SIAM J. Applied Dynamical Systems
Vol. 23, No. 1, pp. 383–409

Identifying Nonlinear Dynamics with High Confidence from Sparse Data*

Bogdan Batko[†‡], Marcio Gameiro[‡§], Ying Hung[¶], William Kalies[‖], Konstantin Mischaikow[‡], and Ewerton Vieira[#]

Abstract. We introduce a novel procedure that, given sparse data generated from a stationary deterministic nonlinear dynamical system, can characterize specific local and/or global dynamic behavior with rigorous probability guarantees. More precisely, the sparse data is used to construct a statistical surrogate model based on a Gaussian process (GP). The dynamics of the surrogate model is interrogated using combinatorial methods and characterized using algebraic topological invariants (Conley index). The GP predictive distribution provides a lower bound on the confidence that these topological invariants, and hence the characterized dynamics, apply to the unknown dynamical system (assumed to be a sample path of the GP). The focus of this paper is on explaining the ideas, thus we restrict our examples to one-dimensional systems and show how to capture the existence of fixed points, periodic orbits, connecting orbits, bistability, and chaotic dynamics.

Key words. sparse data, Gaussian process, nonlinear dynamics, uncertainty quantification

MSC codes. 37M10, 60G15, 37B35, 37C25

DOI. 10.1137/23M1560252

1. Introduction.

We propose a novel framework, combining topological dynamics and statistical surrogate modeling with uncertainty quantification, through which it is possible to

*Received by the editors March 21, 2023; accepted for publication (in revised form) by T. Wanner September 7, 2023; published electronically January 24, 2024.
https://doi.org/10.1137/23M1560252
Funding: The work of the first author was partially supported by DARPA contract HR0011-16-2-0033 and by the Polish National Science Center under Opus grant 2019/35/B/ST1/00874. The work of the second, third, fifth, and sixth authors was partially supported by the National Science Foundation under award DMS-1839294 and HDR TRIPODS award CCF-1934924, by DARPA contract HR0011-16-2-0033, National Institutes of Health award R01 GM126555, and Air Force Office of Scientific Research award FA9550-23-1-0011. The work of the second author was also partially supported by FAPESP grant 2019/06249-7 and CNPq grant 309073/2019-7. The work of the third author was also partially supported by the National Science Foundation under award DMS-2113475. The work of the fourth author was partially supported by Army Research Office award W911NF1810306 and Air Force Office of Scientific Research award FA9550-23-1-0011. The work of the fifth author was also supported by a grant from the Simons Foundation.
[†]Division of Computational Mathematics, Faculty of Mathematics and Computer Science, Jagiellonian University, ul. St. Lojasiewicza 6, 30-348 Kraków, Poland (bogdan.batko@ii.uj.edu.pl).
[‡]Department of Mathematics, Rutgers University, Piscataway, NJ 08854 USA (gameiro@math.rutgers.edu, mischaik@math.rutgers.edu).
[§]Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, São Paulo, Brazil.
[¶]Department of Statistics, Rutgers, The State University of New Jersey, Piscataway, NJ 08854 USA (yhung@stat.rutgers.edu).
[‖]Department of Mathematics and Statistics, University of Toledo, Toledo, OH 43606 USA (william.kalies@utoledo.edu).
[#]DIMACS, Rutgers, The State University of New Jersey, Piscataway, NJ 08854 USA (ewerton.vieira@dimacs.rutgers.edu).

characterize local and global dynamics from data with probability guarantees. Given a data set $\mathcal{T} = \{(x_n, y_n) \in \mathbb{R}^d \times \mathbb{R}^d \mid n = 1, \ldots, N\}$, where it is assumed that the data is generated by a continuous dynamical system on a compact set $X \subset \mathbb{R}^d$, we give rigorous bounds on the probability that the characterization of the dynamics is correct.

To put our results into context, we recall that given a continuous function $f : X \to X$ the traditional focus of dynamical systems has been on understanding the structure of *invariant sets,* i.e., subsets $S \subset X$ such that $f(S) = S$, for which fixed points and periodic orbits are simple examples. On a global level, this is equivalent to understanding the conjugacy classes of f, i.e., the set of $g : Y \to Y$ such that $h \circ f = g \circ h$ for some homeomorphism $h : X \to Y$. This is impossible in general [9]. Even in more restrictive settings, correctly capturing the invariant sets may require correctly identifying the nonlinearity to an extremely high order of precision, which is often impossible from a given finite data set. The logistic map and the associated cascade of period doublings is an archetypal example.

For this reason our approach focuses on coarsely characterizing dynamics rather than identifying the underlying nonlinearity. Our characterization is done via the Conley index, an algebraic topological invariant, from which one can induce the existence of invariant sets and dynamic structure of invariant sets, e.g., existence of fixed points, periodic orbits, heteroclinic orbits, and chaotic dynamics [21].

In essence, the strategy that we propose is straightforward. It involves a fundamental assumption and the three steps encapsulated in Figure 1.

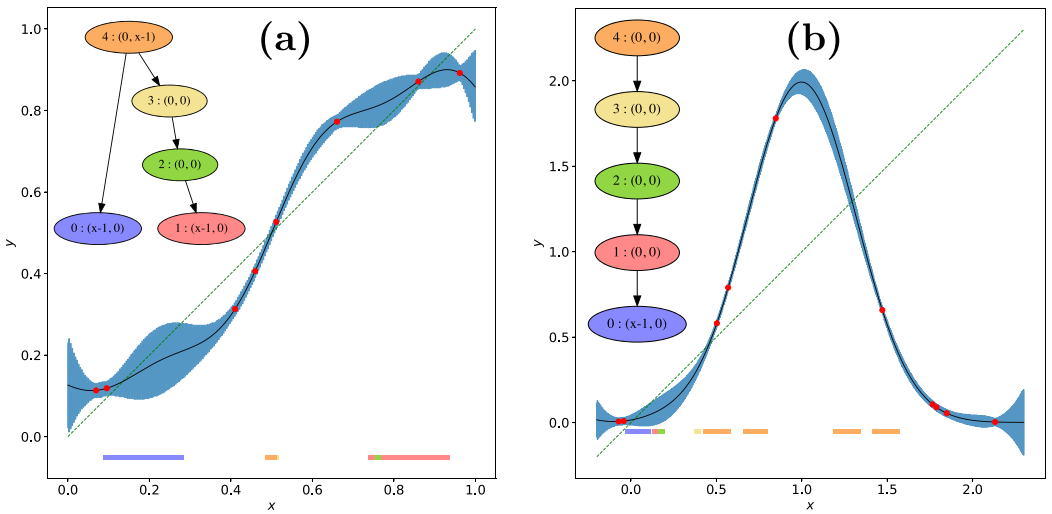

Figure 1. *In both subfigures, elements of $\mathcal{T}$ are indicated in red, the mean function μ is shown in black, and G is shown in blue. The region G is composed of squares of width $2^{-10}\,\mathrm{length}(X)$. The Morse graphs are indicated at the top left, and the corresponding (color coded) regions of phase space $\bar{\nu}(\cdot)$ are indicated at the bottom of each subfigure.* (a) *The Morse graph and the Conley indices of the invariant sets in $\bar{\nu}(\mathcal{M}(0))$ and $\bar{\nu}(\mathcal{M}(1))$—the blue and red regions on the x-axis correspond to the Morse nodes* 0 *(blue) and* 1 *(red)—indicate that bistability is exhibited with* 95% *confidence.* (b) *The invariant set in $\bar{\nu}(\mathcal{M}(4))$—the orange region on the x-axis corresponds to Morse node* 4*—exhibits chaotic dynamics with positive topological entropy with* 95% *confidence.*

A. Assume the observed data is $\mathcal{T} = \{(x_n, y_n) \mid y_n = f(x_n) + \epsilon\}$, where f is an unknown continuous function. Assume also that there is a Gaussian process (GP) with a prespecified semipositive kernel $k(\cdot, \cdot; \theta)$, where θ is a vector of unknown parameters associated with the kernel and ϵ arises from random Gaussian noise, such that f is a realization of this Gaussian process for some value of θ.

Step 1. Given the data set $\mathcal{T}$, estimate the unknown parameters and construct a GP surrogate model (see section 2).

Step 2. Choose a finite cell complex $\mathcal{X}$ [20] whose geometric realization as a regular CW-complex [15] is X. Construct a closed set $G \subset X \times X$ with the following property: G is the geometric realization of products of cells from $\mathcal{X}$, and each fiber $G_x := G \cap (\{x\} \times X)$, $x \in X$, is nonempty and contractible. Use the combinatorial representation of G to identify potential dynamics and compute their associated Conley indexes (see section 3).

The set G above represents a (coarse) combinatorial representation of the dynamics as follows: given a pair of cells $\xi \times \xi' \in \mathcal{X} \times \mathcal{X}$ whose geometric realization is contained in G, we say that ξ maps to ξ' under the combinatorial dynamics (see sections 3 and 4).

Given a GP g, we denote its graph by $\mathsf{G}(g) := \{(x, g(x)) \mid x \in X\}$. The Gaussian predictive distribution determines $\mathbb{P}(\mathsf{G}(g) \subset G)$. It is worth emphasizing that g is the Gaussian process (a random variable) and is not a realization of the Gaussian process. A *sample path* (or a *realization*) of the GP g is a function $h : X \to \mathbb{R}^d$ that is obtained as a realization of g as a random variable (a value of the random variable g). Our goal is to compute dynamics which is valid for all sample paths whose graphs are contained in G, that is, for all functions in the set

$$\mathcal{H} := \{h : X \to \mathbb{R}^d \mid h \text{ is a sample path of } g \text{ and } \mathsf{G}(h) \subset G\}.$$

We denote $\mathbb{P}(\mathcal{H})$ by $\mathbb{P}(\mathsf{G}(g) \subset G)$. Since the dynamics is computed using the combinatorial representation of G, and the Conley index only depends on G, the dynamics computed is valid for all functions in $\mathcal{H}$, and $\mathbb{P}(\mathsf{G}(g) \subset G)$ provides a lower bound on the probability that the dynamics identified in Step 2 occurs for the GP g. From the discussion in section 3, in general, fibers G_x with smaller diameters lead to a greater potential for identifying dynamics.

For many applications, the focus is on particular dynamics and/or specific lower bounds on the confidence of the occurrence of the dynamics. Thus, we introduce a third step.

Step 3. Modify $G \subset X \times X$ to both preserve the dynamics of interest and maximize $\mathbb{P}(\mathsf{G}(g) \subset G)$.

In this paper we construct G as described in section 4. In this case, the probability $\mathbb{P}(\mathsf{G}(g) \subset G)$ provides a *confidence level* that the computed dynamics is valid.

Figure 1 is meant to provide geometric intuition of Steps 1–3. In particular, in Figure 1(a), for any sample path h whose graph lies in the blue region, we can conclude that the global dynamics generated by h exhibits bistability as well as the existence of at least three fixed points. For Figure 1(b) we can conclude the existence of chaotic dynamics. In both cases, because of the application of Step 3 we can conclude that the abovementioned dynamics occurs with a confidence of at least 95%.

There are three natural questions concerning convergence that arise from the success claimed in Figure 1. Recall that f is the unknown continuous function that is assumed to be

a realization of the GP and to have generated the data. The first question is, What dynamics of a given function can be identified via the approximation methods (briefly described in section 3) of Step 2? A precise answer (see [16, Theorem 1.3]) goes beyond the scope of this paper. An imprecise answer is that for many applications, most invariant sets of practical interest are identifiable. The second and third questions are intertwined and address the level of confidence to which our claims on the dynamics can be accepted. That is, how large can we make $\mathbb{P}(\mathsf{G}(g) \subset G)$ in Step 3, and to what level of confidence can we approximate f from data? Theorem 4.1 of this paper indicates that the confidence level for both questions can be made arbitrarily large simultaneously, assuming that $\mathcal{T}$ contains sufficiently many data points and that the diameter of the elements of geometric realization of $\mathcal{X}$ are sufficiently small.

As indicated above, the novelty of our approach arises from the combination of GP surrogate modeling and Conley theory. While each of these topics is well developed, there does not seem to be much overlap of the associated research communities. With this in mind we provide a minimal description of surrogate modeling by GP (section 2) and combinatorial Conley theory (section 3) in the context of maps on $\mathbb{R}^d$. However, the examples are given using maps on $\mathbb{R}$, as this allows us to demonstrate the results using simple figures.

2. Surrogate modeling by Gaussian processes. A GP model (Step 1 in our process), also called kriging in geostatistics, is a widely used surrogate model because of its flexibility, nonlinearity, and capability to perform uncertainty quantification through the predictive distribution [22, 13].

Recall that our data $\mathcal{T}$ is generated by the unknown realization f of the GP in assumption A with random Gaussian noise so that $y_n = f(x_n) + \epsilon$. Let f_ℓ denote the ℓth component of f for $\ell = 1, \dots, d$. Then

$$f_\ell(x) \sim GP(\beta_\ell, \sigma_\ell^2 k(x, x'; \theta_\ell)), \tag{2.1}$$

where β_ℓ and σ_ℓ^2 are the unknown mean and variance, and the correlation is defined by the kernel $k(x, x'; \theta_\ell) = Corr(f_\ell(x), f_\ell(x'); \theta_\ell)$ with $k(x, x'; \theta_\ell) = k(x', x; \theta_\ell)$ for $x, x' \in X$. For simplicity of exposition, we assume the data is noise-free and therefore that $k(x, x; \theta_\ell) = 1$. The results in this paper can be easily extended to noisy data by incorporating nugget effects in the kernel function [13]. There are extensive discussions on correlation functions in the literature [23]. The mean function β can be further extended to include regression terms in the mean function, which is known as universal kriging [23, 5].

Based on (2.1), the maximum likelihood estimators (MLEs) for β_ℓ, σ_ℓ^2, and θ_ℓ can be obtained by

$$\hat{\beta}_\ell = \frac{\mathbf{1}^T K^{-1}(\hat{\theta}_\ell) \mathbf{y}_\ell^T}{\mathbf{1}^T K^{-1}(\hat{\theta}_\ell) \mathbf{1}}, \quad \hat{\sigma}_\ell^2 = \frac{1}{N} \left(\mathbf{y}_\ell^T - \hat{\beta}_\ell\right)^T K^{-1}(\hat{\theta}_\ell) \left(\mathbf{y}_\ell^T - \hat{\beta}_\ell\right),$$

and

$$\hat{\theta}_\ell = \arg\min_\theta \{N \log(\hat{\sigma}_\ell^2) + \log |K(\theta_\ell)|\},$$

where $\mathbf{y}_\ell := (f_\ell(x_1), \dots, f_\ell(x_N))$, $\mathbf{1}$ is a column of 1's with length N, $K(\theta)$ is an $N \times N$ correlation matrix with elements $k(x_i, x_j; \theta)$ for $1 \le i, j \le N$, and $|K(\theta)|$ is the determinant of $K(\theta)$.

Other estimation approaches, such as the restricted maximum likelihood (REML) method and estimations by cross validation, are also applicable [5, 22]. Alternatively, assumption A can be regarded as a Bayesian prior on the unknown function f, and a fully Bayesian approach can be applied to perform estimation and prediction [22, 13]. In this paper, the parameters are estimated by the MLEs.

In Step 1, the prediction for an untried $x \in X$ can be obtained by a d-dimensional multivariate normal distribution, $MN(\mu(x), \Sigma(x))$, where $\mu(x) = (\mu_1(x), \ldots, \mu_d(x))$ is the best linear unbiased predictor (BLUP) with

$$\mu_\ell(x) = E(f_\ell(x)|\mathcal{T}) = \hat{\beta}_\ell + k(x;\hat{\theta}_\ell)^T K^{-1}(\hat{\theta}_\ell)\left(\mathbf{y}_\ell^T - \hat{\beta}_\ell\right),$$

and the covariance matrix $\Sigma(x)$ has diagonal elements

$$Var(f_\ell(x)|\mathcal{T}) = \hat{\sigma}_\ell^2\left(1 - k(x;\hat{\theta}_\ell)^T K^{-1}(\hat{\theta}_\ell)k(x;\hat{\theta}_\ell)\right),$$

where $k(x;\hat{\theta})$ is the correlation between the new observation and the existing data, i.e., $k(x;\hat{\theta}) = (k(x,x_1;\hat{\theta}), \ldots, k(x,x_N;\hat{\theta}))$, and $K(\hat{\theta})$ is an $N \times N$ correlation matrix with elements $k(x_i,x_j;\hat{\theta})$ for $1 \le i,j \le N$. The off-diagonal elements in $\Sigma(x)$ are zeros if the d-dimensional outputs are assumed to be independent. By further assuming some correlation structures among outputs through the kernel function k, the off-diagonal elements can be estimated by techniques such as cokriging [10]. Note that, using different kernel functions, the correlation structure between the d-dimensional outputs can be captured [10, 18], and nugget effects [19, 13] can be included in the kernel function to estimate the sampling error or extrinsic noise associated with the observations.

Recall that for any random variable $Z \sim MN(m,\Lambda)$, the squared Mahalanobis distance

$$\rho_\Lambda^2(Z,m) := (Z-m)^T\Lambda^{-1}(Z-m)$$

has χ^2-distribution with d degrees of freedom [22].

Note that if g is a GP on a parameter space X, then the above applies to $g(x)$ at each point $x \in X$. Let μ and Σ denote the predictive mean and covariance functions of g, respectively. Accordingly, for any $\delta \in (0,1)$ and any fixed $x \in X$, we have

$$\mathbb{P}\left(g(x) \in E_{\Sigma(x)}\left(\mu(x), \chi_d^2(1-\delta)\right)\right) = 1-\delta,$$

where $E_{\Sigma(x)}(\mu(x),c) := \{y \in X \mid \rho_{\Sigma(x)}^2(y,\mu(x)) < c\}$ is the *confidence ellipsoid,* $\chi_d^2(1-\delta)$ stands for a χ^2 quantile of order $1-\delta$ with d degrees of freedom, and $1-\delta$ is the *confidence level.* More generally, if $S \subset X$ is finite and $\delta \in (0,1)$, then there exists a function $r : S \to (0,\infty)$ such that

$$\mathbb{P}\left(g(v) \in E_{\Sigma(v)}(\mu(v), r(v))\ \forall v \in S\right) \ge 1-\delta. \tag{2.2}$$

Observe that the function r is not unique.

To illustrate the aforementioned procedure we present a simple one-dimensional ($d = 1$) example. Figure 2 shows a GP model constructed from five noise-free observations indicated

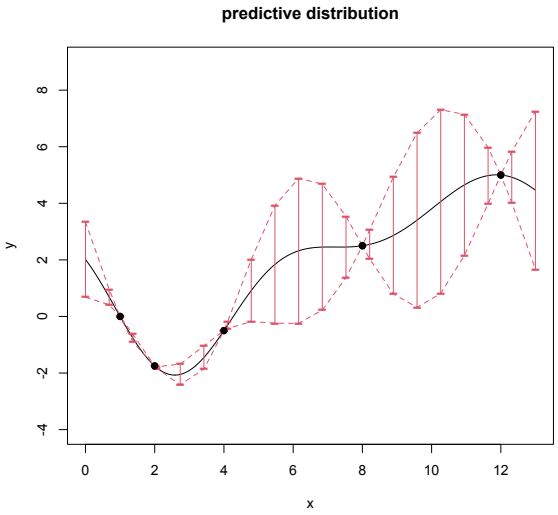

Figure 2. *Based on five observations, illustrated by solid dots, a one-dimensional GP model is fitted. The black curve is the best linear unbiased predictor, and the red bars are the pointwise* 95% *confidence intervals calculated at* 20 *untried points.*

by the solid dots. By utilizing a squared exponential kernel we obtain the BLUP $\mu(x)$, depicted as the black curve, which interpolates the observed data. Additionally, the red bars represent the pointwise 95% confidence intervals ($\delta = 0.05$) evaluated at 20 untried points. These confidence intervals quantify the pointwise prediction uncertainty, which decreases to zero when predicting the observed inputs.

In light of Step 3, to provide a lower bound on the confidence of our characterization of dynamics, it is reasonable to make use of the pointwise bounds of (2.2) and insist that G satisfies the property that $\{y \in X \mid \rho^2_{\Sigma(x)}(y, \mu(v)) \leq r(v)\} \subset G_v$ for each $v \in S$. To obtain appropriate conditions on G_x for $x \in X \setminus S$, in this paper we restrict our attention to kernel functions that are differentiable up to order four, e.g., the squared exponential covariance function or the Matérn kernels with $\nu > 2$ [23], in which case there exist $L_0 > 0$ and constants $a, b > 0$ such that for any $L > L_0$ we have

$$\mathbb{P}\left(\forall x_1, x_2 \in X \ \|g(x_1) - g(x_2)\| \leq L\|x_1 - x_2\|\right) > \left(1 - ae^{-(\frac{L}{b})^2}\right)^{d^2} \tag{2.3}$$

(cf. [12, Theorem 5]).

3. Combinatorial Conley theory and the characterization of dynamics. There are three essential components of combinatorial Conley theory: a finite combinatorial representation of phase space via a cell complex, a combinatorial representation of dynamics via a directed graph, and homological computations. As described at the end of this section, the combinatorial theory is used to characterize the dynamics generated by continuous functions that are sample paths of the GP. To instantiate these ideas throughout this section we describe a particularly simple example that leads to bistability.

Recall [20] that a *cell complex* $\mathcal{X} = (\mathcal{X}, \leq, \dim, \kappa)$ is a finite partially ordered set (poset) $(\mathcal{X}, \leq)$, where the partial order $\leq$ indicates the face relation, together with two associated functions, *dimension,* $\dim : \mathcal{X} \to \mathbb{N}$, and *incidence,* $\kappa : \mathcal{X} \times \mathcal{X} \to \mathbb{F}$, where $\mathbb{F}$ is a principal ideal

domain, subject to the following conditions for all $\xi, \xi', \xi'' \in \mathcal{X}$: (i) $\xi \leq \xi'$ implies $\dim \xi \leq \dim \xi'$, (ii) $\kappa(\xi', \xi) \neq 0$ implies $\xi' \leq \xi$ and $\dim(\xi) = \dim(\xi') + 1$, and (iii) $\sum_{\xi' \in \mathcal{X}} \kappa(\xi, \xi')\kappa(\xi', \xi'') = 0$. A cell complex generates a *chain complex* that we denote by $C_*(\mathcal{X}; \mathbb{F})$. An element $\xi \in \mathcal{X}$ is called a cell, and we denote the maximal elements of $\mathcal{X}$ by $\mathcal{X}^{\text{top}}$. Given $\xi \in \mathcal{X}^{\text{top}}$, define $\downarrow(\xi) = \{\xi' \in \mathcal{X} \mid \xi' \leq \xi\}$. Note that given any subset $\mathcal{A} \subset \mathcal{X}^{\text{top}}$, $\downarrow(\mathcal{A})$ generates a chain complex $C_*(\mathcal{A}; \mathbb{F})$.

Example 3.1. Consider $\mathcal{X} = \{v_i \mid i = -2, -1, 0, 1, 2\} \cup \{e_i \mid i = -2, -1, 0, 1\}$, where we define $\dim(v_i) = 0$ and $\dim(e_i) = 1$, i.e., v_i is a vertex and e_i is an edge (see Figure 3). We set $v_i \leq e_j$ if $i = j$ or $i = j + 1$. We assume that $\mathbb{F}$ is the field $\mathbb{Z}_2$ and set $\kappa(v_i, e_j) = 1$ if and only if $v_i \leq e_j$. This gives rise to the chain complexes $C_0(\mathcal{X}; \mathbb{Z}_2) \cong \mathbb{Z}_2^5$ and $C_1(\mathcal{X}; \mathbb{Z}_2) \cong \mathbb{Z}_2^4$, with a boundary operator $\partial_1 : C_1(\mathcal{X}; \mathbb{Z}_2) \to C_0(\mathcal{X}; \mathbb{Z}_2)$ given by the matrix with entries $\kappa(v_i, e_j)$.

We represent dynamics using a *combinatorial multivalued map* $\mathcal{F} : \mathcal{X}^{\text{top}} \rightrightarrows \mathcal{X}^{\text{top}}$, i.e., for each $\xi \in \mathcal{X}^{\text{top}}$, $\mathcal{F}(\xi) \subset \mathcal{X}^{\text{top}}$. A combinatorial multivalued map is equivalent to a directed graph with vertices $\mathcal{X}^{\text{top}}$ and edges $\xi \to \xi'$ if $\xi' \in \mathcal{F}(\xi)$. To identify the potential recurrent and gradient-like structure of $\mathcal{F}$, we make use of the *condensation graph* of $\mathcal{F}$ obtained by identifying each strongly connected component of $\mathcal{F}$ to a single vertex [4]. As this is a directed acyclic graph, it can be viewed as a poset that we denote by $\mathsf{SC}(\mathcal{F})$. A *recurrent component* is a strongly connected component that contains at least one edge. The *Morse graph* of $\mathcal{F}$, denoted by $\mathsf{M}(\mathcal{F})$, is the subposet of recurrent components of $\mathsf{SC}(\mathcal{F})$. We typically display the Morse graph via the Haase diagram of $\mathsf{M}(\mathcal{F})$. Observe that the order relation on $\mathsf{M}(\mathcal{F})$ provides a combinatorial description of the structure of the gradient-like dynamics.

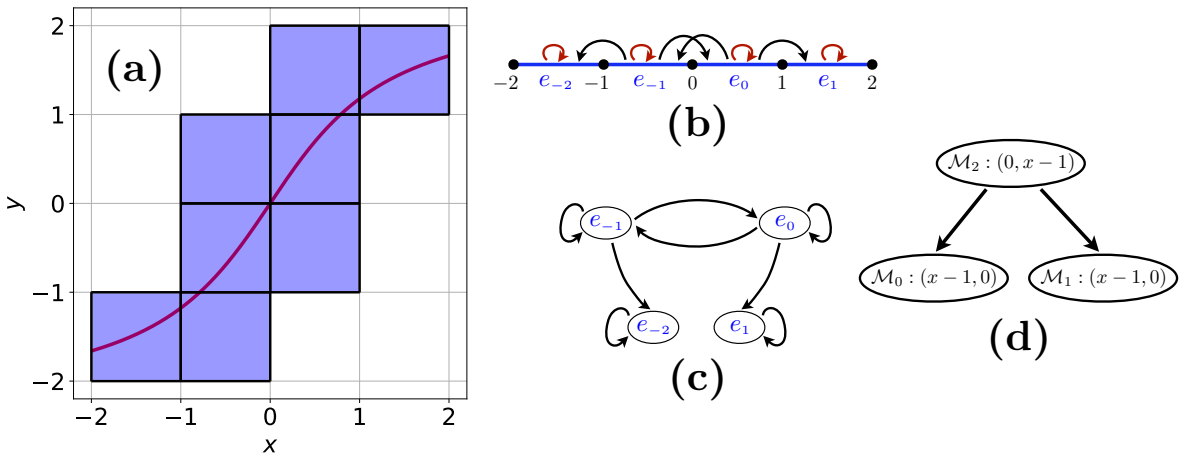

Figure 3. *Simple example exhibiting bistability.* (a) *The map* $f : [-2, 2] \to [-2, 2]$ *given by* $f(x) = \frac{3}{2}\arctan(x)$ *is plotted in red. The domain* $X = [-2, 2]$ *is decomposed into unit intervals* $[i, i+1]$, $i = -2, -1, 0, 1$, *and the graph of* f *is covered by products of these intervals.* (b) *The covering of the graph of* f *in* (a) *can be represented by a (multivalued) cell mapping* $\mathcal{F}$ *on the set of edges* $\{e_{-2}, e_{-1}, e_0, e_1\}$, *where the edge* e_i *represents the interval* $[i, i+1]$. (c) *Directed graph representation of the multivalued map* $\mathcal{F}$. *The nontrivial strongly connected components (recurrent components) of this graph give the nodes of the Morse graph* $\mathcal{M}_0 = \{e_{-2}\}$, $\mathcal{M}_1 = \{e_1\}$, *and* $\mathcal{M}_2 = \{e_{-1}, e_0\}$. (d) *Morse graph with the Conley index of each node.*

Example 3.2. Continuing with Example 3.1, $\mathcal{X}^{\text{top}} = \{e_i \mid i = -2, -1, 0, 1\}$. Set $\mathcal{F} : \mathcal{X}^{\text{top}} \rightrightarrows \mathcal{X}^{\text{top}}$ to be

$$\mathcal{F}(e_{-2}) = \{e_{-2}\}, \quad \mathcal{F}(e_{-1}) = \{e_{-2}, e_{-1}, e_0\}, \quad \mathcal{F}(e_0) = \{e_{-1}, e_0, e_1\}, \quad \text{and} \quad \mathcal{F}(e_1) = \{e_1\}.$$

The strongly connected components are $\mathcal{M}_0 = \{e_{-2}\}$, $\mathcal{M}_1 = \{e_1\}$, and $\mathcal{M}_2 = \{e_{-1}, e_0\}$. The condensation graph has edges $\mathcal{M}_2 \to \mathcal{M}_0$ and $\mathcal{M}_2 \to \mathcal{M}_1$. This is an acyclic directed graph and thus can be thought of as a poset with relations $\mathcal{M}_0 < \mathcal{M}_2$ and $\mathcal{M}_1 < \mathcal{M}_2$. Note that each strongly connected component has at least one edge, and therefore each strongly connected component is a recurrent component. Thus the Morse graph $\mathsf{M}(\mathcal{F})$ is the poset with elements $\{\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2\}$. The dynamics interpretation is that one can move from state $\mathcal{M}_2$ to state $\mathcal{M}_0$ or to state $\mathcal{M}_1$, but one cannot move from state $\mathcal{M}_0$ or state $\mathcal{M}_1$ to any other state.

An alternative perspective for characterizing the dynamics associated with $\mathcal{F}$ is to consider its *attractors* defined by $\mathsf{Att}(\mathcal{F}) := \{\mathcal{A} \subset \mathcal{X}^{\text{top}} \mid \mathcal{F}(\mathcal{A}) = \mathcal{A}\}$. The equivalence arises from the fact that, as shown in [16], $\mathsf{Att}(\mathcal{F})$ is a bounded distributive lattice where the partial order is inclusion. More precisely, if we let $\mathsf{J}(\mathsf{Att}(\mathcal{F}))$ denote the set of join irreducible elements of $\mathsf{Att}(\mathcal{F})$, i.e., those elements of $\mathsf{Att}(\mathcal{F})$ that have a unique immediate predecessor under inclusion, then there exists a poset isomorphism $\nu : \mathsf{M}(\mathcal{F}) \to \mathsf{J}(\mathsf{Att}(\mathcal{F}))$ [17].

Example 3.3. Continuing with Example 3.2,

$$\mathsf{Att}(\mathcal{F}) = \{\emptyset, \{e_{-2}\}, \{e_1\}, \{e_{-2}, e_1\}, \{e_{-2}, e_{-1}, e_0, e_1\}\}.$$

Observe that $\mathsf{J}(\mathsf{Att}(\mathcal{F})) = \{\{e_{-2}\}, \{e_1\}, \{e_{-2}, e_{-1}, e_0, e_1\}\}$, and, using inclusion to define the partial order, we obtain a poset that is isomorphic to $\mathsf{M}(\mathcal{F})$. In this case the poset isomorphism $\nu : \mathsf{M}(\mathcal{F}) \to \mathsf{J}(\mathsf{Att}(\mathcal{F}))$ satisfies $\nu(\mathcal{M}_0) = \{e_{-2}\}$, $\nu(\mathcal{M}_1) = \{e_1\}$, and $\nu(\mathcal{M}_2) = \{e_{-2}, e_{-1}, e_0, e_1\}$.

We use the first perspective (associated with posets, e.g., Morse graphs) for efficient computations and to organize the global information, and use the second perspective (associate with lattices, e.g., attractors) to identify the homological computations that recover nontrivial information about the structure of the dynamics exhibited by the continuous function.

For the sake of simplicity we define an *index pair* for $\mathcal{F}$ to be a pair $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_0)$ where $\mathcal{A}_1, \mathcal{A}_0 \in \mathsf{Att}(\mathcal{F})$ and $\mathcal{A}_0 \subset \mathcal{A}_1$. Observe that $\mathcal{F}(\mathcal{A}_i) \subset \mathcal{A}_i$, $i = 0, 1$. Under rather weak conditions [14] (we return to this point below) $\mathcal{F}$ induces a map on homology, i.e.,

$$\mathcal{F}_* : H_*(\downarrow(\mathcal{A}_1), \downarrow(\mathcal{A}_0); \mathbb{F}) \to H_*(\downarrow(\mathcal{A}_1), \downarrow(\mathcal{A}_0); \mathbb{F}).$$

The *Conley index* of $\mathcal{A}$, denoted by $\text{Con}_*(\mathcal{A}; \mathbb{F})$, is defined to be the shift equivalence class of $\mathcal{F}_*$ (if $\mathbb{F}$ is a field, then this is equivalent to the rational canonical form of the linear map $\mathcal{F}_*$ [3]). In particular, we can assign a Conley index to each $\mathcal{M} \in \mathsf{M}(\mathcal{F})$ by declaring $\text{Con}_*(\mathcal{M}; \mathbb{F}) \cong \text{Con}_*(\downarrow(\nu(\mathcal{M})), \downarrow(\overleftarrow{\nu}(\mathcal{M})); \mathbb{F})$, where $\overleftarrow{\nu}(\mathcal{M})$ is the unique immediate predecessor of $\nu(\mathcal{M})$. Given $\mathcal{X}$ and $\mathcal{F}$ there exists software [11] to compute $\mathsf{M}(\mathcal{F})$, $\mathsf{Att}(\mathcal{F})$, and $\text{Con}_*(\mathcal{M}; \mathbb{F})$ (for this paper we take $\mathbb{F} \cong \mathbb{Z}_5$). This software is based on what are essentially combinatorial algorithms and as a consequence are extremely efficient.

Example 3.4. Continuing with Example 3.3, the set of index pairs is

$$(\{e_{-2}\}, \emptyset), \quad (\{e_1\}, \emptyset), \quad \text{and} \quad (\{e_{-2}, e_{-1}, e_0, e_1\}, \{e_{-2}, e_1\}).$$

The attractors are defined in terms of elements of $\mathcal{X}^{top}$. To compute homology we need to work with the chain complexes associated with the attractors, i.e.,

$$\downarrow(\{e_{-2}\}) = \{e_{-2}, v_{-2}, v_{-1}\}, \quad \downarrow(\{e_1\}) = \{e_1, v_1, v_2\}, \quad \text{and} \quad \downarrow(\{e_{-2}, e_{-1}, e_0, e_1\}) = \mathcal{X}.$$

The relative homology of these index pairs is

$$H_k(\downarrow(\{e_{-2}\}), \emptyset; \mathbb{F}) \cong H_k(\downarrow(\{e_1\}), \emptyset; \mathbb{F}) \cong \begin{cases} \mathbb{Z}_2 & \text{if } k = 0, \\ 0 & \text{otherwise} \end{cases}$$

and

$$\begin{aligned} H_k(\downarrow(\{e_{-2}, e_{-1}, e_0, e_1\}), \downarrow(\{e_{-2}, e_1\}); \mathbb{F}) &\cong H_k(\{e_{-1}, e_0, v_0\}, \emptyset; \mathbb{F}) \\ &\cong H_k(\{e_{-1}\}, \emptyset; \mathbb{F}) \cong \begin{cases} \mathbb{Z}_2 & \text{if } k = 1, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

The induced maps on homology are the identity maps, and thus the rational canonical forms of $(\mathcal{F}_0, \mathcal{F}_1)$ are $(x-1, 0)$, $(x-1, 0)$, and $(0, x-1)$, respectively. These are the Conley indices of the elements $\mathcal{M}_0$, $\mathcal{M}_1$, and $\mathcal{M}_2$ of the Morse graph, respectively.

Before relating the abovementioned combinatorial framework to continuous dynamics we recall the following concepts. Let $g : X \to X$ be a continuous map on a compact space. Given $N \subset X$, the *maximal invariant set* contained in N is given by

$$\mathrm{Inv}(N, g) := \{x \in N \mid \exists\, \sigma : \mathbb{Z} \to N \text{ such that } \sigma(0) = x \text{ and } \sigma(n+1) = g(\sigma(n)) \ \forall\ n \in \mathbb{Z}\}.$$

A compact set $N \subset X$ is an *attracting block* if $g(N) \subset \mathrm{int}(N)$ and an *isolating neighborhood* if $\mathrm{Inv}(N, g) \subset \mathrm{int}(N)$. It is easily checked that an attracting block is an isolating neighborhood.

As suggested above, the phase space X for the dynamics generated by g is represented by the cell complex $\mathcal{X}$. In particular, we assume that X is a regular CW complex [15, 16], and we use the map $|\cdot| : \mathcal{X} \to X$ to identify how the cell complex $\mathcal{X}$ realizes the regular CW complex X, i.e., given $\xi \in \mathcal{X}$, if $\dim(\xi) = n$, then $|\xi|$ represents the corresponding regular closed cell in the n-skeleton of X and $|\mathcal{X}| = X$. In applications, we start with the space X and choose a decomposition $\mathcal{X}$. We define $G = \bigcup_{\xi \in \mathcal{X}} |\xi| \times |\mathcal{F}(\xi)|$.

Example 3.5. Returning to Example 3.1, note that $\mathcal{X}$ represents a decomposition of the interval $[-2, 2] \subset \mathbb{R}$ where $|v_i| = i$ and $|e_i| = [i, i+1]$ for all i. This in turn implies (see Example 3.2) that

$$G = [-2,-1] \times [-2,-1] \cup [-1,0] \times [-2,1] \cup [0,1] \times [-1,2] \cup [1,2] \times [1,2] \subset [-2,2]^2.$$

Since each vertical fiber of G is an interval, $\mathcal{F}$ is acyclic for each ξ.

To relate the combinatorial multivalued map $\mathcal{F} : \mathcal{X}^{\mathrm{top}} \rightrightarrows \mathcal{X}^{\mathrm{top}}$ with the continuous function $g : X \to X$, we make two assumptions. First, we assume that $\mathcal{F}$ is an *outer approximation* of g, that is, $g(|\xi|) \subset \mathrm{int}(|\mathcal{F}(\xi)|)$ for all $\xi \in \mathcal{X}^{\mathrm{top}}$. Second, if we extend $\mathcal{F}$ to all of $\mathcal{X}$ by setting

$\mathcal{F}(\xi) := \downarrow\left(\left\{\mathcal{F}(\xi') \,|\, \xi' \in \mathcal{X}^{\text{top}},\ \xi \le \xi'\right\}\right)$, then $\mathcal{F}(\xi)$ is *acyclic,* i.e., $\bar{H}_*(\downarrow(\mathcal{F}(\xi)) = 0$, where $\bar{H}$ denotes reduced homology. In this case we say that $\mathcal{F}$ an *acyclic outer approximation* of g.

Example 3.6. Consider $f : [-2,2] \to [-2,2]$ defined by $f(x) = \frac{3}{2}\arctan(x)$. Returning to Example 3.5, we let the reader check that (see Figure 3)

$$G(f) = \{(x, f(x) \,|\, -2 \le x \le 2\} \subset \text{int}(G) \subset [-2,2]^2.$$

This implies that $\mathcal{F}$ is an acyclic outer approximation of f.

Under these assumptions, if $\mathcal{M} \in \mathsf{M}(\mathcal{F})$, then

$$\text{Con}_*\left(\text{Inv}\left(|\nu(\mathcal{M}) \setminus \overleftarrow{\nu}(\mathcal{M})|, g\right); \mathbb{F}\right) \sim \text{Con}_*\left(\mathcal{M}; \mathbb{F}\right), \tag{3.1}$$

where Con_* on the left denotes the classical homology Conley index for maps [21], and Con_* on the right is the Conley index defined above. As indicated in the introduction, knowledge of the Conley index provides information about the structure of the dynamics of

$$\text{Inv}\left(|\nu(\mathcal{M}) \setminus \overleftarrow{\nu}(\mathcal{M})|, g\right),$$

i.e., the computations outlined in this section provide information about the invariant dynamics contained in $\bar{\nu}(\mathcal{M}) := |\nu(\mathcal{M}) \setminus \overleftarrow{\nu}(\mathcal{M})|$.

Example 3.7. Combining the information from the previous examples, we have derived the following information concerning the global structure of the dynamics generated by the map $f : [-2,2] \to [-2,2]$. The fact that the Conley indices of $\mathcal{M}_i$ are not trivial (Example 3.4) implies that $\text{Inv}(\bar{\nu}(\mathcal{M}_0), f) = \text{Inv}([-2,-1], f) \neq \emptyset$, $\text{Inv}(\bar{\nu}(\mathcal{M}_1), f) = \text{Inv}([1,2], f) \neq \emptyset$, and $\text{Inv}(\bar{\nu}(\mathcal{M}_2), f) = \text{Inv}([-1,1], f) \neq \emptyset$. Furthermore, the fact that the Conley index is the identity map on a one-dimensional vector space implies that each of these invariant sets contains a fixed point. Finally, the dynamics of f exhibits bistability since the poset structure on the Morse graph implies that if $x \in [-2,-1]$, then $f^n(x) \in [-2,-1]$ for all $n \ge 1$, and similarly if $x \in [1,2]$, then $f^n(x) \in [1,2]$ for all $n \ge 1$.

Remark 3.8. The presentation of Examples 3.1–3.7 clearly was chosen to follow (and hopefully shed light on) the curt review of combinatorial Conley theory. In practice, e.g., in the examples of section 5, the order of development is different (see Figure 3). The first steps involve the choice of the phase space X and the identification of a surrogate model. For this paper, we restrict our attention to $X \subset \mathbb{R}$ (see [25] for applications of these ideas in the context of robotic control where $X \subset \mathbb{R}^n$ with $n > 1$). Furthermore, unlike our choice in Example 3.1 of $\mathcal{X}^{\text{top}}$ consisting of four elements, in the examples of section 5 we choose $\mathcal{X}^{\text{top}}$ containing at least 2^{10} elements. In this setting an explicit list of the rectangular regions that make up the region $G \subset X \times X$ is meaningless; thus we plot G in blue (see Figure 1). We also present the $\mathsf{M}(\mathcal{F})$ as a graph where the relative ordering decreases as one goes from top to bottom, i.e., minimal elements are at the bottom, and we indicate the Conley index within each node.

4. Probabilistic bounds using Gaussian process surrogates. Let $X \subset \mathbb{R}^d$ be a compact regular CW complex indexed by a cell complex $\mathcal{X}$, i.e., if $\xi \in \mathcal{X}$ and $\dim(\xi) = \ell$, then $|\xi|$

is the closure of the ℓ-dimensional cell in X. We assume that Step 1 and Step 2 have been completed, which implies that we have obtained a predictive mean $\mu : X \to \mathbb{R}$ and a predictive covariance function Σ based on the GP, and have identified dynamics via Conley theory. We adopt the approach where we are only interested in dynamics in X that can be obtained with a given amount of confidence quantified by $0 < \delta < 1$.

Let $S = \mathcal{X}^{(0)}$, where $\mathcal{X}^{(0)}$ denotes the set of vertices of $\mathcal{X}$. Choose $r(v) > 0$ for $v \in S$ such that (2.2) is satisfied. We emphasize that there is considerable freedom in the choice of the individual values of $r(v)$. In particular, when $r(v) = \chi^2_d((1-\delta)^{1/\#S})$ for each $v \in S$ we say that we are choosing *pointwise equal confidence*.

To define $\mathcal{F} : \mathcal{X} \rightrightarrows \mathcal{X}$ we make use of the following notation. Let $\operatorname{diam} A = \sup_{x,x' \in A} \|x - x'\|$ for $A \subset X$. Given $\xi \in \mathcal{X}$, set $\operatorname{diam} \xi = \operatorname{diam} |\xi|$ and $\operatorname{diam} \mathcal{X} = \sup_{\xi \in \mathcal{X}} \operatorname{diam} \xi$. Moreover, given $U \subset \mathbb{R}^d$, let $B_z(U) := \{x \in \mathbb{R}^d \mid \inf_{u \in U} \|x - u\| \le z\}$. Choose $L > 0$ sufficiently large (see Lemma 4.3). If $\xi \in \mathcal{X}^{\text{top}}$, set

$$Q(\xi) := \bigcap_{v < \xi,\, v \in S} B_{L \operatorname{diam} \xi}\left(E_{\Sigma(v)}(\mu(v), r(v))\right) \tag{4.1}$$

and define $\mathcal{F} : \mathcal{X}^{\text{top}} \rightrightarrows \mathcal{X}^{\text{top}}$ by $\mathcal{F}(\xi) := \{\xi' \in \mathcal{X}^{\text{top}} \mid |\xi'| \cap Q(\xi) \neq \emptyset\}$. Define

$$G = G_{L,r} := \bigcup_{\xi \in \mathcal{X}^{\text{top}}} |\xi| \times |\mathcal{F}(\xi)| \subset X \times X \tag{4.2}$$

and

$$\tilde{G} = \tilde{G}_{L,r} := G_{L,r} \cup \left(\bigcup_{\xi \in \mathcal{X}^{\text{top}}} |\xi| \times Q(\xi) \right) \subset X \times \mathbb{R}^d, \tag{4.3}$$

where, when convenient, we drop the explicit dependence on L and r. Note that G is a cover by cells of the cell decomposition of $X \times X$ of the confidence sets given by (4.1) restricted to $X \times X$, while $\tilde{G}$ includes the portions of these confidence sets that are not in $X \times X$, and $G = \tilde{G} \cap (X \times X)$.

Theorem 4.1. *Let $\mathcal{T}$ be a data set that satisfies assumption* A *where $\{x_n \mid n = 1, \dots, N\}$ are chosen to be i.i.d. from a uniform distribution, and assume the kernel k satisfies the conditions for (2.3). Let $\alpha > 0$ and $\delta \in (0,1)$. There exist $\varepsilon_0 > 0$ and $n_0 \in \mathbb{N}$ such that the set $\tilde{G}$ given by (4.3) satisfies*

$$\mathbb{P}\left(\sup_{x \in X} \operatorname{diam}(\tilde{G}_x) < \alpha\right) > 1 - \delta \quad \text{and} \quad \mathbb{P}(\mathsf{G}(g) \subset \tilde{G}) > 1 - \delta \tag{4.4}$$

provided that $N > n_0$, g is a GP constructed as in Step 1, *and as in* Step 2, *$X \subset \mathbb{R}^d$ is a compact regular CW complex indexed by a cell complex $\mathcal{X}$ with $\operatorname{diam}(\mathcal{X}) < \varepsilon_0$, where the top dimensional cells $\mathcal{X}^{top}$ are d-dimensional.*

Remark 4.2. Note that $G \subset X \times X$ is defined in terms of the map $\mathcal{F}$, and so the computed dynamics is valid for all sample paths whose graphs are in G, and $\mathbb{P}(\mathsf{G}(g) \subset G)$ gives the confidence level on the dynamics. However, we can only estimate $\mathbb{P}(\mathsf{G}(g) \subset \tilde{G})$, by Theorem

4.1, and hence we adopt the following strategy: If $\tilde{G} \subset X \times X$, and hence $G = \tilde{G}$, then we have the confidence level $\mathbb{P}(\mathsf{G}(g) \subset G) = \mathbb{P}(\mathsf{G}(g) \subset \tilde{G}) > 1 - \delta$ on the dynamics. If, on the other hand, $\tilde{G} \not\subset X \times X$, then we cannot estimate the confidence level $\mathbb{P}(\mathsf{G}(g) \subset G)$, and so we declare failure in identifying the dynamics with confidence level $1 - \delta$, since in this case $\mathbb{P}(\mathsf{G}(g) \subset G) \leq \mathbb{P}(\mathsf{G}(g) \subset \tilde{G})$, and hence the confidence level may be less than $1 - \delta$.

Notice that Theorem 4.1 indicates that if $f(X) \subset \mathrm{int}(X)$, then we should have $\tilde{G} \subset X \times X$ as long as we have enough data points and the grid is sufficiently fine. Therefore a failure suggests that it may be necessary to choose a larger domain X, more data points, or a smaller confidence level $1 - \delta$.

Theorem 4.1 implies that with sufficient data and sufficient computational effort we can obtain the following two fundamental results:

1. Detailed dynamics can be extracted from G via the Conley theory computations, since the sizes of the fibers of G are bounded above by an arbitrarily chosen α with confidence $1 - \delta$.
2. The dynamics identified via G occurs since a given realization h of the GP model is a selector of G, that is, $\mathsf{G}(h) \subset G$, with probability $1 - \delta$. Furthermore, this dynamics is valid with confidence greater than $1 - \delta$.

In the computations in section 5 we fix the data size N, and hence we only give the confidence level of the correctness of the dynamics (item 2 above).

To prove Theorem 4.1 we begin by recalling and establishing the necessary notation, and then we prove a series of lemmas that are used in the proof.

For the remainder of this section we assume that $X \subset \mathbb{R}^d$ is a compact set that is the regular CW complex realization of a cell complex $\mathcal{X}$ where the top dimensional cells $\mathcal{X}^{\mathrm{top}}$ are d-dimensional.

We assume that $\mathcal{T} = \{(x_n, y_n) \mid n = 1, \ldots, N\}$ is a data set that satisfies assumption A, the kernel k satisfies the conditions for (2.3), g is a GP constructed as in Step 1, and $\tilde{G} = \tilde{G}_{L,r} \subset X \times \mathbb{R}$ is constructed according to (4.3), where the conditions on L and r are as described in what follows.

Recall that $\mathsf{G}(g)$ denotes the graph of g. We use the following lemma to quantify the confidence that the graph of the GP g lies in $\tilde{G}$.

For Theorem 4.1 and Lemma 4.5 we need the assumption that $S = \mathcal{X}^{(0)}$; however, for the next two lemmas, we have more flexibility on the choice of S as stated. Note that Lemma 4.3 gives the confidence level for our computations.

Lemma 4.3. *Fix $\delta \in (0,1)$, and let $S \subset \mathcal{X}^{(0)}$ be such that for each $\xi \in \mathcal{X}^{top}$, there exists $v \in S$ such that $v < \xi$. Then, there exist $r : S \to (0, \infty)$ that satisfies (2.2) and $\tilde{L} > 0$ such that $L > \tilde{L}$ implies that*

$$\mathbb{P}\left(\mathsf{G}(g) \subset \tilde{G}_{L,r}\right) > 1 - \delta. \tag{4.5}$$

Proof. Let $L_0, a, b > 0$ be as in (2.3), and let $\delta = \delta_1 + \delta_2$ with $\delta_1, \delta_2 > 0$. Choose $\tilde{L} \geq L_0$ such that $1 - ae^{-(L/b)^2} \geq 1 - \delta_1$ for $L > \tilde{L}$. By (2.3) with probability greater than $1 - ae^{-(L/b)^2}$ we have

$$\|g(x) - g(y)\| \leq L\|x - y\| \ \forall\ x, y \in X. \tag{4.6}$$

Let $\mathcal{X}$ be a cell complex decomposition of X. According to (2.2) we can pick $r : S \to (0, \infty)$ such that with probability greater than $1 - \delta_2$ we have

$$g(v) \in E_{\Sigma(v)}(\mu(v), r(v)) \ \forall \ v \in S. \tag{4.7}$$

It follows that g satisfies (4.6) and (4.7) with probability greater than $1 - \delta$.

Now it suffices to show that if g satisfies (4.6) and (4.7), then $\mathsf{G}(g) \subset \tilde{G}_{L,r}$. If $x \in X$, then $x \in |\xi|$ for some $\xi \in \mathcal{X}^{\text{top}}$. Let $v \in S$ such that $v < \xi$. Then, $\|g(x) - g(v)\| \leq L\|x - v\| \leq L \operatorname{diam}(\xi)$. This, along with (4.7), shows that $g(x) \in B_{L \operatorname{diam}(\xi)}\left(E_{\Sigma(v)}(\mu(v), r(v))\right)$ and therefore, $(x, g(x)) \in \tilde{G}_{L,r}$. ∎

We remark that Lemma 4.3 does not depend on the choice of cell complex $\mathcal{X}$. Thus we exploit the size of the geometric representation of cells to control the size of G. Set

$$\varepsilon := \operatorname{diam}(\mathcal{X}) \quad \text{and} \quad \ell := \max\left\{\operatorname{diam}(E_{\Sigma(v)}(\mu(v), r(v))) \mid v \in S\right\}.$$

Given $\xi \in \mathcal{X}^{\text{top}}$ let $\tilde{Q}(\xi) := |\mathcal{F}(\xi)| \cup Q(\xi) = \left|\left\{\xi' \in \mathcal{X}^{\text{top}} \mid |\xi'| \cap Q(\xi) \neq \emptyset\right\}\right| \cup Q(\xi)$.

Lemma 4.4. *Let $S \subset \mathcal{X}^{(0)}$ be such that for each $\xi \in \mathcal{X}^{top}$, there exists $v \in S$ such that $v < \xi$, and let $\tilde{L}$ be chosen as in Lemma 4.3 and $L > \tilde{L}$. If $\xi \in \mathcal{X}^{top}$, then*

$$\operatorname{diam}(\tilde{Q}(\xi)) < \ell + 2L\varepsilon + 2\varepsilon.$$

Proof. Let $\xi \in \mathcal{X}^{\text{top}}$. Fix $v_0 \in S$ such that $v_0 < \xi$ and observe that, by (4.1),

$$\begin{aligned}\operatorname{diam}(Q(\xi)) &= \operatorname{diam}\left(\bigcap_{v < \xi,\, v \in S} B_{L \operatorname{diam} \xi}\left(E_{\Sigma(v)}(\mu(v), r(v))\right)\right) \\ &\leq \operatorname{diam}\left(B_{L \operatorname{diam} \xi}\left(E_{\Sigma(v_0)}(\mu(v_0), r(v_0))\right)\right) \leq \ell + 2L\varepsilon.\end{aligned}$$

The result follows from the fact that $\mathcal{F}(\xi)$ is obtained by covering $Q(\xi) \cap X$ by elements of $\mathcal{X}^{\text{top}}$. ∎

For the remainder of this section we assume the that $S = \mathcal{X}^{(0)}$.

Lemma 4.5. *Let $\tilde{L}$ be chosen as in Lemma 4.3 and $L > \tilde{L}$, and let $\tilde{G} = \tilde{G}_{L,r}$ be defined as in (4.3). Then,*

$$\operatorname{diam}(\tilde{G}_x) < 2(\ell + 2L\varepsilon + 2\varepsilon).$$

Proof. Note that $\tilde{G}_x = \bigcup\{\{x\} \times \tilde{Q}(\xi) \mid x \in |\xi|, \xi \in \mathcal{X}^{\text{top}}\}$. If there is only one $\xi \in \mathcal{X}^{\text{top}}$ such that $x \in |\xi|$, then the result follows from Lemma 4.4. So assume that there are multiple $\xi \in \mathcal{X}^{\text{top}}$ for which $x \in |\xi|$. Let $(x, y), (x, y') \in \tilde{G}_x$. Then there exists $\xi, \xi' \in \mathcal{X}^{\text{top}}$ such that $x \in |\xi| \cap |\xi'|$, $y \in \tilde{Q}(\xi)$, and $y' \in \tilde{Q}(\xi')$. Since $|\xi| \cap |\xi'| \neq \emptyset$, it follows that there exists $v_0 \in \mathcal{X}^{(0)}$ such that $v_0 < \xi$ and $v_0 < \xi'$. Then, from the definition of $\tilde{Q}$ we have that $\mu(v_0) \in \tilde{Q}(\xi) \cap \tilde{Q}(\xi')$ and hence that $\tilde{Q}(\xi) \cap \tilde{Q}(\xi') \neq \emptyset$. Therefore it follows from Lemma 4.4 that

$$\|y - y'\| \leq \operatorname{diam}(\tilde{Q}(\xi)) + \operatorname{diam}(\tilde{Q}(\xi')) < 2(\ell + 2L\varepsilon + 2\varepsilon),$$

from which the result follows. ∎

Up to this point the construction is valid for any data set $\mathcal{T}$ that satisfies assumption A. To control the size of the fibers of $\tilde{G}$ we recall [8, Proposition 1] that if a kernel k of a 1-dimensional Gaussian process is four times differentiable on the diagonal, and if we sample densely enough, then the posterior predictive variance σ^2 is uniformly bounded from above. More precisely, if the set of sample points is a γ-cover of X, then there exists a constant $Q^2 \leq \sup_X \partial_x^2 \partial_y^2 k(x,y)|_{x=y}$ such that

$$\sup_{x \in X} \sigma(x) \leq \frac{Q\gamma^2}{4}. \tag{4.8}$$

For d-dimensional outputs, by vectorizing the outputs and having a prespecified kernel function k for the Gaussian process, the prediction for an untried point can be obtained by a d-dimensional normal distribution with mean $\mu(x)$ and a covariance matrix $\Sigma(x)$. The maximum eigenvalue of $\Sigma(x)$ is bounded by $\operatorname{trace}(\Sigma(x))$ which is equals to the summation of the variances in each dimension, so the result in (4.8) can be applied to each dimension. Therefore, given $\gamma > 0$ and a set of sample points that is a γ-cover of X, there exists a constant $C = C(\gamma)$ such that

$$\sup \{\lambda(x) \text{ a maximal eigenvalue of } \Sigma(x) \mid x \in X\} \leq C(\gamma). \tag{4.9}$$

Moreover,

$$\lim_{\gamma \to 0} C(\gamma) = 0. \tag{4.10}$$

Proof of Theorem 4.1. Fix $L > \tilde{L}$, where $\tilde{L}$ satisfies the conditions in Lemma 4.3. Fix $\varepsilon_0 > 0$ such that

$$4(L+1)\varepsilon_0 \leq \alpha/2. \tag{4.11}$$

Consider $\mathcal{X}$ a CW structure on X with $\operatorname{diam}(\mathcal{X}) = \varepsilon < \varepsilon_0$ and a map $r : S \to (0, \infty)$ as in Lemma 4.3. Set $R := \max\{r(v) \mid v \in S\}$. By (4.10) we can choose $\gamma_0 > 0$ such that

$$C(\gamma_0) < \frac{\alpha^2}{64R}. \tag{4.12}$$

By [1, Theorem 3.7] there exists an $N_0 \in \mathbb{N}$ such that any sample of size $N > N_0$ is a γ_0-cover of X with probability greater than $1 - \delta$. Thus, we assume that N, the number of data points in $\mathcal{T}$, satisfies $N > N_0$.

Let $\tilde{G}$ be defined as in (4.3). Note that the high probability inclusion (4.5) follows from Lemma 4.3. We conclude the proof by verifying

$$\mathbb{P}\left(\sup_{x \in X} \operatorname{diam}(\tilde{G}_x) < \alpha\right) > 1 - \delta.$$

With probability at least $1 - \delta$, $\mathcal{T}$ is a γ_0-cover of X. Therefore, by Lemma 4.5 and (4.9),

$$\operatorname{diam}(\tilde{G}_x) \leq 2\ell + 4(L+1)\varepsilon \leq 4\sqrt{C(\gamma_0)R} + 4(L+1)\varepsilon < \alpha,$$

where the last inequality follows from (4.12) and (4.11). ∎

5. Examples. We conclude with examples demonstrating that our approach is capable of identifying with high levels of confidence a broad range of dynamics from relatively few data points.

As discussed at the end of section 3, the Conley index computed via the multivalued map $\mathcal{F}$ is valid for any dynamical system for which $\mathcal{F}$ is an outer approximation. Hence the results presented in this section about fixed points, periodic orbits, connecting orbits, and chaotic dynamics occur for any dynamical system for which $\mathcal{F}$ is an outer approximation; that is, the results of each example are valid for any sample path h of the Gaussian process whose graph lies inside the region G defined by (4.2) in terms of $\mathcal{F}$.

As indicated in section 3, the Conley index of $\mathcal{M} \in \mathsf{M}(\mathcal{F})$ in dimension n is the rational canonical form of the linear map

$$\mathcal{F}_n : H_n\left(\downarrow(\nu(\mathcal{M})), \downarrow(\overleftarrow{\nu}(\mathcal{M})); \mathbb{F}\right) \to H_n\left(\downarrow(\nu(\mathcal{M})), \downarrow(\overleftarrow{\nu}(\mathcal{M})); \mathbb{F}\right).$$

Since we are working with 1-dimensional complexes $\mathcal{F}_n = 0$ for all $n \geq 2$–, we can express the Conley index of $\mathcal{M}$ as

$$\mathrm{Con}_*(\mathcal{M}; \mathbb{F}) \cong (p_0(x), p_1(x)),$$

where $p_n(x)$ is a monic polynomial [3]. For the remainder of this discussion, $\mathbb{F} = \mathbb{Z}_5$ (this allows us to distinguish whether the dynamics is orientation preserving or reversing).

We briefly mention a few standard results about the Conley index (see [21] for more details). A *trivial* Conley index in the ith dimension takes the form $p_i(x) = x^k$ for some $k \in \mathbb{N}$. *However, to emphasize the triviality in the figures of this paper we write* $p_i(x) = 0$. If the Conley index is not trivial, i.e., $p_i(x) \neq x^k$, then the maximal invariant set in $|\nu(\mathcal{M})|$ is nonempty (the converse is not true in general). In the examples of this paper we make use of the following facts. If the Conley index has the form $(x \pm 1, 0)$ or $(0, x \pm 1)$, then the maximal invariant set in $|\nu(\mathcal{M})|$ contains a fixed point. If the Conley index has the form $(x^T \pm 1, 0)$ or $(0, x^T \pm 1)$, then the maximal invariant set in $|\nu(\mathcal{M})|$ contains a periodic orbit of period T. Let $\mathcal{M}_i$ and $\mathcal{M}_{i+1}$ be nodes in a Morse graph, and assume that with respect to the poset order, $\mathcal{M}_{i+1}$ covers $\mathcal{M}_i$. If

$$\begin{aligned}\mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_{i+1})), \downarrow(\overleftarrow{\nu}(\mathcal{M}_i)); \mathbb{F}\right) \\ \ncong \mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_{i+1})), \downarrow(\overleftarrow{\nu}(\mathcal{M}_{i+1})); \mathbb{F}\right) \oplus \mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_i)), \downarrow(\overleftarrow{\nu}(\mathcal{M}_i)); \mathbb{F}\right),\end{aligned}$$

then there exists a connecting orbit from $\mathcal{M}_{i+1}$ to $\mathcal{M}_i$.

As indicated above, the goal of this section is to show via examples that our approach can identify, with high probability, interesting dynamics based on relatively few data points. To do this we pick a smooth function f defined on an interval $X \subset \mathbb{R}$ that we know produces dynamics of interest, e.g., fixed points, periodic orbits, and chaotic dynamics. We then randomly (in an i.i.d. way with respect to the uniform measure on X) generate data points $\mathcal{T} = \{(x_n, f(x_n)) \mid n = 1, \ldots, N\}$. In line with the fact that f is smooth, we have chosen to use a smooth kernel, the squared exponential kernel $k(x, x') = \exp\left(-|x - x'|^2/\theta\right)$, and we *assume* that the function f is a sample path of the GP obtained from the data using MLE to estimate θ.

For the examples in this section, we know the function f from which the data is sampled. Therefore, we can verify that indeed $\mathsf{G}(f) \subset G$. As a consequence, the dynamics that we

report is valid for f. Of course, in applications we do not expect to be able to perform this explicit validation. We can only claim that if the Lipschitz constant L used is large enough (see below), and if we *assume* that f is a sample path of the GP obtained from the data, then we have provided a lower bound on the confidence level that the dynamics identified via the homological calculations is valid for f. Without the assumption that f is a sample path of the GP, we can only claim that we have provided a lower bound on the probability that a dynamical system generated by a sample path of the obtained GP will exhibit the dynamics identified by our method.

Our method requires knowledge of the data set $\mathcal{T}$ and the assumption that our choice of L provides a bound for the Lipschitz constant of the unknown function f with confidence $1-\delta/2$; that is, we assume that the probability in (2.3) is at least $1-\delta/2$. We then select the set S and construct the function r such that the probability in (2.2) is also at least $1-\delta/2$. Then we construct the set G in Theorem 4.1 satisfying $\mathbb{P}(\mathsf{G}(g) \subset G) > 1-\delta$ (see Lemma 4.3). Hence the extracted dynamics has a confidence level of at least $1-\delta$.

Although the function f does not need to be specified for our method, in the examples we indicate f to illustrate the fact that, at least in these cases, we can recover the dynamics of f with very few data points. Since we do not know the values of L_0, a, and b necessary for (2.3), in the examples we take L to be at least twice the Lipschitz constant of f. For all the examples presented here we used $L=8$. The cell complex $\mathcal{X}$ is obtained by considering a uniform decomposition of the interval X into 2^B subintervals.

For the results presented here, our primary concern in the choice of G, which, given a cell complex $\mathcal{X}$, is equivalent to a choice of $\mathcal{F}:\mathcal{X}\rightrightarrows\mathcal{X}$, is to emphasize our ability to characterize dynamics with high levels of confidence. Thus, we impose a small δ in (2.2) on S. Unless otherwise stated, we use $\delta=0.05$, and hence we obtain a 0.95 (or 95%) confidence level.

Bistability. To demonstrate that bistability can be identified via the Morse graph, we turn to Figure 1(a) that was generated using $N=8$ data points sampled on the interval $X=[0,1]$ for the function $f(x)=0.3\arctan(8x-4)+0.5$. The set G was constructed using pointwise equal confidence intervals with $B=9$ and $\delta=0.05$. Observe that the Morse graph has two minimal nodes $\mathcal{M}(0)$ and $\mathcal{M}(2)$. Under the poset isomorphism, $\nu(\mathcal{M}(0))$ and $\nu(\mathcal{M}(2))$ are disjoint attracting blocks. Therefore the dynamics of any function h with $\mathsf{G}(h)\subset G$ exhibits at least bistability (it is possible that for a given h there are additional attractors within $\nu(\mathcal{M}(0))$ and/or $\nu(\mathcal{M}(2))$.

The Conley indices for the bistability example are presented in Figure 1(a). The intervals defining the Morse sets are

$$\begin{aligned}
|\nu(M_0)| &= [0.09179688, 0.27929688],\\
|\nu(M_1)| &= [0.76171875, 0.94140625],\\
|\nu(M_2)| &= [0.75976562, 0.76171875],\\
|\nu(M_3)| &= [0.50976562, 0.51171875],
\end{aligned}$$

and

$$\begin{aligned}
|\nu(M_4)| =& [0.49023438, 0.49218750] \cup [0.49414062, 0.49609375] \cup\\
& [0.49804688, 0.50390625] \cup [0.50585938, 0.50781250].
\end{aligned}$$

Periodic orbit. As indicated in section 1, the Conley index can be used to identify periodic orbits of a given period. To demonstrate this, and to emphasize the importance of being able to choose G, we consider the logistic map $f(x) = 3.15x(1-x)$. The global dynamics for f is well understood. The points $x_0 = 0$ and $x_1 \approx 0.68$ are unstable fixed points, and all other initial conditions in $(0,1)$ limit to a stable period-2 orbit γ.

To apply our techniques, set $X = [0,1]$, $N = 4$, $B = 10$, and $\delta = 0.05$ (see Figure 4(a)). Choosing pointwise equal confidence intervals for this value of δ leads to a Morse graph $\mathsf{M}(\mathcal{F}) = \{\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2 \mid \mathcal{M}_0 < \mathcal{M}_1 < \mathcal{M}_2\}$, where $x_0 \in |\nu(M_2)|$ and $|\nu(M_0)|$ contains both x_1 and γ. The Conley index of $\mathcal{M}_0$ identifies the existence of a fixed point. Since $\delta = 0.05$, this description of the dynamics is valid with at least 95% confidence.

What is missing from this description is the identification of a period-2 orbit. Motivated by Theorem 4.1, we repeated the computations with $B = 11$ and detected the existence of a periodic orbit with at least 95% confidence (see Figure 4(b)). We can also become more ambitious and seek a confidence level of 97.5%, i.e., setting $\delta = 0.025$. For $B = 11$ we fail to identify the periodic orbit (see Figure 4(c)). Rather than increase the subdivision, we choose nonuniform confidence intervals, where we make the images of $\mathcal{F}(\xi)$ smaller if $|\xi| \subset |\nu(M_0)|$ and larger elsewhere (see Figure 4 (d)). The resulting Morse graph contains eight nodes where $x_0 \in |\nu(\mathcal{M}_7)|$, $x_1 \in |\nu(\mathcal{M}_6)|$, and the Conley index of $\mathcal{M}_0$ indicates the existence of a period-2 orbit.

The Conley indices for the period-2 orbit example are presented in Figure 4. The intervals defining the Morse sets are as follows: For Figure 4(a),

$$|\nu(M_0)| = [0.45800781, 0.80078125],$$
$$|\nu(M_1)| = [0.03613281, 0.03710938],$$

and

$$|\nu(M_2)| = [0.00000000, 0.03320312] \cup [0.03417969, 0.03515625];$$

for Figure 4(b),

$$|\nu(M_0)| = [0.46972656, 0.58056641] \cup [0.76269531, 0.79638672],$$

$$\begin{aligned}|\nu(M_1)| = {} & [0.58056641, 0.58105469] \cup [0.58154297, 0.58203125] \cup [0.58251953, 0.58300781] \cup \\ & [0.58349609, 0.58398438] \cup [0.58447266, 0.58496094] \cup [0.58544922, 0.76074219] \cup \\ & [0.76123047, 0.76171875] \cup [0.76220703, 0.76269531]\end{aligned}$$

and

$$|\nu(M_2)| = [0.00000000, 0.03808594] \cup [0.03857422, 0.03906250];$$

for Figure 4(c),

$$|\nu(M_0)| = [0.46728516, 0.79687500]$$

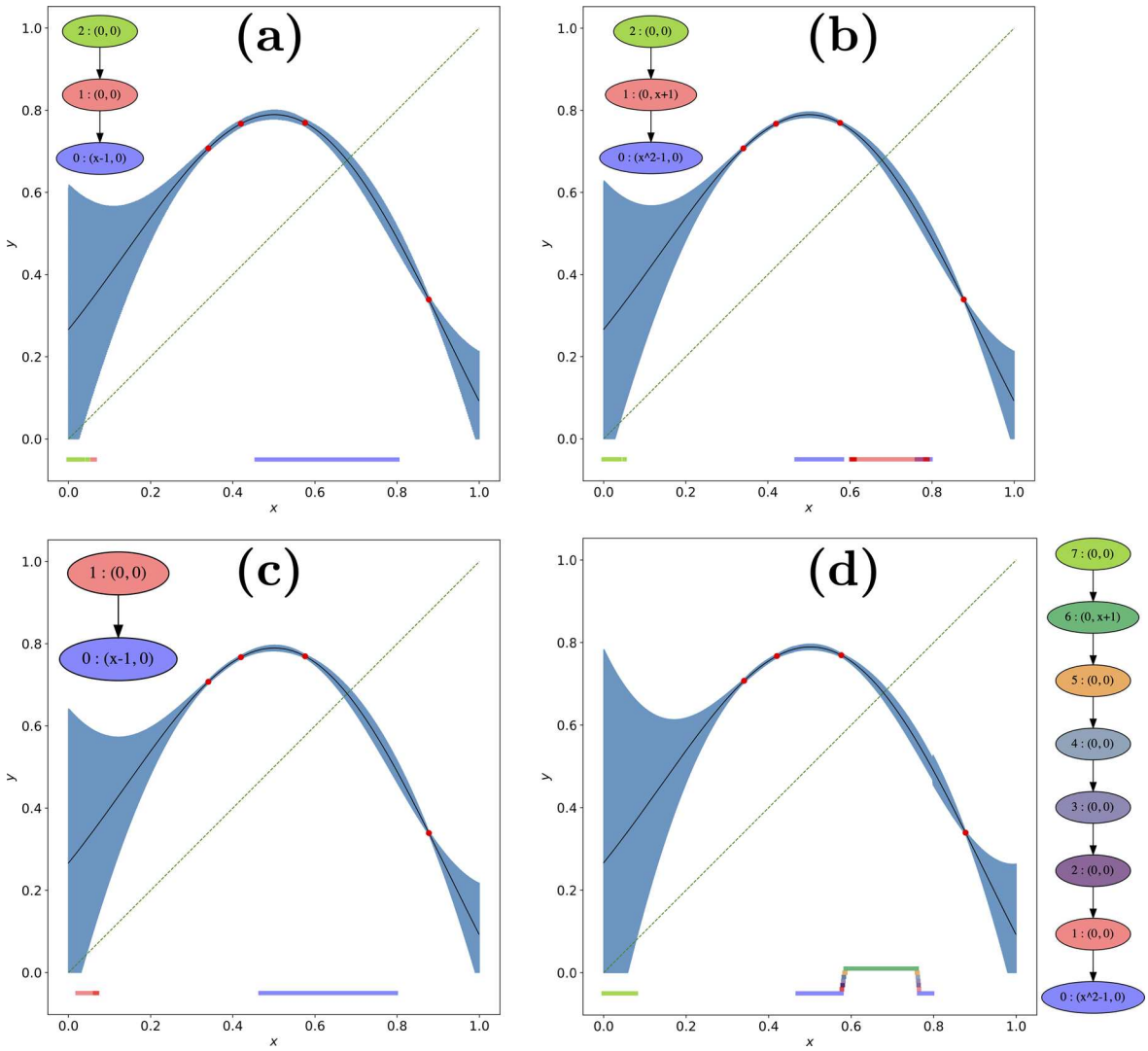

Figure 4. *In all subfigures, elements of $\mathcal{T}$ are indicated in red, the mean function μ is shown in black, and G is shown in blue. The Morse graphs are indicated on the figures, and the corresponding (color coded) regions of phase space $\bar{\nu}(\cdot)$ are indicated at the bottom of each figure.* (a) *The region G is composed of squares of width $2^{-10}\,\mathrm{length}(X)$. The Conley index for $\mathcal{M}_0$ identifies the existence of a fixed point with 95% confidence.* (b) *The region G is composed of squares of width $2^{-11}\,\mathrm{length}(X)$. The Conley index for $\mathcal{M}_0$ indicates the existence of a period-2 orbit with 95% confidence.* (c) *The region G is composed of squares of width $2^{-11}\,\mathrm{length}(X)$, and in this case the Conley index for $\mathcal{M}_0$ identifies the existence of a fixed point with 97.5% confidence, but we do not detect the periodic orbit.* (d) *The region G is composed of squares of width $2^{-11}\,\mathrm{length}(X)$. The Conley index for $\mathcal{M}_0$ indicate the existence of a period-2 orbit with 97.5% confidence.*

and

$$|\nu(M_1)| = [0.00000000, 0.04199219] \cup [0.04248047, 0.04296875];$$

for Figure 4(d),

$$|\nu(M_0)| = [0.47021484, 0.57666016] \cup [0.76464844, 0.79638672],$$

$$|\nu(M_1)| = [0.57666016, 0.57714844] \cup [0.57763672, 0.57812500] \cup [0.76416016, 0.76464844],$$

$$|\nu(M_2)| = [0.57861328, 0.57910156] \cup [0.57958984, 0.58007812] \cup [0.76318359, 0.76367188],$$

$$|\nu(M_3)| = [0.58056641, 0.58105469] \cup [0.76220703, 0.76269531],$$

$$|\nu(M_4)| = [0.58154297, 0.58203125] \cup [0.58251953, 0.58300781] \cup [0.76123047, 0.76171875],$$

$$\begin{aligned}|\nu(M_5)| = {} & [0.58349609, 0.58398438] \cup [0.58447266, 0.58496094] \cup [0.58544922, 0.58593750] \cup \\ & [0.75927734, 0.75976562] \cup [0.76025391, 0.76074219],\end{aligned}$$

$$\begin{aligned}|\nu(M_6)| = {} & [0.58642578, 0.58691406] \cup [0.58740234, 0.58789062] \cup [0.58837891, 0.58886719] \cup \\ & [0.58935547, 0.58984375] \cup [0.59033203, 0.59082031] \cup [0.59130859, 0.75683594] \cup \\ & [0.75732422, 0.75781250] \cup [0.75830078, 0.75878906],\end{aligned}$$

and

$$|\nu(M_7)| = [0.00000000, 0.07812500].$$

Connecting orbits. Consider again the logistic map $f(x) = 3.5x(1-x)$, $X = [0,1]$, $N = 8$, $B = 15$, and $\delta = 0.05$. Using pointwise equal confidence intervals, we obtain the Morse graph in Figure 5. The Conley indices for M_1 and M_0 indicated the existence of period-2 and period-4 orbits, respectively, with at least 95% confidence. The Conley index also indicates the existence of a fixed point for M_3 and a connecting orbit from M_3 to M_1.

The Conley indices for the period-4 orbit and connecting orbit example are presented in Figure 5. The intervals defining the Morse sets are

$$\begin{aligned}|\nu(M_0)| = {} & [0.38165283, 0.38403320] \cup [0.49774170, 0.50427246] \cup \\ & [0.82553101, 0.82818604] \cup [0.87463379, 0.87533569],\end{aligned}$$

$$\begin{aligned}|\nu(M_1)| = {} & [0.42437744, 0.42440796] \cup [0.42443848, 0.42446899] \cup [0.42449951, 0.42453003] \cup \\ & [0.42456055, 0.42459106] \cup [0.42462158, 0.42465210] \cup [0.42468262, 0.42471313] \cup \\ & [0.42474365, 0.42477417] \cup [0.42480469, 0.43246460] \cup [0.43249512, 0.43252563] \cup \\ & [0.43255615, 0.43258667] \cup [0.43261719, 0.43264771] \cup [0.43267822, 0.43270874] \cup \\ & [0.43273926, 0.43276978] \cup [0.43280029, 0.43283081] \cup [0.43286133, 0.43289185] \cup \\ & [0.43292236, 0.43295288] \cup [0.85528564, 0.85531616] \cup [0.85534668, 0.85885620] \cup \\ & [0.85888672, 0.85891724],\end{aligned}$$

$$|\nu(M_2)| = [0.71362305, 0.71365356] \cup [0.71490479, 0.71493530],$$

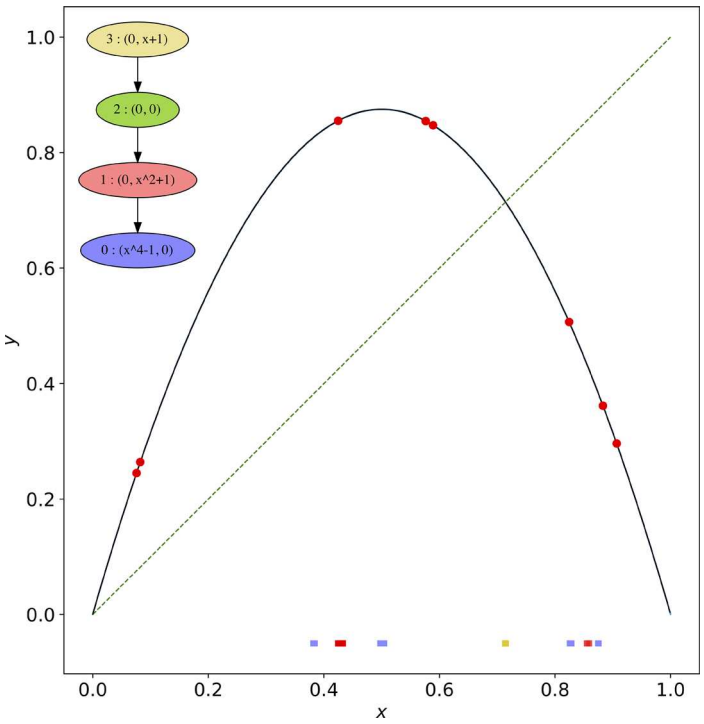

Figure 5. *In the subfigure, elements of $\mathcal{T}$ are indicated in red, the mean function μ is shown in black, and G is shown in blue, where the width of G is approximately similar to the graphical representation of μ. The Morse graphs are indicated on the figure, and the corresponding (color coded) regions of phase space $\bar{\nu}(\cdot)$ are indicated at the bottom of the figure. The region G is composed of squares of width $2^{-15}\,\mathrm{length}(X)$. The Conley indices for $\mathcal{M}_1$ and $\mathcal{M}_0$ indicated the existence of period-2 and period-4 orbits, respectively, with 95% confidence.*

and

$$|\nu(M_3)| = [0.71368408, 0.71371460] \cup [0.71374512, 0.71377563] \cup [0.71380615, 0.71475220] \cup [0.71478271, 0.71481323] \cup [0.71484375, 0.71487427].$$

The direct sum of the Conley indices of $\nu(\mathcal{M}_3)$ and $\nu(\mathcal{M}_1)$,

$$\mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_3)), \downarrow(\overleftarrow{\nu}(\mathcal{M}_3)); \mathbb{F}\right) \oplus \mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_1)), \downarrow(\overleftarrow{\nu}(\mathcal{M}_1)); \mathbb{F}\right),$$

can be represented by

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

The Conley index of the connecting orbit set

$$\mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_3)), \downarrow(\overleftarrow{\nu}(\mathcal{M}_1)); \mathbb{F}\right)$$

can be represented by

$$\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0
\end{bmatrix}.$$

From these computations,

$$\begin{aligned}
&\mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_3)), \downarrow(\overleftarrow{\nu}(\mathcal{M}_1)); \mathbb{F}\right) \\
&\qquad \not\cong \mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_3)), \downarrow(\overleftarrow{\nu}(\mathcal{M}_3)); \mathbb{F}\right) \oplus \mathrm{Con}_*\left(\downarrow(\nu(\mathcal{M}_1)), \downarrow(\overleftarrow{\nu}(\mathcal{M}_1)); \mathbb{F}\right),
\end{aligned}$$

which indicates the existence of a connecting orbit from the Morse set in $|\nu(\mathcal{M}_3)|$ to the Morse set in $|\nu(\mathcal{M}_1)|$.

Chaotic dynamics. Consider the function $f(x) = 2e^{-5(x-1)^2}$, $X = [-0.2, 2.3]$, $N = 10$, $B = 10$, and $\delta = 0.05$. Using pointwise equal confidence intervals, as indicated in Figure 1(b), we obtain a Morse graph with five nodes. While the Conley index of $\mathcal{M}(4)$ is trivial, it consists of multiple disjoint intervals, and the index map $\mathcal{F}_* : H_*(\nu(\mathcal{M}(4)), \overleftarrow{\nu}(\mathcal{M}(4)); \mathbb{Z}_5) \to H_*(\nu(\mathcal{M}(4)), \overleftarrow{\nu}(\mathcal{M}(4)); \mathbb{Z}_5)$ can be used to capture the chaotic dynamics (cf. [6] and references therein).

The Conley indices for the chaotic dynamics example are presented in Figure 1 (b). In this case the Conley index of $\nu(\mathcal{M}(4))$ is trivial; however, the index map

$$\mathcal{F}_1 : H_1\left(\nu(\mathcal{M}(4)), \overleftarrow{\nu}(\mathcal{M}(4)); \mathbb{Z}_5\right) \to H_1(\nu(\mathcal{M}(4)), \overleftarrow{\nu}(\mathcal{M}(4)); \mathbb{Z}_5)$$

represented by

$$\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1
\end{bmatrix}$$

indicates the existence of chaotic dynamics (see [6] and references therein).

The intervals defining the Morse sets are

$$\begin{aligned}|\nu(M_0)| &= [-0.01933594, 0.10517578],\\ |\nu(M_1)| &= [0.10517578, 0.10761719],\\ |\nu(M_2)| &= [0.11005859, 0.11250000],\\ |\nu(M_3)| &= [0.43232422, 0.43476562],\end{aligned}$$

and

$$\begin{aligned}|\nu(M_4)| = {} & [0.43720703, 0.43964844] \cup [0.44208984, 0.47871094] \cup [0.48115234, 0.50312500] \cup \\ & [0.52509766, 0.52753906] \cup [0.52998047, 0.57148437] \cup [0.67158203, 0.67402344] \cup \\ & [0.67646484, 0.71796875] \cup [0.72041016, 0.72285156] \cup [0.73017578, 0.73261719] \cup \\ & [0.73505859, 0.78144531] \cup [0.78388672, 0.78632812] \cup [1.19404297, 1.19648437] \cup \\ & [1.19892578, 1.33320312] \cup [1.42353516, 1.42597656] \cup [1.42841797, 1.46992187] \cup \\ & [1.47236328, 1.47480469] \cup [1.49189453, 1.49433594] \cup [1.49677734, 1.51875000] \cup \\ & [1.52119141, 1.55781250] \cup [1.56025391, 1.56269531].\end{aligned}$$

6. Construction of G for examples in section 5. For the sake of clarity, the discussions of Conley theory in section 3 and the probabilistic bounds in section 4 avoided options that can improve computational efficacy. Turning to the details of the computations for the results reported in section 5, we take explicit advantage of some of these options, and we exploit the fact that we are working with one-dimensional dynamics to pictorially explain the computational enhancements. As a starting point, see Figure 6 where the dashed curve indicates μ.

Throughout this discussion, $X = [\alpha, \beta] \subset \mathbb{R}$. To avoid cumbersome notation we identify the cellular complex $\mathcal{X}$ with a uniform discretization of $[\alpha, \beta]$ into 2^B subintervals. We denote the vertices of $\mathcal{X}$ by $\mathcal{X}^{(0)} = \{v_i \mid i = 0, \ldots, 2^B\}$ and the edges by

$$\mathcal{X}^{(1)} = \{e_i = [v_i, v_{i+1}] \mid i = 0, \ldots, 2^B - 1\}.$$

The set of midpoints of the odd intervals is denoted by

$$S = \{m_{2i+1} = (v_{2i+1} + v_{2i+2})/2 \mid i = 0, \ldots, 2^{B-1} - 1\}.$$

The length of each interval is $\epsilon := 2^{-B}(\beta - \alpha)$. The cubical grid on a subportion of $[\alpha, \beta] \times [\alpha, \beta]$ is indicated via the dotted lines in Figure 6.

Fix $\delta \in (0, 1)$. As discussed in section 2, choose $r : S \to (0, \infty)$ such that (2.2) is satisfied for all $m_i \in S$. For notational convenience we set

$$\underline{w}_{2i+1} := \mu(m_{2i+1}) - r(m_{2i+1})\sigma(m_{2i+1}) \quad \text{and} \quad \overline{w}_{2i+1} := \mu(m_{2i+1}) + r(m_{2i+1})\sigma(m_{2i+1}).$$

We construct a multivalued map $\mathcal{F} : \mathcal{X}^{\text{top}} \rightrightarrows \mathcal{X}^{\text{top}}$ in two steps. Note that $\mathcal{X}^{\text{top}} = \mathcal{X}^{(1)}$ in the settings of this section.

Step 1. Let $\mathcal{X}^{(1)}_{\text{odd}} = \{e_{2i+1} \mid i = 0, \ldots, 2^{B-1} - 1\} \subset \mathcal{X}^{(1)}$. For each $e_{2i+1} \in \mathcal{X}^{(1)}_{\text{odd}}$ define

$$\mathcal{F}(e_{2i+1}) = \left\{e \in \mathcal{X}^{(1)} \mid e \cap [\underline{w}_{2i+1}, \overline{w}_{2i+1}] \neq \emptyset\right\}.$$

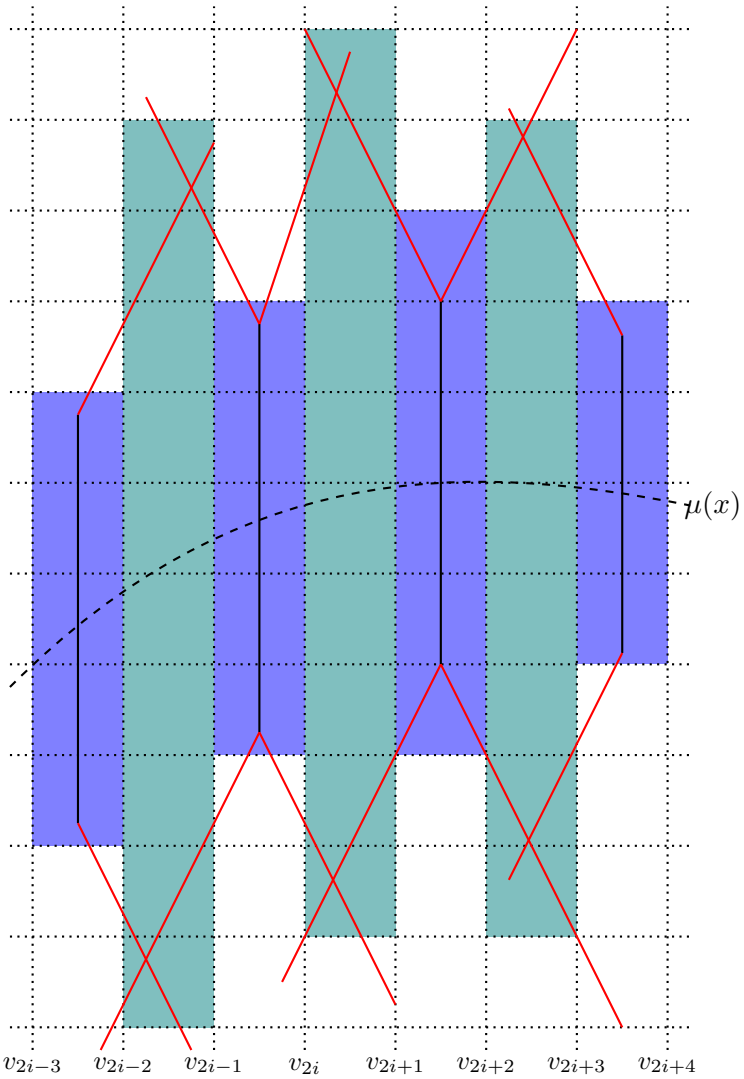

Figure 6. *Construction of $\mathcal{F}$. The dashed line represents $\mu(x)$ derived from the data $\mathcal{T}$. The dotted lines indicate the cell complex $\mathcal{X} \times \mathcal{X}$. The solid dark lines indicate the confidence intervals at the midpoints for the odd numbered intervals. The blue regions indicate the value of $\mathcal{F}$ on odd numbered edges. The red rays have slope $\pm L$ and represent high probability Lipschitz bounds on sample paths. The teal regions indicate the value of $\mathcal{F}$ on even numbered edges.*

In Figure 6 the black lines are used to indicate $[\underline{w}_{2i+1}, \overline{w}_{2i+1}]$, and the light blue shaded regions designate $\mathcal{F}(e_{2i+1})$.

Step 2. Observe that we are guaranteed with probability at least $1-\delta$ that $g(m_{2i+1})$ will lie in the black lines and hence blue regions of Figure 6. To gain control of $\mathsf{G}(g)$ over $X \setminus S$, make use of the bound given by (2.3). As stated in section 5, we assume that L is large enough so that the probability given in (2.3) is at least $(1-\delta)^{1/2}$. We also assume that L is large enough so that the rays defined below intersect. For each point m_{2i+1} we consider the four rays

$$\overline{h}^+_{2i+1}(s) = (m_{2i+1}, \overline{w}_{2i+1}) + s(1, L) \quad \text{and} \quad \overline{h}^-_{2i+1}(s) = (m_{2i+1}, \overline{w}_{2i+1}) + s(-1, L) \tag{6.1}$$

and

$$\underline{h}^+_{2i+1}(s) = \left(m_{2i+1}, \underline{w}_{2i+1}\right) + s(1, -L) \quad \text{and} \quad \underline{h}^-_{2i+1}(s) = \left(m_{2i+1}, \underline{w}_{2i+1}\right) + s(-1, -L) \tag{6.2}$$

parameterized by $s \geq 0$. These are shown as red lines in Figure 6.

Observe that the rays $\overline{h}^+_{2i-1}$ and $\overline{h}^-_{2i+1}$ intersect at

$$\begin{aligned}\overline{h}^+_{2i-1}\left(\epsilon + \frac{\overline{w}_{2i+1} - \overline{w}_{2i-1}}{2L}\right) &= \overline{h}^-_{2i+1}\left(\epsilon - \frac{\overline{w}_{2i+1} - \overline{w}_{2i-1}}{2L}\right) \\ &= \left(m_{2i-1} + \epsilon + \frac{\overline{w}_{2i+1} - \overline{w}_{2i-1}}{2L}, \overline{w}_{2i-1} + \epsilon L + \frac{\overline{w}_{2i+1} - \overline{w}_{2i-1}}{2}\right),\end{aligned}$$

and the rays $\underline{h}^+_{2i-1}$ and $\underline{h}^-_{2i+1}$ intersect at

$$\begin{aligned}\underline{h}^+_{2i-1}\left(\epsilon - \frac{\underline{w}_{2i+1} - \underline{w}_{2i-1}}{2L}\right) &= \underline{h}^-_{2i+1}\left(\epsilon + \frac{\underline{w}_{2i+1} - \underline{w}_{2i-1}}{2L}\right) \\ &= \left(m_{2i-1} + \epsilon - \frac{\underline{w}_{2i+1} - \underline{w}_{2i-1}}{2L}, \underline{w}_{2i-1} - \epsilon L + \frac{\underline{w}_{2i+1} - \underline{w}_{2i-1}}{2}\right).\end{aligned}$$

For $i > 0$ define

$$Q(e_{2i}) = \left[\underline{w}_{2i-1} - \epsilon L + \frac{\underline{w}_{2i+1} - \underline{w}_{2i-1}}{2}, \overline{w}_{2i-1} + \epsilon L + \frac{\overline{w}_{2i+1} - \overline{w}_{2i-1}}{2}\right]$$

and

$$\mathcal{F}(e_{2i}) = \left\{e \in \mathcal{X}^{(1)} \mid e \cap Q(e_{2i}) \neq \emptyset\right\}.$$

For $i = 0$ define

$$Q(e_0) = \left[\underline{w}_1 - \frac{3}{2}\epsilon L, \overline{w}_1 + \frac{3}{2}\epsilon L\right]$$

and

$$\mathcal{F}(e_0) = \left\{e \in \mathcal{X}^{(1)} \mid e \cap Q(e_0) \neq \emptyset\right\},$$

where $\underline{w}_1 - 3/2\epsilon L$ and $\overline{w}_1 + 3/2\epsilon L$ are the second components of the intersections of the line $x = v_0$ with the rays $\underline{h}^-_1(s)$ and $\overline{h}^-_1(s)$, respectively. The teal regions in Figure 6 indicate $\mathcal{F}(e_{2i})$.

Remark. As a consequence of Steps 1 and 2 we have defined the acyclic multivalued map $\mathcal{F} : \mathcal{X}^{\text{top}} \rightrightarrows \mathcal{X}^{\text{top}}$ that is used to identify the Morse graphs and the lattices of attractors and to compute Conley indices.

In the spirit of section 3, define

$$G = \bigcup_{e \in \mathcal{X}^{(1)}} e \times \mathcal{F}(e) \subset X \times X$$

and

$$\tilde{G} = G \cup \left(\bigcup_{e \in \mathcal{X}^{(1)}} e \times Q(e) \right) \subset X \times \mathbb{R},$$

where $Q(e)$ is defined in Step 2 for $e \in \mathcal{X}^{(1)}_{\text{even}}$, and for $e_{2i+1} \in \mathcal{X}^{(1)}_{\text{odd}}$ we define

$$Q(e_{2i+1}) := \left[\underline{w}_{2i+1} - \frac{1}{2}\epsilon L, \overline{w}_{2i+1} + \frac{1}{2}\epsilon L \right],$$

where $\underline{w}_{2i+1} - \frac{1}{2}\epsilon L$ and $\overline{w}_{2i+1} + \frac{1}{2}\epsilon L$ are the second components of the intersections of the rays with the lines $x = v_{2i+1}$ and $x = v_{2i+1}$, respectively.

Observe that by (2.2) and (2.3) and an argument analogous to the proof of Lemma 4.3,

$$\mathbb{P}(\mathsf{G}(g) \subset \tilde{G}) > 1 - \delta.$$

As can be seen from Figure 6, $\mathcal{F}$ is not an outer approximation for every g such that $\mathsf{G}(g) \subset \tilde{G}$ (this could be achieved by enlarging the images of $\mathcal{F}(e_{2i+1})$ but at the risk of losing information about the structure of the dynamics). However, we note that if $(x, y) \in \tilde{G}$, then (x, y) is within distance $\epsilon/2$ from G. Therefore, we can apply the results of [2, section 5] to conclude that the Conley index implications about the dynamics computed using $\mathcal{F}$ are valid for a sample path $h : X \to X$ if $\mathsf{G}(h) \subset \tilde{G}$.

7. Concluding remarks.

As is indicated in section 5, our framework is capable of identifying the fundamental building blocks of traditional nonlinear dynamical systems with high levels of confidence based on a few data points. The most obvious criticism is that we restricted our examples to one-dimensional dynamics. This was for the sake of clarity; the results described in sections 2–4 are dimension independent. Computations of the type described in section 3 using cubical complexes can be done routinely in systems of four dimensions or less [24]. They have also been used for the rigorous analysis of infinite-dimensional systems [7], indicating that, at least conceptually, it is the intrinsic, as opposed to extrinsic dimension, of the dynamics that determines computability.

There are a variety of closely related open problems that arise from our approach. The geometry of isolating blocks in one dimension is reasonably simple: a finite collection of closed intervals. In higher dimensions the geometry can be much more complicated, which raises the question of estimates relating the dynamics, the number of data points, and uncertainty bounds. Even heuristics for optimal sampling methods to identify attractor block lattices is not obvious.

Data availability.

The code to perform the computations and generate the figures is available at https://github.com/marciogameiro/GP_MorseGraph.

Acknowledgment.

The authors thank Cameron Thieme for helpful discussions.

REFERENCES

[1] E. Alvarado, B. Krishnamoorthy, and K. R. Vixie, *Geometry of a Set and Its Random Covers*, preprint, https://arxiv.org/pdf/2112.14979.pdf, 2021.

[2] B. Batko, K. Mischaikow, M. Mrozek, and M. Przybylski, *Conley index approach to sampled dynamics*, SIAM J. Appl. Dyn. Syst., 19 (2020), pp. 665–704, https://doi.org/10.1137/19M1254404.

[3] J. Bush, W. Cowan, S. Harker, and K. Mischaikow, *Conley–Morse databases for the angular dynamics of Newton's method on the plane*, SIAM J. Appl. Dyn. Syst., 15 (2016), pp. 736–766, https://doi.org/10.1137/15M1017971.

[4] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed., MIT Press, Cambridge, MA, 2009.

[5] N. A. Cressie, *Statistics for Spatial Data*, Wiley, New York, 1993.

[6] S. Day and R. Frongillo, *Sofic shifts via Conley index theory: Computing lower bounds on recurrent dynamics for maps*, SIAM J. Appl. Dyn. Syst., 18 (2019), pp. 1610–1642, https://doi.org/10.1137/18M1192007.

[7] S. Day, O. Junge, and K. Mischaikow, *A rigorous numerical method for the global analysis of infinite-dimensional discrete dynamical systems*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 117–160, https://doi.org/10.1137/030600210.

[8] N. De Freitas, A. J. Smola, and M. Zoghi, *Exponential regret bounds for Gaussian process bandits with deterministic observations*, in Proceedings of the 29th International Conference on Machine Learning, ICML'12, Omnipress, Madison, WI, 2012, pp. 955–962.

[9] M. Foreman, D. J. Rudolph, and B. Weiss, *The conjugacy problem in ergodic theory*, Ann. of Math. (2), 173 (2011), pp. 1529–1586, https://doi.org/10.4007/annals.2011.173.3.7.

[10] R. Furrer and M. G. Genton, *Aggregation-cokriging for highly multivariate spatial data*, Biometrika, 98 (2011), pp. 615–631.

[11] M. Gameiro and S. Harker, *CMGDB: Conley Morse Graph Database*, 2022, https://github.com/marciogameiro/CMGDB.

[12] S. Ghosal and A. Roy, *Posterior consistency of Gaussian process prior for nonparametric binary regression*, Ann. Statist., 34 (2006), pp. 2413–4929.

[13] R. Gramacy, *Surrogates: Gaussian Process Modeling, Design and Optimization for the Applied Sciences*, Chapman & Hall/CRC, Boca Raton, FL, 2020.

[14] S. Harker, H. Kokubu, K. Mischaikow, and P. Pilarczyk, *Inducing a map on homology from a correspondence*, Proc. Amer. Math. Soc., 144 (2016), pp. 1787–1801, https://doi.org/10.1090/proc/12812.

[15] A. Hatcher, *Algebraic Topology*, Cambridge University Press, Cambridge, 2002.

[16] W. D. Kalies, K. Mischaikow, and R. Vandervorst, *Lattice structures for attractors* II, Found. Comput. Math., 1 (2015), pp. 1–41.

[17] W. D. Kalies, K. Mischaikow, and R. Vandervorst, *Lattice structures for attractors* III, J. Dynam. Differential Equations, 34 (2022), pp. 1729–1768, https://doi.org/10.1007/s10884-021-10056-8.

[18] L. Le Gratiet and C. Cannamela, *Cokriging-based sequential design strategies using fast cross-validation techniques for multi-fidelity computer codes*, Technometrics, 57 (2015), pp. 418–427.

[19] M. Lee and A. Owen, *Single nugget kriging*, Statist. Sinica, 28 (2018), pp. 649–669.

[20] S. Lefschetz, *Algebraic Topology*, Colloq. Publ. 27, American Mathematical Soc., 1942.

[21] K. Mischaikow and M. Mrozek, *Conley index*, in Handbook of Dynamical Systems, Vol. 2, B. Fiedler, ed., North-Holland, Amsterdam, 2002, pp. 393–460, https://doi.org/10.1016/S1874-575X(02)80030-3.

[22] T. J. Santner, B. J. Williams, and W. I. Notz, *The Design and Analysis of Computer Experiments*, 2nd ed., Springer, New York, 2018.

[23] M. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*, Springer Ser. Statist., Springer, New York, 1999.

[24] E. Vieira, E. Granados, A. Sivaramakrishnan, M. Gameiro, K. Mischaikow, and K. E. Bekris, *Morse graphs: Topological tools for analyzing the global dynamics of robot controllers*, in Algorithmic Foundations of Robotics XV, WAFR 2022, Springer Proc. Adv. Robotics (SPAR) 25, S. M. LaValle, et al., eds., Springer, Cham, 2022, pp. 436–453.

[25] E. Vieira, A. Sivaramakrishnan, Y. Song, E. Granados, M. Gameiro, K. Mischaikow, Y. Hung, and K. E. Bekris, *Data-efficient characterization of the global dynamics of robot controllers with confidence guarantees*, in Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA), London, 2023, pp. 3065–3072.