

# A rate of convergence for a particular estimate of a noise-contaminated chaotic time series

Jason H. Stover

created 1 June 2002

updated Thursday, 8 January 2004

*Copyright © 2002, 2004 by Jason H. Stover. All rights reserved. Permission to copy, store, & view this document unmodified & in its entirety is granted.*

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Lalley's Filter</b>	<b>2</b>
<b>3</b>	<b>Numerical results</b>	<b>4</b>
3.1	Smale's Solenoid map . . . . .	5
3.2	Henon map . . . . .	9
<b>4</b>	<b>Convergence rate</b>	<b>12</b>
<b>5</b>	<b>Conclusion</b>	<b>13</b>
<b>6</b>	<b>Acknowledgments</b>	<b>14</b>
<b>7</b>	<b>Other File Formats</b>	<b>14</b>

### Abstract

We prove a rate of convergence theorem for a consistent estimate of a signal  $X_i$  from a time series of the form  $Y_i = X_i + Z_i$  with  $X_i = F(X_{i-1})$ , where  $F$  is a  $C^2$  Axiom A diffeomorphism and the  $Z_i$ 's are uniformly bounded error terms. The estimate was introduced by Lalley [Lal99]. We also present results of filtered time series from simulated data.

## 1 Introduction

Using both simulated data and a rate of convergence theorem, we show that Lalley's estimate [Lal99] of a chaotic signal in a noise-contaminated time series

performs well. Assume we have a time series  $Y_1, Y_2, \dots, Y_n$  where  $Y_i = X_i + Z_i$ ,  $X_i$  are generated by an Axiom A diffeomorphism  $F$ ; the noise terms  $Z_i$  are independent of each other and the  $X_i$ 's;  $E(Z_i) = 0$ . Lalley [Lal99] developed a method in which noise can be consistently removed from the time series if the  $Z_i$ 's are uniformly bounded by a constant  $\delta$  which is less than a *separation threshold* of  $F$ . This method will yield an estimate of each  $X_i$  for  $i > \kappa_n$  and  $i < n - \kappa_n$  for some pre-chosen  $\kappa_n = o(\log n)$ . The maximum distance between predicted and true values within the  $\kappa_n$  window (i.e., between  $\kappa_n$  and  $n - \kappa_n$ ) will converge almost surely to zero.

In this paper we present numeric results showing “good” estimates obtained using this method, and also show that, for any  $\alpha \in (0, \frac{1}{2})$ , the maximum distance between true and predicted values will almost surely be within  $\frac{1}{n^\alpha}$  of zero eventually. Our examples of noisy time series come from Smale’s Solenoid map and the Henon map, with i.i.d. error terms on the unit ball.

Schreiber [Sch93] presented numerical results for a similar filter for chaotic time series.

For more information on Axiom A diffeomorphisms, see Bowen [Bow75] or Mañe [Mañ87].

## 2 Lalley’s Filter

Our goal here is to recover the  $X_i$ 's from the observed, noisy data. For an Axiom A dynamical system, Lalley [Lal99] developed a method that can be used to recover the  $X_i$ 's if  $|Z_i| < \delta$  uniformly for all  $n$ , and  $5\delta < \Delta$ , where  $\Delta \in (0, 1)$  is called a *separation threshold* for  $F$ ; i.e., there exists a  $C > 0$  such that if  $x$  and  $y$  are in  $\Lambda$  and  $0 < |x - y| < \Delta$  then

$$|F^n(x) - F^n(y)| > \Delta$$

for some  $|n| \leq -C \log |x - y|$ .

We will denote our estimate of  $X_i$  by  $\hat{X}_i$ , which is computed in the following way: Choose some  $\kappa_n = o(\log n)$  such that  $\kappa_n \rightarrow \infty$  (e.g.,  $\kappa_n = (\log n)^{-1}$ ). Let

$$A_i = \left\{ \nu : \max_{-\kappa_n \leq k \leq \kappa_n} |Y_{\nu+k} - Y_{i+k}| < 3\delta \right\}$$

and set

$$\hat{X}_i = \frac{1}{|A_i|} \sum_{\nu \in A_i} Y_\nu$$

**HYPOTHESIS 2.1** *The  $Z_i$  are independent, uniformly bounded by a constant  $\delta$ ,  $5\delta < \Delta$  and  $E(Z_i) = 0$ .*

**HYPOTHESIS 2.2** *The  $Z_i$ 's are independent of the  $X_i$ 's.*

Lalley proved the following theorem, a sketch of whose proof will be necessary to explain the proof of Theorem 4.1:

THEOREM 2.1 (Lalley, 1999) Assume  $X_0$  is chosen at random according to the SRB measure  $\mu$ . Under Hypotheses 2.1 and 2.2,

$$\lim_{n \rightarrow \infty} \max_{\{i=\kappa_n, \dots, n-\kappa_n+1\}} |\hat{X}_i - X_i| = 0$$

almost surely.

This theorem's proof relies on three lemmas from [Lal99], which we present here without proof.

LEMMA 2.1 There exists a constant  $C > 0$  such that if  $\nu \in A_i$  then

$$|X_i - X_\nu| \leq \exp(-C\kappa_n)$$

LEMMA 2.2 For every  $\epsilon > 0$ , all sufficiently large  $n$  and all integers  $i = \kappa_n, \dots, n - \kappa_n$ ,

$$P(|A_i| \leq n^{1-4\epsilon}) \leq \exp(-n^\epsilon)$$

LEMMA 2.3 If  $\xi_1, \xi_2, \dots$  are independent random variables uniformly bounded in modulus by a constant  $\delta < \infty$  and if  $E(\xi_i) = 0$  for every  $i$ , then for every  $\eta > 0$  there exists a  $\gamma = \gamma(\eta, \delta) > 0$  such that for all sufficiently large  $m$ ,

$$P\left[\frac{1}{m} \left| \sum_{i=1}^m \xi_i \right| \geq \eta\right] \leq \exp(-m\gamma)$$

*Sketch of proof of Theorem 2.1.* Write

$$\hat{X}_i = X_i + \frac{1}{|A_i|} \sum_{\nu \in A_i} (X_\nu - X_i) + \frac{1}{|A_i|} \sum_{\nu \in A_i} Z_\nu \quad (1)$$

From Lemma 2.1, the first sum converges to zero uniformly for  $i = \kappa_n, \dots, n - \kappa_n$  as  $n \rightarrow \infty$ . It remains to be shown that the second sum converges to zero. The usual approach would be to invoke the strong law of large numbers, however the  $Z_\nu$ 's are not independent. Lalley's proof hinges on writing  $A_i$  as a union of disjoint sets in which the component sets are independent of some of the random vectors  $Z_\nu$ .

For each  $i$ , define  $A_i^*$  to be the set of indices  $\nu$  in  $A_i$  such that  $|i - \nu| \leq 2\kappa_n$ . Notice that  $|A_i^*| \leq 4\kappa_n + 1 = o(\log n)$ , so when  $|A_i| > n^{1-\epsilon}$  for some  $\epsilon < 1$ , the indices in  $A_i^*$  will have a negligible effect on the average  $\frac{1}{|A_i|} \sum_{\nu \in A_i} Z_\nu$ . For each  $i$  and each  $j = 1, \dots, 2\kappa_n + 1$ , define  $A_i^j$  to be the set of all indices  $\nu$  in  $A_i$  such that  $\nu = j \bmod 2\kappa_n + 1$ . The sets  $A_i^*, A_i^1, \dots, A_i^{2\kappa_n+1}$  are pairwise disjoint and

$$A_i = A_i^* \cup \left( \bigcup_{j=1}^{2\kappa_n+1} A_i^j \right) \quad (2)$$

Now, for each integer  $j \in [1, 2\kappa_n + 1]$ , the set  $A_i^j$  is independent of the set of random vectors  $\{Z_\nu\}$  indexed by integers  $\nu = j \bmod 2\kappa_n + 1$ : Consider an integer  $\nu = j \bmod 2\kappa_n + 1$ . The event  $\nu \in A_i^j$  is completely determined by the values of  $Y_{i+m}$  and  $Y_{\nu+m}$  for  $|m| \leq \kappa_n$ , and no other event  $\{\nu' \in A_i^j\}$  ( $\nu' \neq \nu$ ) is influenced by the values of  $Y_{\nu+m}$  for  $|m| \leq \kappa_n$ . Also, the event  $\{\nu \in A_i^j\}$  is not affected by the value of  $Z_\nu$ : If  $|Y_{\nu+m} - Y_{i+m}| < 3\delta$  for all  $m = 1, \dots, \kappa_n$  then we would have  $|X_i - X_\nu| < \delta/2$  (See the proof of Lemma 2.1 in [Lal99] for a more complete explanation). Thus the set  $A_i^j$  can be determined without knowledge of the random vectors  $\{Z_\nu\}$  where  $\nu = j \bmod 2\kappa_n + 1$ .

Now let  $\mathcal{I}$  be the set consisting the index  $*$  and of indices  $j$  such that  $|A_i^j| < \sqrt{n}$ . Let  $\mathcal{J}$  be the remaining indices of the sets  $A_i^j$ . For each  $j \in \mathcal{J}$ , Lemma 2.3 implies that for any  $\epsilon > 0$  and large  $n$ ,

$$P \left[ \frac{1}{|A_i^j|} \left| \sum_{\nu \in A_i^j} Z_\nu \right| \geq \epsilon \right] \leq \exp(-\gamma|A_i^j|) \leq \exp(-\gamma\sqrt{n}) \quad (3)$$

for some constant  $\gamma > 0$  depending on  $\epsilon$  and  $\delta$  but not on  $n$ . For sufficiently large  $n$ ,

$$\left\{ \left| \sum_{\nu \in A_i} Z_\nu \right| > 2\epsilon|A_i| \right\} \subset \{|A_i| \leq n^{3/4}\} \cup \left( \bigcup_{j \in \mathcal{J}} \left\{ \left| \sum_{\nu \in A_i^j} Z_\nu \right| > \epsilon|A_i^j| \right\} \right)$$

So by Lemma 2.2 and (3), for all large  $n$  and each  $i$  between  $\kappa_n$  and  $n - \kappa_n$ ,

$$P \left[ \left| \sum_{\nu \in A_i^j} Z_\nu \right| > 2|A_i|\epsilon \right] < (2\kappa_n + 1) \exp(-\gamma\sqrt{n}) + \exp(-n^{1/16}) \quad (4)$$

Since the series  $\sum_n n e^{-an^\alpha} < \infty$  for any  $a > 0$ , we have

$$\max_{\{i=\kappa_n, \dots, n-\kappa_n\}} \frac{1}{|A_i|} \left| \sum_{\nu \in A_i} Z_\nu \right| \rightarrow 0$$

by the Borel-Cantelli Lemma.  $\square$

### 3 Numerical results

This section presents the results of smoothing two time series of length 100,000 with Lalley's filter. The large sample sizes are necessary to reveal the intricate structure of the signal; samples with fewer than about 50,000 data do not show the attractors well. We present graphs of noise-free, contaminated and filtered time series, and show the performance of the filter on the Solenoid map for different values of  $\kappa_n$ . The filter performs well for each series, and we will show why in the next section by proving a rate of convergence theorem.

Table 1: Sums of squares for filtering method with  $n = 100,000$ ,  $\kappa_n = 1, \dots, 12$

$\kappa_n$	$\sum_{i=1}^n (X_i - Y_i)^2$	$\sum_{i=1}^n (\hat{X}_i - Y_i)^2$	$\sum_{i=1}^n (\hat{X}_i - X_i)^2$
1	33,208.04	43,430.68	10,545.78
2	33,208.04	35,208.78	2,047.50
3	33,208.04	33,560.26	499.87
4	33,208.04	33,061.06	371.09
5	33,208.04	32,151.31	1,110.41
6	33,208.04	28,964.34	4,274.96
7	33,208.04	18,712.26	14,472.41
8	33,208.04	7,115.43	26,101.77
9	33,208.04	1,987.28	31,215.57
10	33,208.04	500.3	32,707.81
11	33,208.04	115.41	33,092.62
12	33,208.04	26.68	33,181.36

### 3.1 Smale's Solenoid map

Consider the map on the solid torus  $T$  given by

$$F_\beta(\theta, x) = (2\theta, \beta x + \frac{1}{2}e^{i\theta})$$

The map can be thought of intuitively this way: Cut the torus once to get long cylinder. Stretch the cylinder to twice its length while contracting its width by  $\beta$ . Wrap the resulting long, thin cylinder around itself twice, rejoin the ends and replace it inside the original space. Iterating the solenoid map  $n$  times results in a very long, thin tube that winds around the inside of the torus  $2^n$  times. Notice  $F^n(T) = \underbrace{F \circ \dots \circ F}_n(T)$  is a closed set contained completely inside

the interior of  $F^{n-1}(T)$ . Thus  $\bigcap_{n \geq 0} F^n(T) = \Lambda \neq \emptyset$ . It can be shown that  $F$  is a homeomorphism on  $\Lambda$ . Since  $\lim_{n \rightarrow \infty} F^n(x) \in \Lambda$  for any  $x \in T$ ,  $\Lambda$  is the attractor for  $F$ .

100,000 data were generated from Smale's solenoid map with  $\beta = .45$ . A partial plot of the clean data can be seen in figure 1 (plotting all of the data results in a blurry graph). The noisy data are seen in the next figure, and the predicted values in the next. The noise terms  $Z_i$  are distributed uniformly on a ball centered at  $X_i$  with radius 1. From the graph, the predicted values appear close the original values.

Table 1 shows the sums of squares due to total, model and error for the filtered solenoid series for different values of  $\kappa_n$ . From the table we can see why a such a long time series is necessary: we need  $\kappa_n$  to grow more slowly than  $\log n$ , but require it to be large enough to allow several values to be averaged to estimate each  $X_i$ . This is possible only for very large samples.

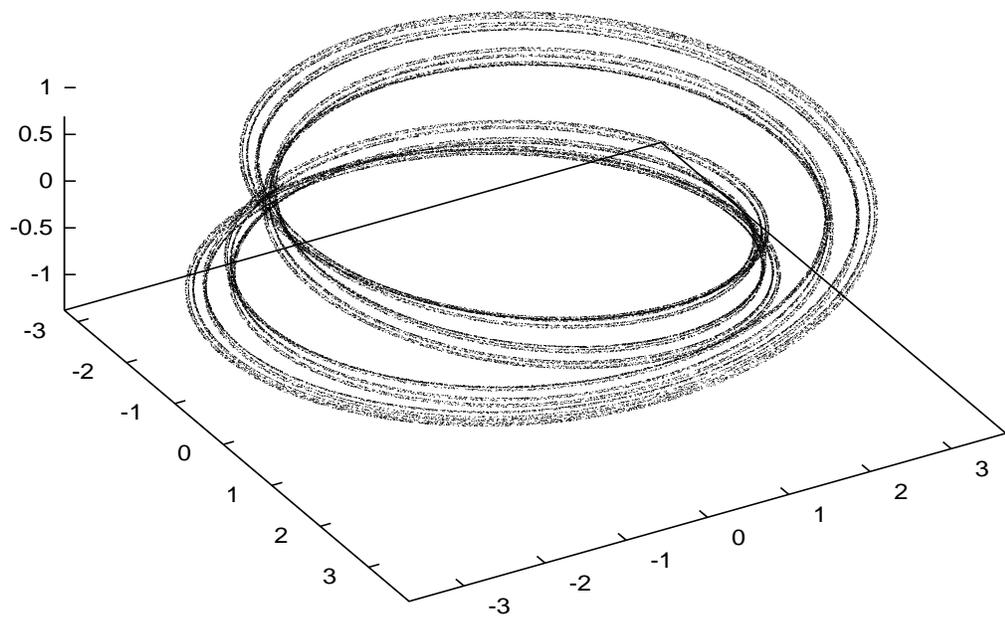


Figure 1: 100,000 points from Smale's solenoid map,  $\beta = .45$ .

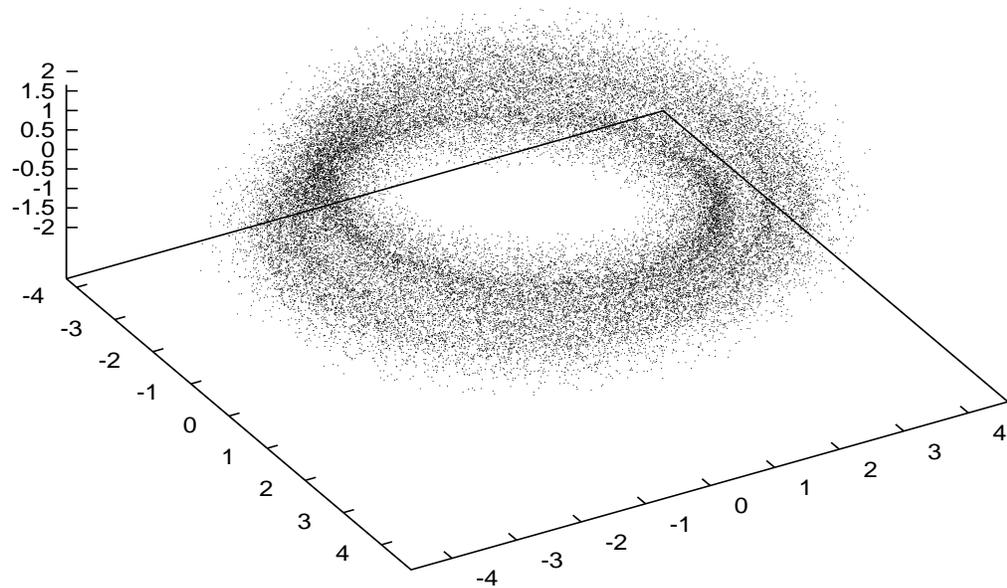


Figure 2: Data from previous figure with independent noise  $Z_i$  distributed uniformly on  $B_{X_i}(1)$

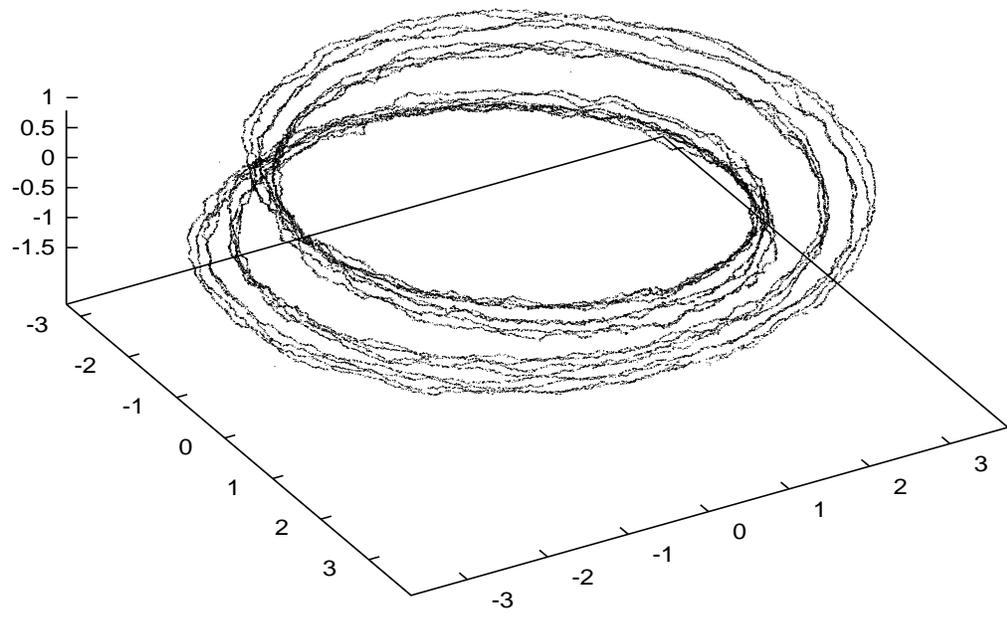


Figure 3: Predicted values with  $\kappa_{100,000} = 4$

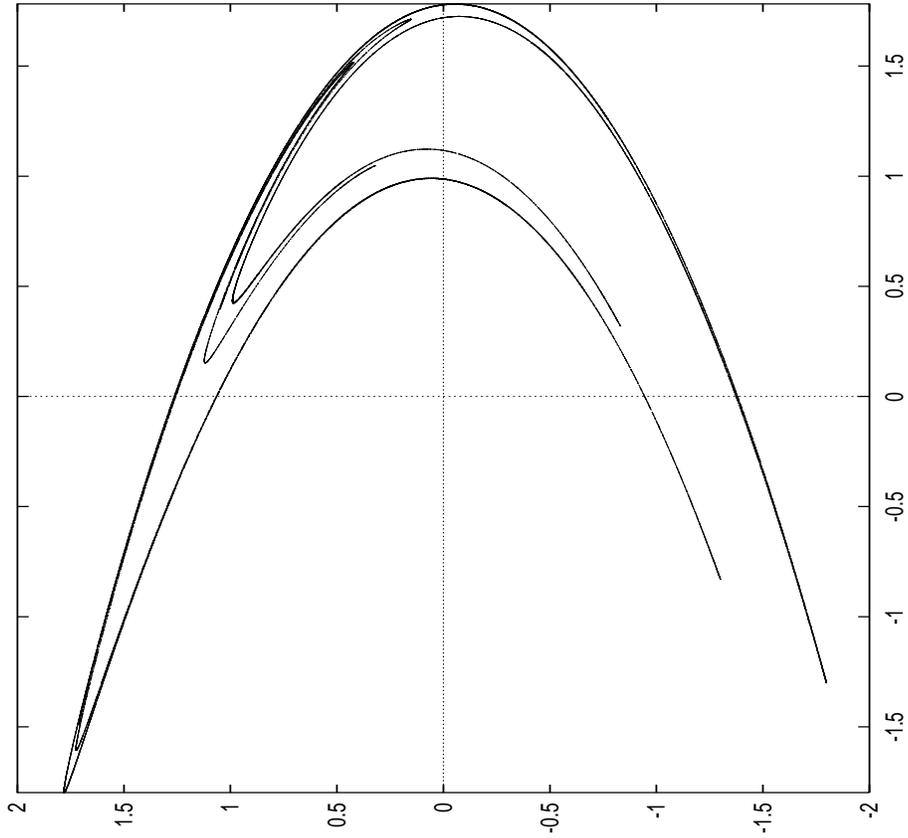


Figure 4: 100,000 points from the Henonmap.

### 3.2 Henon map

The Henon map  $H : \mathbf{R}^2 \rightarrow \mathbf{R}^2$  defined by

$$H(x_1, x_2) = (a - bx_2 - x_1^2, x_1)$$

is a function still not well understood. It is not an Axiom A diffeomorphism. It does possess an invariant set  $\Lambda$  if  $a = 1.4$ ,  $b = .3$  and the initial point is chosen from a quadrilateral with vertices  $(-1.33, 0.42)$ ,  $(1.32, 0.33)$ ,  $(1.245, -0.14)$  and  $(-1.06, -0.5)$ . Despite the possibility that  $H$  may not be an Axiom A diffeomorphism, our filtered time series captures the intricacies of the original series. Figure 4 shows a plot of 100,000 iterations of an initial point  $x$  chosen at random from the aforementioned quadrilateral. Figures 5 and 6 show the noisy and filtered time series respectively.

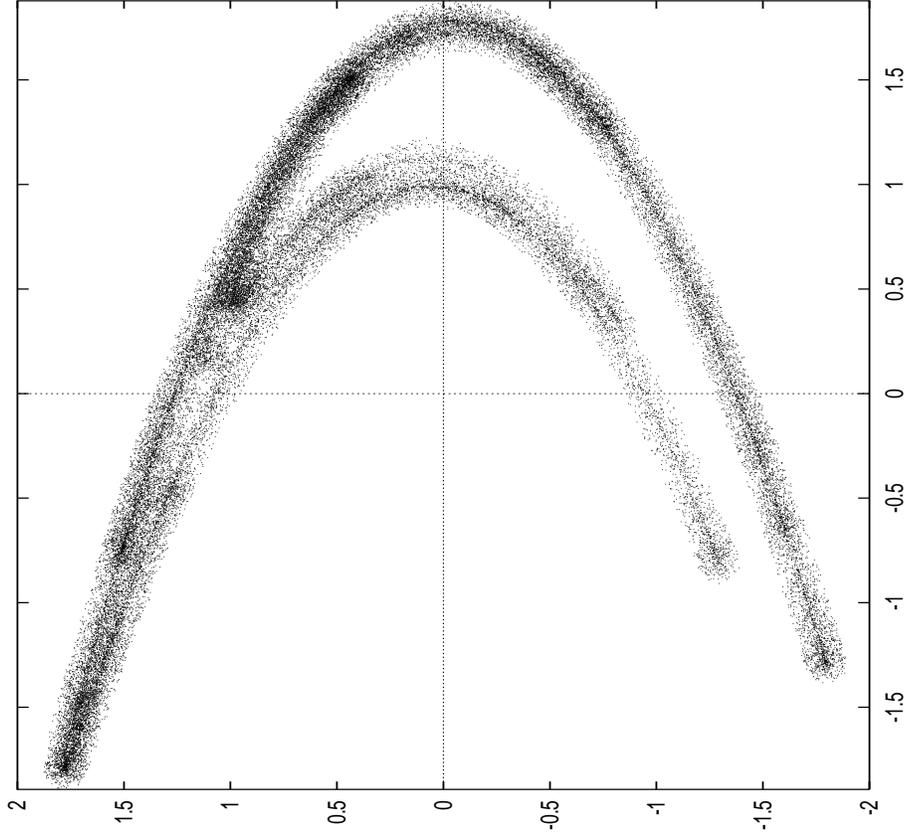


Figure 5: 100,000 points from the Henon map with noise added. Each noise term is uniformly distributed on a ball with radius .11.

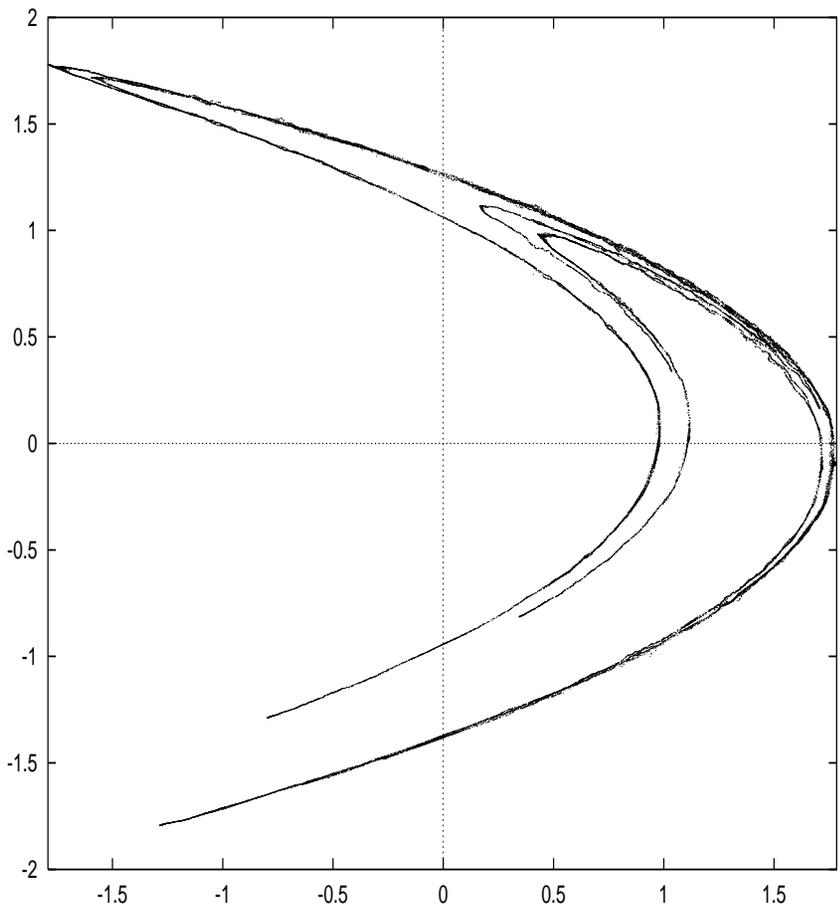


Figure 6: Predicted values with  $\kappa_{100,000} = 4$

## 4 Convergence rate

These estimates perform well because they converge quickly, as our rate of convergence theorem shows. Its proof relies on two lemmas, the first of which was proved by Kolmogorov [Kol29] and is stated here in a slightly more specific form. See [Kol29] for a proof.

LEMMA 4.1 *Let  $S_n = \xi_1 + \dots + \xi_n$  where the  $\xi_i$  are independent,  $E(\xi_i) = 0$  and  $|\xi_i| < \delta$ . If  $x \leq n\delta$  then  $P(S_n > x) < \exp\left(-\frac{x^2}{2n\delta}\left(1 - \frac{x}{2n\delta}\right)\right)$*

Let

$$\bar{Z}_i = \frac{1}{|A_i|} \sum_{j \in A_i} Z_j$$

be the second sum in (1).

THEOREM 4.1 *For any  $\alpha < \frac{1}{2}$  and  $i = \kappa_n, \dots, n - \kappa_n + 1$ ,  $P(\max_{i=\kappa_n, \dots, n-\kappa_n+1} |\bar{Z}_i| > \frac{1}{n^\alpha} \text{ i.o.}) = 0$*

PROOF: As in Theorem 2.1, we may partition  $A_i$  into  $A_i^*$  and  $A_i^j$ ,  $j = 1, \dots, 2\kappa_n + 1$ , where  $A_i^*$  consists of those indices  $m$  such that  $|n - m| \leq 2\kappa_n$  and  $A_i^j$  consists of those remaining indices  $m$  such that  $m = j \pmod{2\kappa_n + 1}$ . Let  $\epsilon > 0$  be given. Also as in Theorem 2.1, for each  $n$  we partition the sets  $A_i^j$  as  $\mathcal{I} \cup \mathcal{J}$ , where  $\mathcal{I}$  consists of the index  $*$  and those indices  $j$  for which  $|A_i^j| < n^{1-\epsilon}$  and  $\mathcal{J}$  consists of the remaining indices. Remember from 2.1 that if  $\nu \in A_i^j$  then  $A_i^j$  and  $Z_\nu$  are independent. From Lemma 4.1, for any  $j \in \mathcal{J}$  we have

$$\begin{aligned} P\left[\left|\sum_{\nu \in A_i^j} Z_\nu\right| > \frac{n^{1-\alpha}}{2\kappa_n + 1} \left|A_i^j\right|\right] &\leq 2 \exp\left(-\frac{n^{2-2\epsilon-2\alpha}}{2n\delta(2\kappa_n + 1)^2} \left(1 - \frac{n^{1-\epsilon-\alpha}}{2n\delta(2\kappa_n + 1)}\right)\right) \\ &\leq 2 \exp\left(-\frac{n^{1-2\epsilon-2\alpha}}{(2\kappa_n + 1)^2} (1 - \gamma)\right) \end{aligned} \quad (5)$$

for some  $\gamma > 0$ .

Notice that when the statements  $|A_i|^{1-\alpha} > n^{1-\epsilon}$  and

$$\left|\sum_{\nu \in A_i^1} Z_\nu + \dots + \sum_{\nu \in A_i^{2\kappa_n+1}} Z_\nu\right| = \left|\sum_{\nu \in A_i} Z_\nu\right| > |A_i|^{1-\alpha}$$

are both true, there must be some  $j \in \mathcal{J}$  such that  $\left|\sum_{\nu \in A_i^j} Z_\nu\right| > \frac{n^{1-\epsilon-\alpha}}{2\kappa_n + 1}$ , and this gives

$$\left\{\left|\sum_{\nu \in A_i} Z_\nu\right| > |A_i|^{1-\alpha}\right\} \subset \{|A_i| \leq n^{1-\epsilon}\} \cup \left(\bigcup_{j \in \mathcal{J}} \left\{\left|\sum_{\nu \in A_i^j} Z_\nu\right| > \frac{n^{1-\epsilon-\alpha}}{2\kappa_n + 1}\right\}\right) \quad (6)$$

Remember  $|\mathcal{J}| \leq 2\kappa_n + 1$ . Using (5) and (6), for  $j \in \mathcal{J}$  we have

$$\begin{aligned} & P \left[ \left| \sum_{\nu \in A_i} Z_\nu \right| > |A_i|^{1-\alpha} \mathbb{1}_{|A_i| > n^{1-\epsilon}} \right] \\ & \leq (2\kappa_n + 1) \max_{j \in \mathcal{J}} P \left[ \left| \sum_{\nu \in A_i^j} Z_\nu \right| > \frac{n^{1-\epsilon-\alpha}}{2\kappa_n + 1} \right] \\ & \leq 2(2\kappa_n + 1) \exp \left( -\frac{n^{1-2\epsilon-2\alpha}}{(2\kappa_n + 1)^2} (1 - \gamma) \right) \end{aligned} \quad (7)$$

By Lemma 2.2,  $P[|A_i| < n^{1-\epsilon}] < \exp(-n^{\epsilon/4})$  so by (7)

$$\begin{aligned} P \left[ \bar{Z}_i > \frac{1}{n^\alpha} \right] & \leq 2(2\kappa_n + 1) \exp \left( -\frac{n^{1-2\epsilon-2\alpha}}{(2\kappa_n + 1)^2} (1 - \gamma) \right) + \exp(-n^{\epsilon/4}) \quad (8) \\ & = \beta_n \quad (9) \end{aligned}$$

Now, when  $\epsilon$  is small, so  $\beta_n$  goes to zero exponentially quickly whenever  $\alpha \in (0, \frac{1}{2})$ . So

$$P \left[ \max_{i=\kappa_n, \dots, n-\kappa_n+1} |\bar{Z}_i| > \frac{1}{n^\alpha} \right] \leq n\beta_n$$

The term on the right hand side is absolutely summable over  $n$ , so by the Borel-Cantelli Lemma,  $P[\max_{i=\kappa_n, \dots, n-\kappa_n+1} |\bar{Z}_i| > \frac{1}{n^\alpha} \text{ i.o.}] = 0$ .  $\square$

Lemmas 2.2 and 2.1 state  $P[|A_i| < n^{1-4\epsilon}] < \exp(-n^\epsilon)$  for any  $\epsilon$ , and that  $|X_i - X_\nu| < \exp(-C\kappa_n)$ . Combining these two, we have

$$P \left[ \max_{i=\kappa_n, \dots, \kappa_n+1} \frac{1}{|A_i|} \left| \sum_{\nu \in A_i} X_i - X_\nu \right| > \frac{1}{n^{1-4\epsilon}} \text{ i.o.} \right] = 0. \quad (10)$$

Thus, from (1), (10) and Theorem 4.1, we have:

**COROLLARY 4.1**  $P[\max_{i=\kappa_n, \dots, n-\kappa_n+1} |\hat{X}_i - X_i| > \frac{1}{n^\alpha} \text{ i.o.}] = 0$  for any  $\alpha \in (0, \frac{1}{2})$

## 5 Conclusion

Lalley's filter converges almost surely at a rate of  $\frac{1}{n^\alpha}$  for any  $\alpha < 1/2$  and thus performs well for time series long enough to make  $\kappa_n$  large. It would a useful tool for experimenters in the field of dynamical systems who have access to such long time series. Numerical results herein suggest the time series should be long enough to ensure  $\kappa_n$  is at least 4 to clearly show the attractor.

## 6 Acknowledgments

Results from this paper were part of the author's Ph.D. thesis, and he would like to thank his advisor Steve Lalley for his guidance and patience. Also, without the help of Doug Crabill and Gene Stover, the author would not have been able to write the computer program that generated and filtered the simulated time series.

## 7 Other File Formats

- This document is available in multi-file HTML format at <http://cybertiggyr.com/ctfs/>.
- This document is available in PDF format at <http://cybertiggyr.com/ctfs/ctfs.ps>.

## References

- [Aba96] H.D.I. Abarbanel. *Analysis of Observed Chaotic Time Series*. Springer-Verlag, New York, 1996.
- [Bow75] R. Bowen. Equilibrium states and the ergodic theory of axiom a diffeomorphisms. *Springer Lecture Notes in Math*, 470, 1975.
- [Kol29] A. N. Kolmogorov. Über das gesetz des iterierten logarithmus. *Mathematische Annalen*, 101:126, 1929.
- [Lal99] S.P Lalley. Beneath the noise, chaos. *Annals of Statistics*, 27:461, 1999.
- [Mañ87] R. Mañé. *Ergodic Theory and Differentiable Dynamics*. Springer-Verlag, Berlin Heidelberg, 1987.
- [Sch93] T. Schreiber. Extremely simple nonlinear noise reduction method. *Physical Review E*, 47:2401, 1993.