

*Delay embedding in the presence of dynamical
noise*

Muldoon, M. R. and Broomhead, D. S. and Huke,
J. P. and Hegger, R.

1998

MIMS EPrint: **2005.13**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

Delay embedding in the presence of dynamical noise

M.R. Muldoon, D.S. Broomhead & J.P. Huke

Mathematics Department, UMIST

P.O. Box 88

Manchester M60 1QD

United Kingdom

R. Hegger

Max-Planck-Institut für Physik komplexer Systeme

Nöthnitzer Str. 38

D 01187 Dresden

Germany

17 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

$$x_{j+1} = \phi(x_j) + \eta_j$$

where ϕ is some fixed map and the η_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 \mathbf{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[18] shows that both these views are correct; the series of delay embedded points is related to the original dynamical system by what

¹See [10] for a more accessible proof, [13, 15, 17] for extensions.

amounts to a nonlinear change of coordinates. Takens' 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 affect the dynamics (and so, indirectly, the signals). Although the only rigorous results about embedding in the presence of noise are fairly recent (see [17] for a review), there is a great deal of practical experience with additive noise on the measured signal (see, *e.g.*, [6, 7, 8, 9, 11, 16]) suggesting that at least some kinds may be overcome.

Dynamical noise has only recently begun to receive attention: again, see [17] 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 [3, 4] 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

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

where the x_j are points in some m -dimensional manifold \mathcal{M} and where the dynamics are given by a sequence of maps $\phi_j : \mathcal{M} \rightarrow \mathcal{M}$; $j \in \mathcal{Z}$. For the moment, we will restrict ourselves to the case of IFS's, 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 $v : \mathcal{M} \rightarrow \mathbf{R}$ 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, Davies, Huke and Stark [5] shows that if one constructs d -dimensional delay vectors

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

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 \mathbf{R}^d :

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

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.* [5] 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) above; 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 $\Phi_{i_1, \dots, i_{d-1}; v} : \mathcal{M} \rightarrow \mathbf{R}^d$:

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

If we now replace the IFS's finite index set with parameters from, say, some n -dimensional manifold \mathcal{N} , then a labeling scheme like the one above 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 properly embed the product 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 $F : \mathcal{M} \times \mathcal{N} \rightarrow \mathcal{M}$;

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

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-parameterised 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 \mathbf{R}^m ; that \mathcal{N} is $B_m(0, \epsilon)$, 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 $\phi_0 : \mathbf{R}^m \rightarrow \mathbf{R}^m$ is the deterministic part of the dynamics and y is a random displacement chosen, say, uniformly from $B_m(0, \epsilon)$. 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 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 $y_j \in \mathcal{N}$, $j \in \mathcal{Z}$ 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 neighborhood 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 theorem below 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 $C^2(\mathcal{M}, \mathcal{M})$, parameterised 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 $v : \mathcal{M} \rightarrow \mathbf{R}^d$ a C^2 function and $d > 2m$;*

it is a generic property that the map $\Phi_{F,v} : \mathcal{M} \times \mathcal{N} \rightarrow \mathbf{R}^{2d}$ 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. $\pi_1(\Phi_{F,v}(\cdot, y)) : \mathcal{M} \rightarrow \mathbf{R}^d$ is an embedding for each $y \in \mathcal{N}$ and

ii. $\pi_2(\Phi_{F,v}(x, \cdot)) : \mathcal{N} \rightarrow \mathbf{R}^d$ 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 \mathbf{R}^{2d} .

Proof: This is really nothing more than a corollary of Whitney's embedding theorem [19, 14]: 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.

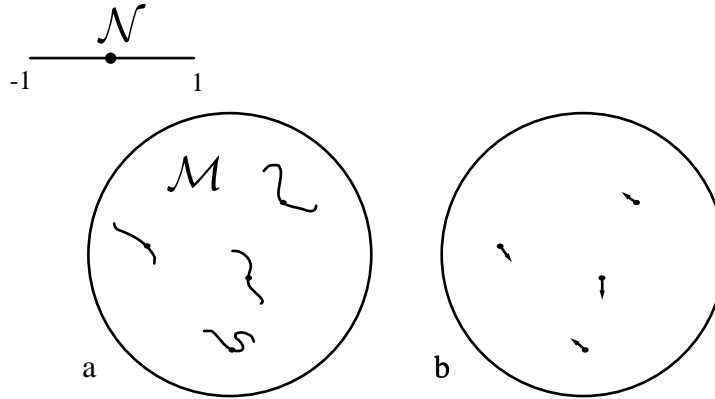


Figure 1: *The hypothesis H2 can never be satisfied if $\mathcal{M} = S^2$ and $\mathcal{N} = [-1, 1]$. Panel (a) shows several images of \mathcal{N} in \mathcal{M} , while panel (b) shows the vector field induced by $D_y F(x, y_0)$.*

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 below, 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. 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, 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 eighties, the dimensions of attractors (see, for example, the bibliography in [2]). As subsequent authors [13] 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 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 examples below are correlation dimension calculations for variously perturbed versions of the celebrated Lorenz system [12]

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

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 \mathbf{R}^3 , which is not a compact manifold. The Lorenz attractor is, however, contained in a compact subset of \mathbf{R}^3 and as Huke [10] 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-parameterised stochastic processes. At one extreme are those that approximate stochastic ODEs—the first two examples discussed below are of this kind—where random perturbation affects the dynamics so profoundly as to require modifications to the notion of an attractor [1]. 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 naively imagine that the separation produced by our two projections is so effective that the embedded system has the structure

$$\text{unperturbed attractor} \times \mathcal{N}. \quad (8)$$

This is not generally the case, though 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 physical system and in these cases the delay-embedded object *is* of the form (8).

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

The system studied in figure 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 $\mathcal{M} = \mathbf{R}^3$, $\mathcal{N} = (-0.1\sigma_x, 0.1\sigma_x)$ and the unperturbed dynamics are the fifth iterate of a map $\phi_{\delta t} : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ that approximates the time- δt evolution operator of the Lorenz system.

The resulting embedded object is four-dimensional. That is, there is a recognizable 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 \vec{v}_j is not without structure: had we not known that the noise manifold \mathcal{N} was 1-dimensional, we could have deduced it from the way in which the slopes of the $D_2(m, \epsilon)$ jump up by 1 when we include data from the second set of measurements.

In the calculation pictured in Figure 3 we consider a perturbation that more closely approximates a classical stochastic process. The noise still affects only the x

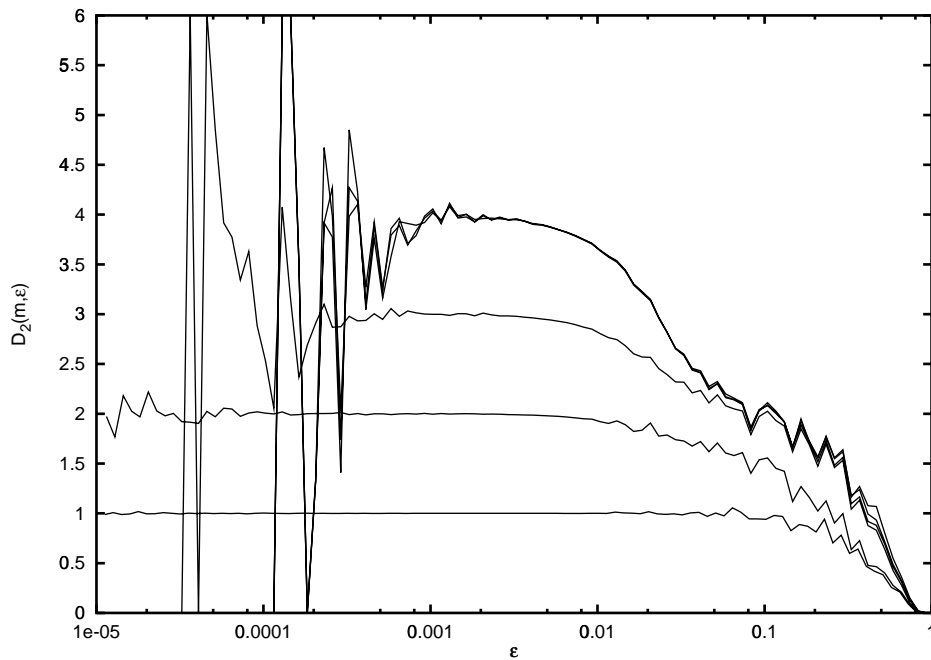


Figure 2: *Correlation dimension estimates for a Lorenz system perturbed by a one-dimensional family of random kicks that acts once per sampling period.*

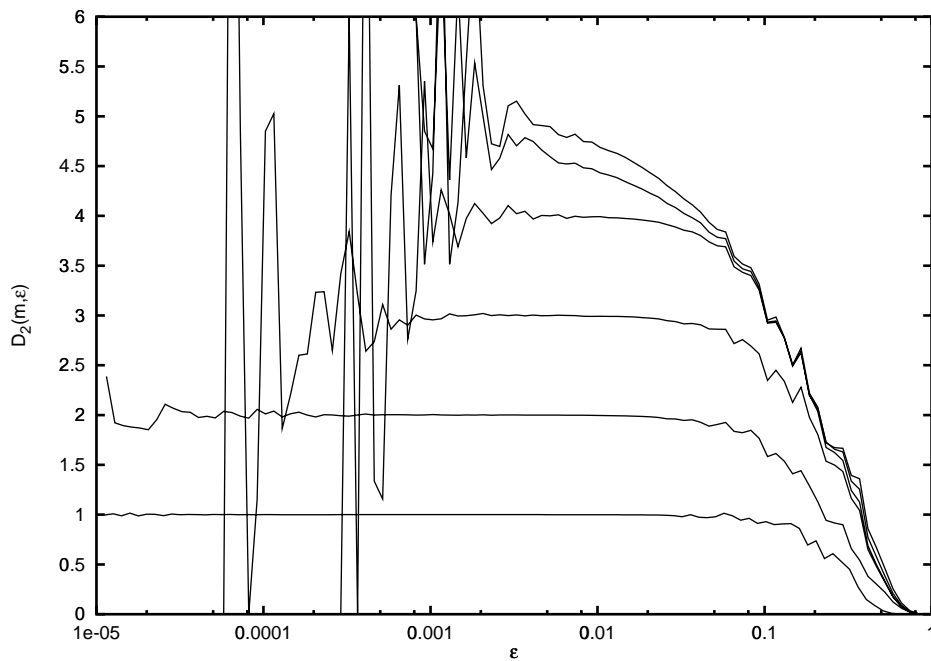


Figure 3: *Correlation dimension estimates for a Lorenz system perturbed by a one-dimensional family of random kicks that acts once per integration time-step: this system is nearer to a diffusion process than to the example in Figure 2.*

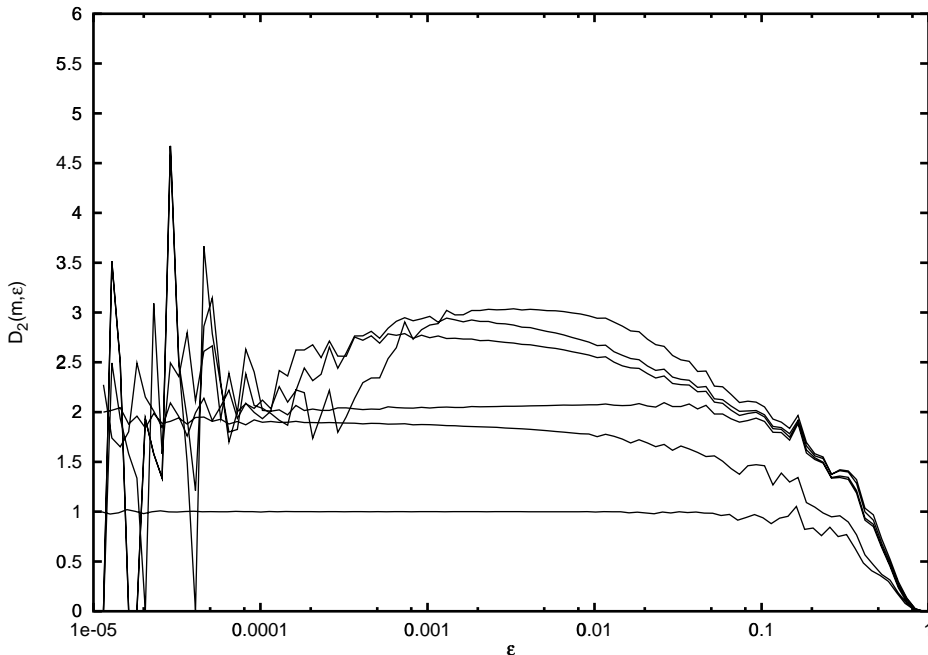


Figure 4: *Correlation dimension estimates for a Lorenz system sampled at irregular time intervals. This one-dimensional family of evolution operators shares the same attracting set, but single-step multivariate delay embedding shows that the dynamics are not the same from step to step.*

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 5 times per sampling interval, so $\dim \mathcal{N} = 5$ and one would not expect the \vec{v}_j to lie on an object of dimension less than $3 + 5 = 8$. The $D_2(m, \epsilon)$ 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 Figure 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 $(0.01, 0.05)$. 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 $D_2(m, \epsilon)$ 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.

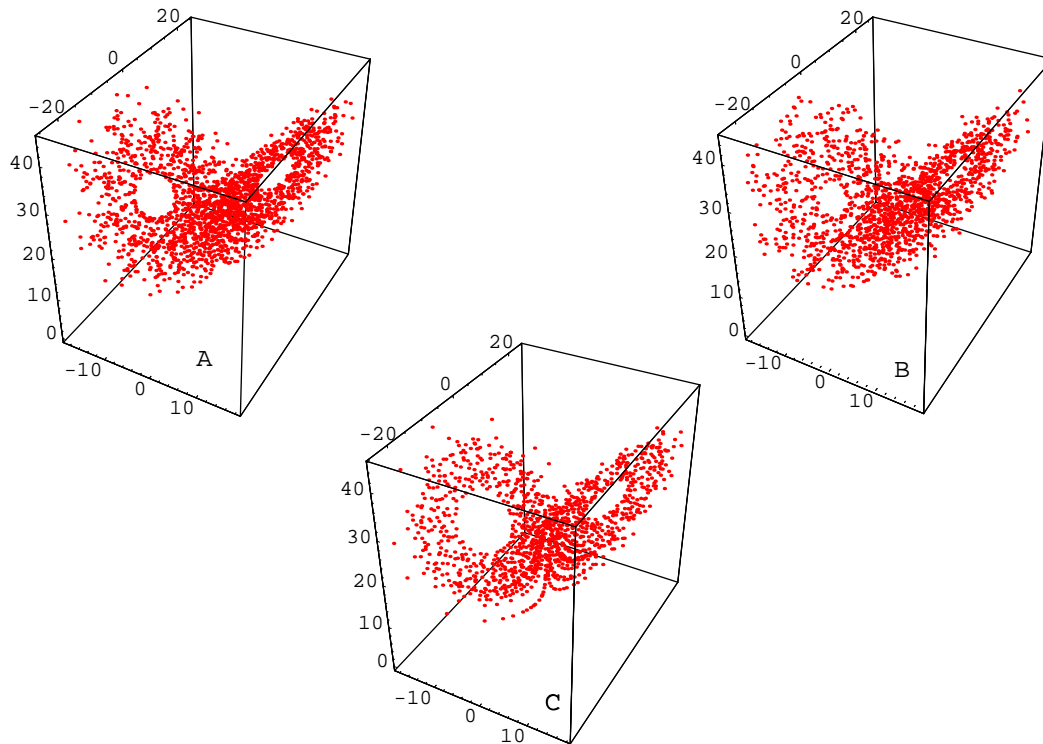


Figure 5: *Two thousand points from each of the trajectories of the stochastic processes discussed above: panels A and B correspond to the cases where the x variable received periodic kicks once (A) or five times (B) per sampling interval. In these cases the fine structure of the Lorenz attractor is blurred. In the system pictured in panel 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.*

3.4 Extensions and applications

Our embedding theorem evades the dimension-counting problem outlined at the end of the 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 Figure 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 above 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 $\mathcal{M} \times \mathcal{N}$ the 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 [6]; combining their ideas and this paper, one could recover the interesting signal cleanly.

4 Conclusion

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.

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

References

- [1] L. Arnold, “Random dynamical systems”, in *Lecture Notes in Mathematics* **1609**, R. Johnson, ed., pp. 1-43, Springer Verlag: Berlin, (1995).
- [2] Hao Bai-Lin, ed., *Chaos*, World Scientific: Singapore, (1984).
- [3] M.F. Barnsley and S. Demko, “Iterated Function Systems and the Global Construction of Fractals”, *Proc. Roy. Soc. of London A*, **339**, pp. 243-375, (1985).
- [4] M.F. Barnsley, *Fractals Everywhere*, Academic Press: London & San Diego, (1988).
- [5] D.S. Broomhead, M.E. Davies, J.P. Huke and J. Stark, unpublished notes.
- [6] D.S. Broomhead, J.P. Huke and M.A.S. Potts, “Canceling deterministic noise by constructing nonlinear inverses to linear filters”, *Physica D*, **89**, pp. 439-458, (1996).
- [7] M. Casdagli, S. Eubank, J. D. Farmer and J. Gibson, “State space reconstruction in the presence of noise”, *Physica D* **51**, pp. 52-98, (1991).
- [8] J.D. Farmer and J.S. Sidorowich, “Exploiting Chaos to Predict the Future and Reduce Noise”, in *Evolution, Learning and Games*, Y.C. Lee, ed., World Scientific: Singapore, (1988).
- [9] H. Kantz, T. Schreiber, I. Hoffman, T. Buzug, G. Pfister, L.G. Flepp, J. Simonet, R. Badii and E. Brun, “Nonlinear noise-reduction — A case-study on experimental data”, *Phys. Rev. E*, **48**, pp. 1529-1538, (1993).
- [10] J.P. Huke, “Embedding Nonlinear Dynamical Systems, A Guide to Takens Theorem”, Internal Report, DRA Malvern, (1993).
- [11] E. J. Kostelich and J. A. Yorke, “Noise Reduction: Finding the Simplest Dynamical System Consistent with the Data”, *Physica D* **41**, pp. 183-196, (1990).

- [12] E.N. Lorenz, “Deterministic Nonperiodic Flow”, *J. Atmospheric Science*, **20**, pp.130-141, (1963).
- [13] T. Sauer, J.A. Yorke and M. Casdagli, “Embedology”, *J. Stat. Phys.* **65**, pp. 579-616, (1991).
- [14] M. Spivak, *Differential Geometry, Vol. 1*, Publish or Perish: Wilmington, (1979).
- [15] J. Stark, “Delay Embeddings for Forced Systems: I. Deterministic Forcing”, submitted to *J. Nonlinear Science*, (1996).
- [16] J. Stark and B.V. Arumugam, “Extracting Slowly Varying Signals from a Chaotic Background”, *Int. J. Bif. Chaos* **2**, pp. 413-419, (1992).
- [17] J. Stark, D.S. Broomhead, M.E. Davies and J.P. Huke, “Takens Embedding Theorems for Forced and Stochastic Systems”, *Nonlinear Analysis, Theory, Methods and Applications* **30**, pp. 5303-5314, (1997)
- [18] F. Takens, “Detecting strange attractors in turbulence”, in *Lecture Notes in Mathematics* **898**, D. A. Rand and L.-S. Young, eds., pp. 366-381, Springer Verlag: Berlin, (1983).
- [19] H. Whitney, “Differentiable Manifolds”, *Ann. of Maths.* **37**, pp. 645-680, (1935).