

Delay embedding in the presence of dynamical noise

Mark R. Muldoon , David S. Broomhead , Jeremy P. Huke & Rainer Hegger

To cite this article: Mark R. Muldoon , David S. Broomhead , Jeremy P. Huke & Rainer Hegger (1998) Delay embedding in the presence of dynamical noise, Dynamics and Stability of Systems, 13:2, 175-186, DOI: [10.1080/02681119808806259](https://doi.org/10.1080/02681119808806259)

To link to this article: <https://doi.org/10.1080/02681119808806259>



Published online: 21 Mar 2007.



Submit your article to this journal [↗](#)



Article views: 78



Citing articles: 18 View citing articles [↗](#)

Delay embedding in the presence of dynamical noise

Mark R. Muldoon¹, David S. Broomhead¹, Jeremy P. Huke¹ and Rainer Hegger²

¹Mathematics Department, UMIST, PO Box 88, Manchester M60 1QD, UK

²Max-Planck-Institut für Physik komplexer Systeme, Nöthnitzer Strasse 38, D-01187 Dresden, Germany

(Received November 1997, final version February 1998)

Abstract. We present a new embedding theorem for time series, in the spirit of Takens's theorem, but requiring multivariate signals. Our result is part of a growing body of work that extends the domain of geometric time series analysis to some genuinely stochastic systems—including such natural examples as $\mathbf{x}_{j+1} = \phi(\mathbf{x}_j) + \boldsymbol{\eta}_j$ where ϕ is some fixed map and the $\boldsymbol{\eta}_j$ are i.i.d. random displacements.

1 Introduction

A natural development of efforts to apply dynamical systems theory to physical systems was a set of techniques—first proposed in the early 1980s—for relating time series of experimental measurements to invariants of a putative underlying dynamical system. In the most popular of these methods, now known as the method of delays, one assembles d -tuples of successive values from a univariate time series, forming a series of points in \mathbb{R}^d . One then regards these points in two ways: as a geometric object with a dimension and topological invariants; and as the orbit of a dynamical system, having, for example, Lyapunov exponents.

Takens's embedding theorem¹ (Takens, 1983) shows that both these views are correct; the series of delay embedded points is related to the original dynamical system by what amounts to a nonlinear change of coordinates. Takens's result is often cited as justification for using the method of delays to do geometric time-series analysis, but the original formulation neglected some important dynamical and signal processing issues, most notably the presence of both 'additive' noise, which corrupts the signals one measures from a dynamical system, but does not affect the dynamics, and 'multiplicative' or 'dynamical' noise, which does effect the dynamics (and so, indirectly, the signals). Although the only rigorous results about embedding in the presence of noise are fairly recent (see Stark *et al.* (1997) for a review), there is a great deal of practical experience with additive noise on the measured signal (see, for example, Broomhead *et al.*, 1996; Casdagli *et al.*, 1991; Farmer & Sidorowich, 1988; Kantz *et al.*, 1993; Kostelich & Yorke, 1990; Stark & Arumugam, 1992) suggesting that at least some kinds may be overcome.

Dynamical noise has only recently begun to receive attention: again, see Stark *et al.* (1997) for a careful discussion of the appropriate framework and a review of recent results. Here, we treat some illustrative examples and formulate a new embedding theorem suited to systems with intrinsically noisy dynamics. We begin by exploring a related problem, that of embedding signals from what Michael Barnsley (Barnsley, 1988; Barnsley & Demko, 1985) has called an iterated function system (IFS). This example, which is a mild form of stochasticity, will both set the stage for our main result and, in the limit of a large IFS, illustrate the ways in which dynamical noise upsets the standard method of delays as applied to univariate signals. We then frame a new result which, by considering multivariate signals, overcomes these problems and offers the prospect of geometric time series analysis for truly stochastic systems.

2 Embedding an IFS

Throughout the paper, we will consider dynamical systems of the form

$$x_{j+1} = \phi_j(x_j) \quad (1)$$

where the x_j are points in some m -dimensional manifold \mathcal{M} and where the dynamics are given by a sequence of maps $\phi_j: \mathcal{M} \rightarrow \mathcal{M}; j \in \mathbb{Z}$. For the moment, we will restrict ourselves to the case of IFSs, so we will imagine that the ϕ_j come from some finite set, $\{f_1, \dots, f_N\}$, and that at each step of the dynamics, one of the f_k is chosen independently with fixed probabilities, $\{p_1, \dots, p_N\}$. Our time series will be the values of some real-valued measurement function $v: \mathcal{M} \rightarrow \mathbb{R}$ recorded along an orbit: $v_j = v(x_j)$.

A complete geometric analysis of such a time series should recover not only the underlying manifold, but also the particular sequence of maps ϕ_j that governed the temporal evolution. Fortunately, the standard method of delays works with only minor modifications: recent work of Broomhead *et al.* (1997) shows that if one constructs d -dimensional delay vectors

$$\mathbf{v}_j = (v_j, v_{j+1}, \dots, v_{j+d-1}) \quad (2)$$

in the usual way, then, provided $d > 2m$, one expects generically to find that the embedded points lie on a finite collection of diffeomorphic images of \mathcal{M} .

To see why, think of the embedding procedure (2) as a map $\Phi_{j,v}$ from \mathcal{M} to \mathbb{R}^d :

$$\begin{aligned} \Phi_{j,v}(x_j) &\equiv \mathbf{v}_j \\ &= (v(x_j), v(x_{j+1}), \dots, v(x_{j+d-1})) \\ &= (v(x), v(\phi_j(x)), v(\phi_{j+1}(\phi_j(x))) \dots) \end{aligned} \quad (3)$$

Clearly, this map depends on the particular sequence of ϕ_j that produced the orbit, but the dependence is mild; $\Phi_{j,v}$ involves only $(d-1)$ successive steps in the dynamics and so, as each step involves only one of the IFS's N maps, there are only $N^{(d-1)}$ distinct possibilities for $\Phi_{j,v}$. In other words, each embedded point \mathbf{v}_j lies in one of finitely many images of \mathcal{M} . An argument similar to that in Takens's original paper shows that each of the delay-embedding maps is generically a diffeomorphism, so providing their images do not overlap—and Broomhead *et al.* (1997) show that generically they do not—then the usual method of delays solves the problem of geometric analysis for time series arising from an IFS.

2.1 Limitations of the IFS model

It is instructive to try to push this idea further: What happens as one increases the number of maps N ? All $N^{(d-1)}$ images of \mathcal{M} must fit into a hypercube

$$[\inf_{x \in \mathcal{M}} v(x), \sup_{x \in \mathcal{M}} v(x)]^d$$

and thus must eventually become so crowded that one could not, in practice, separate them. But an even more drastic problem occurs if one lets $N \rightarrow \infty$ in such a way that one replaces the finite collection of maps in the IFS with a continuously indexed family.

For suppose we wished to enumerate the embedding maps in (3); we could label them with $(d-1)$ -tuples of integers, so that $(i_1, i_2, \dots, i_{d-1})$ would correspond to the map $\Phi_{i_1, \dots, i_{d-1}, v}: \mathcal{M} \rightarrow \mathbb{R}^d$:

$$\Phi_{i_1, \dots, i_{d-1}, v}(x) = (v(x), v(f_{i_1}(x)), v(f_{i_2}(f_{i_1}(x))) \dots)$$

If we now replace the IFS's finite index set with parameters from, say, some n -dimensional manifold \mathcal{N} , then a labelling scheme like the one earlier would require $(d-1)$ -tuples of points in \mathcal{N} . That is, the embedding maps would be indexed by points in $\mathcal{N}^{(d-1)}$. The most we could hope for would be that the union of all the images of \mathcal{M} under all the embedding maps would form a diffeomorphic image of

$$\mathcal{M} \times \underbrace{\mathcal{N} \times \dots \times \mathcal{N}}_{d-1} \tag{4}$$

but this is a manifold of dimension $(m + (d-1)n)$ and so, barring a non-generically fortuitous choice of coordinates, cannot generally be embedded in a space of dimension lower than $2(m + (d-1)n) + 1$. Indeed, if $n > 0$, no choice of embedding dimension d could ever be large enough to embed the product properly in (4).

The usual method of delays thus fails when the underlying dynamical system is anything more random than an IFS. The problem is that the standard approach relies on composition with the dynamics to produce independent coordinate functions on \mathcal{M} . In the stochastic case the dynamics vary from step to step, so that even in the favourable case we have just considered, where the stochastic dynamics are characterized by a finite-dimensional parameter space, adding extra delays does not help unless the parameter space is zero-dimensional (the IFS case). In the next section, we show that an alternative approach, getting more information about \mathcal{M} by taking extra measurements, can resolve this problem.

3 Embedding noisy dynamical systems

The problematic examples at the end of the last section are stochastic dynamical systems on a product manifold whose temporal evolution is governed by F : $\mathcal{M} \times \mathcal{N} \rightarrow \mathcal{M}$:

$$(x_{j+1}, y_{j+1}) = (F(x_j, y_j), \eta_j) \tag{5}$$

Here, the $x_j \in \mathcal{M}$ are points in our system's state-space and one can think of the $y_j \in \mathcal{N}$ as points in an n -dimensional manifold of parameters which determine the action of the dynamics F . At each step, one acts with the map determined by the current parameters, then chooses the next set randomly: the $\eta_j \in \mathcal{N}$ are a sequence of points from \mathcal{N} , chosen independently and distributed identically. The rest of

the paper is devoted to this sort of example, which might be termed ‘finitely parameterized stochasticity’.

Example 1. Our prototype example is a fixed, purely deterministic dynamical system perturbed by small, random displacements. For example, imagine that \mathcal{M} is \mathbb{R}^m ; that \mathcal{N} is $B_m(0, \varepsilon)$, the closed m -dimensional ball of radius ε centred on the origin, and that the map F in equation (5) is

$$F(x, y) = \phi_0(x) + y \quad (6)$$

Here $\phi_0: \mathbb{R}^m \rightarrow \mathbb{R}^m$ is the deterministic part of the dynamis and y is a random displacement chosen, say, uniformly from $B_m(0, \varepsilon)$. This example is a kind of skew product and, bearing it in mind, we will sometimes refer to \mathcal{M} as the base space and \mathcal{N} as the noise manifold. The role of measurements will be played by real-valued functions that depend only on the base space.

3.1 A theorem

A typical realization of a finitely parameterized stochastic process consists of a sequence of i.i.d. parameter values $y_j \in \mathcal{N}$, $j \in \mathbb{Z}$ and the orbit of some point $x_0 \in \mathcal{M}$ under the corresponding dynamics

$$x_{n+1} = F(x_n, y_n)$$

The natural phase-space of such a system is the product manifold $\mathcal{M} \times \mathcal{N}$ and a successful geometric time-series analysis of such a system should be able to reconstruct the product and, ideally, preserve its product structure, so that in problems like (5) one can in some sense tease apart the i.i.d. selection of parameters from the dynamics they induce.

Suppose, for example, that the map ϕ_0 in (6) had a uniformly hyperbolic attracting set. Stochastic perturbation would destroy this attractor, replacing it with, at the simplest, a probability distribution concentrated in a small neighbourhood of the unperturbed system’s attractor. But the dynamics (6) are not structureless: F has two interesting restrictions. In the first, one fixes a set of parameters $y \in \mathcal{N}$ and obtains a diffeomorphism $F(\cdot, y)$ of \mathcal{M} to itself; in the second, one fixes a point $x \in \mathcal{M}$ and considers the image produced by $F(x, \cdot): \mathcal{N} \rightarrow \mathcal{M}$ as the second argument varies over the manifold of parameters. The point of the following theorem is to determine conditions under which this second restriction is experimentally accessible.

The idea is to use a pair of consecutive, multivariate observations to pin down both parts of the Cartesian pair (x_n, y_n) . Where Takens’s original theorem used composition of the dynamics with the measurement function to produce a suite of coordinate functions, here we imagine that we can take so many simultaneous measurements that embedding of the x_n part is guaranteed by Whitney’s embedding theorem. We then combine this multivariate observation with its successor to get information about the y_n part of (x_n, y_n) .

Theorem 1. Let \mathcal{M} and \mathcal{N} be compact manifolds of dimension m and n respectively. For pairs (F, v) with $F: \mathcal{M} \times \mathcal{N} \rightarrow \mathcal{M}$ a C^2 map such that

H1: $F(\cdot, y): \mathcal{M} \rightarrow \mathcal{M}$ is a C^2 diffeomorphism for each $y \in \mathcal{N}$; that is, F is a family of diffeomorphisms in $C^2(\mathcal{M}, \mathcal{M})$, parameterized by $y \in \mathcal{N}$,

H2: $F(x, \cdot): \mathcal{N} \rightarrow \mathcal{M}$ is an embedding for each $x \in \mathcal{M}$,

H3: and $v: \mathcal{M} \rightarrow \mathbb{R}^d$ a C^2 function and $d > 2m$;

it is a generic property that the map $\Phi_{F,v}: \mathcal{M} \times \mathcal{N} \rightarrow \mathbb{R}^{2d}$ defined by

$$\Phi_{F,v}(x, y) = (v(x), v(F(x, y)))$$

is an embedding. Further, $\Phi_{F,v}$ preserves the product structure of $\mathcal{M} \times \mathcal{N}$ in the sense that

- (i) $\pi_1(\Phi_{F,v}(\cdot, y)): \mathcal{M} \rightarrow \mathbb{R}^d$ is an embedding for each $y \in \mathcal{N}$;
- (ii) $\pi_2(\Phi_{F,v}(x, \cdot)): \mathcal{N} \rightarrow \mathbb{R}^d$ is an embedding for each $x \in \mathcal{M}$;

where the operators π_1 and π_2 project out either the first, or last d -tuple of components in \mathbb{R}^{2d} .

Proof. This is really nothing more than a corollary of Whitney's embedding theorem (Spivak, 1979; Whitney, 1935): the main conclusion, that $\Phi_{F,v}$ is an embedding, follows from the numbered conclusions about the preservation of product structure. And the first of these is an immediate consequence of Whitney's embedding theorem and *H3* while the second follows from the first and *H2*. ■

3.2 Remarks on the hypothesis *H2*

The most difficult part of applying this theorem will be to find examples for which the hypothesis *H2*, that the dynamics $F(x, \cdot)$ embed \mathcal{N} in \mathcal{M} , can be verified. Here we discuss some topological issues related to this problem. As we will show later, it is not always possible to satisfy *H2*. Indeed, there are pairs of manifolds \mathcal{N} and \mathcal{M} for which it can never be satisfied.

But first, we consider the elementary arguments related to dimension counting. At the very least, embedding requires $n \leq m$ and could in principle be much more restrictive—the Whitney bound suggests that generically one needs $2n < m$, but this seems extreme. Here is an example where $n = m$.

Example 2. Consider a variation on the system on Example 1 where we take \mathcal{M} to be a circle and the noise manifold \mathcal{N} to be an interval of length 2ε . The dynamics (6) become the composition of a random rotation through a small angle with some fixed map $\phi_0: S^1 \rightarrow S^1$. Provided ε is small enough, *H2* will be satisfied even though $\dim \mathcal{N} = \dim \mathcal{M}$.

In general, however, there can be subtle topological obstructions, as the following example shows.

Example 3. Take $\mathcal{M} = S^2$, the two-sphere, and $\mathcal{N} = [-1, 1]$, an interval (see Fig. 1). It is impossible for this system to satisfy *H2*; that is, it is impossible for all the maps $F(x, \cdot): \mathcal{N} \rightarrow \mathcal{M}$ to be embeddings of \mathcal{N} into \mathcal{M} . To see why, fix a value of $y \in \mathcal{N}$, say $y = y_0$, and consider the derivatives $D_y F(x, y_0): T\mathcal{N}_{y_0} \rightarrow T\mathcal{M}_{F(x, y_0)}$. As \mathcal{N} is just an interval, the tangent space $T\mathcal{N}_{y_0}$ is spanned by a single tangent vector that we shall call e_{y_0} . The image of this vector under $D_y F$ is

$$D_y F(x, y_0)e_{y_0} \in T\mathcal{M}_{F(x, y_0)}$$

a vector tangent to \mathcal{M} at the point $F(x, y_0)$. By varying x in the expression above,

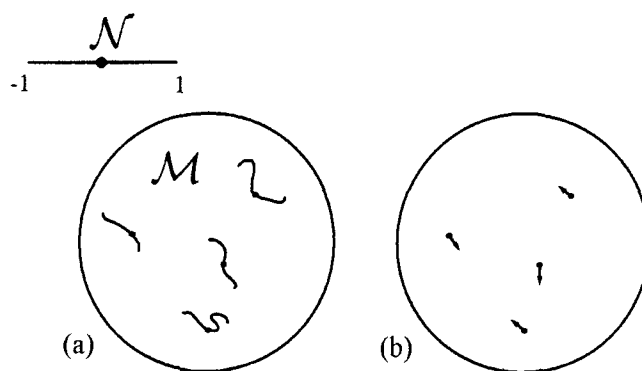


Fig. 1. The hypothesis $H2$ can never be satisfied if $\mathcal{M} = S^2$ and $\mathcal{N} = [-1, 1]$. (a) shows several images of \mathcal{N} in \mathcal{M} , while (b) shows the vector field induced by $D_y F(x, y_0)$.

we obtain such a tangent vector at every point in \mathcal{M} (here we use $H2$: the fact that $F(\cdot, y_0): \mathcal{M} \rightarrow \mathcal{M}$ is a diffeomorphism). That is to say, the images of the tangent vector e_{y_0} form a vector field on $\mathcal{M} = S^2$. But any such vector field must vanish in at least two places, hence $F(x, \cdot)$ must fail to be an immersion in at least two places and so certainly cannot be an embedding.

3.3 Numerical examples

Takens's original paper addressed a large body of already existing practice: many scientists had been using delay embedding to measure such dynamical invariants as Lyapunov exponents and, especially in the early 1980s, the dimensions of attractors (see, for example, the bibliography in Hao (1984)). As subsequent authors (Sauer *et al.*, 1991) have emphasized, failure of the delay-map $\Phi_{F,v}$ to be an embedding would not necessarily affect the dimension of the image of $\mathcal{M} \times \mathcal{N}$; indeed, a rather unlikely and spectacular failure of embedding would be required. None the less, such calculations remain popular and, in the case at hand, informative.

In Section 3.1, we said that the point of our theorem was to exploit multiple simultaneous observations to make the restriction $F(x, \cdot)$ experimentally accessible—here we discuss that notion more thoroughly. The idea is to consider delay-embedded points of the form

$$(v(x_0), v(F(x_0, y)))$$

where F and v are as in our theorem and $x_0 \in \mathcal{M}$ is some particular point in the base space. Ideally, one would examine points sharing a common value of the projection π_1 , that is, points whose first d -coordinates were the same. One could then, for example, compute the dimension and topology of \mathcal{N} using standard tools from geometric time series analysis. In practice, this would require very long time series and even then one would probably have to consider delay-embedded points whose first d -coordinates were close, perhaps inside some small ball, rather than strictly identical.

In the rest of this section, we will discuss some less demanding calculations bearing on the same issues. All the following examples are correlation dimension

calculations for variously perturbed versions of the celebrated Lorenz system (Lorenz, 1963):

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz\end{aligned}\tag{7}$$

with the canonical parameters $\sigma = 10$, $r = 28.0$ and $b = (8/3)$. These systems are not, of course, strictly within the domain of our results: the phase-space of the Lorenz system is \mathbb{R}^3 , which is not a compact manifold. The Lorenz attractor is, however, contained in a compact subset of \mathbb{R}^3 and as Huke (1993) has pointed out, there is an analogue of the standard Takens theorem suited to this case.

The framework set out in Section 3 encompasses a wide variety of finitely parameterized stochastic processes. At one extreme are those that approximate stochastic ordinary differential equations—the first two examples discussed are of this kind—where random perturbation affects the dynamics so profoundly as to require modifications to the notion of an attractor (Arnold, 1995). Those experienced with the standard Takens theorem might imagine that the addition of this kind of dynamical noise, which in one sense makes the system infinite dimensional, must preclude any sort of dimension calculations: this is certainly not the case.

At the other extreme, one might naïvely imagine that the separation produced by our two projections is so effective that the embedded system has the structure

$$\text{unperturbed attractor} \times \mathcal{N}\tag{8}$$

This is not generally the case, although it does happen in the third of our examples. Here the sampling period varies randomly, but the underlying continuous process is an unmodified deterministic flow. Similar examples arise quite naturally in the measurement of the physical system and in these cases the delay-embedded object is of the form (8).

To address the issues just discussed, we subjected the Lorenz system to three kinds of random perturbation. All three examples involve simultaneously recorded time series of the three standard Lorenz variables, x , y and z . In the spirit of the theorem, we investigated the dimension of the collection of points \vec{v}_j formed by concatenating two consecutive multivariate measurements:

$$\vec{v}_j = (x_j, y_j, z_j, x_{j+1}, y_{j+1}, z_{j+1})$$

The results are shown in Figs 2–4. Each figure shows six curves, one each for embeddings based on the first $m = 1, 2, 3, \dots, 6$ of the coordinates of 500 000 points. The curves are numerical derivatives, $d_2(m, \varepsilon)$ of the log of the m -dimensional correlation integral, $C_2(m, \varepsilon)$, with respect to $\log \varepsilon$. A scaling region appears here as a range of $\log \varepsilon$ over which the slope $D_2(m, \varepsilon)$ appears constant.

The system studied in Fig. 2 exhibits a straightforward sort of dynamical noise: we integrated the equations of motion (7) with a fourth-order Runge–Kutta scheme and a time step of $\delta t = 0.01$. Every five time steps we recorded x, y and z , then applied a random displacement to the x -coordinate. The displacements were distributed uniformly over the interval $\pm 0.1\sigma_x$, where σ_x^2 is the variance of the unperturbed Lorenz x signal. Thus, $\mathcal{M} = \mathbb{R}^3$, $\mathcal{N} = (-0.1\sigma_x, 0.1\sigma_x)$ and the unperturbed dynamics are the fifth iterate of a map $\phi_{\delta t}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that approximates the time- δt evolution operator of the Lorenz system.

The resulting embedded object is four-dimensional. That is, there is a recogniz-

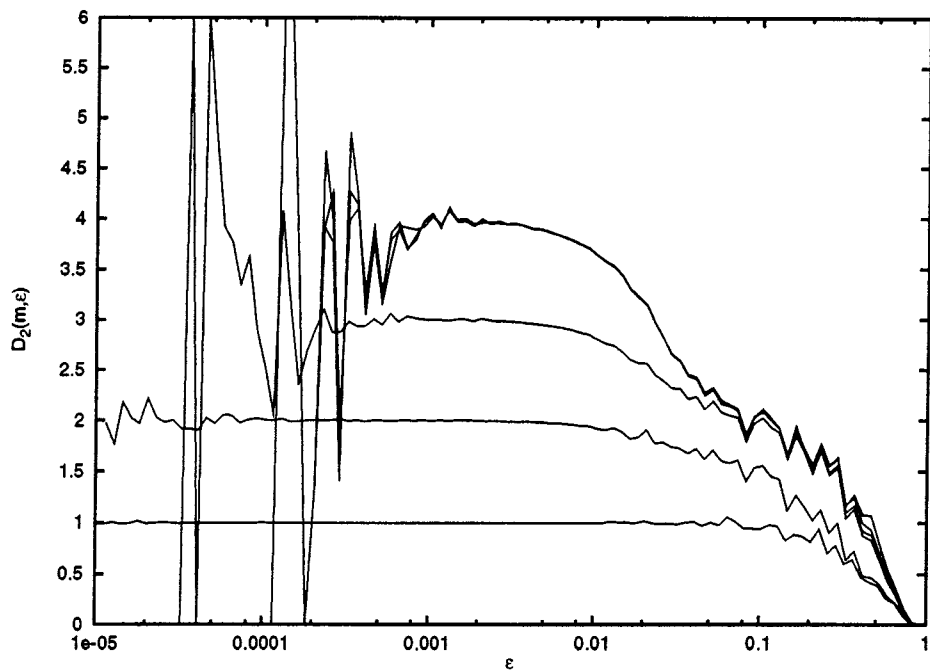


Fig. 2. Correlation dimension estimates for a Lorenz system perturbed by a one-dimensional family of random kicks that acts once per sampling period.

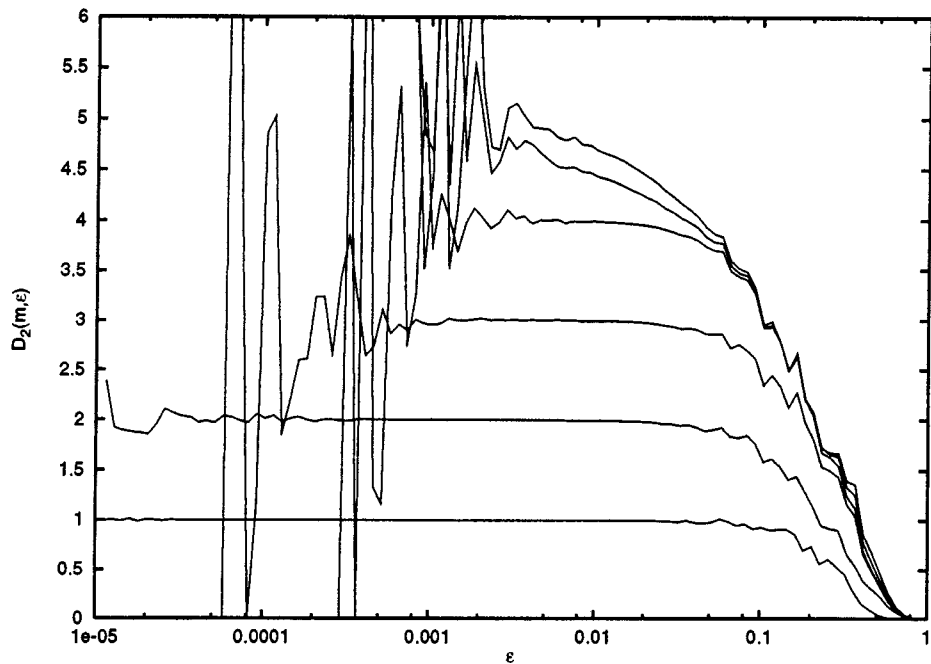


Fig. 3. Correlation dimension estimates for a Lorenz system perturbed by a one-dimensional family of random kicks that acts once per integration time-step: this system is nearer to a diffusion process than to the example in Fig. 2.

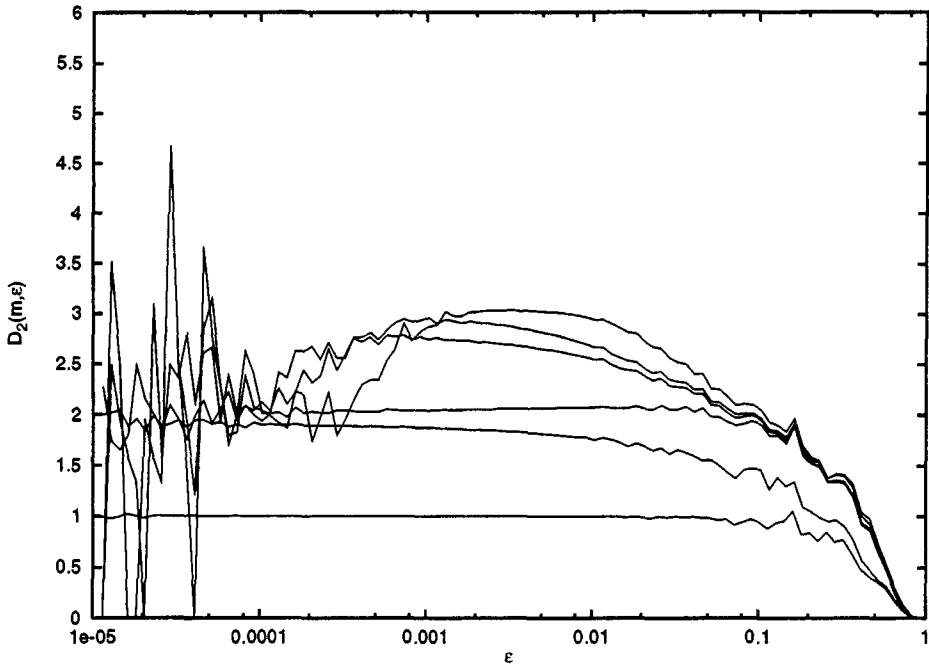


Fig. 4. Correlation dimension estimates for a Lorenz system sampled at irregular time intervals. This one-dimensional family of evolution operators shares the same attracting set, but single-step multivariate delay embedding shows that the dynamics are not the same from step to step.

able scaling region for all six curves and although the curves for dimensions 1 through 4 suggest that the data fill out those embedding spaces, the curves for dimensions 5 and 6 lie almost directly on top of that for dimension 4. Clearly the addition of noise changes the dynamics dramatically; as viewed in the system's phase-space, the noise destroys the original Lorenz attractor (which has a dimension slightly larger than 2) and replaces it with a blurred object that, at least on small length scales, is volume filling. But the distribution of the \bar{v}_j is not without structure: had we not known that the noise manifold \mathcal{N} was one-dimensional, we could have deduced it from the way in which the slopes of the $D_2(m, \varepsilon)$ jump up by 1 when we include data from the second set of measurements.

In the calculation shown in Fig. 3, we consider a perturbation that more closely approximates a classical stochastic process. The noise still affects only the x -coordinate, but now we permit it to act once per integration step rather than once per sampling interval. The noise manifold is thus higher dimensional; the random kicks act five times per sampling interval, so $\dim \mathcal{N} = 5$ and one would not expect the \bar{v}_j to lie on an object of dimension less than $3 + 5 = 8$. The $D_2(m, \varepsilon)$ bear this out: those for dimensions 1 through 4 suggest that the data fill out the embedding space. But in contrast to the previous example, the curves for dimensions 5 and 6 do not lie on top of that for dimension 4; they do not offer an estimate of the dimension of \mathcal{N} .

Finally, we examine a case where the perturbation does not change the appearance of the attractor at all (see also Fig. 5), but does affect the dynamics. The idea is to perturb the sampling rate, so that rather than recording the data at regular intervals, we choose successive sampling intervals independently from the interval

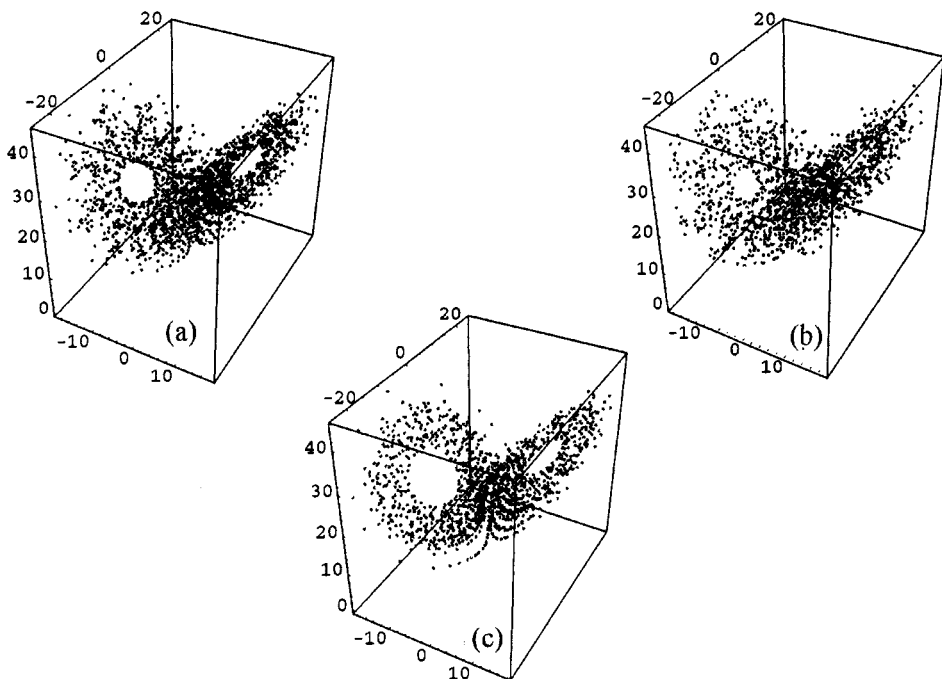


Fig. 5. Two thousand points from each of the trajectories of the stochastic processes discussed in the text: (a) and (b) correspond to the cases where the x variable received periodic kicks once (a) or five times (b) per sampling interval. In these cases the fine structure of the Lorenz attractor is blurred. In the system shown in (c), the sampling interval varies randomly, but as these variations cause displacements only along an orbit of the Lorenz flow, the sampled points all lie on the unperturbed attractor.

$(0.01, 0.05)$. Once again, \mathcal{M} is a compact subset of \mathbb{R}^3 and \mathcal{N} is an interval, but this time the attractor's structure in \mathcal{M} is preserved: the first three $D_2(m, \varepsilon)$ curves (which all depend on data gathered at the same instant in the Lorenz attractor's natural phase space) accumulate on a dimension slightly larger than 2. When one adds the data from the next measurement, the apparent dimension of the embedded object jumps to something nearer 3, but no higher, indicating that \mathcal{N} , the manifold of parameters, has dimension 1. In this case the product structure of $\mathcal{M} \times \mathcal{N}$ is especially straightforward: the embedded data lie on a kind of sheaf of Lorenz attractors, each leaf associated with a fixed value of the sampling interval and containing a complete copy of the attractor.

3.4 Extensions and applications

Our embedding theorem evades the dimension-counting problem outlined at the end of Section 2 by adding extra observations, so many extra that the measurements taken at a single instant are sufficient to embed \mathcal{M} . As a result, the product manifold $\mathcal{M} \times \mathcal{N}$ gets embedded in a space of dimension higher than is strictly necessary. The excess—always at least one—is even larger if $n < m$. This means that some of the measurements are redundant: it would be interesting to know if this redundancy has any structure that one could exploit.

More generally, one would like to know how to exploit the product structure of $\mathcal{M} \times \mathcal{N}$. In the irregularly sampled series of Fig. 4 the situation is fairly clear and,

provided one had sufficient data, one could hope to deduce the sequence of sampling intervals. A related problem arises in the analysis of signals contaminated by cross-talk in noisy, nonlinear channels. In that case, the underlying dynamical system is the internal state of the channel; the noise process is the contaminating cross-talk and our goal is to recover the contaminating signal with an eye to cancelling it. Another promising class of applications involves the processing of spatio-temporal signals. Here the kind of multi-channel observations required by Theorem 1 arise naturally.

One might, in this setting, want to use the sorts of processes discussed earlier as a more sophisticated class of ‘noise’ model. Here ‘noise’ has the sense of ‘some uninteresting signal’ as opposed to the more technical sense, ‘the stochastic component of some mixed stochastic/deterministic process’, that it has had in the rest of the paper. Imagine that some interesting signal is superimposed on a spatio-temporal noise process which satisfies the hypotheses of Theorem 1 and that it is possible to recombine the outputs of the measurement channels in such a way as to cancel the interesting signal; naturally this would also distort the noise. But provided that the recombined measurements still satisfied the hypotheses of the theorem, we could use them to construct an embedding of $\mathcal{M} \times \mathcal{N}$. This embedding would then be related invertibly (indeed, diffeomorphically) to the embedding of $\mathcal{M} \times \mathcal{N}$ that would have been constructed from the original measurements in the absence of the interesting signal. This structure is a sort of spatio-temporal analogue to that required for the signal-separation protocol of Broomhead *et al.* (1996); combining their ideas and this paper, one could recover the interesting signal cleanly.

4 Conclusions

We have introduced a new theorem for the geometric analysis of time series measured from systems perturbed by dynamical noise. It is worth emphasizing how our result differs from Takens’s original. To begin with, we have given up the most remarkable aspect of the first delay-embedding theorem: Takens showed that one can, in a rigorous sense, learn much of what there is to know about a dynamical system by studying a single univariate time series measured from it. One can calculate the dimension of attractors, estimate such dynamical invariants as the Lyapunov spectrum and even make predictive models.

By contrast, we require so many simultaneous measurements that an embedding of the original state-space becomes automatic. And we have also, by turning our attention to genuinely stochastic systems, given up many of the dynamical invariants and much of the predictive power that one associates with geometric time series analysis applied to deterministic systems. What we have gained is a framework in which to analyse much larger classes of processes and signals. Our result is part of a larger mathematical enterprise whose aim is to extend geometric time series analysis from the domain of low-dimensional, deterministic dynamical systems into the kingdom of stochastic processes and classical time series analysis.

Acknowledgements

We would like to thank James Theiler, Lenny Smith, Christine Ziehmann, Holger Kantz and Mike Davies for enjoyable discussions and bracing skepticism.

Note

- [1] See Huke (1993) for a more accessible proof; Sauer *et al.* (1991), Stark (1996) and Stark *et al.* (1997) for extensions.

References

- Arnold, L. (1995) Lecture Notes in Mathematics 1609 (ed. R. Johnson; Springer-Verlag, Berlin), pp. 1–43.
- Bai-Lin, Hao (ed.) (1984) *Chaos* (World Scientific, Singapore).
- Barnsley, M. F. (1988) *Fractals Everywhere* (Academic Press, London).
- Barnsley, M. F. and Demko, S. (1985) Iterated function systems and the global construction of fractals. *Proceedings of the Royal Society London A* **339**, 243–375.
- Broomhead, D. S., Davies, M. E., Huke, J. P. and Stark, J. (1997) Unpublished notes.
- Broomhead, D. S., Huke, J. P. and Potts, M. A. S. (1996) Canceling deterministic noise by constructing nonlinear inverses to linear filters. *Physica D* **89**, 439–458.
- Casdagli, M., Eubank, S., Farmer, J. D. and Gibson, J. (1991) State space reconstruction in the presence of noise. *Physica D* **51**, 52–98.
- Farmer, J. D. and Sidorowich, J. S. (1988) Exploiting chaos to predict the future and reduce noise. *Evolution, Learning and Games* (ed. Y. C. Lee; World Scientific, Singapore).
- Huke, J. P. (1993) Embedding nonlinear dynamical systems, a guide to Takens theorem. Internal Report, DRA Malvern.
- Kantz, H., Schreiber, T., Hoffman, I., Buzug, T., Pfister, G., Flepp, L. G., Simonet, J., Badii, R. and Brun, E. (1993) Nonlinear noise-reduction—A case-study on experimental data. *Physics Review E* **48**, 1529–1538.
- Kostelich, E. J. and Yorke, J. A. (1990) Noise reduction: finding the simplest dynamical system consistent with the data. *Physica D* **41**, 183–196.
- Lorenz, E. N. (1963) Deterministic nonperiodic flow. *Journal of Atmospheric Science* **20**, 130–141.
- Sauer, T., Yorke, J. A. and Casdagli, M. (1991) Embedology. *Journal of Statistical Physics* **65**, 579–616.
- Spivak, M. (1979) *Differential Geometry* (Publish or Perish, Wilmington), Vol. 1.
- Stark, J. (1996) Delay embeddings for forced systems: I. Deterministic forcing. *Journal of Nonlinear Science*.
- Stark, J. and Arumugam, B. V. (1992) Extracting slowly varying signals from a chaotic background. *International Journal of Bifurcation & Chaos* **2**, 413–419.
- Stark, J., Broomhead, D. S., Davies, M. E. and Huke, J. P. (1997) Takens embedding theorems for forced and stochastic systems. *Nonlinear Analysis, Theory, Methods and Applications* **30**, 5303–5314.
- Takens, F. (1983) *Detecting Strange Attractors in Turbulence*, 898 (eds D. A. Rand and L.-S. Young; Springer-Verlag, Berlin), pp. 366–381.
- Whitney, H. (1935) Differentiable manifolds. *Annals of Mathematics* **37**, 645–680.