_k$ ,  $Q(x_k(0)) := Q_k$ , and so forth. Thanks to this particular transformation, the parametric gradient is automatically orthonormalized, because the chain rule implies that

$$(4.7) \quad \nabla_{\xi_{k+1}} x_{k+1} = \nabla_{\xi_k} x_{k+1} R_{k+1}^{-1} = Q_{k+1},$$

or, equivalently,

$$(4.8) \quad \partial_{\xi_{k+1}^{(i)}} x_{k+1} = \partial_{\xi_k^{(j)}} x_{k+1} (R_{k+1}^{-1})^{(ji)} = Q_{k+1}^{(i)}.$$

It means that the parametric gradient of the chart has an orthonormal basis of the column space in the updated coordinate system. Note the  $R$  matrix represents the Jacobian of the step-to-step parametric transformation, i.e.,  $R_{k+1} = \partial \xi_{k+1} / \partial \xi_k$ . In an analogous manner, we can derive a similar relation for the Hessian of  $x_{k+1}$ , represented by an  $n \times m \times m$  tensor,

$$(4.9) \quad \partial_{\xi_{k+1}^{(i)}} \partial_{\xi_{k+1}^{(j)}} x_{k+1} = \partial_{\xi_k^{(p)}} \partial_{\xi_k^{(q)}} x_{k+1} (R_{k+1}^{-1})^{(pi)} (R_{k+1}^{-1})^{(qj)}.$$

The major benefit of the variable change is a dramatic simplification of Eq. 4.5. Indeed, in the orthonormalized system, the  $R$  matrix reduces to the identity matrix, while the norm of each column of the parametric gradient equals 1. This gives rise to the following expression for  $g$ ,

$$(4.10) \quad \begin{aligned} g_{k+1}^{(i)} &= -\text{tr} \left( Q_{k+1}^T \partial_{\xi_k^{(i)}} \nabla_{\xi_k} x_{k+1} \right) = \\ &= -Q_{k+1}^{(i,j)} \cdot \partial_{\xi_{k+1}^{(i)}} \partial_{\xi_{k+1}^{(j)}} x_{k+1} := -Q_{k+1}^{(j)} \cdot a_{k+1}^{(i,j)}. \end{aligned}$$

Consequently, only two ingredients are necessary to compute the density gradient function at  $\xi = 0$ . First, we need the orthogonal basis of the column space of the parametric gradient  $\nabla_{\xi_k} x_{k+1}$ . A recursive formula for the basis can be obtained by differentiating the system in Eq. 4.6 and performing QR factorization, i.e.,

$$(4.11) \quad \nabla_{\xi_k} x_{k+1} = D\varphi_k \nabla_{\xi_k} x_k = Q_{k+1} R_{k+1}.$$

Therefore, per Eq. 4.7, we automatically obtain the orthonormal parametric gradient at the next time step without the need of inverting  $R_{k+1}$ . Since the orthonormalization is performed in a recursive manner,  $\nabla_{\xi_k} x_k = Q_k$  by construction. To complete the algorithm, we also need a recursion for  $a$ . This equation can be naturally derived by differentiating Eq. 4.6 twice, which gives rise to

$$(4.12) \quad \begin{aligned} \partial_{\xi_k^{(i)}} \partial_{\xi_k^{(j)}} x_{k+1} &= D^2 \varphi_k (\partial_{\xi_k^{(i)}} x_k, \partial_{\xi_k^{(j)}} x_k) + D\varphi_k \partial_{\xi_k^{(i)}} \partial_{\xi_k^{(j)}} x_k \\ &= D^2 \varphi_k (Q_k^{(i)}, Q_k^{(j)}) + D\varphi_k a_k^{(i,j)}. \end{aligned}$$

Note that in order to compute the SRB density gradient at step  $k+1$ , we need to apply the Hessian re-scaling described by Eq. 4.9 to retrieve  $a_{k+1}$ . We summarize this algorithm and carefully analyse its computational properties in Section 4.2.

The procedure in Eq. 4.11 reduces to the recursion in Eq. 3.9 if  $m = 1$ . Regardless of the choice of initial condition  $Q_0$ , the column vectors of  $Q_k$  rigorously converge to backward Lyapunov vectors as  $k \rightarrow \infty$  [18]. The set of these column vectors is in fact an orthonormal basis of the unstable (expanding) subspace  $E_k^u$  of the tangent space  $TM_k$ . Specific directions of backward Lyapunov vectors, however, depend on the choice of  $Q_0$ . Therefore, in this case, the ‘‘convergence’’ should be understood that, for any orthonormal  $Q_0$ , the column space of  $Q_k$  is guaranteed to coincide with some orthonormal basis of  $E_k^u$  if  $k \rightarrow \infty$ . A similar procedure can be used to compute all  $n$  Lyapunov vectors, including those corresponding to the negative LEs, spanning the stable (contracting) subspace  $E_k^s$ . In uniformly hyperbolic systems,  $TM_k = E_k^u \oplus E_k^s$

at every  $k$ , and both the subspaces are  $D\varphi$ -invariant (or *covariant*). The covariance property implies that the product  $D\varphi_k Q_k$ , which we compute in Eq. 4.11, outputs  $m$  vectors that belong to the unstable subspace of the tangent space at the next time step,  $TM_{k+1}$ . In general, the new vectors are not orthonormal. By performing the QR factorization, however, we obtain an orthonormal basis of the unstable subspace at  $k+1$ . Therefore, the components of  $R_{k+1}$  contain projections of the column vectors of  $D\varphi_k Q_k$  onto the basis vectors of  $E_{k+1}^u$ .

We also observe the general recursion for the acceleration vector  $a$  in Eq. 4.12 can be simplified to its one-dimensional counterpart in Eq. 3.8 if  $m = 1$ . Using the properties of uniform hyperbolicity, the authors of [10, 9] rigorously show the recursion for  $a$  (a.k.a. unstable manifold curvature equation) in systems with one positive LE rigorously converges at an exponential rate. To the best of our knowledge, no rigorous results for higher-dimensional cases exist. The proof of convergence for systems with one-dimensional unstable subspaces uses the fact  $a_k$  can be expressed as  $C_k + D\varphi_{k-1} \dots D\varphi_0 a_0 / \prod_{i=0}^{k-1} \alpha_{k-1}^2$ , where  $C_k$  does not depend on  $a_0$  (see Section 3 for the notation explanation). By the uniform expansion property, the  $a_0$ -dependent term exponentially converges to zero if  $k \rightarrow \infty$ . In case of general systems, we find similar dependencies between  $a_k^{i,j}$  and all initial conditions for the second-order tangent equation. Here, instead of re-scaling with respect to the length of the projection of  $D\varphi_k q_k$  onto  $q_{k+1}$ , we are re-scaling with respect to the collection of projections included in the  $R_{k+1}$  matrix. Moreover, the process of computing  $a$  in the general case involves inverting  $R_{k+1}$ , not just a scalar, which makes it hard to directly apply the properties resulting from the uniform hyperbolicity assumption. Therefore, in this paper, we resort to an empirical study of the convergence of our algorithm (see Section 4.2). Note that even if the recursion converges, the specific direction of  $a$  is not unique at any point on  $M$ , because  $Q$  is also not unique as discussed above. Their product, however, that equals the SRB density gradient  $g$  is unique by construction.

**4.2. General algorithm for high-dimensional systems.** We provide a practicable algorithm based on the derivation presented in the previous section. In addition, we carefully analyse its computational cost, memory requirements, and numerically investigate its convergence. Algorithm 1 summarizes all steps necessary to numerically compute the SRB density gradient at  $N$  points along a trajectory initiated at  $x_0 \in M \subset \mathbb{R}^n$ . The only optional step is included in Line 1; this step is meant to compute the dimension of the unstable subspace/manifold  $m$ . For many chaotic maps, this parameter is known *a priori* and therefore Line 1 can be skipped. If this is not the case, however, one can apply Benettin et al.'s numerical procedure [5] to approximate a subset of the spectrum of Lyapunov exponents. This procedure requires solving  $i \in \mathbb{Z}^+$  homogeneous tangent equations to identify  $i$  largest LEs. The parameter  $T$  represents the trajectory length and affects the accuracy of LE approximation. If the LE spectrum is evidently separated from the origin (i.e., the value of 0), then  $T$  does not need to be large. Lines 3-22 of Algorithm 1 represent the main time for-loop that computes the  $g$  vector at one point on the manifold per iteration. Inside this loop, we distinguish five major stages: 1) advancing first-order tangent equation and QR factorization (Eq. 4.11), 2) advancing second order tangent equations (Eq. 4.12), 3) inverting the  $R$  matrix and rescaling the acceleration vector  $a$  (Eq. 4.9), 4) evaluating  $g$  (Eq. 4.10), and 5) transitioning to the next time step; updating the Jacobian and Hessian.

Table 1 summarizes the computational cost of Algorithm 1. The third column of this table includes the number of the floating point operations required in each stage

**Algorithm 1:** SRB density gradient

---

```

Input :  $N, T, x_0, n = \text{size}(x_0)$ 
1  $m = \text{Benettin}(T)$  if  $m$  unknown;
2 Randomly generate  $Q_0, a_0^{(i,j)}$  such that  $\text{ncol}(Q_0) = m$ ,
    $\text{nrow}(Q_0) = \text{size}(a_0^{(i,j)}) = n, Q_0^T Q_0 = I$ , and  $i, j = 1, \dots, m$ ;
3 for  $k = 0, \dots, N - 1$  do // main time loop
4    $S_k = D\varphi_k Q_k$ ;
5   QR-factorize:  $Q_{k+1} R_{k+1} = S_k$ ;
6   Invert  $R_{k+1}$ ;
7   for  $i = 1, \dots, m$  do // 2nd-order tangent equations
8     for  $j = 1, \dots, i$  do
9        $\tilde{a}_{k+1}^{(i,j)} = D^2\varphi_k(Q_k^{(i)}, Q_k^{(j)}) + D\varphi_k a_k^{(i,j)}$ ;
10    end
11  end
12  for  $i = 1, \dots, m$  do // re-scaling
13    for  $j = 1, \dots, i$  do
14       $a_{k+1}^{(i,j)} = \tilde{a}_{k+1}^{(p,q)} (R^{-1})_{k+1}^{(pi)} (R^{-1})_{k+1}^{(qj)}$ ;
15    end
16  end
17  for  $i = 1, \dots, m$  do // evaluating  $g$ 
18     $g_{k+1}^{(i)} = -Q_{k+1}^{(i,j)} \cdot a_{k+1}^{(i,j)}$ ;
19  end
20   $x_{k+1} = \varphi(x_k)$ ;
21  Evaluate:  $D\varphi_{k+1}$  and  $D^2\varphi_{k+1}$ ;
22 end

Output:  $g_k^{(i)}, i = 1, \dots, m, k = 1, \dots, N - 1$ 

```

---

as a function of the trajectory length ( $N$  or  $T$ ), system dimension  $n$ , and unstable manifold dimension  $m$ . Note the third column includes only the leading term of the flop count. The final two stages involve evaluations of nonlinear equations defined by  $\varphi$  and thus their computational cost is problem-dependent. In many physics-inspired chaotic systems the cost of Lines 20-21 is relatively low. Consider the Lorenz '63 system, for example. In this case, we can think of  $\varphi$  as a time discretization operator of the continuous-in-time system. For Lorenz '63, the Jacobian  $D\varphi$  involves a collection of linear terms proportional to the coordinates of  $x$ , while the Hessian  $D^2\varphi$  is constant. In many scientific/engineering applications, PDE models are discretized in space using schemes with local support (such as the finite element method), which implies the resulting Jacobians and Hessians of the fully-discretized system are sparse. Therefore, in these special cases, the cost of the most expensive stage of Algorithm 1, which involves second-order tangent equations, can potentially be reduced to  $N n m^2$ . Table 1, however, reflects the worst-case scenario in which no sparsity patterns occur. We also highlight the fact that in many high-dimensional chaotic systems  $m \ll n$  [6]. Thus, if  $n$  is large, the re-scaling stage (Lines 12-16) is rather cheaper than the second-order tangent equation stage (Lines 7-11).

We conclude that the leading term of the total flop count of Algorithm 1 is proportional to  $N n^3 m^2$  in a general chaotic system. In many real-world problems,

TABLE 1  
Computational cost of Algorithm 1.

Stage Name	Line No.	Total Cost
Computing $m$ (Benettin's algorithm)	1	$T n^2 m$
Generating initial conditions	2	–
Advancing first-order tangent equations	4	$N n^2 m$
QR factorization (Householder)	5	$N n m^2$
Inverting $R$	6	$N m^3$
Advancing second-order tangent equations	7-11	$N n^3 m^2$
Re-scaling $a$	12-16	$N n m^4$
Computing $g$	17-19	$N n m^2$
Advancing primal equation	20	Varies
Evaluating Jacobian and Hessian	21	Varies

however, the final cost can be significantly reduced if one takes the advantage of a system's special structure. Our algorithm is moderately cheap in terms of the memory requirements. The most memory-consuming structure is the Hessian which, in the worst-case scenario, requires storing  $n^3$  floats. As we pointed out above, however, in practical high-dimensional models, the actual "size" of the Hessian might be dramatically smaller. Note also that, in order to advance tangent equations, we need to store  $m$   $n$ -dimensional basis vectors (i.e., column vectors of  $Q$ ) and  $\sim 1/2 m^2$  acceleration vectors. The  $1/2$  factor is a consequence of the assumed smoothness of the coordinate chart, which implies  $a^{(i,j)} = a^{(j,i)}$  everywhere on the manifold. Notice also that the our procedure is in fact a one-step method, which means that all quantities at step  $k + 1$  require data only from step  $k$ . We do not need to store data generated at previous time steps.

The final aspect of our algorithm is its convergence. As we mentioned in the previous sections, the convergence of the Lyapunov vector equation and the second-order tangent (unstable manifold curvature) equation is rigorously guaranteed in uniformly hyperbolic systems if  $m = 1$ . Moreover, the convergence rate is exponential in such systems. Since we lack generalization of these theoretical studies, we perform a numerical test to investigate convergence properties of Algorithm 1. For this purpose, we use the Baker's map introduced in Eq. 3.11, as well as its 3D version  $\varphi : [0, 2\pi]^3 \rightarrow [0, 2\pi]^3$  defined as

$$(4.13) \quad x_{k+1} = \varphi(x_k) = \left( \begin{array}{c} \left[ \begin{array}{c} 2 x_k^{(1)} \\ 3 x_k^{(2)} \\ x_k^{(3)}/6 + \pi \lfloor x_k^{(1)}/\pi \rfloor + \pi/3 \lfloor x_k^{(2)}/(2\pi/3) \rfloor \end{array} \right] \\ + \left[ \begin{array}{c} s_1 \sin(2x_k^{(1)}) \sin(3/2 x_k^{(2)}) \\ s_2 \sin(x_k^{(1)}) \sin(3 x_k^{(2)}) \\ s_3 \sin(6 x_k^{(3)}) \end{array} \right] \end{array} \right) \bmod 2\pi,$$

which we shall refer to as the 3D Baker's map. This is an invertible chaotic map with two positive and one negative LEs, and seemingly hyperbolic behavior (see Appendix B for more details). This map has two expanding directions, along the  $x^{(1)}$  and  $x^{(2)}$  axes, and one contracting direction along the third axis. Analogously to its 2D counterpart, this map models the kneading operation. The dough is extended by the factor 2 and 3 along the two orthogonal directions on the  $x^{(1)} - x^{(2)}$  plane,

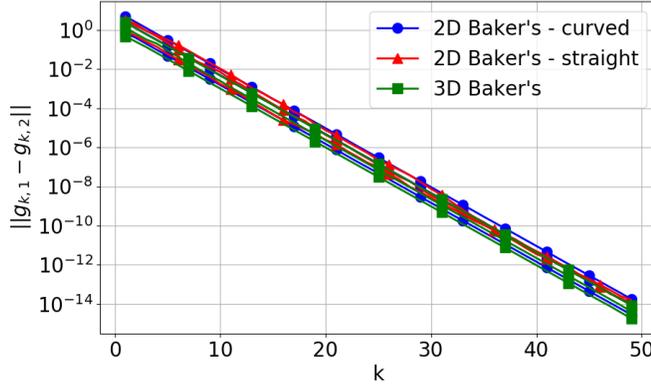


FIG. 5. Relation between  $\|g_{k,1} - g_{k,2}\|$  and the time step  $k$  in the semilogarithmic scaling. This plot contains nine curves of three different colors. Each color corresponds to a different map: 2D Baker's map with curved unstable manifolds (blue), 2D Baker's map with straight unstable manifolds (red), and 3D Baker's map with  $s_1 = 0$ ,  $s_2 = 0.9$ ,  $s_3 = 0.1$  (green). In case of the 2D Baker's map, the parameter values are the same as those in Figure 3.

cut into  $2 \cdot 3 = 6$  squares, which are subsequently stacked in the order defined by the floor functions. These history-dependent floor functions are used to guarantee the invertibility of the nonlinear transformation by periodically distributing the third component of  $x_{k+1}$  across  $[0, 2\pi]$ . Higher-dimensional Baker's maps have been widely used in image encryption as a convenient generalization of Bernoulli shifts [25, 21].

To analyse the convergence, we generate three sufficiently long trajectories started at randomly chosen initial conditions  $x_0$ . For each of these trajectories, we run two independent simulations with different, randomly chosen initial conditions for the tangent equations (see Line 2 of Algorithm 1). Motivated by the rigorous studies, we investigate if (and how) the difference between the SRB density gradients computed along a single trajectory but using different initial conditions for tangent equations decreases in norm as we advance the iteration. In particular, we compute  $\|g_{k,1} - g_{k,2}\|$ ,  $k = 0, 1, 2, \dots$  for two random initial condition choices for tangent equations per trajectory, labelled as 1 and 2. The relation between this norm and time step  $k$  for three different chaotic models is plotted in Figure 5. The  $g$  function is generated using Algorithm 1 for the 2D Baker's map with  $m = 1$  (Eq. 3.11), as well as the 3D Baker's map with  $m = 2$  (Eq. 4.13).

We observe the norm-versus- $k$  relation is linear in the semilogarithmic scale, which clearly indicates an exponential convergence of our algorithm if applied to the Baker's map. This result implies that a relatively small number of steps ( $k \approx 50$ ) is required to obtain the machine-precision value of the norm. Note also that the choice of trajectory ( $x_0$ ) or model has a negligible effect on the error.

**4.3. Numerical example: Monte Carlo integration.** To validate Algorithm 1, we consider a square-integrable function  $f(x) \in L^2(\mu)$  and integrate it with respect to the SRB measure  $\mu$  using a Monte Carlo procedure. By the Central Limit Theorem, this integral can be approximated by taking the average of the sample distributed according to  $\mu$ , while the approximation error is upper-bounded by  $\sqrt{\text{Var}(f)/N}$ , i.e.,

$$(4.14) \quad \left| \int_M f(x) d\mu(x) - \frac{\omega(M)}{N} \sum_{k=0}^{N-1} f(x_k) \right| \leq C \sqrt{\frac{\text{Var}(f)}{N}},$$

where  $C > 0$  and  $x_{k+1} = \varphi(x_k) \in M$ . Therefore, by generating a sufficiently long trajectory and evaluating  $f$  at every point along it, we gradually approach the sought-after solution. Motivated by particular applications of the SRB density gradient function (see Sections 1-2), we consider  $f(x) := \sum_{j=1}^m \partial_{Q^{(j)}} v(x)$ , where  $v(x) : M \rightarrow \mathbb{R}$  is some smooth function. In other words, we strive to integrate a sum of  $m$  directional derivatives along  $m$ -dimensional unstable manifolds of the scalar function  $v(x)$ . Note integrals of this type are critical in the sensitivity computation using, for example, the general S3 method [8, 10]. Thanks to the partial integration (see Eq. 2.5-2.11), we can apply Monte Carlo to two alternative versions of the same integral, since

$$(4.15) \quad \int_M \sum_{j=1}^m \partial_{Q^{(j)}} v(x) d\mu(x) = I = - \int_M \sum_{j=0}^m g^{(j)}(x) v(x) d\mu(x).$$

Using this equation, we validate Algorithm 1 for  $g$  by comparing numerical approximations of the LHS and RHS. Due to its trajectory-driven structure, Algorithm 1 is naturally compatible with the Monte Carlo procedure.

Two different maps shall be tested. First, we shall consider the 2D Baker's map (Eq. 3.11) with  $s_4 = 0.4$  and  $s_1 = s_2 = s_3 = 0$ . As illustrated in Figure 3, its unstable manifolds are curved and therefore the simplified version of the recursion for  $g$  (Eq. 3.10) cannot be used. In this particular case,  $q$  has in fact two nonzero components. Indeed, we numerically estimate that

$$\max_{k \in \{1, 2, \dots, N\}} \arctan \left| \frac{q_k^{(2)}}{q_k^{(1)}} \right| \approx 0.24 \text{ rad} \approx 14^\circ,$$

which is consistent with the illustration of unstable manifolds in Figure 3. The second map is the 3D Baker's map (Eq. 4.13) with  $s_1 = 0, s_2 = 0.9, s_3 = 0.1$ . One can easily verify unstable manifolds of this map are flat surfaces aligned with the  $x^{(1)}-x^{(2)}$  plane. These expanding surfaces could be curved by adding an  $x^{(3)}$ -dependent perturbation term to the third component of the map.

Figure 6 includes results of the integration test. Our primary conclusion is that the Monte Carlo approximations of the LHS and RHS of Eq. 4.15 approach each other as  $N \rightarrow \infty$  with the rate  $\mathcal{O}(1/\sqrt{N})$ , which directly confirms the correctness of Algorithm 1. Recall we require  $g$  to regularize the linear response formula, as it involves derivatives of strongly-oscillatory functions (see Section 2). The examples presented in this section, however, include mildly-oscillatory functions  $v(x)$  with derivatives that behave similarly (note they involve a combination of trigonometric, exponential and linear functions). Nevertheless, we observe significantly smaller errors of the RHS approximation in the 2D Baker's map case. Note the approximation error of Monte Carlo integration also depends on the variance of the integrand, which can be upper-bounded by a quantity proportional to the  $L^2(\mu)$ -norm of the SRB density gradient  $g$ , denoted by  $\|g\|_{L^2(\mu)}$ . Indeed,  $\|g\|_{L^2(\mu)}$  equals  $\mathcal{O}(10^{-2})$  and  $\mathcal{O}(10^1)$  for the 2D and 3D Baker's map, respectively. This explains the significantly better performance of the Monte Carlo procedure in the former case. Therefore, if  $\|g\|_{L^2(\mu)}$  exists and is sufficiently small, Monte Carlo integration might be significantly cheaper if applied to the regularized integrals of this type, regardless of the behavior of  $v(x)$ . If  $g$  is not even Lebesgue-integrable, i.e.  $g \notin L^1(\mu)$ , the integrals in Eq. 4.15 do not converge, as showed in [30].

**5. Conclusions.** Ruelle's linear response formula is fundamental in the construction of numerical methods for sensitivity analysis of  $n$ -dimensional hyperbolic

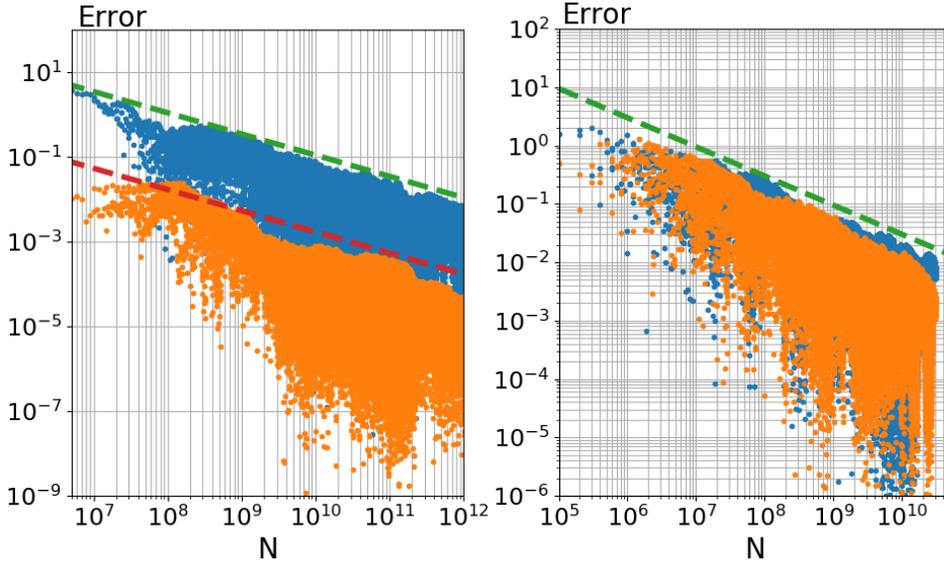


FIG. 6. Error of the Monte Carlo approximation of the LHS (blue dots) and RHS (orange dots) of Eq. 4.15 versus the amount of data  $N$  used. Left: 2D Baker's map. Here, we compute the relative error with respect to the reference value  $-1.05335809$  (which equals the approximation of the RHS integral at  $N = 10^{13}$ ) for  $v(x) = \sin(x^{(1)}) \exp(x^{(2)})$ . Right: 3D Baker's map. Here, we compute the absolute error with respect to the reference value of 0 for  $v(x) = \sin(x^{(1)}) \sin(3/2 x^{(2)}) x^{(3)}$ . The dashed lines represent the slope  $-1/2$  in the logarithmic scaling.

chaotic systems. Its original form, however, is impractical for direct computation due to the presence of derivatives of composite functions that grow exponentially in time. Fortunately, it is possible to easily regularize this expression through partial integration. In case of nonuniform measures describing the statistics of chaos, the by-product of the integration by parts, per the generalized fundamental theorem of calculus, involves the SRB density gradient  $g$  defined as the directional derivative of conditional SRB density on  $m$ -dimensional unstable manifolds. Computation of  $g$  is the price that must be paid for a computable version of Ruelle's formula.

Using the measure-based coordinate parameterization, the time evolution of the measure gradient is rigorously derived by applying the measure preservation property, differentiating the coordinate charts with the chain rule on smooth manifolds. Indeed,  $g$  can be computed in a recursive manner by solving a set of  $\mathcal{O}(m)$  first- and  $\mathcal{O}(m^2)$  second-order tangent equations, as well as step-by-step QR-factorization and inversion of  $n \times m$  and  $m \times m$  matrices, respectively. While the total cost of approximating  $g$  at  $N$  consecutive points along a trajectory is  $\mathcal{O}(Nn^3m^2)$  in the worst-case scenario, the actual computational cost may scale linearly with the dimension of the system in many real-world models due to their sparse structure. Moreover, this procedure requires storing  $\mathcal{O}(m^2)$   $n$ -dimensional vectors only from the current time step to advance the iteration in time. Therefore, in terms of the hardware requirements, our algorithm would definitely be a reasonable choice for high-dimensional physical systems since  $m \ll n$ .

The algorithm we propose is compatible with existing methods for sensitivity analysis that stem from the linear response theory, including the space-split sensitivity (S3) and FDT-based methods. Many of them approximate sensitivities through an ergodic-averaging Monte Carlo procedure and require knowledge of the directional

derivative of conditional SRB measures. Moreover,  $g$  can be used to assess the differentiability of statistical quantities in hyperbolic systems, which is a recurring theme in theoretical studies of chaos. Thus, we believe our method provides a new major tool for both rigorous analysis and applied studies of large chaotic systems.

**Acknowledgments.** The authors acknowledge the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing HPC resources that have contributed to the research results reported within this paper.

### Appendix A. Applying the simplified recursive formula for SRB density gradient to 1D non-injective maps.

Throughout this paper, we assume  $\varphi$  is an invertible map. Based on this assumption, we directly use the measure preservation property to derive a recursive formula for  $g$ , including the simplified version for maps with straight unstable manifolds, as described in Section 3.2. However, in the literature, one can find several one-dimensional maps such as the sawtooth/Bernoulli map [29], cusp map [22], logistic map [33], onion map [30], tent map [4], and so forth. All of them are scientifically relevant, as they represent some simplified physics or feature interesting mathematical properties. However, most of them are non-injective, which violates the basic assumption of our derivation. In this section, however, we argue that Eq. 3.10 can still be used to compute  $g$  for such maps.

Many of the popular 1D chaotic maps (such as those listed above) are two-to-one. Thus, we assume  $\varphi$  satisfies this condition; however, the argument we present can be naturally extended to other types of surjection. Let us also assume, without loss of generality,  $\varphi : [0, 1] \rightarrow [0, 1]$  and  $\varphi$  is monotonic in  $[0, 0.5]$  and  $(0.5, 1]$ . Let us now define a two-dimensional analog of  $\varphi$ , denoted by  $\varphi_{2D} : [0, 1]^2 \rightarrow [0, 1]^2$  and satisfying

$$(A.1) \quad x_{k+1} = \varphi_{2D}(x_k) = \begin{bmatrix} \varphi(x_k^{(1)}) \\ x_k^{(2)}/2 + 0.5 \lfloor 2x_k^{(1)} \rfloor \end{bmatrix}.$$

Note  $\varphi_{2D}$  is invertible and resembles the 2D Baker's map (see Eq 3.11). The invertibility is guaranteed by adding the floor function in  $\varphi_{2D}^{(2)}$ . Analogously to the 2D/3D Baker's map, here the discontinuity point is located at  $x^{(1)} = 0.5$ , which means that the value of 0.5 is added to  $x_k^{(2)}/2$  if  $x_k^{(1)} > 0.5$ . If the monotonicity breaking point was different, then the coefficients of the floor function would need to be modified accordingly. One of the main messages of this example is to point out that any surjective 1D map can be represented as a higher-dimensional invertible map with one positive Lyapunov exponent.

Note 1D unstable manifolds of  $\varphi_{2D}$  are aligned with the first phase space coordinate, per the argument given in Section 3.2. Thus, its SRB distribution is similar to the one of Baker's map presented in Figure 3. Note also that the horizontal deformation of the trajectory of  $\varphi_{2D}$  is determined solely by  $\varphi$ . This implies that the SRB distribution of  $\varphi$  is in fact an integral of SRB distributions of  $\varphi_{2D}$  restricted to single unstable manifolds over all values of  $x^{(2)}$ . In other words,  $\varphi_{2D}$  scatters the SRB measure of  $\varphi$  (which is supported on  $[0, 1]$ ) over an infinite set of vertically stacked intervals  $[0, 1]$  (which geometrically coincide with unstable manifolds of  $\varphi_{2D}$ ). This further implies the SRB density of  $\varphi$  equals the SRB distribution of  $\varphi_{2D}$  integrated with respect to the vertical (second) coordinate.

In case of a map defined by Eq. A.1, the simplified recursive formula for  $g$  can be

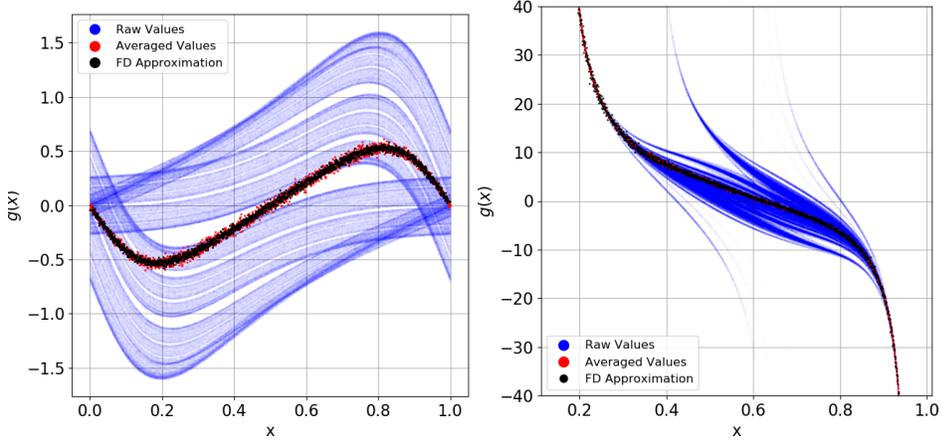


FIG. 7. SRB density gradient generated for the sawtooth map  $x_{k+1} = 2x_k + s \sin(2\pi x_k) \bmod 1$  at  $s = 0.1$  (left) and the onion map  $x_{k+1} = 0.97\sqrt{1 - |1 - 2x_k|^\gamma}$  at  $\gamma = 0.4$  (right). The averaged values (red dots) were computed by averaging the raw values (blue dots) in each of 2048 bins. The FD Approximation data points represent the central finite difference approximation of the SRB density gradient using the definition of  $g$  and empirically computed SRB densities. We generated a trajectory of length  $N = 10^6$  to compute the raw/averaged values of  $g$ .

expressed in terms of phase space derivatives of  $\varphi$  (see Section 3.2 for the derivation),

$$(A.2) \quad g(\varphi(x)) = \frac{g(x)}{\varphi'(x)} - \frac{\varphi''(x)}{\varphi'(x)^2}.$$

Here, the prime symbol ( $'$ ) denotes differentiation with respect to phase space. Let  $\rho(x)$  be the SRB density of  $\varphi$ . The  $g(x)$  function that satisfies Eq. A.2 is not the SRB density gradient of  $\varphi$ , defined as  $g^\varphi(x) := \rho'(x)/\rho(x)$ . According to our discussion in Section 3.2,  $g(x)$  is in fact a conditional SRB density gradient of  $\varphi_{2D}$  associated with the unstable manifold parameterized by  $x^{(2)}$ . However, as we discussed in the previous paragraph, the SRB measure of  $\varphi$  can be computed by integrating “slices” of the SRB measure of  $\varphi_{2D}$  parallel to  $x^{(1)}$ . This implies that, given the definition of the SRB density gradient,  $g^\varphi(x)$  can be computed by Lebesgue-integrating the SRB density gradients obtained in the above iteration along the vertical axis.

In practice, to construct a trajectory-based algorithm for  $g^\varphi$ , we can directly use the recursion in Eq. A.2. The algorithm we propose is the following. Divide the phase space  $[0, 1]$  in  $K \in \mathbb{Z}^+$  bins of equal width. Generate a sufficiently long sequence  $\{g_0, g_1, g_2, \dots\}$  using Eq. A.2 starting from a random initial condition  $g_0$ . For each bin, take the average of the members of the sequence that correspond to one bin. Based on our discussion above, the obtained average value converges to  $g^\varphi$ . This algorithm in fact provides a piecewise constant approximation of  $g^\varphi$ .

To verify our argument, we present a numerical experiment in which we apply the algorithm to two different 1D maps, the sawtooth map and onion map. Both of them are two-to-one and piecewise smooth. Figure 7 shows raw values of the sequence  $\{g_0, g_1, g_2, \dots\}$  obtained using Eq. A.2, their averaged values, and finite-difference (FD) approximation of  $g^\varphi$  using empirical SRB densities of these maps. We observe there is a good agreement between the averaged values and FD approximations in both cases.

Finally, we perform the relative error convergence test of the averaged values

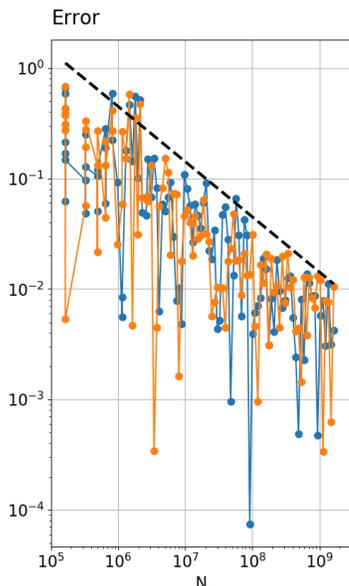


FIG. 8. Relative error of the approximation of  $g^\varphi(x)$  versus the trajectory length  $N$ . The error was computed for the sawtooth map at  $s = 0.1$  at two phase space coordinates,  $x \approx 0.4$  (blue curve) and  $x \approx 0.6$  (orange curve). All error values were computed with respect to the reference value generated using  $N = 3.3 \cdot 10^{11}$  samples. The reference dashed line represents the slope  $-1/2$  in the logarithmic scaling.

with respect to the trajectory length  $N$ . We focus on two different bins and compute the relative error with respect to a reference value generated using significantly more samples. Our results generated for the sawtooth map are shown in Figure 8. As expected, the error decays and is upperbounded by  $\mathcal{O}(1/\sqrt{N})$ , which is a consequence of the Lebesgue-integration (or, equivalently, weighted averaging) of (conditional) SRB density gradients. This example shows that a trajectory of minimum length  $N = 10^9$  should be generated in order to obtain an approximation with a relative error smaller than 1%.

**Appendix B. Probing the hyperbolicity of the Baker’s map.** Hyperbolicity guarantees the tangent space can be decomposed into two  $D\varphi$ -invariant subspaces, where one is asymptotically expanding (unstable), while the other one is asymptotically contracting (stable). If the expansion/contraction is uniform, then such systems are uniformly hyperbolic. Hyperbolicity is the major assumption for the dynamical systems we consider in this paper. Indeed, if the system is hyperbolic and has absolutely continuous conditional measures on unstable manifolds, then the SRB measure exists [11]. It is not always possible to analytically verify that a particular map is hyperbolic. Fortunately, there exist numerical procedures allowing for an efficient assessment of hyperbolicity [18]. Most of them test the two basic criteria of hyperbolicity: 1) No zero LEs, and 2) Strict separation of the stable and unstable subspaces. Here, we apply the method proposed in [19], which computes the basis vectors of the two subspaces and approximates the smallest angle between them at different points of the manifold. If any of these angles is close to zero, then the stable and unstable subspaces are (almost) tangent, which implies the systems is likely to be non-hyperbolic. In Figure 9, we compute the PDF of  $d \in [0, 1]$ , which is a normalized

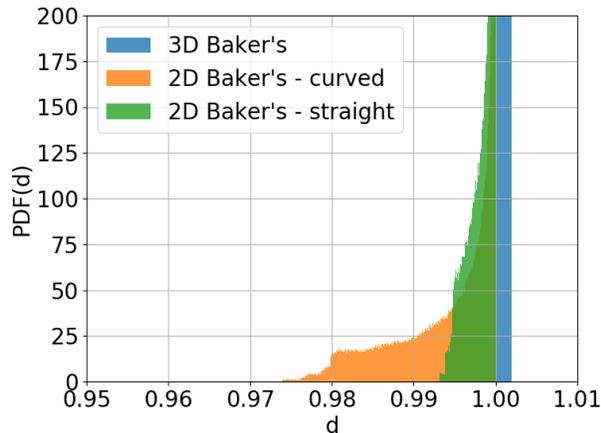


FIG. 9. *Hyperbolicity test performed for the 2D and 3D Bakers map. The parameter values are the same as the ones used in the numerical examples in Section 3 (2D Baker's) and Section 4 (3D Baker's). To generate the PDF, we computed  $N = 10^6$  samples of  $d$  along a trajectory.*

quantity associated with the smallest principal angle between the stable and unstable subspace (our  $d$  equals  $k!d_k$ ; see the above reference for a rigorous definition of  $d_k$ ). If the distribution is evidently separated from the origin ( $d = 0$ ), then it is highly likely there are no tangencies between the two subspaces. We observe the normalized parameter  $d$  is highly unlikely to drop below the value of 0.97. As a by-product of the applied algorithm, we computed the spectrum of Lyapunov exponents (alternatively, one can use Benettin et al.'s algorithm [5]). The LEs approximately equal:  $0.69 \approx \log(2)$ ,  $-0.69$  (2D Baker's with straight unstable subspaces),  $0.69 \approx \log(2)$ ,  $-0.71$  (2D Baker's with curved unstable subspaces),  $1.09 \approx \log(3)$ ,  $0.69 \approx \log(2)$ ,  $-1.16$  (3D Baker's). Although a small change in the parameter value does not significantly impact the LE values, it may move the PDF of  $d$  closer to the origin. Based on the empirical evidence presented in this section, we conclude the 2D/3D Baker's map is clearly hyperbolic at the chosen parameter values.

#### REFERENCES

- [1] R. V. ABRAMOV AND A. J. MAJDA, *Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems*, *Nonlinearity*, 20 (2007), <https://doi.org/10.1088/0951-7715/20/12/004>.
- [2] R. V. ABRAMOV AND A. J. MAJDA, *New Approximations and Tests of Linear Fluctuation-Response for Chaotic Nonlinear Forced-Dissipative Dynamical Systems*, *Journal of Nonlinear Science*, 18 (2008), pp. 303–341, <https://doi.org/10.1007/s00332-007-9011-9>.
- [3] J. ALVES, C. BONATTI, AND M. VIANA, *SRB measures for partially hyperbolic systems whose central direction is mostly expanding*, *Inventiones Mathematicae*, 140 (2000), pp. 351–398, <https://doi.org/10.1007/s002220000057>.
- [4] V. BALADI, *Linear response, or else*, arXiv e-prints, arXiv:1408.2937 (2014).
- [5] G. BENETTIN, L. GALGANI, A. GIORGILLI, AND J.-M. STRELCYN, *Lyapunov Characteristic Exponents for smooth dynamical systems and for hamiltonian systems; A method for computing all of them. Part 2: Numerical application*, *Meccanica*, 15 (1980), pp. 21–30, <https://doi.org/10.1007/BF02128237>.
- [6] P. BLONIGAN, *Least Squares Shadowing for Sensitivity Analysis of Large Chaotic Systems and Fluid Flows*, PhD thesis, Massachusetts Institute of Technology, 2016.
- [7] K. BURNS, D. DOLGOPYAT, Y. PESIN, AND M. POLLICOTT, *Stable ergodicity for partially hyperbolic attractors with negative central exponents*, *Journal of Modern Dynamics*, 2 (2008), pp. 63–81, <https://doi.org/10.3934/jmd.2008.2.63>.

- [8] N. CHANDRAMOORTHY AND Q. WANG, *A computable realization of Ruelle's formula for linear response of statistics in chaotic systems*, arXiv e-prints, arXiv:2002.04117 (2020).
- [9] N. CHANDRAMOORTHY AND Q. WANG, *An ergodic-averaging method to differentiate covariant Lyapunov vectors*, *Nonlinear Dynamics*, (2021), <https://doi.org/10.1007/s11071-021-06478-0>.
- [10] N. CHANDRAMOORTHY AND Q. WANG, *Efficient computation of linear response of chaotic attractors with one-dimensional unstable manifolds*, arXiv e-prints, arXiv:2103.08816 (2021).
- [11] V. CLIMENHAGA, D. DOLGOPYAT, AND Y. PESIN, *Non-stationary non-uniform hyperbolicity: SRB measures for dissipative maps*, *Communications in Mathematical Physics*, 346 (2016), pp. 553–602, <https://doi.org/10.1007/s00220-016-2710-z>.
- [12] A. CRUZ AND P. VARANDAS, *SRB measures for partially hyperbolic attractors of local diffeomorphisms*, *Ergodic Theory and Dynamical Systems*, 40 (2020), pp. 1545–1593, <https://doi.org/10.1017/etds.2018.115>.
- [13] S. FIORI, *Neural systems with numerically-matched input-output statistic: variate generation*, *Natural Processing Letters*, 23 (2006), pp. 143–170, <https://doi.org/10.1007/s11063-005-4016-6>.
- [14] G. GALAVOTTI AND E. G. D. COHEN, *Dynamical ensembles in stationary states*, *Journal of Statistical Physics*, 80 (1995), pp. 931–970, <https://doi.org/10.1007/BF02179860>.
- [15] A. GRITSUN AND G. BRANSTATOR, *Climate response using a three-dimensional operator based on the fluctuation-dissipation theorem*, *Journal of Atmospheric Sciences*, 64 (2016), pp. 2558–2575, <https://doi.org/10.1175/JAS3943.1>.
- [16] A. JAMESON, *Aerodynamic design via control theory*, *Journal of Scientific Computing*, 3 (1988), pp. 233–260, <https://doi.org/10.1007/BF01061285>.
- [17] R. KUBO, *The fluctuation-dissipation theorem*, *Reports on Progress in Physics*, 29 (1966), <https://doi.org/10.1088/0034-4885/29/1/306>.
- [18] P. KUPTSOV AND U. PARLITZ, *Theory and Computation of Covariant Lyapunov Vectors*, *Non-linear Science*, 22 (2012), pp. 727–762, <https://doi.org/10.1007/s00332-012-9126-5>.
- [19] P. V. KUPTSOV, *Fast numerical test of hyperbolic chaos*, *Physical Review E*, 85 (2012), p. 015203, <https://doi.org/10.1103/PhysRevE.85.015203>.
- [20] J. LARSSON, *Grid-adaptation for chaotic multi-scale simulations as a verification-driven inverse problem*, *AIAA*, 2018, pp. 1–17, <https://doi.org/10.2514/6.2018-0371>.
- [21] Y. MAO, G. CHEN, AND S. LIAN, *A novel fast image encryption scheme based on 3D chaotic Baker maps*, *International Journal of Bifurcation and Chaos*, 14 (2004), pp. 3613–3624, <https://doi.org/10.1142/S021812740401151X>.
- [22] M. MEHTA, A. K. MITTAL, AND S. DWIVEDI, *The double-cusp map for the forced Lorenz system*, *International Journal of Bifurcation and Chaos*, 13 (2003), pp. 3029–3035, <https://doi.org/10.1142/S0218127403008429>.
- [23] G. NAGY, *Radon-Nikodym theorems*, Kansas State University, 2021, pp. 300–321, <https://www.math.ksu.edu/~nagy/real-an/4-04-rn.pdf>.
- [24] A. NI, *Fast linear response algorithm for differentiating stationary measures of chaos*, arXiv e-prints, arXiv:2009.00595 (2021).
- [25] F. PICHLER AND J. SCHARINGER, *Finite dimensional generalized baker dynamical systems for cryptographic applications*, in *Computer Aided Systems Theory — EUROCAST '95*, F. Pichler, R. M. Díaz, and R. Albrecht, eds., Berlin, Heidelberg, 1996, Springer Berlin Heidelberg, pp. 465–476, <https://doi.org/10.1007/BFb0034782>.
- [26] D. RUELE, *A measure associated with Axiom-A attractors.*, *American Journal of Mathematics*, 98 (1976), <https://doi.org/10.2307/2373810>.
- [27] D. RUELE, *Differentiation of SRB states*, *Communications in Mathematical Physics*, 187 (1997), pp. 227–241, <https://doi.org/10.1007/s002200050134>.
- [28] D. RUELE, *Differentiation of SRB states: correction and complements*, *Communications in Mathematical Physics*, 234 (2003), pp. 185–190, <https://doi.org/10.1007/s00220-002-0779-z>.
- [29] A. A. ŚLIWIAK, N. CHANDRAMOORTHY, AND Q. WANG, *Ergodic sensitivity analysis of one-dimensional chaotic maps*, *Theoretical and Applied Mechanics Letters*, 10 (2020), <https://doi.org/10.1016/j.taml.2020.01.058>.
- [30] A. A. ŚLIWIAK, N. CHANDRAMOORTHY, AND Q. WANG, *Computational assessment of smooth and rough parameter dependence of statistics in chaotic dynamical systems*, *Communications in Nonlinear Science and Numerical Simulation*, 101 (2021), <https://doi.org/10.1016/j.cnsns.2021.105906>.
- [31] A. A. ŚLIWIAK AND Q. WANG, *Differentiating densities on smooth manifolds*, *Applied Mathematics and Computation*, 410 (2021), <https://doi.org/10.1016/j.amc.2021.126444>.
- [32] Q. WANG, *Uncertainty quantification for unsteady fluid flow using adjoint-based approaches*,

- PhD thesis, Stanford University, 2009.
- [33] C. L. WORMELL AND G. A. GOTTWALD, *Linear response for macroscopic observables in high-dimensional systems*, *Chaos*, 29 (2019), pp. 21–30, <https://doi.org/10.1063/1.5122740>.
- [34] L.-S. YOUNG, *What Are SRB Measures, and Which Dynamical Systems Have Them?*, *Journal of Statistical Physics*, 108 (2002), pp. 733–754, <https://doi.org/10.1023/A:1019762724717>.