Physica D 325 (2016) 126–145

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Wavelet shrinkage of a noisy dynamical system with non-linear noise impact

Matthieu Garcin [a,b,*], Dominique Guégan [a]

[a] *Université Paris 1 Panthéon-Sorbonne, MSE - CES, 106 boulevard de l'hôpital, 75013 Paris, France*

[b] *Natixis Asset Management, 21 quai d'Austerlitz, 75013 Paris, France*

HIGHLIGHTS

- Wavelet shrinkage of a noisy chaos.
- Nonlinear noise influence and nonequispaced sample: a new kind of signal processing.
- Noise described by alpha-stable random variables.
- Robustness of the threshold filters to leptokurtic noise.
- Application to simulated logistic and Lorenz chaos and to financial data.

ARTICLE INFO

Article history:
Received 20 February 2015
Received in revised form
29 February 2016
Accepted 18 March 2016
Available online 28 March 2016
Communicated by J. Garnier

Keywords:
Wavelets
Dynamical systems
Thresholding
Nonequispaced design
Non-linear noise impact

ABSTRACT

By filtering wavelet coefficients, it is possible to construct a good estimate of a pure signal from noisy data. Especially, for a simple linear noise influence, Donoho and Johnstone (1994) have already defined an optimal filter design in the sense of a minimization of the error made when estimating the pure signal. We set here a different framework where the influence of the noise is non-linear. In particular, we propose a method to filter the wavelet coefficients of a discrete dynamical system disrupted by a weak noise, in order to construct good estimates of the pure signal, including Bayes' estimate, minimax estimate, oracular estimate or thresholding estimate. We present the example of a logistic and a Lorenz chaotic dynamical system as well as an adaptation of our technique in order to show empirically the robustness of the thresholding method in presence of leptokurtic noise. Moreover, we test both the hard and the soft thresholding and also another kind of smoother thresholding which seems to have almost the same reconstruction power as the hard thresholding. Finally, besides the tests on an estimated dataset, the method is tested on financial data: oil prices and NOK/USD exchange rate.

1. Introduction

Donoho and Johnstone (1994) have developed a theory of signal denoising using wavelets [1]. Their optimal filtering method has been used for many applications. However, for some signals, the noise has a non-linear influence and therefore, the classical theory of wavelet-based denoising has to be adapted. This is the aim of the present paper in the particular framework of dynamical systems.

Dynamical systems are used to depict non-linearity by a deterministic way [3,4]. They differ from other kinds of non-linear models relying on a stochastic description, like heteroskedastic processes [5–7], jump processes [8] or also long-memory processes [9–14]. Dynamical systems are particularly relevant in some applications, like in video processing [15], in natural sciences [16] as well as in finance [17–19].

We consider a dynamical system, defined in discrete time, x_t. Two successive states of that dynamical system are linked by an evolution function z [19]:

$$\forall t \in \{1, \dots, T\}, \quad x_{t+1} = z(x_t). \tag{1}$$

However, some measurement noise may perturb the observation of that dynamical system [20,19]. In that case, we do not observe directly the state of the system x_t, but rather a noisy observation u_t,

* Corresponding author at: Université Paris 1 Panthéon-Sorbonne, MSE - CES, 106 boulevard de l'hôpital, 75013 Paris, France. Tel.: +33 178403789.
E-mail addresses: matthieu.garcin@polytechnique.edu (M. Garcin), dominique.guegan@univ-paris1.fr (D. Guégan).

http://dx.doi.org/10.1016/j.physd.2016.03.013

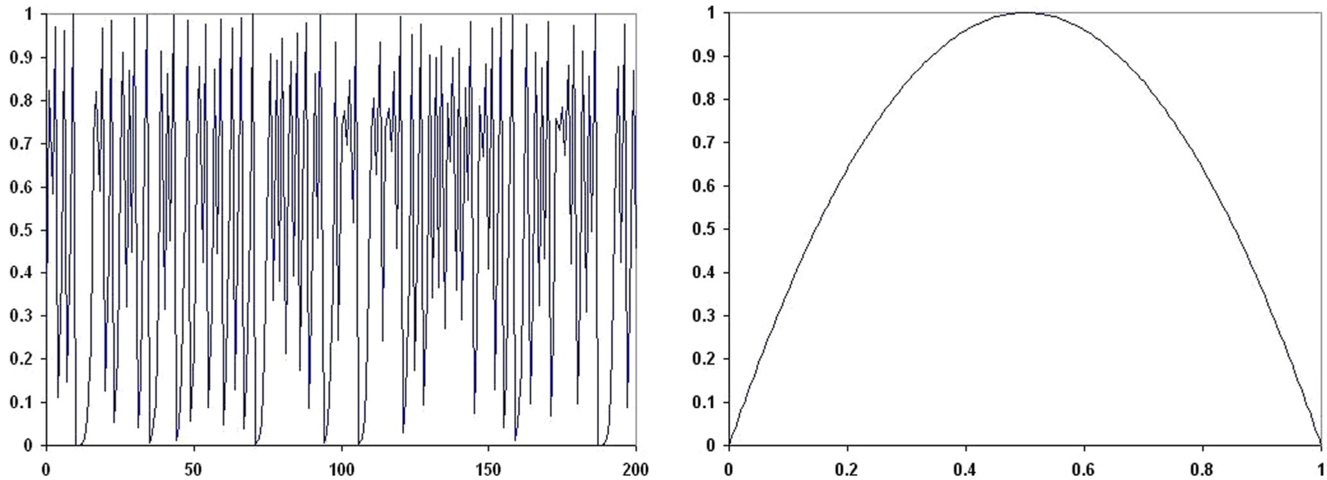

Fig. 1. Time trajectory (on the left) and evolution function (on the right) of the logistic map of parameter 4: $z : x \mapsto 4x(1-x)$.

which consists in an alteration of the state x_t by an additive random variable ε_t:

$$\forall t \in \{1, \dots, T\}, \quad u_t = x_t + \varepsilon_t,$$

where $\varepsilon_1, \dots, \varepsilon_T$ are independent identically distributed random variables. Therefore, the observed evolution function is not z but the function z^ε that links two successive noisy observations:

$$\forall t \in \{1, \dots, T\}, \quad u_{t+1} = z^\varepsilon(u_t) = z(u_t - \varepsilon_t) + \varepsilon_{t+1}. \tag{2}$$

If we observe N states of the noisy system, we can sort all the observations and therefore we get a discretization of the state space of the dynamical system: $u_{1:N} \leq \cdots \leq u_{N:N}$, respectively noted for simplicity $u_1 \leq \cdots \leq u_N$. Hence, we have N discrete observations, $z^\varepsilon(u_1), \dots, z^\varepsilon(u_N)$, of the noisy dynamical system:

$$\forall n \in \{1, \dots, N\}, \quad z^\varepsilon(u_n) = z(u_n - \varepsilon_n^\star) + \varepsilon_n, \tag{3}$$

where $\varepsilon_1, \dots, \varepsilon_N$, $\varepsilon_1^\star, \dots, \varepsilon_N^\star$ are $2N$ independent identically distributed random variables [21].[1] That noisy evolution function, z^ε, is a non-linear function of the noise.

To sum up the problem, we have sparse observations of a noisy evolution function, whereas we are mostly interested in the knowledge of the pure dynamical system. The aim of the present paper is then to present a method to denoise such a noisy signal (the function z^ε) and therefore to estimate the true or pure evolution function z introduced in Eq. (1).

Trajectories of dynamical systems are often very erratic, like, for instance, in Fig. 1, where we represent a logistic chaos. That erratic nature of the pure trajectory makes the denoising of many noisy trajectories very challenging, even though a few methods have already been tested to denoise noisy trajectories of dynamical systems using linear wavelet filtering [22] or other smoothing techniques [23]. Instead, evolution functions are often smoother than time trajectories and their denoising is therefore more feasible. Fig. 1 attests the smoothness of the same logistic chaos if we consider its evolution function in the phase space rather than its time trajectory. Therefore, we do not intend to denoise directly the trajectory of a dynamical system in the time domain but in the phase space. Both problems are linked and if we estimate accurately z then we get an estimate of the pure time trajectory.

We can use several methods to denoise z^ε: local methods and singular value decomposition [24–26], maximum-likelihood-based techniques [20,27], methods based on correlation observation [28], kernel-based non-parametric estimates [29] or methods using radial basis functions [30,31]. For a review, we refer to [4,32].

We are mostly interested in the wavelet shrinkage, because this technique of analysis of a signal into localized elements, which is very popular for spatially inhomogeneous signals in which the noise influence is linear, allows a good accuracy and a parsimonious representation [33]. Some empirical papers have already studied that method applied to dynamical systems [23,34,35]. The big challenge – that we investigate in this paper – in such a method is the non-linear influence of the noise on the signal. Indeed, the literature mainly deals with the denoising based on wavelets for signals with linear noise influence, and in the present article we adapt those classical methods to the specific case of signals with non-linear noise influence.

The irregular observation grid is another specificity of our framework. Indeed, as we are interested in the evolution function, the step size between two consecutive observations in the phase space is non-constant, whereas the time trajectory is discretized by a regular observation grid. We can then make some remarks about that specificity:

- The observation grid is not only irregular, it is also stochastic, and then all the following developments are conditional to a set of observations.
- Irregular grids make the computation of classical empirical wavelet coefficients biased and time-consuming. On the one hand, the slow computation is due to the fact that fast empirical wavelet transform algorithms are designed for regular observation grids. A solution has been proposed to avoid this computational drawback. Indeed *second-generation* wavelets allow fast empirical wavelet transform using what is known as *lifting* [36–39]: the main difference with classical wavelets is that wavelets are not built anymore by dilatations and translations of a unique mother wavelet. Second-generation wavelets are particularly pertinent in multiresolution analysis and in the definition of the nested subspaces representing the different scales. Nevertheless in our framework we prefer to use first-generation wavelets because multiresolution and fast computation are not our goals. On the other hand, there is a bias when one erroneously uses in the non-equispaced design the classical empirical wavelet coefficient formula,

$$\sum_{n=1}^{N} z^\varepsilon(u_n)\psi_{j,k}(u_n),$$

[1] We note $\varepsilon_1, \dots, \varepsilon_N$, $\varepsilon_1^\star, \dots, \varepsilon_N^\star$ the noise in Eq. (3). These variables are not identically related to the $\varepsilon_1, \dots, \varepsilon_T$ appearing in Eq. (2). As well as u_{n+1} is not $z^\varepsilon(u_n)$, $z(u_n - \varepsilon_n^\star)$ is not disrupted in Eq. (3) by a variable ε_{n+1} but by another variable noted ε_n.