Real Time Series analysis and modelling

Aid machine learning with dynamical insight

Hailiang Du
Department of Mathematical Sciences, Durham University
hailiang.du@durham.ac.uk
Data Science Institute, London School of Economics and Political Science
h.l.du@lse.ac.uk

All theorems are true, All models are wrong. All data are inaccurate. What are we to do?

The aim of this course is to teach you how to deal with real data, to increase your scepticism regarding reliable modelling in practice, and to expand the tool box you carry to include nonlinear techniques, both deterministic and stochastic with the aid of dynamical insight.

In short: to get you to think before you compute (and perhaps afterwards too.)

Lecture 3

Fundamental aspects of predictability

$$x_{i+1} = x_i + x_{i-1}$$

Fibonacci (Leonardo of Pisa 1202)

1, 1, 2, 3, 5, 8, …

Question: What would the forecast error be if there had been two pairs on day one? (Measurement error). [“two” is not enough information]

Mature	Young	Total
0	2	2
2	0	2
2	2	4
4	2	6
6	4	10
10	6	16

$$s_0 = x_0 + e$$

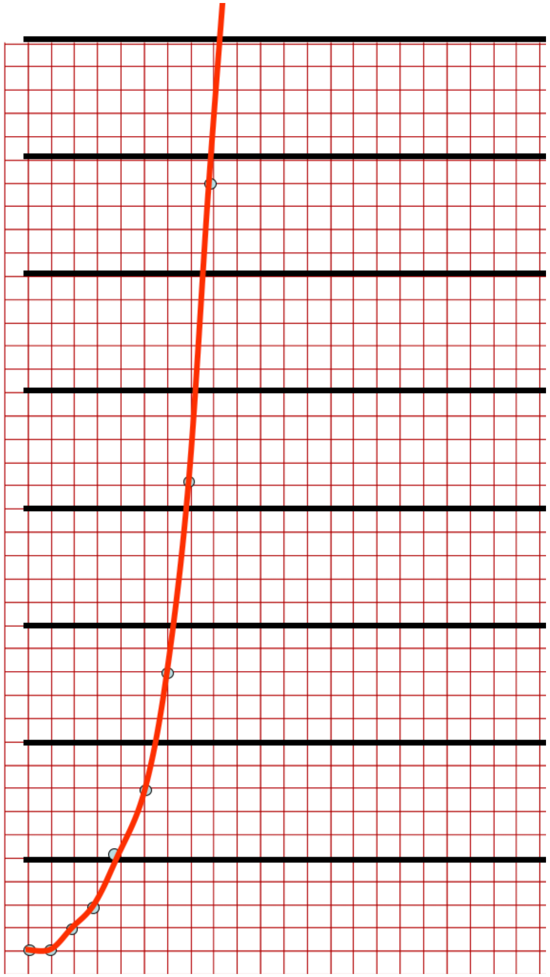
Mature	Young	Total
0	1	1
1	0	1
1	1	2
2	1	3
3	2	5
5	3	8

$$x_{i+1} = x_i + x_{i-1}$$

Fibonacci (Leonardo of Pisa 1202)

1, 1, 2, 3, 5, 8, …

Question: What would the forecast error be if there had been two pairs on day one? (Measurement error).



"effective exponential"

Question: If $x_i = \alpha x_{i-1}$ what is α?

$$\alpha = (1 + \sqrt{5})/2$$

"The golden mean"

Average (effective) exponential growth

Recall what we did with the Rabbit Map: assume exponential growth of an initial uncertainty: how would one quantify it give |ε(t)| ?

ε_0

$\varepsilon_1 = 2\varepsilon_0$

$\varepsilon_2 = 3\varepsilon_1$

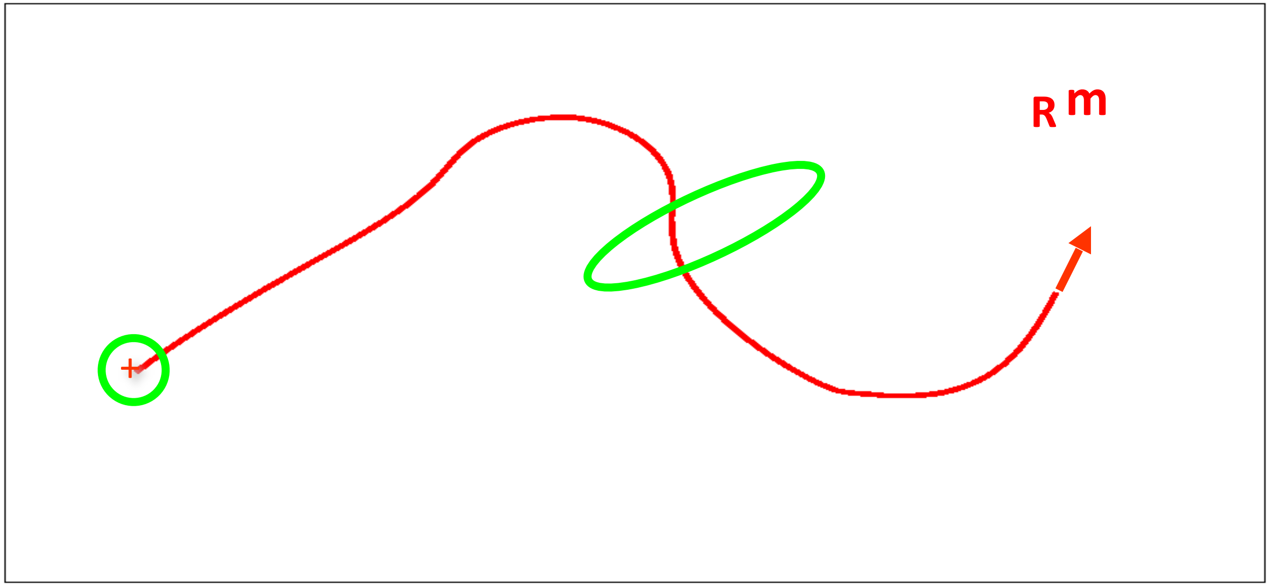
$\varepsilon_3 = 4\varepsilon_2$

$\varepsilon_4 = \dfrac{1}{3}\varepsilon_3$

$\varepsilon_5 = 4\varepsilon_4$

Linear Algebra in 10^7 Dimensions

Suppose we have already estimated the current state of the atmosphere



…and a trajectory (forecast) from that model initial condition.

Under a smooth dynamic, an infinitesimal sphere of uncertainty will evolve into an ellipse; we might sample preimages of the "leading" axes of this ellipse.

Growth of Infinitesimal无穷小 Uncertainties
(or: Why care about linear algebra?)

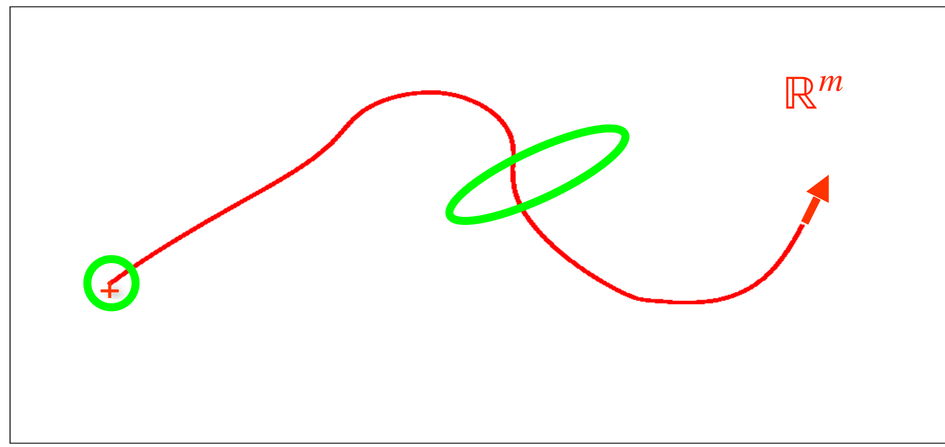
$$
\begin{aligned}
\epsilon_{i+1} &= F(x_i + \epsilon_i) - F(x_i) \\
&= [F(x_i) + J(x_i)\epsilon_i + o(\epsilon_i)] - F(x_i) \\
&\approx J(x_i)\epsilon_i
\end{aligned}
$$

$\epsilon, x \in \mathbb{R}^m$

Taylor expansion

Similarly

$$
\begin{aligned}
\epsilon_{i+2} &\approx J(x_{i+1})\epsilon_{i+1} \\
&\approx J(x_{i+1})J(x_i)\epsilon_i
\end{aligned}
$$



Define the Linear Propagator M to be $M^{(k)} = J(x_{k-1}) \cdot J(x_{k-2}) \ldots J(x_1) \cdot J(x_0)$

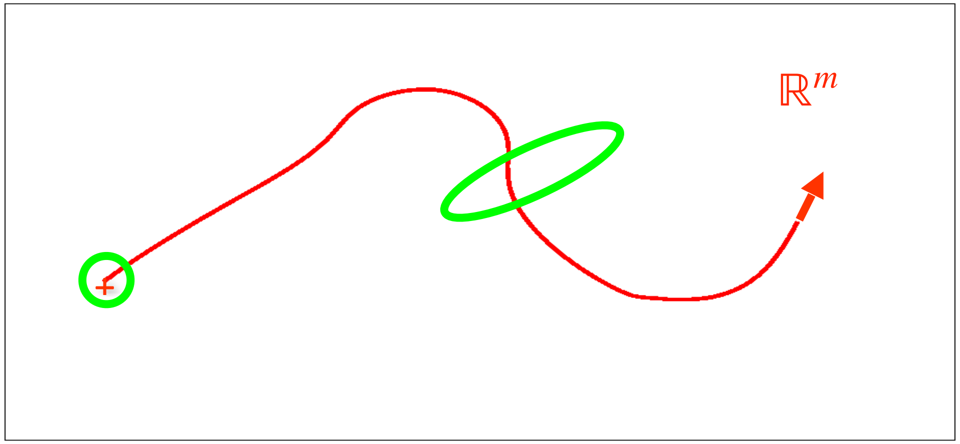
$$
\epsilon_k \approx M^{(k)}\epsilon_0 = J(x_{k-1}) \cdot J(x_{k-2}) \ldots J(x_1) \cdot J(x_0)\epsilon_0
$$

Growth of Infinitesimal Uncertainties
(or: Why care about linear algebra?)

$$\epsilon_k \approx M^{(k)}\epsilon_0 = J(x_{k-1}) \cdot J(x_{k-2}) \ldots J(x_1) \cdot J(x_0)\epsilon_0$$

Three points to remember:

1. As long as $\epsilon \approx 0$ it is not a problem.
   If $\epsilon > 0$ this linearization idea fells apart.
2. Matrix multiplication does not commute.
3. Model error $F_{model}(x) \neq F_{system}(x)$



Singular value decomposition

We have for any $n \times m$ matrix A

$$A = U \Sigma V^{\mathrm{T}},$$

where

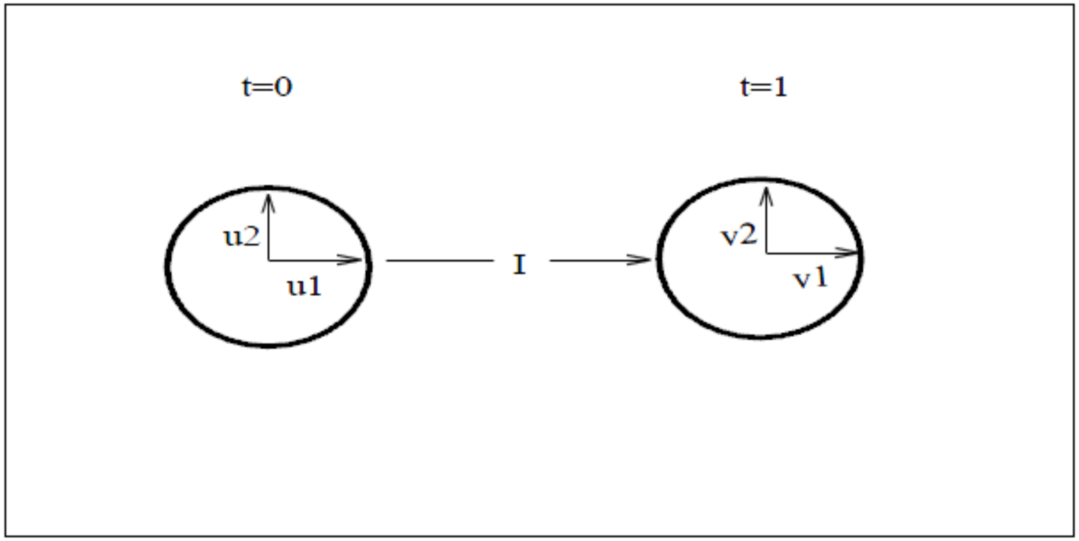
- U is $n \times n$ and orthogonal, so that $U^{\mathrm{T}}U = I$
- Σ is $n \times m$ and diagonal, with elements $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_j$, where $j = min(n, m)$
- V is $m \times m$ and orthogonal, so that $V^{\mathrm{T}}V = I$

The U and V matrices correspond to changes of coordinate and Σ to stretching/shrinking.

What linear propagator M does to the infinitesimal uncertainty is simply rotating, stretching and shrinking.

Singular value decomposition

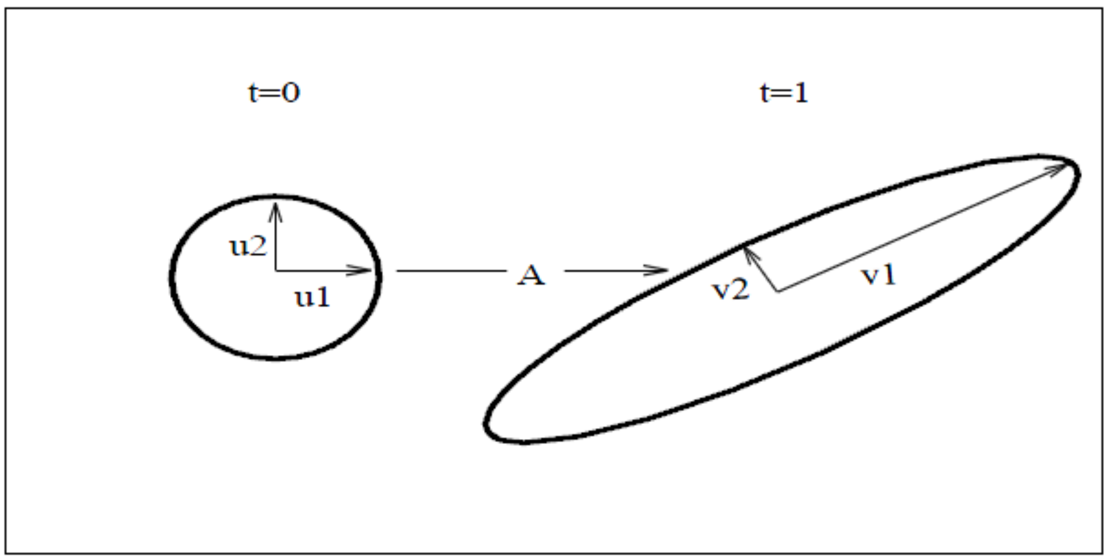
Suppose A is the identity matrix: $A = I$, then the axes and the noise ball are unchanged, as shown in figure .



Evolution of a unit ball under the identity propagator I. The directions u1 and u2 are actually degenerate since there is no natural orientation with the ball.

Singular value decomposition

Under a general linear propagator A we observe the following (see Figure). Here the vectors $\mathbf{u}_i$ form a basis for the initial space and the $\mathbf{v}_i$ span the target space. The linear propagator A essentially maps $\mathbf{u}_i \rightarrow \mathbf{v}_i$ and the σ_i correspond to the rate of expansion/contraction.



Evolution of a unit ball under the linear propagator A

rotating, stretching and shrinking!

What does a Singular-Value Decomposition tell us about Dynamics?

Matrix: linear propagator M

Singular vectors: v_i , u_i

Singular values: σ_i

Orthogonality: $v_i \perp v_j$ and $u_i \perp u_j$, $i \neq j$

Rank ordering: $\sigma_i \geq \sigma_j \geq 0, i \geq j$

Geometry: $Mu_i = \sigma_i v_i$

If $\sigma_1 < 1$ then no uncertainty can grow

Linear Propagator —> Lyaponuv exponent

When J is constant, we can contrast instantaneous growth rates and effective rates. On Henon, of course, J is a function of location!

Linear system: $M^{(k)} = J \cdot J \cdot J \ldots J$ (same J)
For Henon: $M^{(k)} = J(x_{k-1}) \cdot J(x_{k-2}) \ldots J(x_1) \cdot J(x_0)$ follow a trajectory.

$$\epsilon_k \approx M^{(k)}\epsilon_0 = M^{(k)}(\sum \alpha_i u_i) = \sum \sigma_i^{(k)} \sigma_i v_i \approx \sigma_1^{(k)} \alpha_1 v_1$$

(if we only consider the largest uncertainty growing direction and assume $\alpha_1 \neq 0$)

$$\frac{|\epsilon_k|}{|\epsilon_0|} \approx \frac{\sigma_1^{(k)}\alpha_1}{|\epsilon_0|} \propto \sigma_1^{(k)} \equiv e^{\lambda^{(k)}k}$$

$$\lambda^{(k)} = \frac{1}{k} \log \sigma_1^{(k)}$$ so called finite time Lyaponuv exponent

Note $\lambda^{(k)}$ represents effective growth rate!

Linear Propagator —> Lyaponuv exponent

$$\epsilon_k \approx M^{(k)}\epsilon_0 = M^{(k)}(\sum \alpha_i u_i) = \sum \sigma_i^{(k)}\sigma_i v_i \approx \sigma_1^{(k)}\alpha_1 v_1$$

(if we only consider the largest uncertainty growing direction and assume $\alpha_1 \neq 0$)

$$\frac{|\epsilon_k|}{|\epsilon_0|} \approx \frac{\sigma_1^{(k)}\alpha_1}{|\epsilon_0|} \propto \sigma_1^{(k)} \equiv e^{\lambda^{(k)}k}$$

$$\lambda^{(k)} = \frac{1}{k}\log \sigma_1^{(k)}$$ so called finite time Lyaponuv exponent

Note $\lambda^{(k)}$ is a function of x_0

In the limit $k \to \infty$, approaches the largest global Lyapunov exponent, Λ_1

Oseledec's (Multiplicative Ergodic) Theorem

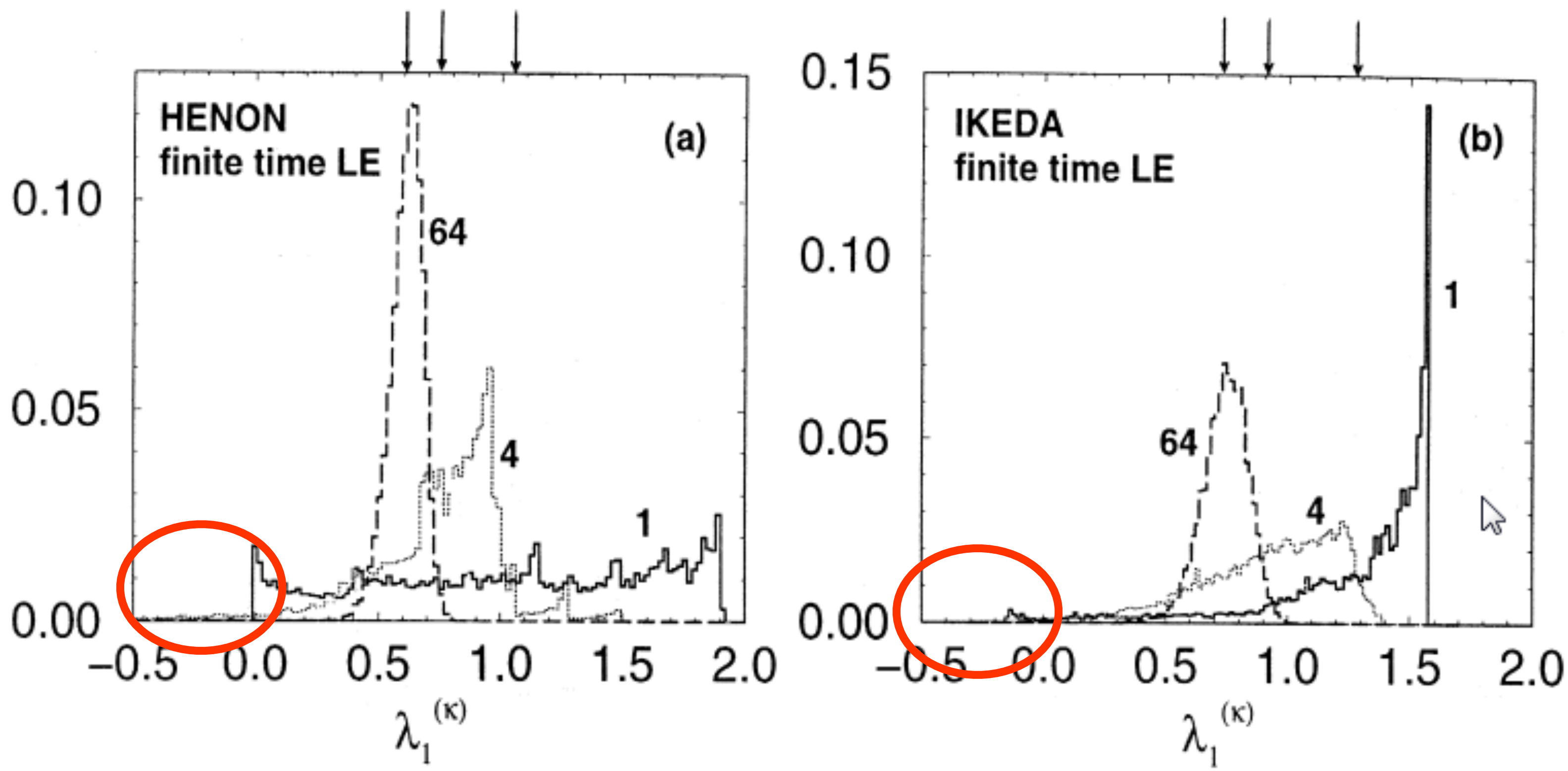
Oseledec proved that Λ depends on the orientation of ϵ_0 but not on x_0. Given that matrix multiplication does not commute, this results is impressive (and holds well beyond chaotic dynamics)

Having a positive global Lyapunov exponent is a necessary condition for chaos

What does a Singular-Value Decomposition tell us about Dynamics?

$$\lambda^{(k)} = \frac{1}{k} \log \sigma_1^{(k)}$$ so called finite time Lyaponuv exponent

Distributions of finite time Lyaponov exponent：



Growth differs with time and location

Ziehmann, C., Smith, L.A. and Kurths, J. (2000) 'Localized Lyapunov exponents and the prediction of predictability', Physics Letters A, 271 (4): 237-251.

What does a Singular-Value Decomposition tell us about Dynamics?

$$M^{(k)} = J(x_{k-1}) \cdot J(x_{k-2}) \ldots J(x_1) \cdot J(x_0)$$

$$\lambda^{(k)} = \frac{1}{k} \log \sigma_1^{(k)}$$ so called finite time Lyaponuv exponent

In Henon, does it matter whether or not we follow a trajectory?

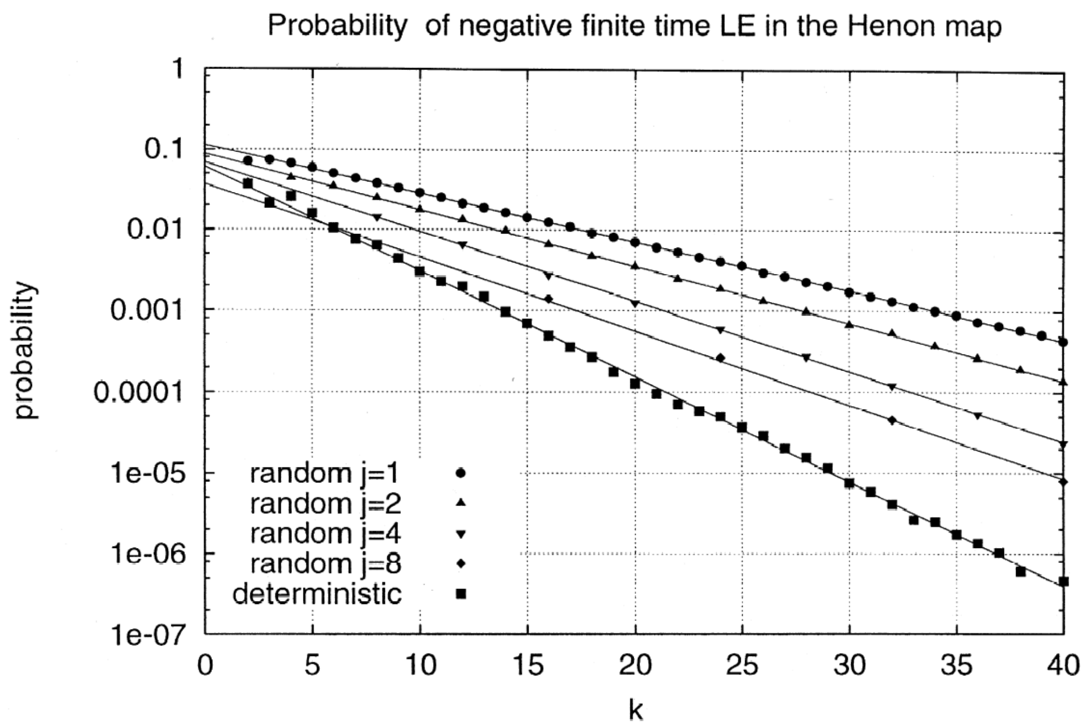


Fig. 6. For the Henon system ($a = 1.4, b = 0.3$), the fraction of points with $\lambda_1^{(k)} < 0$ as a function of k in the deterministic case (squares). Also shown are the results for random matrices, where the matrices are drawn from the distribution of the linear propagators of the Henon map, that is $\mathscr{M}^{(j)}$, for $j = 1,2,4$ and 8. The solid lines reflect the best fit to an exponential decay over the range $8 \le k \le 40$. For large k, about 2^{30} iterations of the map where considered in the deterministic case. Note, that determinism is a strong constraint reducing the likelihood of finding negative finite time exponents.

In the Ikeda Map, there is a regions where $\sigma_1 < 1$ what does this mean?

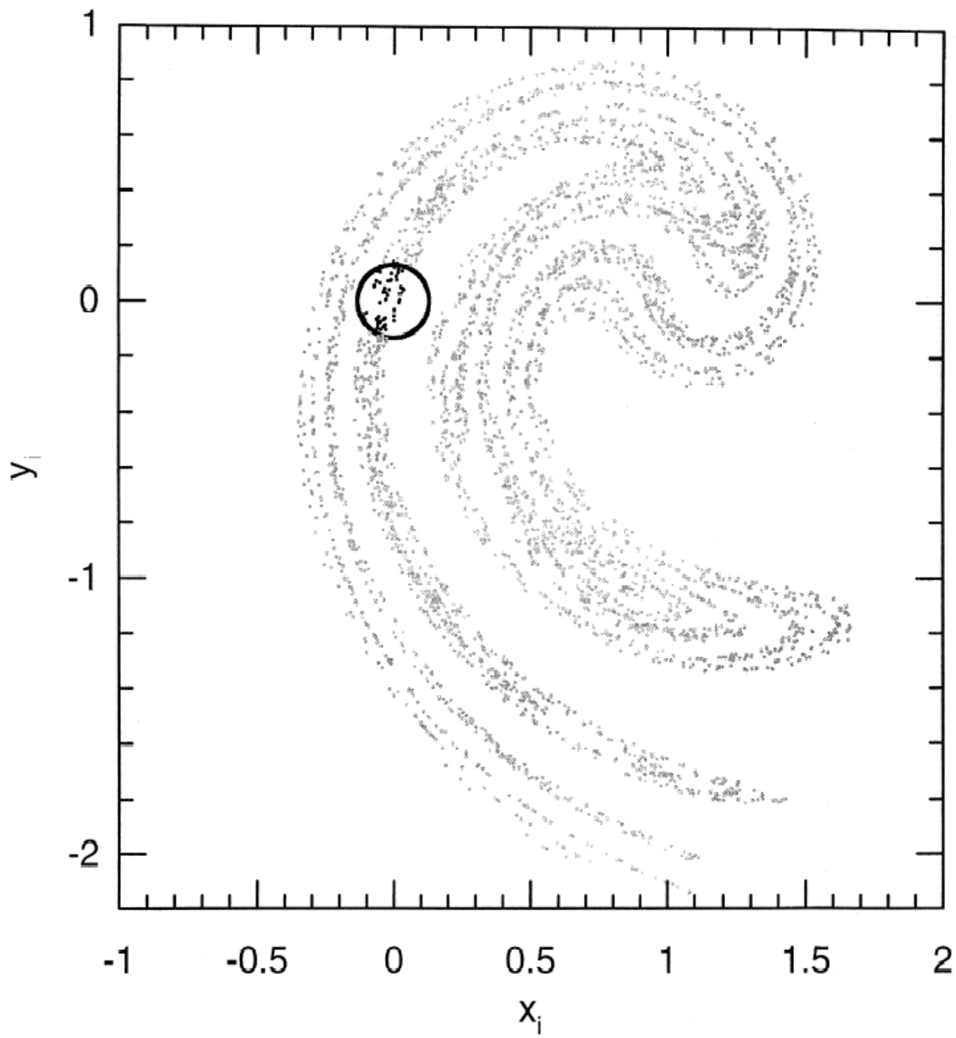


Fig. 3. Regions of decreasing uncertainty in the Ikeda system. Points on the attractor are colored grey if $\lambda_1^{(1)}(\boldsymbol{x}) > 0$, black otherwise. Within the circular region near the origin, $\lambda_1^{(1)}(\boldsymbol{x}) < 0$ for all $\boldsymbol{x}$.

Lyaponuv exponent can be very misleading

Lyaponuv exponent can be very misleading

9.83s

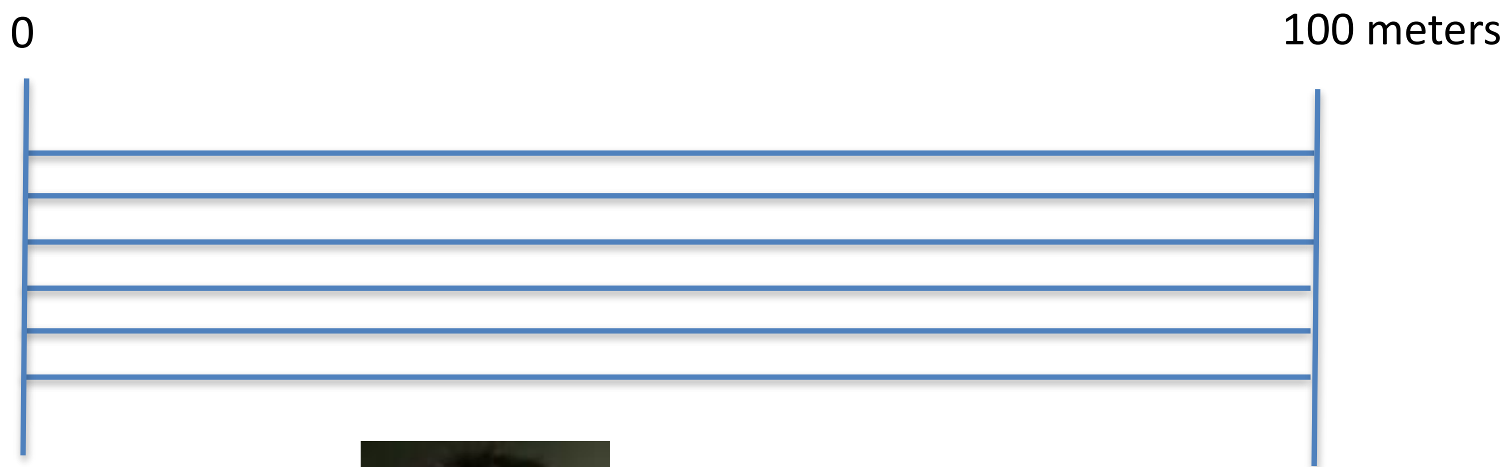
Lyaponuv exponent can be very misleading

0

100 meters

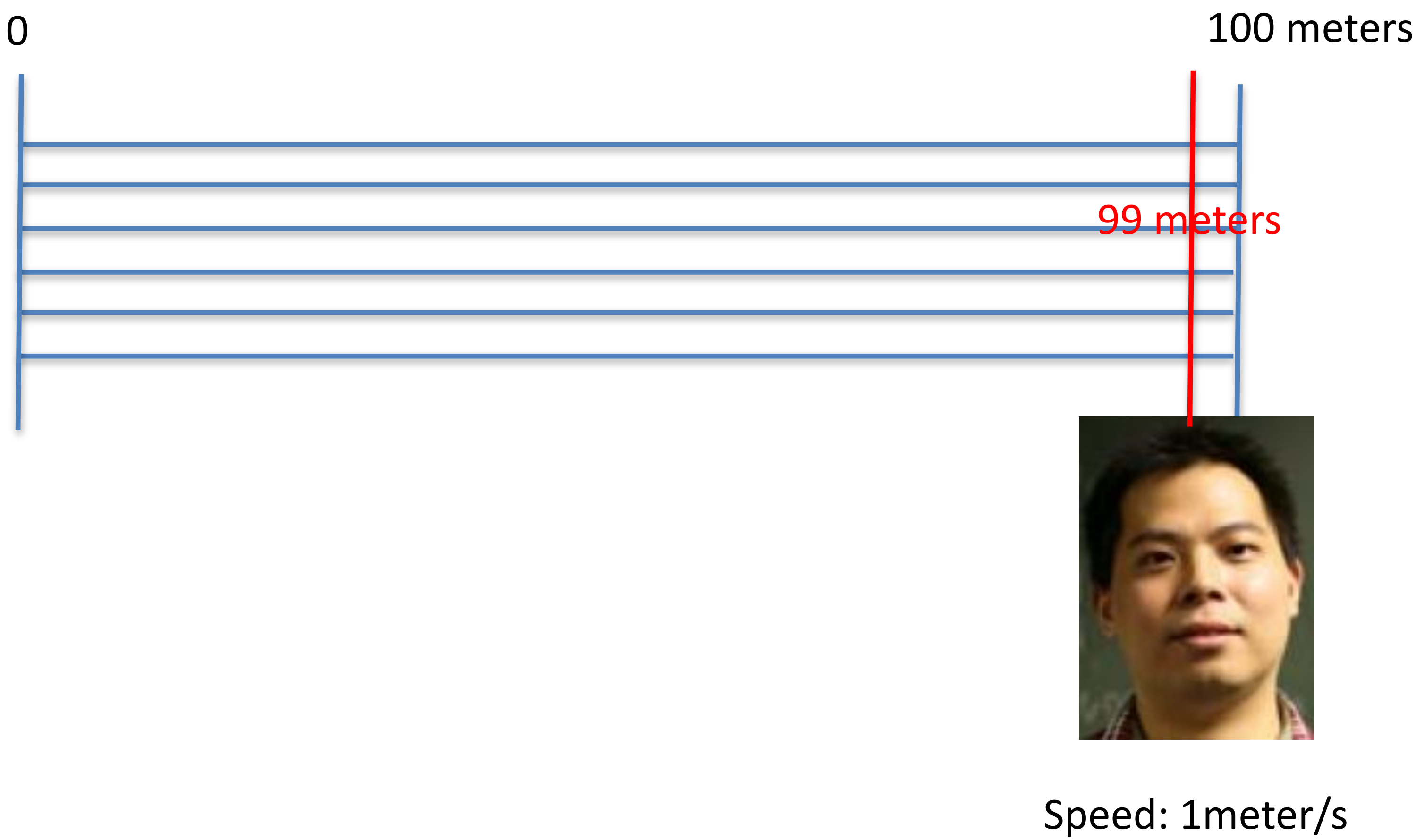
Speech: 1meter/s

Lyaponuv exponent can be very misleading



Speed: 1meter/s

Lyaponuv exponent can be very misleading



Lyaponuv exponent can be very misleading

Speed: 300000000 meter/s

Lyaponuv exponent can be very misleading

My average speed is much faster than Su. But it took me 10times longer to finish the race!

Speed: 300000000 meter/s

9.83s

August 20, 2021
1:58 PM BST
Last Updated 8 hours ago

United Kingdom

England's COVID R number estimate rises to 0.9 to 1.2

1 minute read

Reuters

q-pling time

Alternatively, we define the q-pling time $\tau_q(x_0, \epsilon_0)$ as at which the initial uncertainty ϵ_0 about x_0 has increased by

$$\tau_q(x_0, \epsilon_0) = \min_{t>0}\{t \,|\, \| F_t(x_0 + \epsilon_0) - F_t(x_0) \| \geq q \| \epsilon_0 \|\}$$

Whenever an initial orientation of ϵ is well defined for each x, we have

$$\tau_q(\| \epsilon_0 \|) = \langle \tau_q(x, \epsilon) \rangle_x$$

where the average is taken over all points x on the attractor.

q-pling time provides a more direct measure of a prediction time scale through the average of the minimum time required for an uncertainty to increase by a factor of q.

when $q = 2$, doubling time

Most people think chaos is unpredictable, but exponential-on-average does not mean exponential!

Red: $\tau_2 < 0.15$

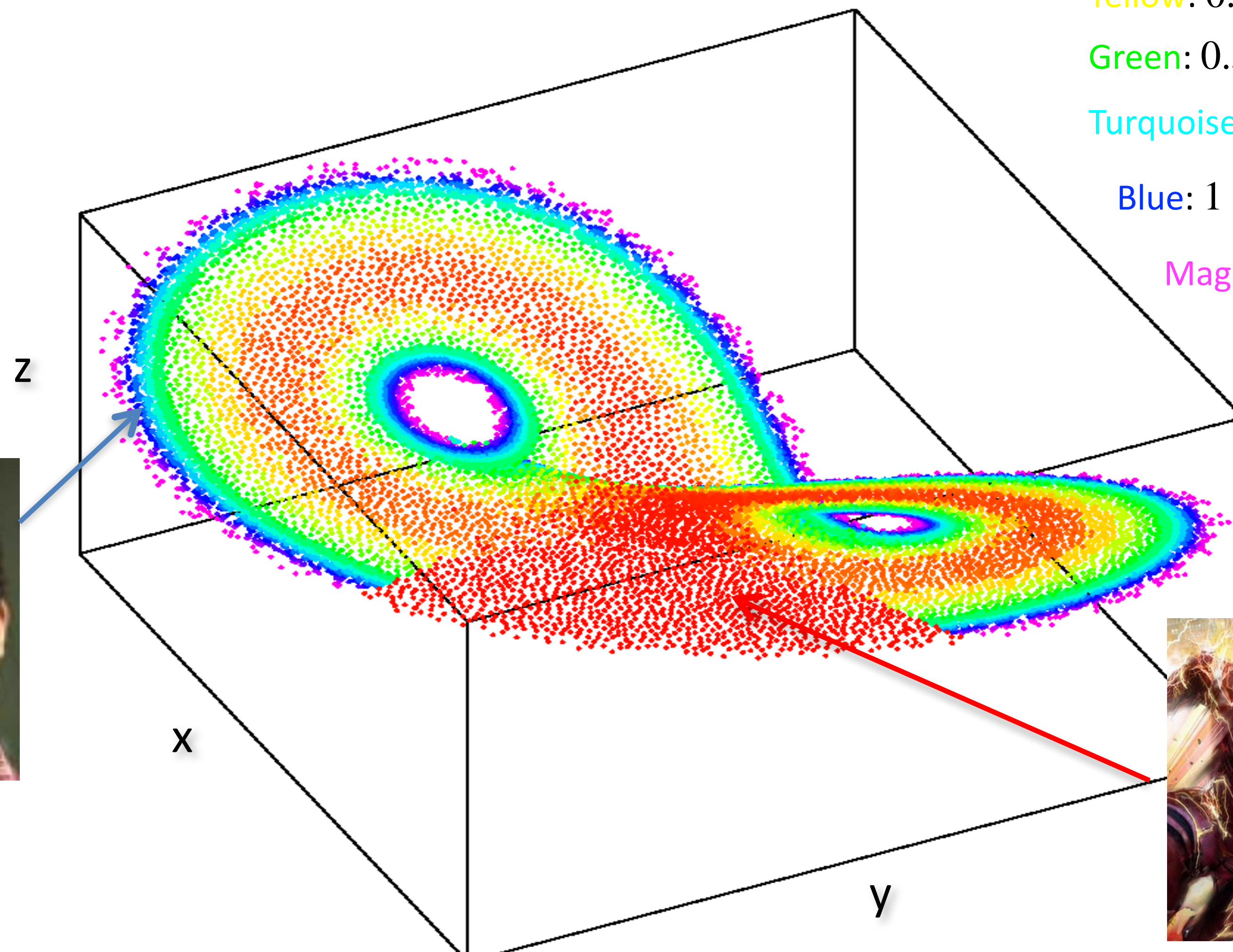
Yellow: $0.15 \leq \tau_2 < 0.5$

Green: $0.5 \leq \tau_2 < 0.8$

Turquoise: $0.8 \leq \tau_2 < 1$

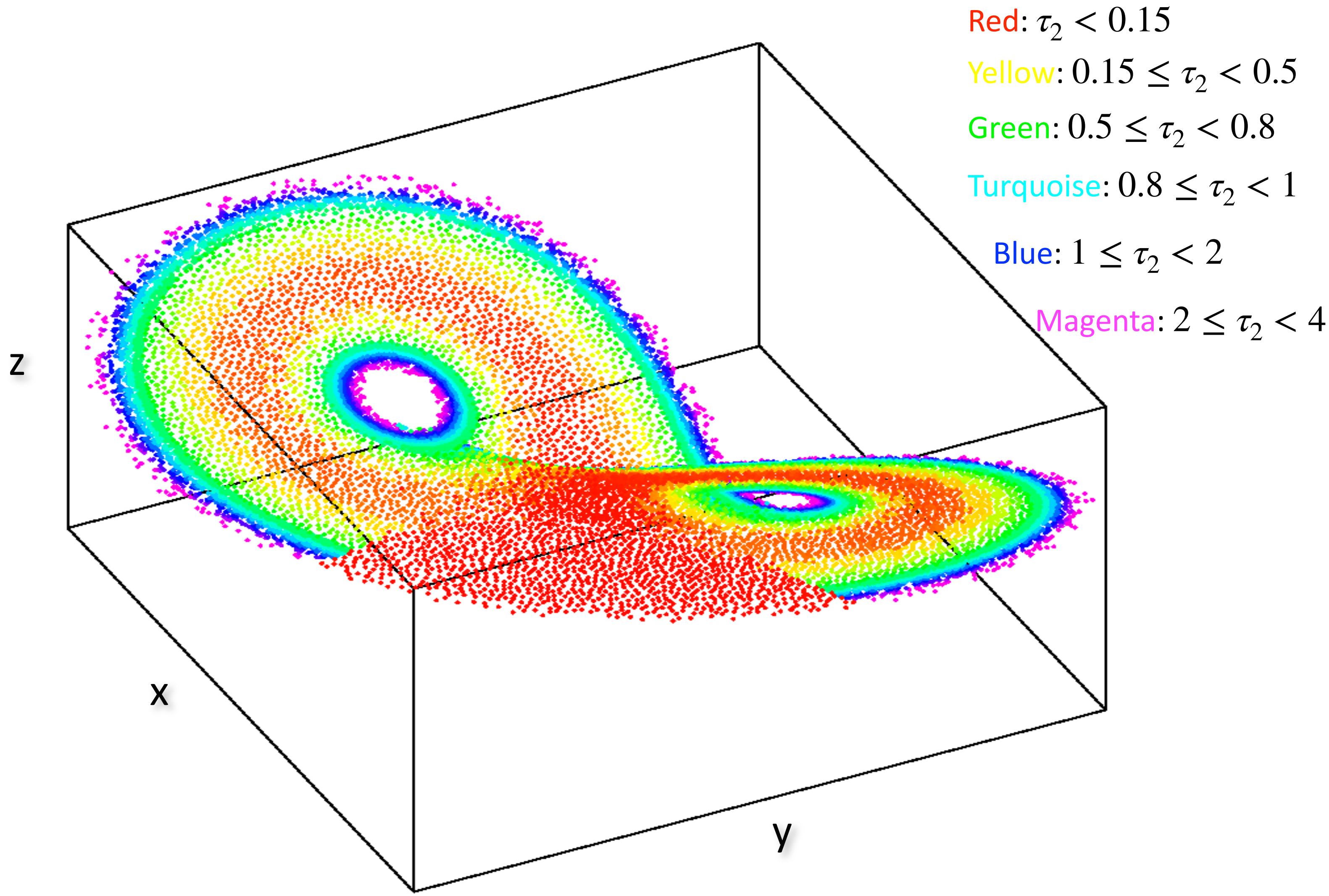
Blue: $1 \leq \tau_2 < 2$

Magenta: $2 \leq \tau_2 < 4$



Local Doubling Times on Lorenz '63

Most people think chaos is unpredictable, but exponential-on-average does not mean exponential!



Local Doubling Times on Lorenz '63

Of course it is hard to get people to bet with you on the Lorenz Differential Equations!
And how would we know if "today" was a good day to bet?

And what do doubling time tell us in these systems?

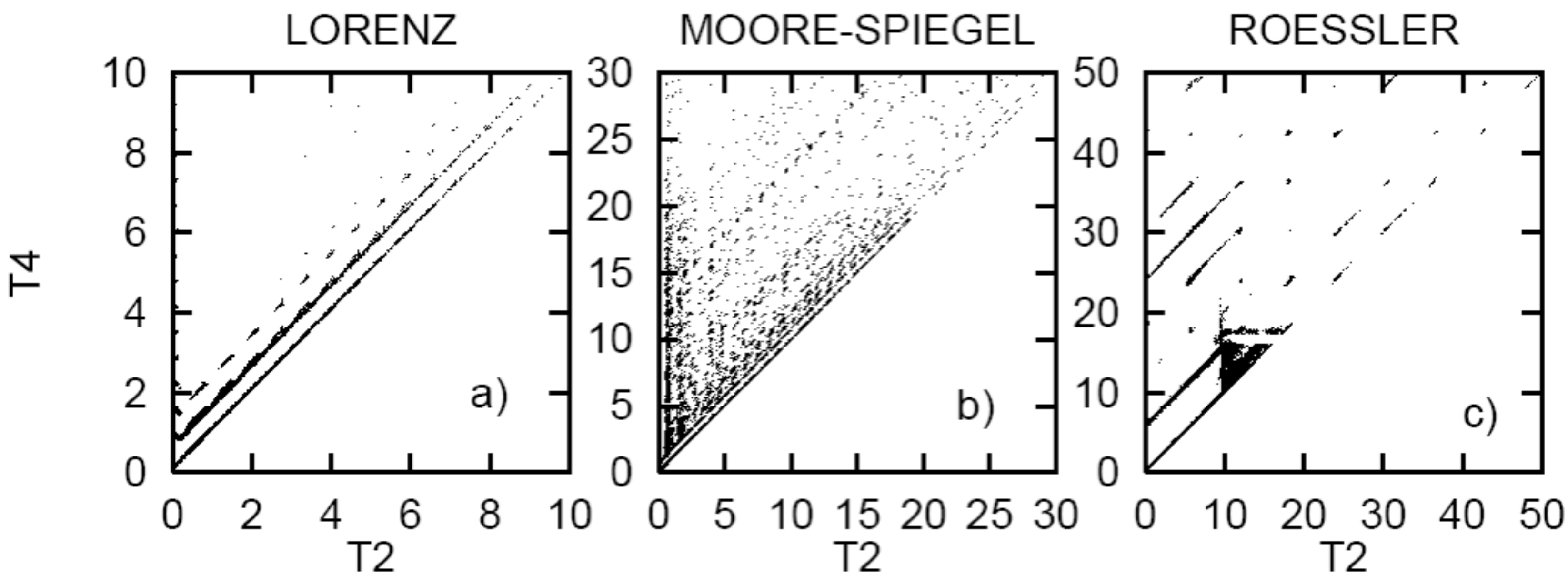
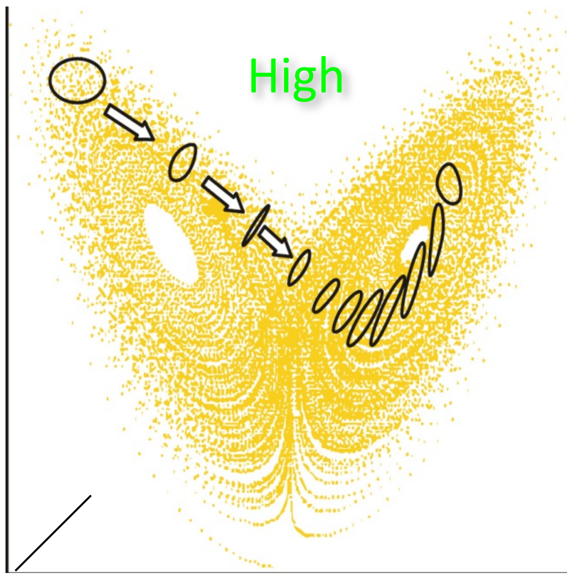


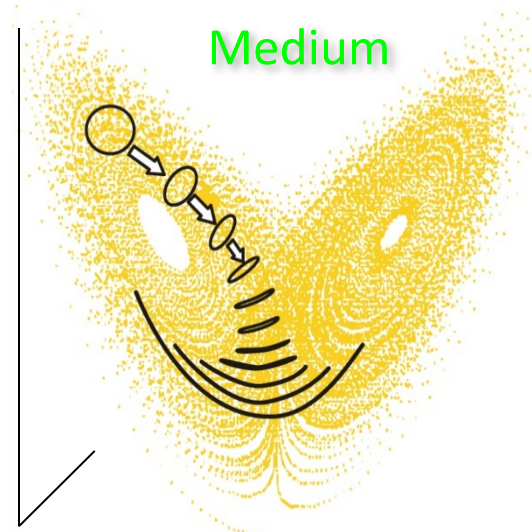
Figure 8. A comparison of the doubling time and quadrupling time of $N = 2^{16}$ initial conditions for the Lorenz, the Moore-Spiegel and the Rössler systems. In each case, the initial uncertainty is in the Lyapunov orientation.

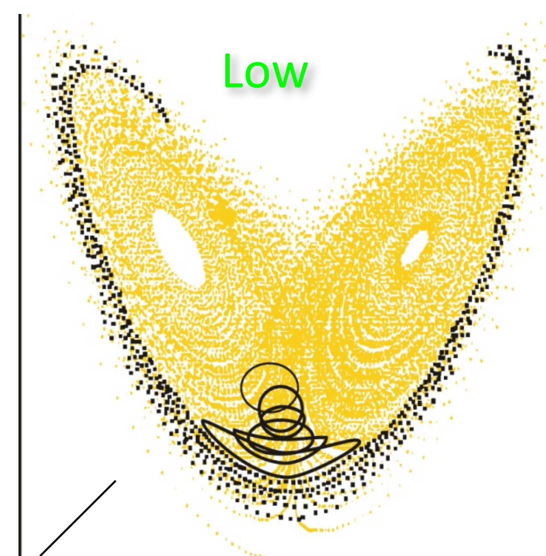
Predictability

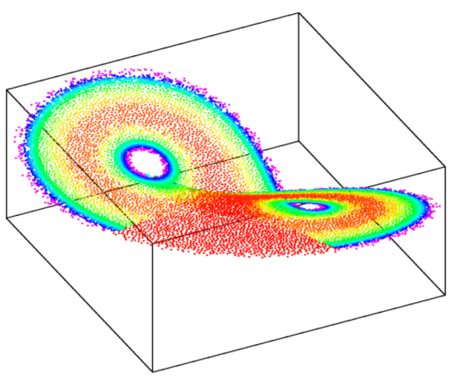
CAN YOU SEE THIS 3-D PICTURE?

Scientific Basis for Ensemble Prediction









We would like to quantify day to day variations in predictability…

Scientific Basis for Ensemble Prediction