IL NUOVO CIMENTO Vol. 25 C, N. 3 Maggio-Giugno 2002

Identification of a coupled dynamical system(*)

L. Battiston(1) and M. Rossi(2)

(1) *ISDGM-CNR - S. Polo 1364, Venezia 30125, Italia*
(2) *Laboratoire de Modélisation en Mécanique, Université de Paris VI 4, Place Jussieu, F-75252 Paris cedex 05, France*

(ricevuto il 12 Marzo 2001; approvato il 15 Novembre 2001)

Summary. — An identification problem for a coupled dynamical system is addressed. More specifically, the system, known from measurements of a scalar quantity, is governed by a set of Langevin equations coupled to a deterministic forcing evolving in a much slower fashion. A statistical method is presented which identifies the deterministic forcing without assuming any parameterization for both sub-systems. This procedure, which is based on a proper orthogonal decomposition applied on probability density functions, works when measurement sampling times remain much smaller than the characteristic time of the forcing. Several test cases are performed.

PACS 05.40.-a – Fluctuation phenomena, random processes, noise, and Brownian motion.
PACS 47.20.Ky – Nonlinearity (including bifurcation theory).
PACS 47.52.+j – Chaos.

1. – Introduction

Time series methods [1-3] aim at extracting, from an experimental data set, a dynamical system which reproduces such a data and is capable to forecast future measurements. The difficulties of identification problems clearly depend on the *a priori* knowledge one assumes on the system. Most methods presently available in the tool box of the nonlinear dynamicist, are derived under the assumption of determinism. In this context, the so-called observational noise, due to measurement errors, may be taken into account even though it deteriorates the efficiency of classical procedures [3, 4] such as fractal dimension computations algorithms. By contrast, dynamical noise [5-7] that characterises stochastic systems, has been less considered. This case however drastically modifies the identification procedure and classical algorithms should then be thought over anew.

(*) The authors of this paper have agreed to not receive the proofs for correction.

Coupled deterministic systems may also be rather challenging for nonlinear time series methods [8,9] in particular when different characteristic times are simultaneously present, *e.g.*, when a fast and a slow dynamical system are coupled. In such a context, *i.e.* standard procedures which compute embedding dimensions, fractal dimensions could be *a priori* directly applied. However they may slowly converge with the total number of measurement points, or may suffer from a high sensitivity to additive noise determining their failure in the presence of a very small amount of observational noise.

In the present study, we consider both effects: coupled interactions and stochasticity. More specifically, measurements are assumed to be produced by a stochastic system with a white Gaussian noise, which is forced by a deterministic system evolving in a much slower fashion. No peculiar forms are given for both dynamical systems and sampling measurement time may be greater than the correlation time of the Langevin sub-system. In such a case, it is hopeless to use deterministic aspects such as orbit tracking methods [10, 11]: only a statistical approach is then pertinent to identify the slow deterministic behaviour from the knowledge of the observed time series of the full system. The reconstruction approach presented below is based on a set of probability density functions (p.d.f.). Each p.d.f. is obtained by collecting data during a period which is large compared to the correlation time of the Langevin system but small compared to the characteristic time of the forcing. The set of these probability density functions is then decomposed on a Proper Orthogonal Basis (POD) [12]. It is shown that the resulting components are governed by a set of deterministic equations which contains the dynamical system describing the forcing as a sub-system. In section **2**, we first recall some particulars from stochastic equation theory. The connection between the deterministic system and the POD components is then derived and discussed. In sect. **3**, this procedure is implemented on two examples of coupled systems: a stochastic Lorenz system when forced by another chaotic Lorenz system or by a two-dimensional oscillator. Finally the procedure is tentatively applied to the case of two coupled deterministic systems.

2. – The statistical approach

Let us consider a system defined as two coupled sub-systems. The first one is characterised by N stochastic variables $\mathbf{x} \equiv (x_1, ..., x_N)$ satisfying a set of N nonlinear Langevin equations([1])

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = F_i(\mathbf{x}; \mu) + \sigma_i \theta_i(t), \tag{1}$$

where $\mu \equiv (\mu_1, .., \mu_P)$ are P forcing parameters and σ_i denote dynamical noise amplitudes. Quantities θ_i stand for normalised uncorrelated white Gaussian noise functions such that $\langle \theta_i(t)\theta_j(t')\rangle = 2\delta_{ij}\delta(t-t')$ with δ_{ij} the Kronecker symbol and $\delta(t-t')$ a Dirac distribution.

The second sub-system

$$\frac{\mathrm{d}\mu_i}{\mathrm{d}t} = \epsilon G_i(\mu) \tag{2}$$

([1]) For the sake of clarity, we associate to eq. (1) a companion deterministic system $\mathrm{d}x_i/\mathrm{d}t = F_i(\mathbf{x}; \mu)$ which will be denoted by CDS.

is autonomous, deterministic and governs the dynamics of the forcing parameters μ of eq. (1). Parameter ϵ in (2) stands for a dimensionless constant which governs the ratio $t_{\rm D}/t_{\rm L}$ between the characteristic times $t_{\rm D}$ and $t_{\rm L}$ (see below for a definition) of the deterministic and Langevin sub-systems, respectively. When ϵ is small enough, the deterministic system (2) evolves with a much slower time dependence than the full coupled system (1), (2). In such a case, a method is proposed to retrieve some of the basic characteristics of sub-system (2), from measurements of a unique scalar variable of the full system

$$v(t) = \mathcal{V}[\mu(t); \mathbf{x}(t)]. \tag{3}$$

More specifically, the full data set

$$\mathcal{S} = (v(t_k);\ t_k = kT, k = 1.., M) \tag{4}$$

is produced by M such measurements sampled every time period T.

2˙1. *The p.d.f. time evolution for* $\epsilon = 0$. – It is known from stochastic equation theory [13,14], that, for Langevin processes (1) the p.d.f. $P(\mathbf{x}, t)$ of finding at time t the orbit at position $\mathbf{x}$ is governed by the Fokker-Planck equation

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = L_\mu P(\mathbf{x}, t), \tag{5}$$

$$L_\mu P(\mathbf{x}, t) \equiv -\sum_{i=1}^{N} \nabla_i.[F_i(\mathbf{x}, \mu(t))P(\mathbf{x}, t)] + \sum_{i=1}^{N} \sigma_i^2 \frac{\partial^2 P(\mathbf{x}, t)}{\partial x_i^2}. \tag{6}$$

The second term of the r.h.s. in (6), which contains the noise amplitude σ_i, may be interpreted as a diffusion process which smoothes the probability density. When parameters μ are constant—*e.g.*, when $\epsilon = 0$ or when the orbit $\mu(t)$ has reached a fixed point of sub-system (2)—the p.d.f. $P(\mathbf{x}, t)$ is attracted towards a stationary solution $P_0(\mathbf{x}; \mu)$ which satisfies $L_\mu P_0(\mathbf{x}; \mu) = 0$. This occurs for times t large enough to ensure that transients have elapsed. In the absence of dynamical noise, the function $P_0(\mathbf{x}; \mu)$ is known to be of fractal type for chaotic attractors. On the contrary, when noise amplitudes are not zero, the p.d.f. is smooth. It is thus possible to get a sufficient estimate with a finite set of data points. Note, however, that evaluating integrals of p.d.f. might be done even for pure chaotic system.

In the present work, the eigenvectors $(\Phi_j(\mathbf{x}, \mu),\ j = 1, \cdots)$ of the linear operator L_μ, are assumed to form an infinite, complete and discrete set [13,14] on which to expand any p.d.f. The stationary solution $P_0(\mathbf{x}; \mu)$ is associated with the first one $\Phi_1(\mathbf{x}, \mu)$ which has hence a zero eigenvalue $\lambda_1(\mu) = 0$. Each eigenvector $\Phi_j(\mathbf{x}, \mu)$ with $j \neq 1$ is characterised by an eigenvalue $\lambda_j(\mu)$ the real part of which is always strictly negative (no ordering is assumed for the eigenvalues). As a consequence, any p.d.f. $P(\mathbf{x}, t)$ is attracted towards the stationary solution $P_0(\mathbf{x}; \mu)$ and the characteristic time $t_{\rm L}$ necessary to reach this equilibrium then scales as $t_{\rm L} = 1/\beta$ with $\beta = -{\rm Min}({\rm Re}(\lambda_j),\ j = 2, \cdots)$ (Re means that the real part is taken). In addition, a set $(\Phi_k^*(\mathbf{x}, \mu(t)),\ k = 1, \cdots)$ of eigenfunctions of the

backward Fokker-Planck equation

$$\sum_{i=1}^{N} F_i(\mathbf{x},\mu(t)).\nabla_i[\Phi_k^*] + \sum_{i=1}^{N} \sigma_i^2 \frac{\partial^2 \Phi_k^*}{\partial x_i^2} = \lambda_k(\mu)\Phi_k^* \tag{7}$$

may be associated [13] to the previous eigenvectors and the following bi-orthogonal relation holds:

$$(\Phi_k^*, \Phi_j) = \delta_{kj}, \tag{8}$$

where the scalar product reads

$$(\Phi_k^*, \Phi_j) \equiv \int \Phi_k^*(\mathbf{x},\mu)\Phi_j(\mathbf{x},\mu)\mathrm{d}\mathbf{x}. \tag{9}$$

Note that, for standard non-reflecting conditions, the adjoint eigenvector associated with the stationary solution reads $\Phi_1^* = 1$.

2˙2. *The p.d.f. time evolution when the forcing is slowly evolving.* – Let us now assume that the forcing $\mu(t)$ slowly evolves with respect to the stochastic system, *i.e.* parameters μ change according to eq. (2) with a characteristic time t_{D} much larger than the typical time t_{L}. In such a case, we show that a quasi-static approximation may be found for the probability density function. This is done as follows: first, at each time t, an eigenvector basis is defined which is associated with the operator $L_{\mu(t)}$. Let us expand the p.d.f. $P(\mathbf{x},t)$ over this instantaneous eigenvector basis

$$P(\mathbf{x},t) = \sum_{j=1}^{\infty} a_j(t)\Phi_j(\mathbf{x},\mu(t)). \tag{10}$$

By construction of operator (6) and a normalisation prescription, eigenvectors $\Phi_j(\mathbf{x},\mu)$ are such that

$$(\Phi_1^*, \Phi_j) = \int \Phi_j(\mathbf{x},\mu)\mathrm{d}\mathbf{x} = \delta_{1j}. \tag{11}$$

This equality ensures the conservation of the total probability, *i.e.* the integral of $P(\mathbf{x},t)$ over the entire phase space. Normalisation of Φ_1 then imposes that $a_1(t) = 1$ in (10). Introducing expansion (10) in the Fokker-Planck equation (5), one gets

$$\sum_{p=1}^{P} \frac{\mathrm{d}\mu_p}{\mathrm{d}t}\frac{\partial \Phi_1}{\partial \mu_p}(\mathbf{x},\mu(t)) + \sum_{j=2}^{\infty}\left[\frac{\mathrm{d}a_j}{\mathrm{d}t}\Phi_j(\mathbf{x},\mu(t)) + a_j(t)\sum_{p=1}^{P}\frac{\mathrm{d}\mu_p}{\mathrm{d}t}\frac{\partial \Phi_j}{\partial \mu_p}(\mathbf{x},\mu(t))\right] = \sum_{j=2}^{\infty} a_j(t)\lambda_j(\mu(t))\Phi_j(\mathbf{x},\mu(t)). \tag{12}$$

By taking, for $k \neq 1$, the scalar product of eq. (12) with the adjoint eigenvector $\Phi_k^*(\mathbf{x},\mu(t))$ and by using eq. (2), one gets the infinite linear system

$$\frac{\mathrm{d}a_k}{\mathrm{d}t} = \mathcal{B}_{k1}(\mu) + \sum_{j=2}^{\infty}\mathcal{L}_{kj}(\mu)a_j \tag{13}$$

for variables a_k with $k \neq 1$, where

$$\mathcal{L}_{kj}(\mu) = \lambda_k(\mu)\delta_{kj} + \mathcal{B}_{kj}(\mu), \tag{14}$$

$$\mathcal{B}_{kj}(\mu) = -\epsilon \sum_{p=1}^{P} G_p(\mu)(\Phi_k^*(\mathbf{x},\mu), \frac{\partial \Phi_j}{\partial \mu_p}(\mathbf{x},\mu)) \tag{15}$$

are nonlinear functions of μ. It is easily seen that the scalar product of eq. (12) with the adjoint eigenvector $\Phi_1^* = 1$ does not provide any constraint on variables a_k.

Within the quasi-static approximation, nonlinear functions $\lambda_k(\mu)$, $\mathcal{B}_{kj}(\mu)$ may be considered as constant coefficients and eq. (13) thus becomes a nonhomogeneous linear system with constant coefficients. Its solutions either rapidly diverge—an unrealistic case which would lead to inconsistencies—or converge, on a time scale t_{L}, towards a "fixed" point $a_k(t) = A_k(\mu(t))$. Within this approximation, the p.d.f. entirely depends on time through forcing $\mu(t)$

$$P(\mathbf{x},t) = \Phi_1(\mathbf{x},\mu(t)) + \sum_{j=2}^{\infty} A_j(\mu(t))\Phi_j(\mathbf{x},\mu(t)). \tag{16}$$

In the extreme case in which the linear damping term $\lambda_k a_k$ in eq. (13) is predominant with respect to $\mathcal{B}_{kj}$, variables a_k for $k > 1$ rapidly converge towards

$$a_k \sim -\epsilon \frac{\mathcal{B}_{k1}(\mu)}{\lambda_k(\mu)} + O(\epsilon^2). \tag{17}$$

When μ is fixed, one recovers the standard result $a_k(t) = 0$ for $k > 1$. A straightforward consequence follows from the above analysis: any variable extracted at any given time t from the p.d.f. $P(\mathbf{x},t)$, exhibits a slow deterministic dynamics that only depends on $\mu(t)$.

The validity of the above analysis clearly depends on the value of $t_{\mathrm{D}}(\epsilon)$. More specifically, $t_{\mathrm{L}} \ll t_{\mathrm{D}}$ must be satisfied (this is true for $\epsilon \ll 1$). When $t_{\mathrm{L}} \sim t_{\mathrm{D}}$, there exists a set $\mathcal{J}$ of $\mathcal{N}$ indices j such that $1/\mathrm{Re}(\lambda_j)$ is of the same order of magnitude than t_{D}. The above analysis then fails. The problem, however, may be reformulated by separating the set of modes a_k with k belonging to $\mathcal{J}$ from the others. The latter modes are slaved to $\mu(t)$ and a_j with j in $\mathcal{J}$. The center manifold theory [15] may then be invoked to ensure that the time evolution of the p.d.f. is described by a deterministic system of dimension $\mathcal{N} + N_P$. Consider, for instance, the Langevin bistable potential model ($s > 0$)

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = (2sx_1 - 3x_1^3) + \sigma\theta_1(t) \tag{18}$$

which possesses a stationary probability

$$P(x_1,t) \sim \exp\left[\frac{sx_1^2 - x_1^4}{\sigma^2}\right]. \tag{19}$$

When s/σ is large([2]), orbits may remain trapped for a long time, close to one of the two deterministic fixed points and t_{L} thus increases. This feature may be put on quantitative terms for a bistable model with a rectangular potential well [14]. For this analytically solvable model, the lowest nonzero eigenvalue decreases to zero, with increasing H/σ^2, as $\sigma^2 \exp[-(H/\sigma^2)]$. For weak noise, the above analysis thus fails since time t_{L} becomes large. Conversely, the method should work for large noise amplitudes since the lowest nonzero eigenvalue linearly increases as σ^2. When the CDS is a chaotic system, *i.e.* a system with highly mixing properties, the time t_{L} does not unboundedly increase in the weak noise limit and the method presented here is more robust.

The above results have been established in the framework of a stochastic sub-system with an uncorrelated white noise. For other cases, *e.g.*, a pure chaotic deterministic sub-system ($\sigma_i = 0$) or a stochastic sub-system with coloured noise, it is argued that this method, which relies on general properties of the spectrum of operator $L_{\mu(t)}$, may be tentatively used (see below for numerical results of the pure deterministic case). The p.d.f. of a stochastic process with coloured noise satisfies an equation like (5) with very similar properties [16] though the operator L_μ must be modified. Consider, for instance, that functions θ_i in eq. (1), are normalised coloured noises with a decay time t_c, $\langle \theta_i(t)\theta_j(t')\rangle = (\delta_{ij}/t_c)\exp\left[\frac{-|t-t'|}{t_c}\right]$. This noise is produced [14] by equation

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = -\frac{\theta_i}{t_c} + \frac{\gamma_i(t)}{t_c} \tag{20}$$

with γ_i a normalised uncorrelated white noise. As a consequence, the coloured stochastic system may be written in a system of the form (1) with i) greater dimension since θ_i become additional stochastic variables and with ii) some noise amplitudes put to zero.

2˙3. *Extracting the data: theoretical aspects*. – In this subsection, we show how to obtain from the scalar measurements $\mathcal{S}$ of the stochastic quantity v, a deterministic time variable $b(t)$. This latter quantity is then used to identify the deterministic part of the system.

Assume that variable v varies in $\mathcal{S}$ between $v_{\min} \leq v \leq v_{\max}$. The total range of variation $[v_{\min}, v_{\max}]$ is then divided into D non-overlapping intervals Δ_k ($k = 1, ..., D$). By virtue of definition (3), each Δ_k corresponds to a specific phase space region for $\mathbf{x}$. The probability $p_v(k,t)$ of measuring v inside Δ_k during the time interval $\left[t - \frac{\Delta T}{2}, t + \frac{\Delta T}{2}\right]$, where $\Delta T = 2ZT$ corresponds to the time between $Q \equiv 2Z+1$ consecutive measurements, reads

$$p_v(k,t) = \frac{1}{\Delta T}\int_{t-\frac{\Delta T}{2} \leq s \leq t+\frac{\Delta T}{2}} \mathrm{d}s \left(\int H_{\Delta_k}(v(\mathbf{x}, \mu(s)))P(\mathbf{x}, s) \prod_{r=1}^{N} \mathrm{d}x_r \right), \tag{21}$$

where $H_{\Delta_k}(v)$ equals 1 if v is contained in the interval Δ_k and 0 otherwise. Within the quasi-static approximation (16), the phase space integral

$$\mathcal{P}_k(\mu(t)) = \int H_{\Delta_k}(v(\mathbf{x}, \mu(t)))P(\mathbf{x}, t) \prod_{r=1}^{N} \mathrm{d}x_r \tag{22}$$

([2]) This may occur because the noise amplitude σ^2 is small or the potential barrier height H between the two fixed points $x_1 = \pm\sqrt{2s/3}$ is large.

is only a function of μ. In this case, the relation

$$\frac{\mathrm{d}p_v}{\mathrm{d}t}(k,t) = \frac{1}{\Delta T}\left[\mathcal{P}_k\left(\mu\left(t+\frac{\Delta T}{2}\right)\right) - \mathcal{P}_k\left(\mu\left(t-\frac{\Delta T}{2}\right)\right)\right] \tag{23}$$

holds. If parameters μ slightly evolve during the time interval ΔT between Q consecutive measurements, eq. (23) reads

$$\frac{\mathrm{d}p_v}{\mathrm{d}t}(k,t) = \sum_{p=1}^{P}\frac{\mathrm{d}\mu_p}{\mathrm{d}t}\frac{\partial \mathcal{P}_k}{\partial \mu_p} = \epsilon\sum_{p=1}^{P} G_p(\mu)\frac{\partial \mathcal{P}_k}{\partial \mu_p}(\mu). \tag{24}$$

As a consequence, any linear combination of variables $p_v(k,t)$ $(k = 1, ..., D)$ defines a scalar $b(t)$, the dynamics of which is described by a $P+1$ deterministic system emulating eqs. (2) and (24). Moreover, since eq. (2) does not depend on the continuous variable $b(t)$, the correlation dimension of $b(t)$ is less or equal than Dim + 1, where Dim denotes the correlation dimension of the attractor of (2).

2˙4. *Extracting the data: practical aspects.* – This subsection is divided into two distinct parts. In the first place, various vectors are computed from the data set. The reconstructed time series $b(t)$ is then extracted by applying a POD method to such vectors.

Given a time $t = t_m$, the quantity $p_v(k, t_m)$ can be numerically estimated as $p_v(k, t_m) \sim N_k/Q$, where N_k denotes the number of $v(t_s)$ with $m - Z \leq s \leq m + Z$ which lie inside Δ_k. A D-vector $(p_v(1, t_m), ..., p_v(D, t_m))$ which depends on $\mu(t_m)$ can be thus generated. From M measurements in $\mathcal{S}$, $q \sim M/S$ overlapping such intervals $t_{m-Z} \leq t_s \leq t_{m+Z}$, may be defined each one being translated of S measurements from the previous one $(S \leq Q)$. Finally this procedure yields q consecutive D-vectors $(p_v(1, t_m), \ldots, p_v(D, t_m))$ separated by a time interval ST, *i.e.* $t_m = mST$.

A coherence does exist between such D-vectors since they are produced by the system (1), (2). This coherence however may be hidden because of obvious statistical limitations arising from the data set $\mathcal{S}$. A Proper Orthogonal Decomposition (POD) [12] determines which linear combination $b(t) = \sum \alpha_j p_v(j, t)$ contains the greatest part of the coherence. This method, which separates the time evolution from the "phase space" coherence, is technically identical to the singular value decomposition (SVD) used for classical deterministic cases. However the present method should not to be confused with SVD (when applied to stochastic data $\mathcal{S}$, SVD actually fails to retrieve any features of the deterministic sub-system) since we work on vectors extracted from probability density functions and not from any time delay algorithms. The POD, which solves the eigenvalue problem [12] of the auto-correlation matrix built from the D-vectors $(p_v(1, t_m), ..., p_v(D, t_m))$, provides an empirical basis of D-vectors $(K_n(1), ..., K_n(D))$ with $n = 1, ..., D$ such that the expansion

$$p_v(k, t_m) = \sum_{n=1}^{D} b_n(t_m) K_n(k) \tag{25}$$

is optimal. This means that POD coefficients $|b_n(t_m)|$ decrease faster with n than coefficients of other possible expansions. The first coefficient $b_1(t_m)$, *i.e.* the scalar product of $(p_v(1, t_m), ..., p_v(D, t_m))$ with $(K_1(1), ..., K_1(D))$, contains the greatest coherent part of

the signal. As a consequence, we use, in the sequel, the reconstructed variable $b(t) \equiv b_1(t)$ sampled every time ST.

The value of Q is a major factor to get an adequate $b(t)$: if Q is too small, the number of data points is insufficient to estimate the quantity $p_v(k, t_m)$ though the first coefficient of the POD does filter out part of fluctuations due to finite measurement sets. This effect can be directly seen on the convergence of the empirical functions and eigenvalues coefficients with respect to the number of points Q as well as on the projection $b(t) \equiv b_1(t)$ itself. Conversely if Q becomes too large, parameters μ significantly evolve during the time interval $\Delta T = (Q-1)T$ and eq. (24) is then no more valid. One should check *a posteriori* that such constraints are verified.

From the results of subsect. **2**˙3, variable $b(t)$ sampled each time interval ST may be used to characterise and emulate eqs. (2) and (24) with the aid of suitable chaotic indicators such as minimal sufficient embedding dimension (here the false nearest-neighbours method) or mutual information. This is done in the next section. We did not attempt to compute the correlation dimension since measurement noise in the present calculations is too large for this algorithm to be effective with a realistic number of points M (correlation dimension algorithms are known to be very sensitive to additive noise [4]).

3. – Application to two cases of coupled systems

The above technique is now applied to two cases of coupled systems. For the first one (hereafter denoted case A), the forcing $\mu(t)$ is governed by a Lorenz system in a chaotic regime [17]

$$\frac{\mathrm{d}\mu_1}{\mathrm{d}t} = \epsilon[a(\mu_2 - \mu_1)], \quad \frac{\mathrm{d}\mu_2}{\mathrm{d}t} = \epsilon[b\mu_1 + \mu_2 - \mu_1\mu_3], \quad \frac{\mathrm{d}\mu_3}{\mathrm{d}t} = \epsilon[-c\mu_3 + \mu_2\mu_1]\,, \tag{26}$$

while, in the second case (hereafter denoted case B), $\mu(t)$ behaves like a linear oscillator

$$\frac{\mathrm{d}\mu_1}{\mathrm{d}t} = -\epsilon\mu_2, \qquad \frac{\mathrm{d}\mu_2}{\mathrm{d}t} = \epsilon\mu_1\,. \tag{27}$$

In both cases, the stochastic sub-system satisfies a Lorenz system

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = a(x_2 - x_1) + \sigma\theta_1(t), \quad \frac{\mathrm{d}x_2}{\mathrm{d}t} = f(\mu_2)x_1 + x_2 - x_1x_3 + \sigma\theta_2(t), \tag{28}$$

$$\frac{\mathrm{d}x_3}{\mathrm{d}t} = -cx_3 + x_2x_1 + \sigma\theta_3(t) \tag{29}$$

with white noise of amplitude σ. In the above equations, parameters a, b and c are taken to be equal to $a = 10$, $b = 28$, $c = 2.666$. The coupling between the two sub-systems is implemented through the Rayleigh parameter. This coefficient is given by a suitable function of the slow variable μ_2. For case A, $f(\mu_2) = 40 - 0.5\mu_2$ is used, while, for case B, $f(\mu_2) = 40 - 15\mu_2$. In both instances, such a transformation maintains the CDS of the stochastic sub-system (1) in the chaotic range.

The full system is simulated with the one time step numerical algorithm explained in [18] with time step $ht = 0.001$. For most cases considered below, the algorithm is tested using a synthetic data produced with noise amplitude $\sigma = 5$. However a single

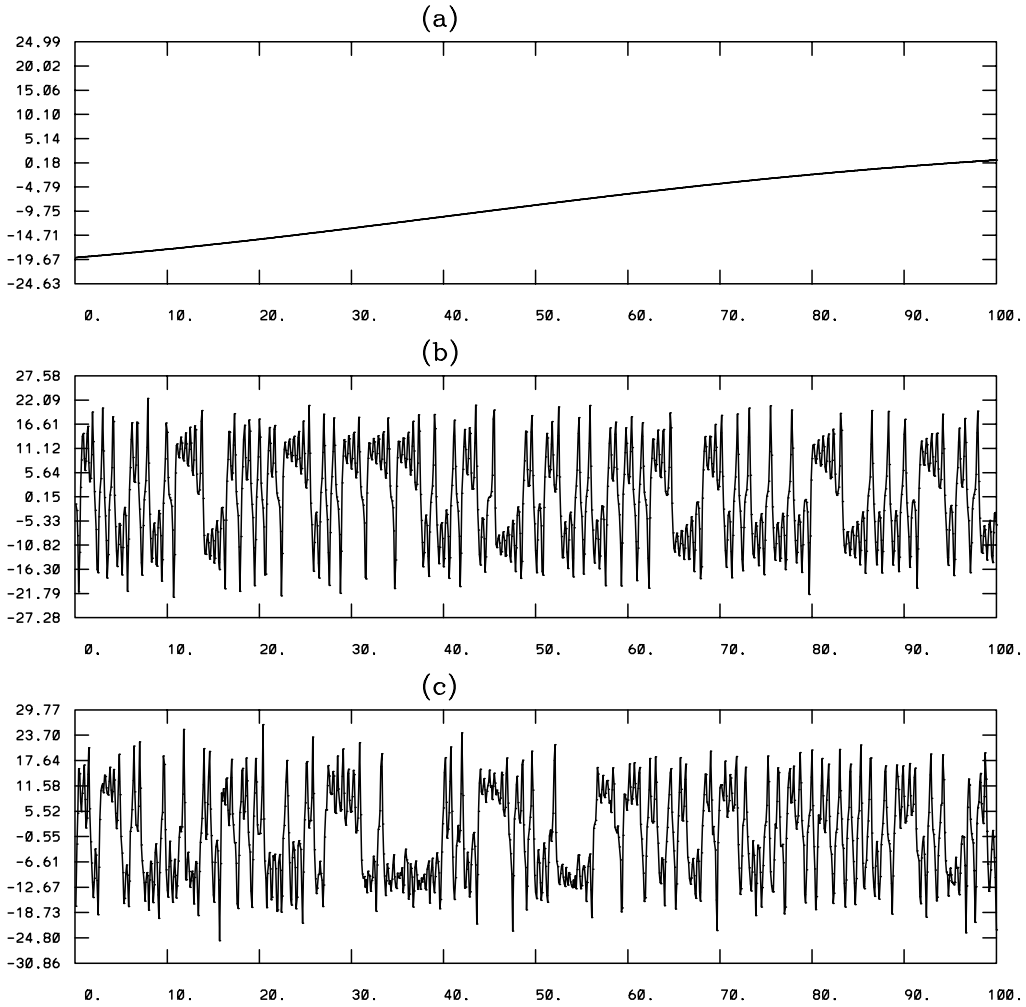

Fig. 1. – Case A: (a) behaviour of variable $\mu_2(t)$ of the deterministic sub-system (2); (b) behaviour of variable $x_1(t)$ of the coupled system (1), (2) for $\sigma = 0$ and (c) for $\sigma = 5$.

reconstruction is attempted for a system (1) with $\sigma = 0$, *i.e.* two coupled deterministic Lorenz sub-systems. The parameter ϵ in eq. (2) is typically of the order 10^{-3} for the Lorenz system and 10^{-2} for the regular system. For case A, fig. 1 provides a hint about the respective characteristic times of eq. (2) (fig. 1a), eq. (1) with $\sigma = 0$ (fig. 1b) and $\sigma = 5$ (fig. 1c). The data set $\mathcal{S}$ is obtained by $M = 10^6$ scalar measurements([3]) sampled every time period $T = 0.1$. The scalar variable v is here equal to the fast variable x_1.

3˙1. *Results of classical methods.* – Standard methods, such as Fourier analysis or nonlinear time series methods, are first directly implemented on series $v(t)$. A visual inspec-

([3]) We also tried for some cases a data set with $M = 10^7$ and checked the persistence of the reconstruction results.

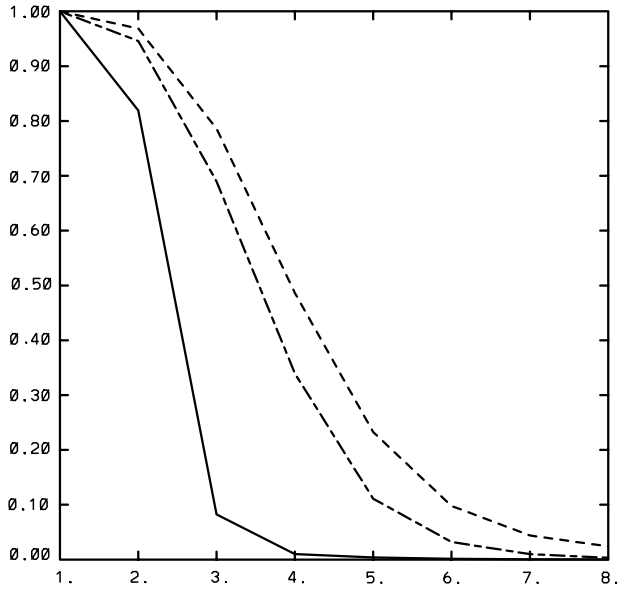

Fig. 2. – Case A with $M = 10^6$: the percentage of false nearest neighbours of v as a function of the embedding dimension for $\sigma = 5$ (dashed), $\sigma = 0$ (solid) without observational noise, $\sigma = 0$ (dot-dashed) with an observational noise of two percent in r.m.s. units.

tion of the data set $v(t)$ for a long time period shows, for case B, some regular behaviour on the amplitude. However the Fourier spectrum is incapable to find a distinct feature at forcing frequency ϵ: in the complete system, this component is hidden by the presence of noise or chaotic dynamics of the Langevin sub-system (1) which generates a broad spectrum. Filtering out by a low-pass filter this component is thus helpless. For case A, a visual inspection shows no trend of an underlying deterministic or large-scale behaviour.

The series v can be analysed using nonlinear time series analysis softwares such as those given in the package [19]. The time delay is determined by the minimum of the mutual information (program mutual of the package [19]). For case A with $\sigma = 5$ or

TABLE I. – *Case A, $\sigma = 5$: Eigenvalues and eigenvectors of the first, second and fifth components for $S = 100$, $D = 10$, $Q = 2000$, with $M = 10^6$.*

0.16227	0.00076	0.00011
−0.00073	0.00314	0.00371
−0.03852	0.05886	−0.00910
−0.24234	−0.09368	−0.14794
−0.45932	−0.66466	0.16934
−0.45725	−0.23623	−0.59878
−0.45272	0.10173	0.70668
−0.47350	0.59264	−0.26955
−0.29480	0.35756	0.12299
−0.06061	0.02743	0.06012
−0.00185	0.00083	0.00211

TABLE II. – *Case A, $\sigma = 5$: the first eigenvector and eigenvalue ($S = 100, D = 10$) for various Q ($Q = 100, Q = 500, Q = 1000, Q = 2000, Q = 3000, Q = 10000$), with $M = 10^6$.*

0.16232	0.16228	0.16228	0.16227	0.16226	0.16223
−0.00072	−0.00073	−0.00073	−0.00073	−0.00073	−0.00074
−0.03836	−0.03847	−0.03849	−0.03852	−0.03856	−0.03871
−0.24173	−0.24216	−0.24224	−0.24234	−0.24242	−0.24270
−0.46006	−0.45940	−0.45933	−0.45932	−0.45932	−0.45935
−0.45677	−0.45728	−0.45732	−0.45725	−0.45718	−0.45691
−0.45209	−0.45261	−0.45270	−0.45272	−0.45270	−0.45269
−0.47438	−0.47368	−0.47356	−0.47350	−0.47347	−0.47338
−0.29451	−0.29467	−0.29471	−0.29480	−0.29488	−0.29514
−0.06029	−0.06050	−0.06054	−0.06061	−0.06068	−0.06088
−0.00185	−0.00185	−0.00185	−0.00185	−0.00185	−0.00187

$\sigma = 0$, the mutual information fails when applied on the data set $\mathcal{S}$ since it does not show any clear minimum. This may be explained since, for a deterministic Lorenz system with parameters $a = 10$, $b = 28$, $c = 2.666$, the mutual information shows a minimum for $t = 0.16$ a value of the same order of sampling measurement time $T = 0.1$ of data set $\mathcal{S}$. The time delay 0.1 is thus introduced in the false-nearest algorithm (false-nearest program of the package) which provides the minimal sufficient embedding dimension. For case A and noise amplitude $\sigma = 5$, the percentage of false nearest neighbours does not go to zero at the embedding dimension $3 + 3$ (dashed line in fig. 2). Even though, for the deterministic case A with $\sigma = 0$ and no observational noise (solid line in fig. 2), the correct value is recovered, an addition of a very small amount of additive noise is sufficient to deteriorate this result (dot-dashed line in fig. 2). This sensitivity is due to the disparity in time scales between the two coupled sub-systems and to the lack of measurement points to catch the longest time scales. Similar results are obtained for case B where $3 + 2$ is an expected value for the embedding dimension. Finally it was checked that the SVD technique applied to v does not indicate any trend whatsoever of a large-scale dynamics.

3˙2. *Results of the statistical method.* – The data set $\mathcal{S}$ obtained for noise amplitude $\sigma = 5$ is now analysed using the method of sect. **2**. The following parameters have been chosen: temporal windows are translated of $S = 100$ measurements and the range of variation of v is divided into $D = 10$ intervals([4]). The POD representation hence generates ten empirical vectors $(K_n(1), ..., K_n(D))$ $(n = 1, ..., 10)$. We have used different numbers Q of consecutive measurements: $Q = 100$, $Q = 500$, $Q = 1000$, $Q = 2000$, $Q = 3000$, $Q = 10^4$. For $Q = 2000$ and both cases, the spectrum of the correlation matrix, which estimate the relative contribution of each POD component in expansion (25), indicates that the first one contains most part of vector $(p_v(1, t_m), ..., p_v(D, t_m))$ (see table I for case A). This result is confirmed by other values of Q. The first eigenvector and eigenvalue for case A and $\sigma = 5$ are shown in table II for various window widths Q ($Q = 100$, $Q = 500$, $Q = 1000$, $Q = 2000$, $Q = 3000$, $Q = 10^4$). A comparison of the first eigenvector ensures that results are almost convergent. For eigenvectors $(K_n(1), ..., K_n(D))$

([4]) We have tried $D = 30$. For cases considered, we did not find any significant change.

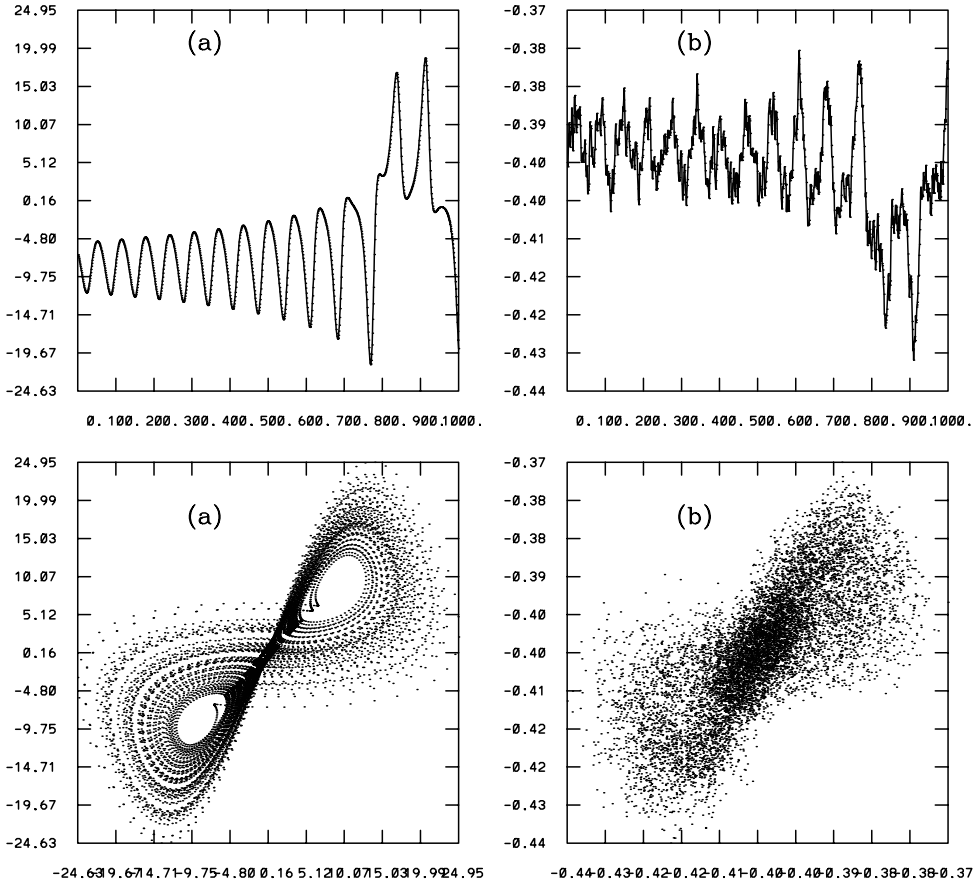

Fig. 3. – Case A, $\sigma = 5$ with $S = 100$, $D = 10$. The orbit (top) and 2D phase plot $(\mu_2(t_j), \mu_2(t_{j+10}))$ or $(b(t_j), b(t_{j+10}))$ (bottom) for (a) the true forcing series $\mu_2(t)$, (b) the reconstructed variable $b(t)$ with $Q = 500$.

with $5 \leq n$, results do not converge due to lack of points. However such eigenvectors are not significant since only a small fraction of the total vector $(p_v(1, t_m), ..., p_v(D, t_m))$ in expansion (25) is contained in these higher empirical modes.

For each time series $v(t)$, one now compares two deterministic series: the exact forcing $\mu_2(t_j)$ and the reconstructed series $b(t_j)$ for $t_j = jST$. Note that no noise reduction is applied on the reconstructed data. A visual inspection of $b(t_j)$ ascertains that the statistical method is capable to grasp some dynamical features of $\mu_2(t_j)$: in fig. 3, a phase space plot $(\mu_2(t_j), \mu_2(t_{j+10}))$ for the true forcing or $(b(t_j), b(t_{j+10}))$ for the reconstructed series and a time evolution curve are shown for case A and window widths $Q = 500$. In fig. 4, the same is done for $Q = 2000$ and $Q = 3000$. Equivalent plots for $Q = 2000$ and $Q = 10^4$ are presented in fig. 5 for the linear oscillator (case B). For small Q, the error on the probability distribution induces on the reconstructed value $b(t_j)$ an observational noise. A region may then be identified where the value of Q is not significant and the measurement error seems to reach a minimum. Note that, when Q becomes too large, the theory should be modified as mentioned in subsect. **2**˙4. This problem appears in

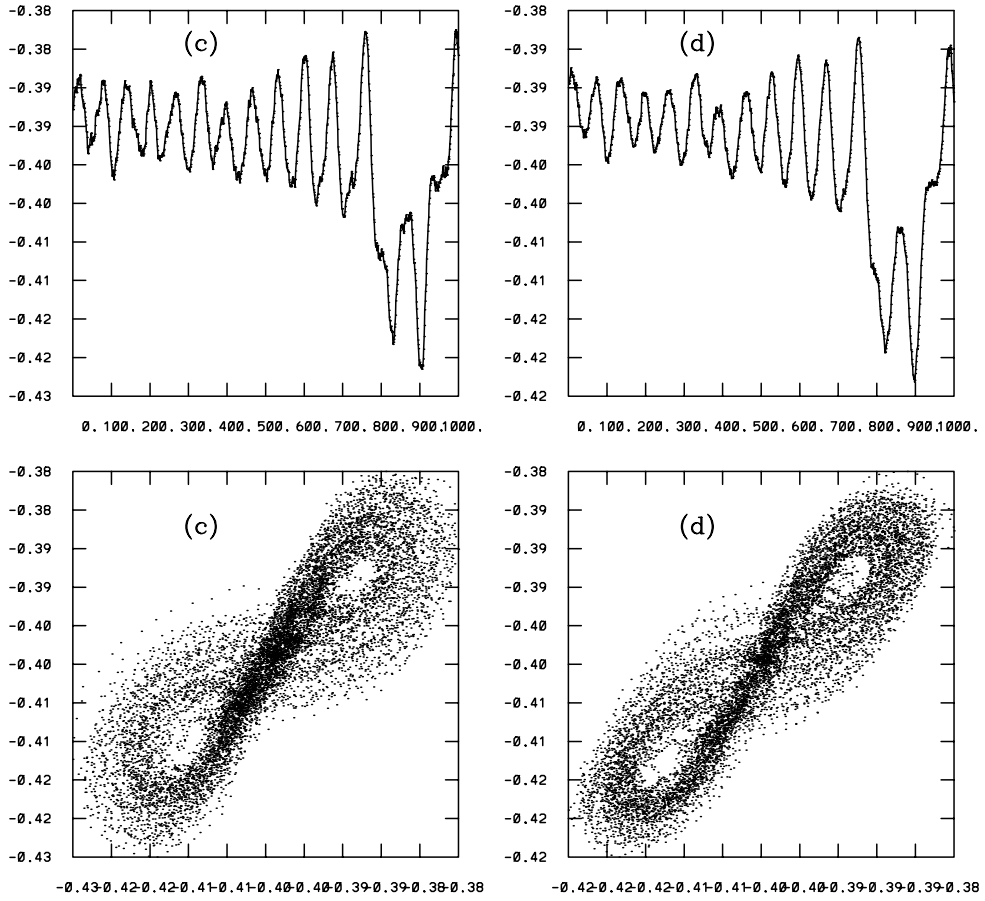

Fig. 4. – Case A, $\sigma = 5$ with $S = 100$, $D = 10$. The orbit (top) and 2D phase plot $(\mu_2(t_j), \mu_2(t_{j+10}))$ or $(b(t_j), b(t_{j+10}))$ (bottom) for (c) the reconstructed variable $b(t)$ with $Q = 2000$, (d) the reconstructed variable $b(t)$ with $Q = 3000$.

fig 5 for the phase plot of case B $Q = 10^4$.

Using the software given in the package [19], we quantitatively analyse the true forcing $\mu_2(t_j)$ with $t_j = jST$ and the reconstructed series $b(t_j)$ for cases A and B with $\sigma = 5$. We set the parameters for the false nearest-neighbours algorithm, so that the analysis of the true series $\mu_2(t_j)$ provides the known correct values 2 for oscillator (see table III) and 3

TABLE III. – *Case B with* $Q = 2000$*, percentage of false nearest neighbours for increasing embedding dimensions* (*ED*) *for* $b(t)$ *with* $\sigma = 5$ *and* $\sigma = 0$ *and for the true variable* $\mu_2(t)$.

ED	$\sigma = 5$	$\sigma = 0$	$\mu_2(t)$
1	0.937	0.887	0.694
2	0.0177	0.00813	0.0
3	0.0	0.0	0.0

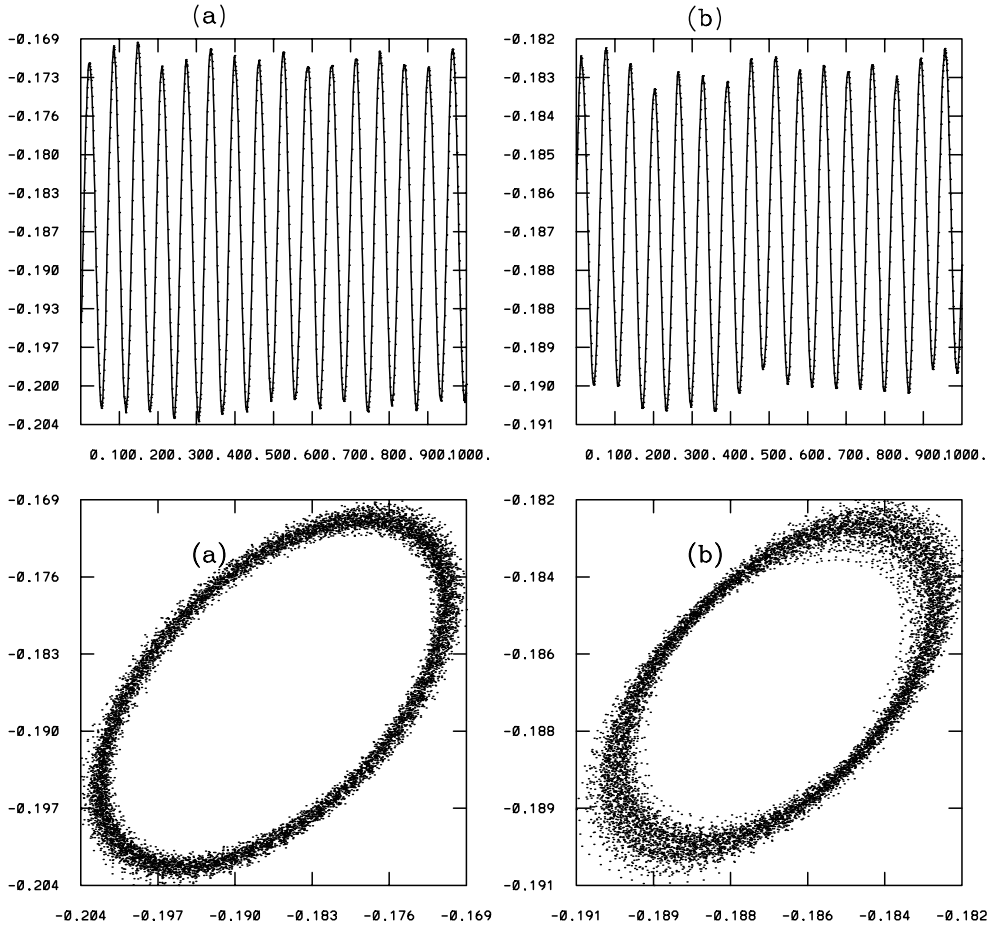

Fig. 5. – Case B, $\sigma = 5$ with $S = 100$, $D = 10$. The orbit (top) and a 2D phase plot $(\mu_2(t_j), \mu_2(t_{j+10}))$ or $(b(t_j), b(t_{j+10}))$ (bottom) for (a) the reconstructed $b(t)$ with $Q = 2000$, (b) the reconstructed $b(t)$ with $Q = 10^4$.

for Lorenz (see table IV). For case A and $\sigma = 5$, the time delay was obtained to be very stable with Q and equal to $t = 190$ which favourably compares with the expected value $t = 160$ (since 0.16 should be divided by a renormalized time factor $\epsilon = 10^{-3}$). For case B and $\sigma = 5$, the time delay was very stable with Q and equal to $t = 160$. This result is close to the value $t = 150$ corresponding to one fourth of the period with $\epsilon = 10^{-2}$. Using the time delay computed, table III shows, for Case B and $Q = 2000$, the percentage of false nearest neighbours of $b(t)$. The minimal sufficient embedding dimension for $b(t_j)$ is obtained to be $3 = 2 + 1$ which, according to the previous analysis, correctly implies that the forcing μ lives in a 2-dimensional phase space. For case A and $Q = 2000$, table IV indicates that the minimal sufficient embedding dimension takes the value $4 = 3 + 1$. It is known that the correlation dimension of the Lorenz attractor is 2.05. If we increase too much the heuristic threshold of the false nearest-neighbours algorithm, we underestimate the correct minimal sufficient embedding dimension of one unit: for the real forcing μ we get 2 and for the reconstructed variable b we obtain $2 + 1$.

TABLE IV. – *Case A with $Q = 2000$, percentage of false nearest neighbours for increasing embedding dimensions (ED) for $b(t)$ with $\sigma = 5$ and $\sigma = 0$ and for the true variable $\mu_2(t)$.*

ED	$\sigma = 5$	$\sigma = 0$	$\mu_2(t)$
1	0.951	0.942	0.893
2	0.0474	0.0211	0.00358
3	0.000445	0.000625	0.0
4	0.0	0.0	0.0

3˙3. *Case of two deterministic coupled sub-systems.* – The reconstruction method has been applied on data produced by two coupled deterministic sub-systems, namely we set $\sigma = 0$ for case A and case B. Though this case is *a priori* out of the range of our assumptions, results are nonetheless satisfactory. As stated previously, this is due to two ingredients: a) the Lorenz sub-system (1) always remains in the chaotic regime which ensures that t_{L} does not unboundedly increases, and b) the integral of the p.d.f. $P(\mathbf{x}, t)$ over an interval might be correctly evaluated with a reasonable number of points. Note that, if system (1) with $\sigma = 0$ were a regular deterministic nonlinear system, the p.d.f. would be a singular delta-function associated with a fixed point or periodic or quasi-periodic orbits. The integration over an interval of such a p.d.f. leads to a drastic loss of information determining the failure of the present method.

For case A (respectively, case B), the time delay was obtained to be very stable with Q and equal to $t = 170$ (respectively, $t = 160$). For case A and $S = 100$, $D = 10$, table V shows, for various Q, the first eigenvalue and eigenvector containing most part of the coherence. The orbit and 2D phase plot for the reconstructed variable $b(t)$ with $Q = 2000$ are then presented in fig. 6. All these results are quite satisfactory when compared to the true forcing. Similarly the percentage of false nearest neighbours of $b(t)$ shown in table III (respectively, table IV) is correct for Case B (respectively, Case A) with $Q = 2000$.

TABLE V. – *Case A, $\sigma = 0$: the first eigenvector and eigenvalue ($S = 100, D = 10$) for various Q ($Q = 100, Q = 500, Q = 1000, Q = 2000, Q = 3000, Q = 10000$), with $M = 10^6$.*

0.14908	0.14907	0.14907	0.14906	0.14905	0.14902
−0.00371	−0.00370	−0.00370	−0.00371	−0.00372	−0.00376
−0.09466	−0.09483	−0.09487	−0.09496	−0.09505	−0.09534
−0.32132	−0.32149	−0.32151	−0.32152	−0.32156	−0.32163
−0.42248	−0.42266	−0.42265	−0.42261	−0.42262	−0.42253
−0.45669	−0.45691	−0.45692	−0.45688	−0.45681	−0.45654
−0.45731	−0.45746	−0.45748	−0.45744	−0.45737	−0.45712
−0.42873	−0.42813	−0.42807	−0.42810	−0.42814	−0.42841
−0.31749	−0.31726	−0.31725	−0.31733	−0.31740	−0.31766
−0.08353	−0.08371	−0.08375	−0.08383	−0.08391	−0.08418
−0.00282	−0.00281	−0.00282	−0.00282	−0.00283	−0.00286

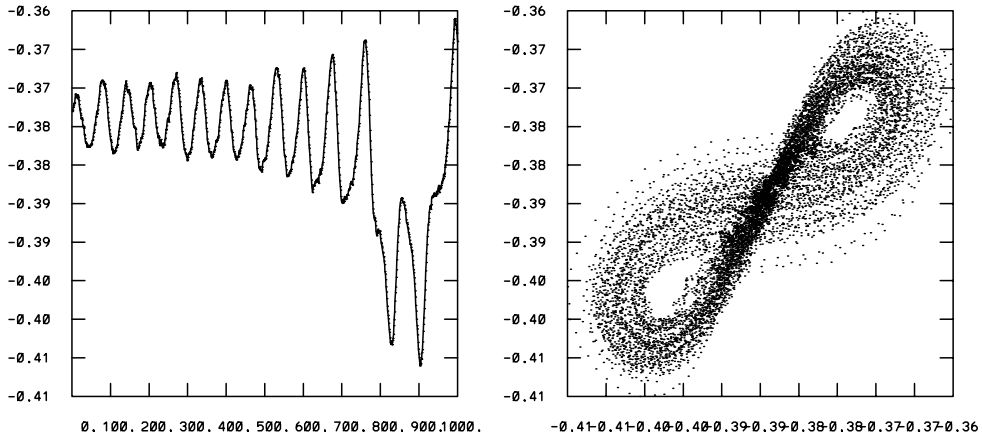

Fig. 6. – Case A, $\sigma = 0$ with $S = 100$, $D = 10$ and $Q = 2000$. The orbit and a 2D phase plot $(b(t_j), b(t_{j+10}))$ for the reconstructed $b(t)$ with $Q = 2000$.

4. – Conclusion

We have presented a new method which can retrieve a deterministic forcing from scalar measurements of a forced stochastic or deterministic system. Classical procedure fails in case considered because of dynamical noise or because of the presence of various characteristic times. The statistical method presented, which computes integrals of p.d.f. and their projections on a POD basis, produces a reconstructed deterministic quantity. Examples indicate that qualitative as well as some quantitative features can be correctly obtained from this approach.

* * *

The authors acknowledge the financial support of a CNR-CNRS agreement of Scientific Cooperation (Project 8006).

REFERENCES

[1] Ott E., Sauer T. and Yorke J. G., *Coping with Chaos* (John Wiley, New York) 1994.
[2] Weigend A. S. and Gershenfeld N. A., *Time Series Prediction* (Santa Fe Institute, Addison-Wesley) 1994.
[3] Kantz H. and Schreiber Th., *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge) 1997.
[4] Kostelich E. J. and Schreiber T., *Phys. Rev. E*, **3** (1993) 1752.
[5] Paparella F. and Provenzale A., *Distinguishing chaos from noise: the role of embedding delay*, in *Forecasting and Modelling for Chaotic and Stochastic System*, edited by Bellacicco A. *et al.* (FrancoAngeli, Milano) 1995.
[6] Von Hardenberg J. *et al.*, *Ann. N. Y. Acad. Sci.*, **808** (1997) 79.
[7] Battiston L. and Rossi M., *Int. J. Chaos Theory Appl.*, **5** (2000) 3.
[8] Von Hardenberg J. and Provenzale A., *Dynamics of forced and coupled systems*, in *Proceedings of the International School of Physics Enrico Fermi* (IOS Press Amsterdam) 1997.
[9] Boffetta G. *et al.*, *Physica D*, **116** (1998) 301

[10] Chavent G., *Identification of distributed parameter systems*, in *IFAC Symposium on Identification and System Parameters Estimations*, **85** (1979) 97.
[11] Fullana J. M., *Identification d'équations modèles décrivant un écoulement de Bénard-Von-Karman*. Thèse de l'Université de Paris VI (1997).
[12] Holmes P., Lumley J. L. and Berkooz G., *Turbulence, Coherent Structures, Dynamical Systems, and Symmetry* (Cambridge University Press, Cambridge) 1996.
[13] Gardiner C.W., *Handbook of Stochastic Methods* (Springer-Verlag, Berlin) 1983.
[14] Risken H., *The Fokker-Planck Equation* (Springer-Verlag, Berlin) 1989.
[15] Guckenheimer J. and Holmes P., *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer-Verlag, Berlin) 1983.
[16] Moss F. and McClintock P. V. E., *Noise in Nonlinear Dynamical Systems*, Vol. **3** (Cambridge University Press, Cambridge) 1989.
[17] Lorenz E. N., *J. Atmos. Sci.*, **20** (1963) 130.
[18] Mannella R., *Computer experiments in non-linear stochastic physics*, in *Noise in Nonlinear Dynamical Systems*, Vol. **3**, edited by Moss F. and McClintock P. V. E. (Cambridge University Press, Cambridge) 1989.
[19] Hegger R., Kantz H. and Schreiber T., *CHAOS*, **9** (1999) 413.