ESTIMATION OF CHAOTIC STATES USING EXTENDED KALMAN FILTERS

Miguel Almeida[1], Paulo Oliveira[2]

[1] *Instituto de Telecomunicações*
[2] *Instituto de Sistemas e Robótica*
[1, 2] *Instituto Superior Técnico, Lisbon, Portugal*

Abstract: The knowledge on the state of chaotic models for the human neuron axon voltage activity is important to understand brain activity. An approach that consists of estimating the state of individual oscillators (individual neurons or populations of synchronous neurons), from sensor measurements providing combinations of state variables, can be chosen to tackle that problem. The number of sensors is typically much smaller than the number of underlying oscillators. In this paper a solution to this problem is exploited, resorting to Extended Kalman Filters, when one or more chaotic systems (Lorentz attractors) are considered.

Keywords: chaotic behaviour, estimators, oscillators, attractors, extended Kalman filters, synchronization, observability

1. INTRODUCTION

1.1 Motivation

In recent years the neurons of the human brain have been studied from a dynamic system perspective. There is evidence suggesting that a neuron's output, defined as the voltage difference between the outside and the inside of its axon, can be described by continuous-time chaotic systems (Izhikevich, 2007).

In some cases, direct measurements of these signals can be obtained from electrodes placed directly in the neurons. However, in a typical situation, only indirect measurements corrupted by noise are available. For example, in EEG, these are obtained from electrodes in the individual's scalp. Moreover, these measurements typically contain only a fraction of signals when compared to the number of underlying oscillators (the typical number of sensors is 64 or 128 in EEG, and up to 300 in MEG, while the brain has approximately 10^{11} neurons).

It is important to estimate the state of the underlying oscillators when the number of measurements is much smaller than the number of state variables assumed in the brain, i.e. the product of the number of the state variables of the neuron by the number of neurons. Even when the connectivity between neurons is discarded, this problem is very hard because the number of available measurements is always a small fraction of the number of state variables in the brain.

In this work we aim to provide some insight into a possible solution for this problem using Extended Kalman Filters (EKFs). This work is pedagogical in nature, and several possible improvements are listed later (see Section 6).

1.2 Objective

In this work the objectives are twofold. First, we aim to demonstrate that it is possible to use an EKF to successfully estimate the state of a chaotic oscillator (the Lorentz attractor), even when the measurements are not invertible, i.e., there are less measurements than state variables. This is done in Section 4. Second, we will show that, again with EKFs, it is possible to estimate the state of two chaotic oscillators (Lorentz attractors), when the measurements are not invertible and also not separable, meaning that each sensor measures linear combinations of variables coming from both oscillators. This is done in Section 5.

1.3 Notation

The notation used in this text is typical in the literature on dynamic systems and is reviewed here for completeness. Throughout this text, we will use the following system of equations:

$$\begin{cases} \dot{x} = f(x) + \xi \\ z = g(x) + \theta \end{cases}$$

where x denotes the n-dimensional state of the system, $f(x)$ is a non-linear function, ξ is the zero mean white Gaussian plant noise, z is the m-dimensional measurement, $g(x)$ is a non-linear function and θ is white noise (called the sensor noise). Note that the possible time dependence of all these variables and functions has been omitted for clarity.

1.4 Paper structure

We begin by introducing two concepts that are important for this work: the Lorentz attractor in Section 2, and Kalman filters in Section 3. In this last section we introduce both the Kalman filter for linear systems and the Extended Kalman filter for non-linear systems. Section 4 summarizes the methods and results obtained for a simulated dataset consisting of one oscillator, and Section 5 for two oscillators. We discuss the results in Section 6, where possible improvements and extensions of this work to be addressed in the future are outlined, followed by some conclusions in Section 7.

2. THE LORENTZ ATTRACTOR

The Lorentz attractor is one of the most widely known continuous-time deterministic dynamical systems. It has a three-dimensional state variable, governed by the following state equation,

$$\begin{cases} \dot{x}_1 = \sigma(x_1 - x_2) \\ \dot{x}_2 = x_1(\rho - x_3) - x_2 \\ \dot{x}_3 = x_1 x_2 - \beta x_3 \end{cases} \qquad (1)$$

parametrized by three parameters (β,ρ,σ). It is known that for some choices of these parameters the system exhibits chaotic behaviour, e.g. for β = 8/3, ρ = 28, σ = 10.

This system is called an *attractor* because its orbits are attractive (Devaney, 2003), which means that for any initial point, the trajectory of the system is bounded. After a transient period, the trajectory of the Lorentz attractor falls into one of two “surfaces”, as shown in Fig. 1.

3. KALMAN FILTERING

3.1 Linear Kalman filter

The optimal solution to estimate the state of linear dynamic systems is the Kalman Filter, introduced 50 years ago. A detailed description of this technique can be found in (Anderson and Moore, 1979); here we present a more intuitive view for brevity.

Given the current state estimate $\hat{x}(t)$, the Kalman filter for linear and time-invariant systems (systems where $f(x) = \mathbf{A}x,\ g(x) = \mathbf{C}x$, where **A** and **C** are constant matrices) provides estimates on what the measurements should be ($z_{predicted}(t) = \mathbf{C}\hat{x}(t)$) and compares this with the actual measurement *z*(*t*). If $z_{predicted} < z$, the measurements are being underestimated and the filter should correct the state estimate derivatives to increase the predicted measurements. In mathematical terms, the evolution of the state should be corrected according to a law of the following type:

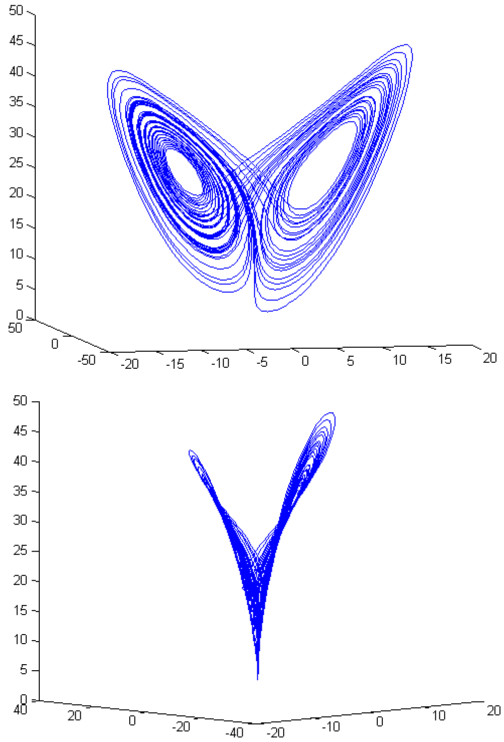

Fig. 1. Plot of the trajectory of the Lorentz attractor from two different viewpoints. The bottom figure shows the two “surfaces”.

$$\dot{\hat{x}}(t) = \mathbf{A}\hat{x}(t) + \mathbf{H}(t)\left[z(t) - z_{predicted}(t)\right].$$

It remains to be seen how to compute the gain matrix **H**(*t*) (called the Kalman gain). If Gaussian distributions are assumed for the plant and sensor noises, and for the distribution of the initial state of the system, and if all these quantities are independent from each other, then it can be shown that to minimize the expected square error

$$\mathrm{E}\left[\left\| x(t) - \hat{x}(t) \right\|^2 / Z(t)\right],$$

where *Z*(*t*) is the set of all measurements up to and including time *t*, the Kalman gain should be computed as

$$\mathbf{H}(t) = \mathbf{\Sigma}(t)\mathbf{C}\mathbf{\Theta}^{-1},$$

where **Θ** is the covariance of the sensor noise and **Σ**(*t*) is the covariance of the state estimate at time t, which follows the following propagation rule:

$$\dot{\mathbf{\Sigma}}(t) = \mathbf{A}\mathbf{\Sigma}(t) + \mathbf{\Sigma}(t)\mathbf{A} + \mathbf{\Xi} - \mathbf{\Sigma}(t)\mathbf{C}^{\mathrm{T}}\mathbf{\Theta}^{-1}\mathbf{C}\mathbf{\Sigma}(t).$$

Note that the equation for the propagation of **Σ** can be solved off-line. The Kalman filter is, then, the best estimator in the sense that it is unbiased and has the least mean squared error.

3.2 Extended Kalman filter (for nonlinear systems)

The extended Kalman filter (EKF) is the generalization of the Kalman filter for nonlinear systems. Suppose that the functions f and g are differentiable in the system model introduced in Section 1.3, such that the model can be approximated to a linearized version that can be written as

$$\dot{x} = \hat{\mathbf{A}}(x)x + \xi$$
$$z = \hat{\mathbf{C}}(x)x + \theta$$

where

$$\hat{\mathbf{A}}(x) = \frac{\mathrm{d}f}{\mathrm{d}x}, \quad \hat{\mathbf{C}}(x) = \frac{\mathrm{d}g}{\mathrm{d}x}$$

If the state x was known exactly, this would be entirely equivalent to the real model. In practice, the real state x is unknown, but its estimate $\hat{x}$ is known. This leads to the following approximation:

$$\hat{\mathbf{A}}(\hat{x}) = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x=\hat{x}}, \quad \hat{\mathbf{C}}(\hat{x}) = \left.\frac{\mathrm{d}g}{\mathrm{d}x}\right|_{x=\hat{x}}$$

The EKF uses this assumption. The equations that govern the EKF are exactly equal to the ones for the KF, replacing everywhere $\mathbf{A}$ and $\mathbf{C}$ by $\hat{\mathbf{A}}(\hat{x})$ and $\hat{\mathbf{C}}(\hat{x})$. The propagation of $\mathbf{\Sigma}$ must in this case be solved on-line.

Contrary to the KF case, there is no proof that the EKF is the best estimator in any sense. It is not guaranteed to be stable, robust, or to have good performance. Its advantages are its simplicity, the strong motivation (as a generalization of the KF) and the light computational demand compared to other observers for nonlinear systems.

4. PART 1: ONE OSCILLATOR

4.1 Methodology

A Lorentz attractor with parameters $\beta = 8/3$, $\rho = 28$ and $\sigma = 10$ was implemented in SIMULINK 7. With these parameters, the Lorentz attractor exhibits chaotic behaviour as previously mentioned. An EKF was designed and implemented to model the exact nonlinearity (i.e., no approximation for $f(x)$ is used). The attractor and EKF were simulated for t in [0,10]. The measurement was arbitrarily chosen as $g(x) = x_1$ (the first state variable).

Although no plant noise was implemented, the EKF assumes a small value (10^{-6}) for the variance of each state variable to work correctly.

4.2 Results

The above method was repeated for several initial conditions of the attractor and EKF, and for several noise levels. Generally, the EKF performs well for low levels of noise (variance of the sensor noise below ~10^{-4}, although this threshold seems to vary with the initial condition), and is able to track the attractor's state with minimal error. As an illustration, we present the case with initial state (-6, -6, 20), initial estimate (-6, -6, 20.1) and sensor noise variance 10^{-6}.

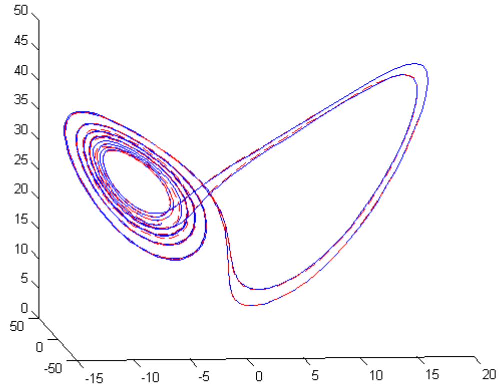

Fig. 2: 3D plot of the real state (blue) and EKF estimate (red, dashed). It is clear that the EKF manages to follow the real system quite well, despite the initial offset.

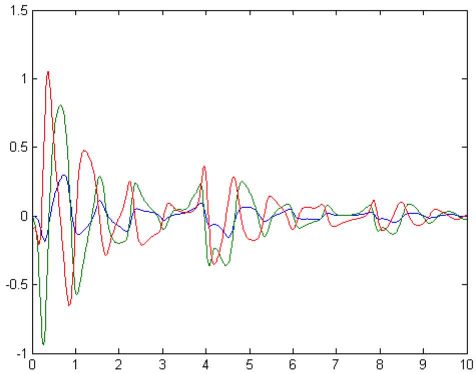

Fig. 3: Estimation errors for the 3 state variables The fact that the initial state is different in the real system and in the EKF causes some initial large errors, but the EKF manages to "catch up" with the real system and after some transient behaviour it performs well.

5. PART 2: TWO OSCILLATORS

5.1 Methodology

Two Lorentz attractors, with the same parameters as above, were implemented in SIMULINK 7 as one single system with six state variables. An EKF with six state variables was also implemented with no approximation on $f(x)$. Given the limited computation power available, the system was simulated for t in [0,5]. The measurement was chosen as having three scalar values, each being a fixed, known linear combination of all six state variables with random coefficients (uniform between 0 and 1). In other words, $g(x) = \mathbf{C}x$ where $\mathbf{C}$ is a 3x6 random matrix.

Just like in Part 1, a small value (10^{-6}) is assumed in the EKF for the plant noise, but not actually implemented, in order to make the estimator work correctly.

5.2 Results

Contrary to Part 1, in this case the performance of the EKF depends heavily on the initial conditions of the state variables and their estimates, and on the number of sensors. The EKF will sometimes perform poorly even for very low levels of noise. For this reason, we chose to present a case where the EKF performs badly and one where it performs well to illustrate the kind of results that can be obtained. The main difference between these two cases is the number of sensors: there are two in the "bad example" and three in the "good example".

Bad performance example: In this case it will be seen that the EKF can perform poorly, and even diverge (despite the fact that the original system is bounded).

For this case the initial state was randomly generated as (-1.7681, -3.8994, 20.1202, -0.5350, -7.8059, 8.9165), the initial state estimate was taken exactly equal to the initial state, and the mixing matrix C was randomly generated as

$$\begin{bmatrix} 0.6261 & 0.0194 & 0.1334 & 0.8934 & 0.2704 & 0.6550 \\ 0.2341 & 0.1454 & 0.3737 & 0.6015 & 0.3395 & 0.4326 \end{bmatrix}$$

The noise covariance was chosen diagonal with diagonal elements equal to 10^{-7} for generating the noise, but the EKF assumed it as 10^{-6} (numerical issues arise if 10^{-7} is used here, but using 10^{-6} for the actual noise causes the EKF to not be able to follow the attractor).

The plots below (Fig. 4) show the trajectory of the first and second oscillators and the point where the EKF diverges from the real state. In this case, it can be seen that this separation between the real system and the EKF occurs when one of the individual attractors is passing from one of the "surfaces" to the other one (in this case it was the first attractor). In our simulations, we have verified that this is a general property, i.e, **whenever the EKF diverges from the real state, it does so when one of the two attractors is passing from one "surface" to the other one**. The converse is not true, i.e., the attractors can switch from one "surface" to the other one without causing the EKF to diverge (see next example).

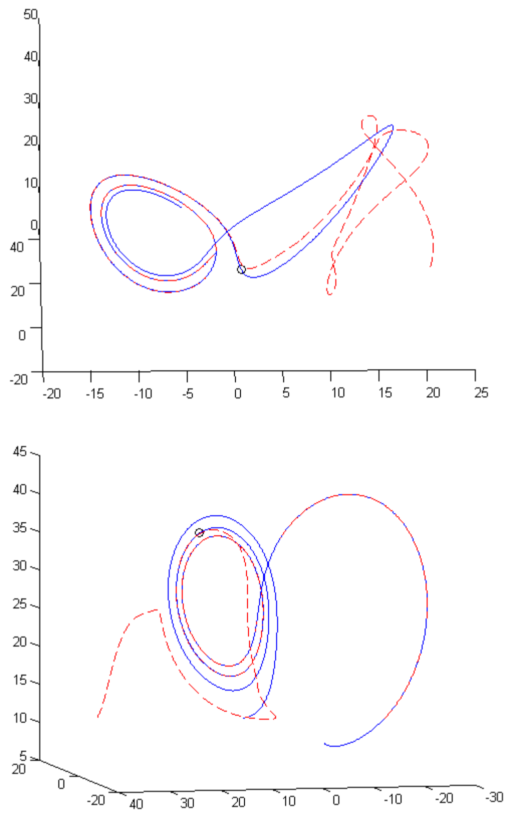

Fig. 4: 3D plot of the real state (blue) and EKF estimate (red, dashed). The first three variables (i.e. the first Lorentz attractor) are on the left, the last three (i.e. the second Lorentz attractor) on the right. The black circle marks approximately the point where the EKF "loses" the real state (t = 1.6). It can be seen that one of the attractors (the first one, on the left) was passing from one surface to the other one at that time. This is a general finding in our simulations.

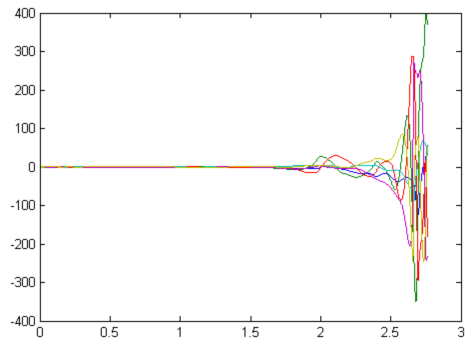

Fig. 5: Plot of the difference between the real state and the EKF estimate. It is clear that after t = 1.6 the EKF starts to perform very poorly and eventually diverges. Each different color represents one of the six state variables.

Good performance example: In this case the initial state was randomly generated as (-3.1286, -1.1406, 5.7074, -0.5008, -6.6860, 13.4181), the initial state estimate was taken exactly equal to the initial state, and the mixing matrix **C** was randomly generated as

$$
\begin{bmatrix}
0.9787 & 0.4711 & 0.0424 & 0.0967 & 0.7224 & 0.5186 \\
0.7127 & 0.0596 & 0.0714 & 0.8181 & 0.1499 & 0.9730 \\
0.5005 & 0.6820 & 0.5216 & 0.8175 & 0.6596 & 0.6490
\end{bmatrix}
$$

The noise covariance matrix was chosen diagonal with diagonal elements equal to 10^{-7} for generating the noise, but the EKF assumed it as 10^{-6} (numerical issues arise if 10^{-7} is used here, but using 10^{-6} for the actual noise causes the EKF to not be able to follow the attractor). This is a very small value, and our simulations show that greater values cause the EKF to lose track of the real state and eventually diverge.

With these choices, the EKF performs well and manages to follow the attractor with a maximum absolute error in each component below 0.2, and averaging 0.02. Note that the trajectory of the Lorentz attractor has typical values of each coordinate between 5 and 25, so this is a small error. This does not happen every time (it depends on the starting point for the attractors).

Below we present some plots to illustrate the performance of the EKF (figs. 6 and 7).

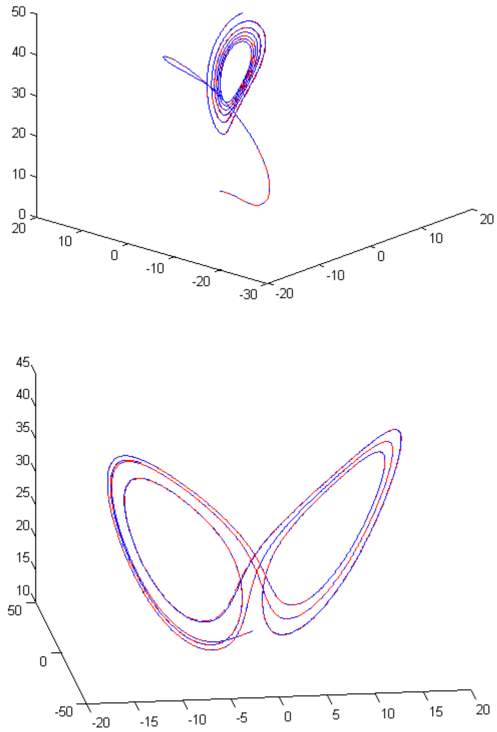

Fig. 6: 3D plot of the real state (blue) and EKF estimate (red). The first three variables (i.e. the first Lorentz attractor) are on the left, the last three (i.e. the second Lorentz attractor) on the right. It can be seen that the EKF manages to follow the real system very well, despite the fact that the second attractor switches surfaces a few times.

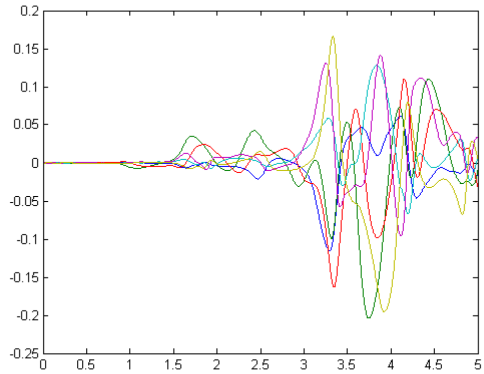

Fig. 7: Plot of the components of the estimation error (real state minus estimate). Although there are some spontaneous "outbursts" of larger error, they are always under 0.2 in absolute value.

6. DISCUSSION

The work presented here shows that the use of dynamic filters such as the EKF to estimate the state of a chaotic system is possible, however in the class of problems addressed here, this strategy can fail. Also, it shows that it is possible to estimate the state of chaotic systems whose measurements are mixed. There are, however, several improvements possible.

The **Lorentz attractor** is, among the span of continuous-time chaotic systems, a tough one to follow. This happens because there are very strong bifurcations, where points very close to each other will go to different "surfaces" (see Fig.1). If near one of these points the sensor noise causes the EKF to estimate that the correct state will lead to one surface when in fact it will lead to the other, the estimation can become very poor. This is what is demonstrated in the *Bad Example* above.
There are plenty more continuous-time chaotic systems that could be used. In fact, the Lorentz attractor does not intend to model the neuron activity (see (Izhikevich, 2007) for some chaotic systems that aim to do that). To use systems with few variables that model the neuron reasonably well, one could for example use the Fitzhugh oscillator (Fitzhugh, 1961), which is related to the more well-known Van der Pol oscillator.

Also, the **EKF** is not the right tool for this estimation problem. Recently, some groups have developed observer design techniques which specialize in estimating the state of chaotic systems. For example, Xiao *et. al.* (2003) have found a technique to develop observers for the Van der Pol system which are globally stable. The problem we saw in the *Bad Example* would never occur, because a stable observer coupled with a bounded system can never diverge.

In linear systems, the concept of **observability** is used to classify systems as observable or not observable. Intuitively, an observable system is one in which knowing a finite number of past measurements in a deterministic setting allows us to infer what the state of the system was at the start of those measurements (see (Anderson and Moore, 1979) for a rigorous definition). For non-linear systems there is a reasonably similar concept introduced by Takens (1981). **Takens' theorem** gives conditions on a chaotic system which allow for its state to be reconstructed from a series of measurements. In more intuitive terms, it is the analog for chaotic systems of the observability concept. This theorem is considerably less intuitive than the simple observability concept and will not be detailed nor used here. In principle, the conditions mentioned in this theorem could be used to better understand under which conditions a pair of Lorentz attractors can be followed by an EKF or not. This is a topic for further research.

7. CONCLUDING REMARKS

We demonstrated that an EKF can correctly estimate the state of a chaotic system (the Lorentz attractor) under low-noise conditions. We also showed that the state variables of a pair of such systems providing measurements which mix the variables of both systems can be correctly estimated under low-noise conditions for some sensor matrices while for others it can diverge. The study of which sensor matrices allow correct estimation is a topic for further work.

REFERENCES

B. Anderson, J. Moore, *Optimal Filtering*, Prentice-Hall, 1979.

R. Devaney, *An Introduction to Chaotic Dynamical Systems*, 2nd edition, Westview, 2003.

R. FitzHugh, Impulses and physiological states in theoretic models of nerve membrane, *Biophysical Journal*, **1**:145 - 446, 1961.

E. Izhikevich, *Dynamic Systems in Neuroscience*, MIT Press, 2007. Available online at http://nsi.edu/users/izhikevich/publications/dsn.pdf

F. Takens, Detecting strange attractors in turbulence, *Dynamical Systems and Turbulence*, *Lecture Notes in Mathematics*, **898**: 366 – 381, Springer-Verlag, 1981.

M. Xiao, N. Kazantzis, C. Kravaris, and A. Krener, Nonlinear Discrete-time Observer Design with Linearizable Error Dynamics, *IEEE Trans. Automatic Control*, **48**, n.4: 622-626, 2003.

ACKNOWLEDGEMENTS

Miguel Almeida is funded by the Portuguese Science and Technology Foundation (FCT) grant SFRH / BD / 28834 / 2006. This work was supported by the FCT (ISR/IST pluri-annual funding) through the PIDDAC Program funds.