



HAL
open science

Contraction analysis of nonlinear random dynamical systems

Nicolas Tabareau, Jean-Jacques Slotine

► **To cite this version:**

Nicolas Tabareau, Jean-Jacques Slotine. Contraction analysis of nonlinear random dynamical systems. [Research Report] RR-8368, INRIA. 2013, pp.17. hal-00864079v2

HAL Id: hal-00864079

<https://inria.hal.science/hal-00864079v2>

Submitted on 26 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Contraction analysis of nonlinear random dynamical systems

Nicolas Tabareau, Jean-Jacques Slotine

**RESEARCH
REPORT**

N° 8368

September 2013

Project-Teams Ascola



Contraction analysis of nonlinear random dynamical systems

Nicolas Tabareau*, Jean-Jacques Slotine†

Project-Teams Ascola

Research Report n° 8368 — September 2013 — 14 pages

Abstract: In order to bring contraction analysis into the very fruitful and topical fields of stochastic and Bayesian systems, we extend here the theory describes in [Lohmiller and Slotine, 1998] to random differential equations. We propose new definitions of contraction (almost sure contraction and contraction in mean square) which allow to master the evolution of a stochastic system in two manners. The first one guarantees eventual exponential convergence of the system for almost all draws, whereas the other guarantees the exponential convergence in L_2 of the system to a unique trajectory. We then illustrate the relative simplicity of this extension by analyzing usual deterministic properties in the presence of noise. Specifically, we analyze stochastic gradient descent, impact of noise on oscillators synchronization and extensions of combination properties of contracting systems to the stochastic case. This is a first step towards combining the interesting and simplifying properties of contracting systems with the probabilistic approach.

Key-words: random differential equations, contraction theory

* Inria, Nantes, France

† Non Linear System Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA

**RESEARCH CENTRE
RENNES – BRETAGNE ATLANTIQUE**

Campus universitaire de Beaulieu
35042 Rennes Cedex

Analyse de la contraction pour les systèmes dynamique aléatoires non-linéaires

Résumé : La théorie de la contraction ([Lohmiller and Slotine, 1998]) sert à l'étude de la stabilité des systèmes dynamiques non-linéaires. Dans ce rapport, nous étendons la théorie de la contraction aux cas des équations différentielles aléatoires. Nous proposons deux nouvelles définitions de la contraction dans un cadre aléatoire (contraction presque sûre, et contraction aux moindres carrés) qui permettent de contrôler l'évolution d'un système stochastique de deux manières. La première garantie la convergence exponentielle pour presque toutes les réalisations du système, tandis que la deuxième garantie la convergence dans L_2 à une trajectoire unique. Nous illustrons enfin la relative simplicité de cette extension en analysant des propriétés déterministes bien connu dans un cadre bruité. Plus spécifiquement, nous analysons la descente de gradient stochastique, l'impact du bruit sur la synchronisation d'oscillateurs et l'extension des propriétés de combinaison de systèmes contractant aux cas stochastique.

Mots-clés : équations différentielles aléatoires, théorie de la contraction

Contents

1	Introduction	3
2	Part 1 : State’s dependency	4
2.1	Nonlinear random system : the discrete time case	4
2.1.1	Almost sure contraction : asymptotic exponential convergence a.s.	4
2.1.2	Contraction in mean square: asymptotic exponential convergence in mean square	6
2.2	Stochastic gradient	7
2.3	Nonlinear random system : the continuous time case	7
2.3.1	Almost sure contraction	8
2.3.2	Contraction in mean square	9
3	Part 2 : Noise’s dependency	11
3.1	Study of the average trajectory	12
3.2	Study of deviation	12
3.3	Oscillator Synchronization	13

1 Introduction

The concept of random dynamical systems appeared to be a useful tool for modelers of many different fields, especially in computational neuroscience. Indeed, introducing noise in a system can be either a way to describe reality more sharply or to generate new behaviors absent from the deterministic world. Naively, the definition of such systems generalized the usual deterministic one as follows,

$$\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i, i, \xi_i(\omega)) \quad \text{or} \quad \dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, \xi_t(\omega), t) \tag{1}$$

where ξ_i (resp. ξ_t) defines a discrete (resp. continuous) stochastic process.

Unfortunately, it comes out that whereas the extension of the definition in the discrete time case is rather straightforward, things get much harder for continuous time. Indeed, equation 1 does not make sense for every continuous stochastic process. There are basically two different ways of fixing this problem :

1. One can restrict the definition to sufficiently smooth stochastic processes, namely those whose trajectories are *cadlag* (see subsection 2.3). This gives rise to *random differential equations*.
2. Or, as important stochastic processes like *White noise* do not satisfy this property, one can switch to stochastic differential equations and Itô calculus.

When first looking at it, Itô calculus seems more accurate as it just generalized the usual differential calculus. But the problem of this formalism lies in the absence of “realisability” of such defined systems. By realisability, we mean the ability to simulate the behavior of a system by mean of a computer. Hence, all result we can obtain are purely theoretic, and it is hard to make up an intuition on how such a system evolves. From a modeling point of view, this lack of manageability make this formalism unusable in most fields as we are hardly able to find explicit solution of a huge dynamical system as encountered in biology, and thus even in the deterministic case. It follows that we need our system to be realizable, and so we must restrict our attention to random differential equations.

Once this restriction performed, we can really solve our system with the traditional tools of differential analysis. Indeed, it now makes sense to fix a draw in the probability space and solve the equation as the noise being an external input. We can thus describes the trajectory of the system given a sample path

of the stochastic process. It results that we are also able to use deterministic contraction theory to study stability when the noise never put the system out of contraction bounds. But this implies that every trajectory is contracting, and so noise does not really matter.

In this paper, we will investigate new definitions of stochastic stability together with sufficient condition to guarantee that a random differential system has a nice behavior even if the noise can induce partial divergence in the trajectories.

The notations of this paper follow those of the keystone [Arnold, 1998].

2 Part 1 : State's dependency

2.1 Nonlinear random system : the discrete time case

2.1.1 Almost sure contraction : asymptotic exponential convergence a.s.

As a first step, we define the stochastic contraction in the field of discrete system. In that case, there is no problem regarding the equation generating the dynamic system as it is just the iterated application of possibly different functions. We are dealing with stochastic processes of the form :

$$\xi : \Omega \times \mathbb{N} \rightarrow \mathbb{R}^n$$

ie. we assume ξ to be equivalent to a sequence of random variable $\xi_i : \Omega \rightarrow \mathbb{R}^n$.

The definition of contraction in discrete-time case is a direct extension of the deterministic case and makes use of the notion of *discrete stochastic differential systems* ([Jaswinski, 1970]) of the form

$$\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i, i, \xi_i(\omega))$$

where the \mathbf{f}_i are continuously differentiable (a condition that will be assumed in the rest of this paper).

This definition is the natural extension of the definition of [Lohmiller and Slotine, 1998], which assessed that the the difference between two trajectories tends to zero exponentially fast. In the stochastic case, we look for similar conditions satisfy almost surely.

First, we have to reformulate the traditional property satisfied by the metric allowing a space deformation. In the deterministic case, the metric $M_i(\mathbf{x}_i, i)$ has to be uniformly positive definite with respect to \mathbf{x}_i and i . This property of uniformity for a metric $M_i(\mathbf{x}, i, \xi_i)$ depending also on noise can be written

$$\exists \lambda \quad \forall \mathbf{x}, i \quad M(\mathbf{x}, i, \xi_i) \geq \lambda \mathbf{I} \quad \text{a.s.}$$

But as we want the noise to introduce local bad behaviors, we need to relax the property in a sense that contraction can only been guaranteed asymptotically. This introduces a slightly difficulty as the naive formulation

$$\exists 0 \leq \alpha < 1 \quad \lim_{n \rightarrow \infty} (\|\delta \mathbf{z}\| - \|\delta \mathbf{z}_0\| \alpha^n) = 0$$

just says that $\|\delta \mathbf{z}\|$ tends to zero, which is not what we want. We thus have to switch to the following refined formulation.

Definition 1 *The random system $\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i, i, \xi_i(\omega))$, is said to be almost surely contracting if there exists a uniformly positive definite metric $M_i(\mathbf{x}, i, \xi_i)$ and $\eta < 0$ such that:*

$$\mathbb{P}\{\omega \in \Omega, \limsup_{n \rightarrow \infty} \frac{1}{n} \log(\|\delta \mathbf{z}\|) \leq \eta\} = 1$$

ie. the difference between two trajectories tends almost surely to zero exponentially.

Remark 1 1. The notion of contracting region cannot be extended to stochastic case as it is hardly possible to guarantee that a stochastic trajectory stay in a bowl without requiring strong bound on the noise. And in that case, the noise can be treated as a bounded perturbation by analyzing the worst case.

2. The definition makes use of logarithm whereas $\|\delta \mathbf{z}\|$ can be equal to 0. Nevertheless, the reader should not be deterred and each time such a case appears, the equation is also satisfy if we allow infinite value and basic analytic extension (for example, $\mathbb{E}(\log \alpha) = -\infty$ as soon as $\alpha(\omega) = 0$ for some ω).

We can now state the first theorem of this paper. Remark that in the definition of the system below, the dependence on the stochastic perturbation is almost linear, as we can not master a malicious non-linear system which strongly use the “bad draw” of the stochastic process to diverge. In a sense, the system must satisfy a notion of uniformity with respect to the stochastic process.

Theorem 1 Given the random system $\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i, i, \boldsymbol{\xi}_i(\omega))$, note $\sigma_f(\mathbf{x}_i, i, \boldsymbol{\xi}_i)$ the largest singular value of the generalized Jacobian of \mathbf{f}_i at time i according to some metric $M_i(\mathbf{x}, i, \boldsymbol{\xi}_i)$.

A sufficient condition for the system to be almost surely contracting is that

- the random process $\sigma_{f_i}(\mathbf{x}, i, \boldsymbol{\xi}_i)$ can be bounded independently from \mathbf{x} , ie there exists a stochastic process η_i such that

$$\forall \mathbf{x} \quad \sigma_{f_i}(\mathbf{x}, i, \boldsymbol{\xi}_i) \leq \eta_i \quad a.s.$$

- the stochastic process $\log(\eta_i)$ follows the strong law of large number (eg. i.i.d.)

$$\frac{1}{n} \sum_{i=1}^n \log(\eta_i) \rightarrow \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\log(\eta_i)) \quad a.s.$$

- the expectation of the random variables $\log(\eta_i)$ can be uniformly bounded by some $\eta < 0$

$$\forall i \quad \mathbb{E}(\log(\eta_i)) \leq \eta$$

Proof 1 Proof.

We make a strong use of the basic idea of the original proof.

Note F_i the discrete generalized Jacobian of $f_i : F_i(\mathbf{x}, i, \boldsymbol{\xi}_i) = \Theta_{i+1} \frac{\partial}{\partial \mathbf{x}} \mathbf{f}_i(\mathbf{x}, i, \boldsymbol{\xi}_i) \Theta_i^{-1}$ we have:

$$\begin{aligned} \delta \mathbf{z}_{i+1}^T \delta \mathbf{z}_{i+1} &= \delta \mathbf{z}_i^T (F_i^T F_i) \delta \mathbf{z}_i \\ \Rightarrow \delta \mathbf{z}_{i+1}^T \delta \mathbf{z}_{i+1} &\leq \sigma_{f_i}^2(\mathbf{x}_i, i, \boldsymbol{\xi}_i) \delta \mathbf{z}_i^T \delta \mathbf{z}_i \quad a.s. \end{aligned}$$

and hence,

$$\|\delta \mathbf{z}_n\|^2 \leq \|\delta \mathbf{z}_0\|^2 \prod_{i=0}^n \sigma_{f_i}^2(\mathbf{x}_i, i, \boldsymbol{\xi}_i) \quad a.s.$$

So by monotony of logarithm and the two required properties, we can deduce that for almost every ω

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log(\|\delta \mathbf{z}_n\|) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log(\prod_{i=0}^n \sigma_{f_i}(\mathbf{x}_i, i, \boldsymbol{\xi}_i)) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n \log(\eta_i) \\ &= \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n \mathbb{E}(\log(\eta_i)) \\ &\leq \eta \end{aligned}$$

That is

$$\mathbb{P}\{\omega \in \Omega, \limsup_{n \rightarrow \infty} \frac{1}{n} \log(\|\delta \mathbf{z}\|) \leq \eta\} = 1$$

□

2.1.2 Contraction in mean square: asymptotic exponential convergence in mean square

We have seen in the subsection above sufficient conditions to guarantee almost sure asymptotic exponential convergence. But we could also be interested in looking for conditions to guarantee exponential convergence in mean square. That's what we are trying to capture with the notion of *contraction in mean square*.

Definition 2 *The random system $\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i, i, \boldsymbol{\xi}_i(\omega))$, is said to be contracting in mean square if there exists an uniformly positive definite metric $M_i(\mathbf{x}, i, \boldsymbol{\xi}_i)$ such that:*

$$\exists 0 \leq \eta < 1 \quad \mathbb{E}(\|\delta \mathbf{z}(i, \boldsymbol{\xi}_i)\|^2) \leq \|\delta \mathbf{z}_o\|^2 \eta^i$$

Theorem 2 *Given the random system , $\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i, i, \boldsymbol{\xi}_i(\omega))$ note $\sigma_{f_i}(\mathbf{x}, i, \boldsymbol{\xi}_i)$ the largest singular value of the generalized Jacobian of \mathbf{f}_i at time i according to some metric M .*

A sufficient condition for the system to be contracting in mean square is that

- *the random process $\sigma_{f_i}(\mathbf{x}, i, \boldsymbol{\xi}_i)$ can be bounded independently from \mathbf{x} , ie there exists a stochastic process η_i such that*

$$\forall \mathbf{x} \quad \sigma_{f_i}(\mathbf{x}, i, \boldsymbol{\xi}_i) \leq \eta_i \quad a.s.$$

- *the stochastic process η_i is constituted of independent random variables*
- *the expectation of the random variables η_i^2 can be uniformly bounded by some $0 \leq \eta < 1$*

$$\forall i \quad \mathbb{E}(\eta_i^2) \leq \eta$$

Proof 2 *Proof. Note F_i the discrete generalized Jacobian of f_i , we have again*

$$\delta \mathbf{z}_{i+1}^T \delta \mathbf{z}_{i+1} \leq \sigma_{f_i}^2(\mathbf{x}_i, i, \boldsymbol{\xi}_i) \delta \mathbf{z}_i^T \delta \mathbf{z}_i \leq \eta_i^2 \delta \mathbf{z}_i^T \delta \mathbf{z}_i \quad a.s.$$

Introducing the expectation value of $\|\delta \mathbf{z}\|^2$, we use that independence between X and Y is defined as the uncorrelation of $f(X)$ and $g(Y)$ for all mesurable functions f and g .

$$\mathbb{E}(\|\delta \mathbf{z}_{i+1}\|^2) \leq \mathbb{E}(\eta_i^2 \|\delta \mathbf{z}_i\|^2) = \mathbb{E}(\eta_i^2) \mathbb{E}(\|\delta \mathbf{z}_i\|^2)$$

and hence

$$\mathbb{E}(\|\delta \mathbf{z}_i\|^2) \leq \|\delta \mathbf{z}_o\|^2 \eta^i$$

□

2.2 Stochastic gradient

Let us have a look to a stochastic way of minimizing a function, highly used in computational neuroscience community, called *stochastic gradient*.

The idea is to use the traditional minimization by *gradient descent* but we want to avoid explicit computation of the gradient as it is generally of high cost or even infeasible. For that, we introduced a stochastic perturbation which has the role of a “random differentiation”.

Let $\hat{P} = P + \Pi$ be the perturbation of the state P with respect the vector Π of stochastic processes Π_i . Define the discrete system

$$P_{n+1} = P_n - \mu \cdot (\mathcal{E}(\hat{P}_n) - \mathcal{E}(P_n))\Pi$$

with $\mu > 0$.

Providing that the Π_i 's are mutually uncorrelated and of auto-correlation σ^2 (ie. $\mathbb{E}(\Pi_i \cdot \Pi_j) = \sigma^2 \cdot \delta_{i,j}$), the system satisfies:

$$\delta P_{n+1} = \delta P_n - \mu \cdot \left(\sum_k \frac{\partial^2 \mathcal{E}}{\partial p_k \partial p_i} (P^*) \Pi_k \cdot \Pi_j \right)_{i,j} \cdot \delta P_n$$

where P^* is given by the finite difference theorem. So, by taking the expectation:

$$\mathbb{E}(\delta P_{n+1}) = (I - \mu \cdot \sigma^2 \cdot \frac{\partial^2 \mathcal{E}}{\partial P_n^2}) \cdot \mathbb{E}(\delta P_n)$$

So the system is contracting in mean square if

- $\frac{\partial^2 \mathcal{E}}{\partial P_n^2} > 0$ that is \mathcal{E} is strictly convex.
- $\mu \cdot \sigma^2 \frac{\partial^2 \mathcal{E}}{\partial P_n^2} < I$ (that is $\mu \cdot \sigma^2$ sufficiently small)

2.3 Nonlinear random system : the continuous time case

We have seen in subsection 2.1 that the notion of contraction for discrete-time varying systems harmonizes well with stochastic calculus. Unfortunately the story is less straightforward in the continuous time case. Nevertheless, as outlined in introduction, for some practical reasons, we can restrict our intention to the case of random differential systems as define in [Arnold, 1998]. Let us briefly summarize the technical background.

We want to define the stochastic extension of deterministic differential systems as follows.

$$\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t, \boldsymbol{\xi}_t(\omega)) \tag{2}$$

where $\boldsymbol{\xi}_t$ is a continuous stochastic process and \mathbf{f} is a sufficiently smooth function, namely continuously differentiable with respect to \mathbf{x}_t (the condition on \mathbf{f} can be reduced to a lipschitz condition, but as we need differentiability in the rest of this paper, we prefer to assume it right now).

But this formulation does not make sense for every kind of continuous processes. Typically, when dealing with White noise process, the right-hand part of equation 2 does not present finite variation. In order to overcome this difficulty, we will assume that $\boldsymbol{\xi}_t$ is a “nice” stationary stochastic process whose trajectories are cadlag (for the french “continue à droite et avec des limites à gauche”), ie. are right continuous with left-hand limits.

Arnold proved in [Arnold, 1998] that under some assumption on \mathbf{f} , equation 2 admits a unique solution which is a global flow, whereas in general it is just a local flow.

Theorem 3 ([Arnold, 1998]) Suppose that ξ is cadlag and $\mathbf{f} \in C^1$. Then, equation 2 admit a unique maximal solution which is a local random dynamical system continuously differentiable. If furthermore \mathbf{f} satisfies

$$\|\mathbf{f}(\xi, \mathbf{x})\| \leq \alpha(\xi)\|\mathbf{x}\| + \beta(\xi)$$

where $t \rightarrow \alpha(\xi_t(\omega))$ and $t \rightarrow \beta(\xi_t(\omega))$ are locally integrable, then the solution is a global RDS.

Thus, we cannot assume that random differential system defines a unique continuous trajectory for every ω . This problem is also known in the deterministic case where Vidyasagar has shown the prevalence of differential equations despite our knowledge of only very restrictive characterization. Indeed, the set of equations admitting a unique solution is non-meager, whereas the set of equations we are able to exhibit is meager. That is, “practically all” equations admit a unique solution whereas we can characterize “practically none” of them ! So we will assume in the rest of this paper that the solution of the differential equation exists and is a unique continuously differentiable RDS.

All those restriction are rather technical and we refer the interesting reader to [Arnold, 1998] for further explanations. We can yet give two slogans reformulating intuitions coming from those restrictions :

The perturbation is memoryless The noise appearing in the right-hand side of equation 2 is *memoryless* in the sense that only the value of the perturbation at time t enters into the generator \mathbf{f} .

The perturbation do have small variations The cadlag condition is a nice way to avoid problems generated by dramatically varying processes like White noise process while allowing interesting discontinuous processes such as jump Markov processes.

From now on, when we will talk about *random differential system*, we assume that the solution of equation 2 exists and is a unique continuously differentiable RDS. We also suppose that all the processes we are dealing with are stationary and have cadlag trajectories.

2.3.1 Almost sure contraction

We can now define the notion of almost sure contraction for the continuous-time case.

Definition 3 A random differential system $\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t, \xi_t(\omega))$ is said to be almost surely contracting if there exists an uniformly positive definite metric $M(\mathbf{x}, t, \xi_t)$ and $\eta < 0$ such that:

$$\mathbb{P}\{\omega \in \Omega, \limsup_{t \rightarrow \infty} \frac{1}{t} \log(\|\delta \mathbf{z}\|) \leq \eta\} = 1$$

ie. the difference between two trajectories tends almost surely to zero exponentially.

We now state conditions for a system to be almost surely contracting. The traditional contraction analysis requires that the largest eigenvalue of the general Jacobian is uniformly bounded by a negative constant. The stochastic version of it mainly requires that the largest eigenvalue which define a process, is bounded by a process which follows the law of large number and of expectation uniformly negative.

Theorem 4 Given the system equations , $\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t, \xi_t(\omega))$ note $\lambda_f(\mathbf{x}, t, \xi_t)$ the largest eigenvalue of the generalized Jacobian of \mathbf{f} at time t according to some metric M .

A sufficient condition for the system to be almost surely contracting is that

- the random process $\lambda_f(\mathbf{x}, t, \xi_t)$ can be bounded independently from \mathbf{x} , ie there exists a stochastic process η_t such that

$$\forall \mathbf{x} \quad \lambda_f(\mathbf{x}, t, \xi_t) \leq \eta_t \quad a.s.$$

- the stochastic process η_t follows the strong law of large number

$$\frac{1}{t} \int_0^t \eta_t \rightarrow \frac{1}{t} \int_0^t \mathbb{E}(\eta_t) \quad a.s.$$

- We can uniformly bound the expectation of the η_t with some $\eta < 0$

$$\forall t \quad \mathbb{E}(\eta_t) \leq \eta$$

Proof 3 *Proof.* We make a strong use of the basic idea of the original proof.

$$\frac{1}{2} \frac{d}{dt} (\delta \mathbf{z}^T \delta \mathbf{z}) \leq \lambda_f(\mathbf{x}_t, t, \boldsymbol{\xi}_t) \delta \mathbf{z}^T \delta \mathbf{z} \quad a.s.$$

and hence

$$\|\delta \mathbf{z}\| \leq \|\delta \mathbf{z}_o\| e^{\int_0^t \eta_t dt} \quad a.s.$$

Since η_t verifies the law of large numbers, we have almost surely

$$\begin{aligned} \limsup \frac{1}{t} \log \|\delta \mathbf{z}\| &\leq \limsup \frac{1}{t} \int_0^t \eta_t dt \\ &= \limsup \frac{1}{t} \int_0^t \mathbb{E}(\eta_t) dt \\ &\leq \eta \end{aligned}$$

□

Remark 2 *It is reassuring that if we take a continuous random system satisfying conditions above, then the “discrete envelop” defined by $X_{n+1} = \exp[\int_{P_n} \lambda_f(\mathbf{x}_t, t, \boldsymbol{\xi}_t)] X_n$ is a discrete almost surely contracting system.*

2.3.2 Contraction in mean square

As it is the case in discrete-time case, we would like to find sufficient that guarantee the contraction in mean square of our system

$$\mathbb{E}(\|\delta \mathbf{z}(t, \boldsymbol{\xi}_t)\|^2) \leq \|\delta \mathbf{z}_o\|^2 e^{\eta t} \quad (3)$$

Unfortunately, we have seen that this property required discrete independent stochastic processes, whose continuous counterpart are processes like the White noise process. As we have refused to deal with that kind of processes, we need to find a stronger condition that yet ensure a similar constraint on the average trajectory.

That’s why we are moving to coarse-grained version of equation 3, namely where the property is guaranteed only for a discrete sample of the average trajectory. This property will be assessed when dealing with stochastic process which are coarse-grain independent, as define below.

Definition 4 *A random differential system $\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t, \boldsymbol{\xi}_t(\omega))$ is coarse-grain contracting if there exists a metric $M(\mathbf{x}, t, \boldsymbol{\xi}_t)$ and a partition $t_1 < t_2 < \dots$ such that:*

$$\exists 0 \leq \eta < 1, \forall i \quad \mathbb{E}(\|\delta \mathbf{z}_{t_i}\|^2) \leq \|\delta \mathbf{z}_o\|^2 \eta^i$$

To guarantee this property, we have to deal with particular kind of continuous stochastic processes which satisfy a condition of independence in a coarse-grain scale.

Definition 5 (coarse-grain independence) A continuous stochastic process η_t is said to be coarse grain independent with respect to a partition (P_i) of \mathbb{R}^+ if

$$\text{the random variables } \eta_i = \int_{P_i} \eta_t dt \text{ are independent}$$

Remark 3 1. By a partition (P_n) , we mean equivalently a strictly increasing infinite sequence $t_1 < t_2 < \dots$ or a sequence of intervals $P_0 = [0, t_1], P_i = (t_i, t_{i+1}]$ for $i \geq 1$.

2. In case of Gaussian or uniform random variables, the condition is satisfied if two random variables lying in two different sets of the partition are always independent.

Example of coarse grain independent process We will now define the typical type of coarse grain independent process we have in mind. Take a partition (P_n) of \mathbb{R}^+ and an independent stochastic process $G_n(\omega)$.

Define the process $\gamma_t(\omega) = G_n(\omega)$ for $t \in P_n$.

Then $\gamma_t(\omega)$ is a coarse grain independent process. In that case, each trajectory is piecewise constant and we have

$$\int_{P_n} \gamma_t dt = |P_n|.G_n \text{ define an independent stochastic process}$$

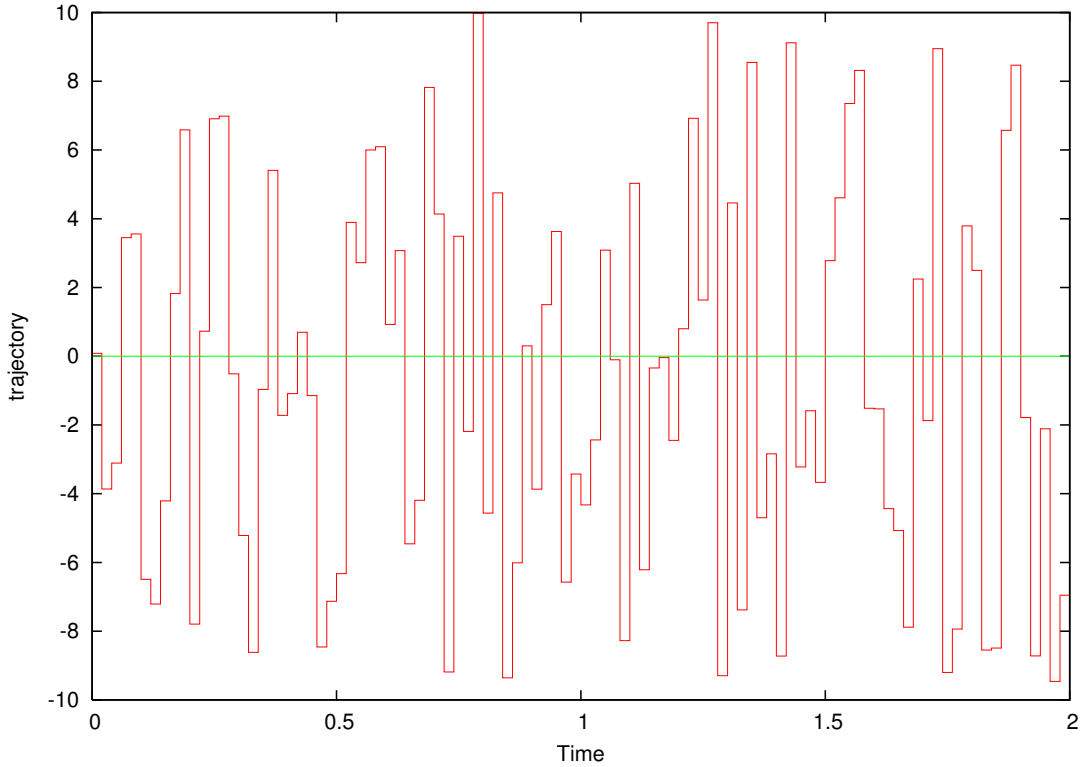


Figure 1: Example of a trajectory of a coarse grain independent process

Theorem 5 Given the system equations , $\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t, \boldsymbol{\xi}_t(\omega))$ note $\lambda_f(\mathbf{x}, t, \boldsymbol{\xi}_t)$ the largest eigenvalue of the generalized Jacobian of \mathbf{f} at time t according to some metric M .

A sufficient condition for the system to be coarse-grain contracting is that

- the random process $\lambda_f(\mathbf{x}, t, \boldsymbol{\xi}_t)$ can be bounded independently from \mathbf{x} , ie there exists a stochastic process η_t such that

$$\forall \mathbf{x} \quad \lambda_f(\mathbf{x}_t, t, \boldsymbol{\xi}_t) \leq \eta_t \quad a.s.$$

- the process η_t is a coarse-grain independent stochastic process with respect to a partition (P_n) .
- We can uniformly bound the expectation of the $e^{\int_{P_n} \eta_t}$ with some $0 \leq \eta < 1$

$$\forall n \quad \mathbb{E}((e^{\int_{P_n} \eta_t})^2) \leq \eta$$

Proof 4 Proof.

$$\frac{1}{2} \frac{d}{dt} (\delta \mathbf{z}^T \delta \mathbf{z}) \leq \lambda_f(\mathbf{x}_t, t, \boldsymbol{\xi}_t) \delta \mathbf{z}^T \delta \mathbf{z} \leq \eta_t \delta \mathbf{z}^T \delta \mathbf{z} \quad a.s.$$

Which leads to

$$\|\delta \mathbf{z}\| \leq \|\delta \mathbf{z}_0\| e^{\int_0^t \eta_t dt} \quad a.s.$$

Thus, we can define the system $Z_{n+1} = e^{\int_{t_n}^{t_{n+1}} \eta_t dt} Z_n$ and $Z_0 = \|\delta \mathbf{z}_0\|$, which satisfies

$$\|\delta \mathbf{z}_{t_n}\| \leq Z_n$$

By definition of coarse-grain stochastic process and as $\mathbb{E}((e^{\int_{P_n} \eta_t})^2) \leq \eta$, we can applied theorem 2 to conclude on the contraction in mean square of Z_n with rate η

$$\mathbb{E}(\|\delta \mathbf{z}_{t_i}\|^2) \leq \mathbb{E}(Z_i^2) \leq \|\delta \mathbf{z}_0\|^2 \eta^i$$

□

Remark 4 The condition imposed on the process η_t , namely $\mathbb{E}((e^{\int_{P_n} \eta_t})^2) \leq \eta < 1$, is really different from the condition we have seen for the almost sure contraction $\mathbb{E}(\eta_t) \leq \eta < 0$.

3 Part 2 : Noise's dependency

Let us now turn on a very special case of perturbed systems, namely when the impact of the noise does not depend on the state space.

Proposition 1 Consider a contracting system $\hat{\mathbf{x}} = \mathbf{f}(\hat{\mathbf{x}}, t)$ and take a perturbed version of it $\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t) + \boldsymbol{\xi}_t(\omega)$. The system is automatically both contracting on average and almost surely contracting.

The proof is obvious as it is the case mention above of a system which is contracting for every ω . Let us now study the mean and the variance of the unique solution.

3.1 Study of the average trajectory

Consider a contracting system $\dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}, t)$ in the metric $M = I$, take a perturbed version of it $\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t) + \boldsymbol{\xi}_t(\omega)$. Then, assuming that

- $\mathbb{E}(\boldsymbol{\xi}_t) = 0$
- $\forall t \|\boldsymbol{\xi}_t\| < \alpha$ almost surely with $\mathbb{E}(\alpha) < \infty$

we have that

$$\forall t \quad \mathbb{E}(\mathbf{x}_t) = \hat{\mathbf{x}}_t$$

Proof 5 *Proof.* Let $\mathbf{x}_\Delta = \hat{\mathbf{x}} - \mathbb{E}(\mathbf{x}_t)$ and consider

$$\begin{aligned} \mathbf{x}_\Delta^T \frac{d}{dt}(\mathbf{x}_\Delta) &= \mathbf{x}_\Delta^T (\mathbf{f}(\hat{\mathbf{x}}, t) - \mathbb{E}(\mathbf{f}(\mathbf{x}_t, t))) \quad \text{by dominated convergence theorem} \\ &= \mathbf{x}_\Delta^T \mathbb{E}(\mathbf{f}(\hat{\mathbf{x}}, t) - \mathbf{f}(\mathbf{x}_t, t)) \\ &= \mathbf{x}_\Delta^T \mathbb{E} \left(\int_0^1 \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\hat{\mathbf{x}} + c\mathbf{x}_\Delta, t) dc \mathbf{x}_\Delta \right) \\ &\leq \mathbb{E} \left(\int_0^1 \lambda_f^{max} \mathbf{x}_\Delta^T \mathbf{x}_\Delta \right) \\ &= \lambda_f^{max} \|\mathbf{x}_\Delta\|^2 \end{aligned}$$

So we have that $\|\mathbf{x}_\Delta\| \leq \|\mathbf{x}_{\Delta_0}\| e^{2\lambda_f^{max} t}$. But $\|\mathbf{x}_{\Delta_0}\| = 0$. \square

3.2 Study of deviation

Theorem 6 *Take random system $\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t) + \boldsymbol{\xi}_t(\omega)$ satisfying the conditions of the subsection 3.1. Suppose now that the deviations of the $\boldsymbol{\xi}_t$ are uniformly bounded*

$$\mathbb{E}(\|\boldsymbol{\xi}_t\|) \leq \sigma$$

The deviation of \mathbf{x}_t is then majored by the maximum deviation σ in the following way:

$$\mathbb{E}(\|\mathbf{x}_1 - \mathbf{x}_2\|) - \mathbb{E}(\|\mathbf{x}_1 - \mathbf{x}_2\|_0) e^{\lambda_f^{max} t} \leq \frac{2\sigma}{|\lambda_f^{max}|} (1 - e^{\lambda_f^{max} t})$$

Proof 6 *Proof.* Let $\tilde{\mathbf{x}} = \mathbf{x}_1 - \mathbf{x}_2$ Let us look at $\|\tilde{\mathbf{x}}\|$

$$\left. \begin{aligned} \dot{\mathbf{x}}_1 &= \mathbf{f}(\mathbf{x}_1, t) + \boldsymbol{\xi}_1 \\ \dot{\mathbf{x}}_2 &= \mathbf{f}(\mathbf{x}_2, t) + \boldsymbol{\xi}_2 \end{aligned} \right\} \Rightarrow \frac{d}{dt} \tilde{\mathbf{x}} = \int_0^1 \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_2 + c\tilde{\mathbf{x}}, t) dc \tilde{\mathbf{x}} + \boldsymbol{\xi}_1 - \boldsymbol{\xi}_2$$

Multiply by $\tilde{\mathbf{x}}^T$, it becomes:

$$\tilde{\mathbf{x}}^T \frac{d}{dt} \tilde{\mathbf{x}} = \tilde{\mathbf{x}}^T \int_0^1 \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_2 + c\tilde{\mathbf{x}}, t) dc \tilde{\mathbf{x}} + \tilde{\mathbf{x}}^T (\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)$$

$$\frac{1}{2} \frac{d}{dt} \|\tilde{\mathbf{x}}\|^2 \leq \lambda_f^{max} \|\tilde{\mathbf{x}}\|^2 + \|\tilde{\mathbf{x}}\| \|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\|$$

So, by dominated convergence theorem again

$$\frac{d}{dt} \mathbb{E}(\|\tilde{\mathbf{x}}\|) \leq \lambda_f^{max} \mathbb{E}(\|\tilde{\mathbf{x}}\|) + \mathbb{E}(\|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\|) \leq \lambda_f^{max} \mathbb{E}(\|\tilde{\mathbf{x}}\|) + 2\sigma$$

Solving $\frac{d}{dt} \mathbb{E}(\|\tilde{\mathbf{x}}\|) = \lambda_f^{max} \mathbb{E}(\|\tilde{\mathbf{x}}\|) + 2\sigma$ and using the positiveness of all terms in the equation (which means that replacing $=$ by \leq just make the slope of increasement smaller), we have :

$$\mathbb{E}(\|\tilde{\mathbf{x}}\|) - \mathbb{E}(\|\tilde{\mathbf{x}}\|_0) e^{\lambda_f^{max} t} \leq \frac{2\sigma}{|\lambda_f^{max}|} (1 - e^{\lambda_f^{max} t})$$

□

3.3 Oscillator Synchronization

Consider two identical Van der Pol oscillator couples as

$$\ddot{\mathbf{x}}_1 + \alpha(\mathbf{x}_1^2 - 1)\dot{\mathbf{x}}_1 + w^2 \mathbf{x}_1 = \alpha \epsilon_1 (\dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1)$$

$$\ddot{\mathbf{x}}_2 + \alpha(\mathbf{x}_2^2 - 1)\dot{\mathbf{x}}_2 + w^2 \mathbf{x}_2 = \alpha \epsilon_2 (\dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_2)$$

where

- $\alpha > 0, w > 0$
- ϵ_1, ϵ_2 are stationary processes

Using (Combescot, Slotine 2000), we can show that $x_1 \xrightarrow[t \rightarrow \infty]{} x_2$ when

$$\mathbb{E}(\epsilon_1) + \mathbb{E}(\epsilon_2) > 1$$

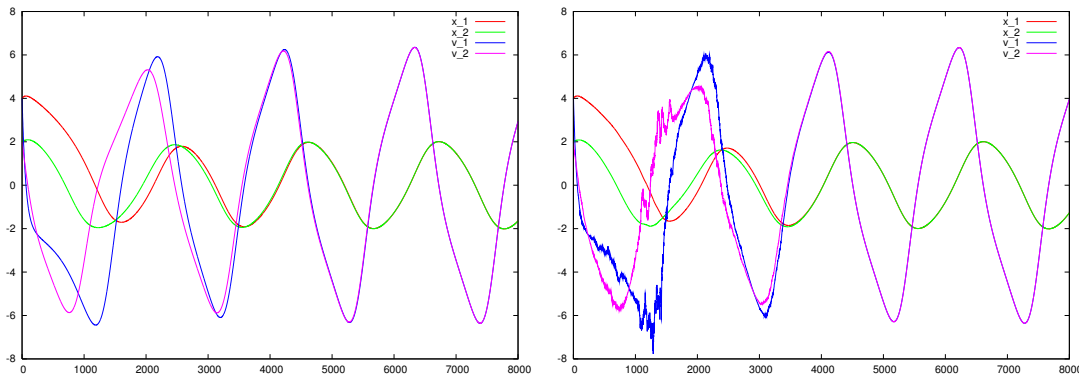


Figure 2: A comparison between stochastic synchronization of two Van der Pol oscillator with and without noise ($\epsilon_1, \epsilon_2 \in [-40, 40]$)

Remark that we can add noise in the input of both oscillators, the synchronization still occurs on average (fig. 3.3).

References

[Arnold, 1998] Arnold, L. (1998). *Random Dynamical Systems*. Springer-Verlag.
 [Jaswinski, 1970] Jaswinski, A. (1970). *Stochastic Processes and Filterig Theory*. Academic Press, New York.

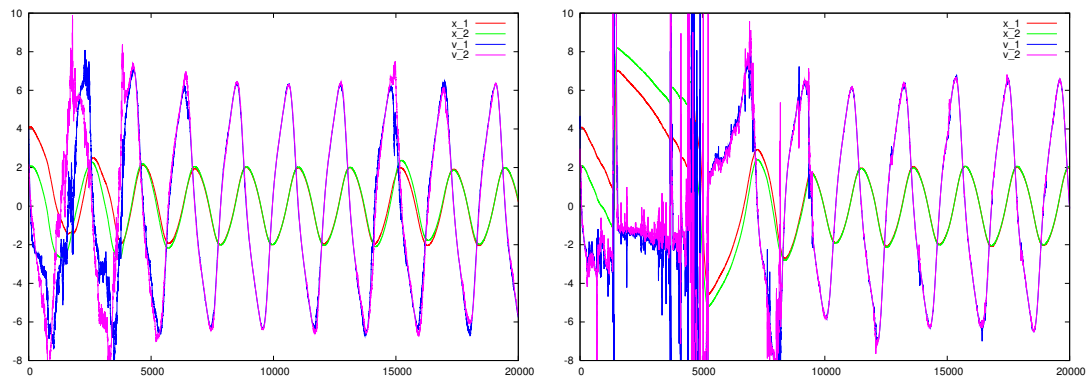


Figure 3: Stochastic synchronization with noise in input. Left part obtained with $\epsilon_1, \epsilon_2 \in [-80, 80]$, right part with $\epsilon_1, \epsilon_2 \in [-400, 400]$

[Lohmiller and Slotine, 1998] Lohmiller, W. and Slotine, J. (1998). Contraction analysis for nonlinear systems. *Automatica*, 34(6):683–696.



**RESEARCH CENTRE
RENNES – BRETAGNE ATLANTIQUE**

Campus universitaire de Beaulieu
35042 Rennes Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399