**FP6-IST-027140**

**BACS**
Bayesian Approach to Cognitive Systems

Instrument: IP  
Thematic Priority: IST

**D3.4: Contraction Analysis of Nonlinear Stochastic Systems, Theory and Applications**

Due date of deliverable: June 30, 2008  
Actual submission date: July 8, 2008

Start date of project: 01.01.2006  
Duration: 50 month

Organisation name of lead contractor for this deliverable: CNRS – LPPA

Revision 1

| Dissemination Level                                      | PU | PP                                      | RE                                      | CO |
|----------------------------------------------------------|----|-----------------------------------------|-----------------------------------------|----|
| Project co-funded by the European Commission within the  |    | Restricted to other programme participants (including the Commission Services) | Restricted to a group specified by the consortium (including the Commission Services) |    |
| Sixth Framework Programme (2002–2006)                    |    |                                        |                                         |    |
| PU | Public                  |    |                                        |                                         |    |
| PP | Restricted to other programme participants (including the Commission Services) |    |                                        |                                         |    |
| RE | Restricted to a group specified by the consortium (including the Commission Services) |    |                                        |                                         |    |
| CO | Confidential, only for members of the consortium (including the Commission Services) |    |                                        |                                         |    |
This deliverable contains the two following articles:

- “A Contraction Theory Approach to Stochastic Incremental Stability” (pp. 3-12). This manuscript is a revised version of our previous deliverable and presents the theoretical background of stochastic contraction theory. It has been accepted for publication in *IEEE Transaction on Automatic Control*.

- “Analysis of Discrete and Hybrid Stochastic Systems by Nonlinear Contraction Theory” (pp. 13-18). This manuscript extends stochastic contraction theory (which was developed for continuous systems) to the case of discrete and hybrid systems. It has been accepted at the 10th International Conference on Automation, Robotics, Control and Vision (ICARCV 2008) to be held in December 2008.
A Contraction Theory Approach to Stochastic Incremental Stability
Quang-Cuong Pham, Nicolas Tabareau and Jean-Jacques Slotine

Abstract

We investigate the incremental stability properties of Itô stochastic dynamical systems. Specifically, we derive a stochastic version of nonlinear contraction theory that provides a bound on the mean square distance between any two trajectories of a stochastically contracting system. This bound can be expressed as a function of the noise intensity and the contraction rate of the noise-free system. We illustrate these results in the contexts of nonlinear observers design and stochastic synchronization.

Index Terms
Stochastic stability, incremental stability, nonlinear contraction theory

I. INTRODUCTION

Nonlinear stability properties are often considered with respect to an equilibrium point or to a nominal system trajectory (see e.g. [1]). By contrast, incremental stability is concerned with the behavior of system trajectories with respect to each other. From the triangle inequality, global exponential incremental stability (any two trajectories tend to each other exponentially) is a stronger property than global exponential convergence to a single trajectory.

Historically, work on deterministic incremental stability can be traced back to the 1950's [2; 3; 4] (see e.g. [5; 6] for a more extensive list and historical discussion of related references). More recently, and largely independently of these earlier studies, a number of works have put incremental stability on a broader theoretical basis and have clarified the relations with more traditional stability approaches [7; 8; 9; 10]. Furthermore, it has been shown that incremental stability is especially relevant in the study of such problems as state detection [10], observer design, or synchronization analysis.

While the above references are mostly concerned with deterministic stability notions, stability theory has also been extended to stochastic dynamical systems, see for instance [11; 12]. This includes important recent developments in Lyapunov-like approaches [13; 14], as well as applications to standard problems in systems and control [15; 16; 17]. However, stochastic versions of incremental stability have not yet been systematically investigated.

The goal of this paper is to extend some concepts and results in incremental stability to stochastic dynamical systems. More specifically, we derive a stochastic version of contraction analysis in the specialized context of state-independent metrics.

We prove in section II that the mean square distance between any two trajectories of a stochastically contracting system is upper-bounded by a constant after exponential transients. In contrast with previous works on incremental stochastic stability [18], we consider the case when the two trajectories are affected by distinct and independent noises, as detailed in section II-B. This specificity enables our theory to have a number of new and practically important applications. However, the fact that the noise does not vanish as two trajectories get very close to each other will prevent us from obtaining asymptotic almost-sure stability results (see section III-B). In section III-D, we show that results on combinations of...
deterministic contracting systems have simple analogues in the stochastic case. Finally, as illustrations of our results, we study in section IV the convergence of contracting observers with noisy measurements, and the synchronization of noisy FitzHugh-Nagumo oscillators.

II. THE STOCHASTIC CONTRACTION THEOREM

A. Background: nonlinear contraction theory

Contraction theory [8] provides a set of tools to analyze the incremental exponential stability of nonlinear systems, and has been applied notoriously to observer design [8; 19; 20; 21; 22], synchronization analysis [23; 24] and systems neuroscience modelling [25]. Nonlinear contracting systems enjoy desirable aggregation properties, in that contraction is preserved under many types of system combinations given suitable simple conditions [8].

While we shall derive global properties of nonlinear systems, many of our results can be expressed in terms of eigenvalues of symmetric matrices [26]. Given a square matrix \( A \), the symmetric part of \( A \) is denoted by \( A_s \). The smallest and largest eigenvalues of \( A_s \) are denoted by \( \lambda_{\min}(A) \) and \( \lambda_{\max}(A) \). Given these notations, a matrix \( A \) is positive definite (denoted \( A > 0 \)) if \( \lambda_{\min}(A) > 0 \). Finally, a time- and state-dependent matrix \( \mathbf{A}(x, t) \) is uniformly positive definite if

\[
\exists \beta > 0 \quad \forall x, t \quad \lambda_{\min}(\mathbf{A}(x, t)) \geq \beta
\]

The basic theorem of contraction analysis, derived in [8], can be stated as follows:

**Theorem 1 (Deterministic contraction):** Consider, in \( \mathbb{R}^n \), the deterministic system

\[
\dot{x} = f(x, t)
\]

where \( f \) is a smooth nonlinear function satisfying standard conditions for the global existence and uniqueness of solutions (for instance: for all \( T \in [0, \infty) \), there are constants \( M \) and \( L \) such that \( \forall t \in [0, T], \forall x \in \mathbb{R}^n : ||f(x, t)|| \leq M + L||x|| \) [4]).

Denote the Jacobian matrix of \( f \) with respect to its first variable by \( \frac{\partial f}{\partial x} \). If there exists a square matrix \( \Theta(x, t) \) such that \( \mathbf{M}(x, t) = \Theta(x, t)^T \Theta(x, t) \) is uniformly positive definite and the matrix

\[
\mathbf{F}(x, t) = \left( \frac{d}{dt} \Theta(x, t) + \Theta(x, t) \frac{\partial f}{\partial x} \right) \Theta^{-1}(x, t)
\]

is uniformly negative definite, then all system trajectories converge exponentially to a single trajectory, with convergence rate \( |\sup_{x,t} \lambda_{\max}(\mathbf{F})| = \lambda > 0 \). The system is said to be contracting, \( \mathbf{F} \) is called its generalized Jacobian, \( \mathbf{M}(x, t) \) its contraction metric and \( \lambda \) its contraction rate.

B. Settings

Consider a noisy system described by an Itô stochastic differential equation

\[
\begin{align*}
    da &= f(a, t)dt + \sigma(a, t)dW^d \\
    a(0) &= \xi
\end{align*}
\]

where \( f \) is a \( \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^n \) function, \( \sigma \) is a \( \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^{nd} \) matrix-valued function, \( W^d \) is a standard \( d \)-dimensional Wiener process and \( \xi \) is a random variable independent of the noise \( W^d \). To ensure existence and uniqueness of solutions to equation (II.2), we assume that for all \( T \in [0, \infty) \)

**Lipschitz condition** there exists a constant \( K_1 > 0 \) such that \( \forall t \in [0, T], \forall a, b \in \mathbb{R}^n \)

\[
||f(a, t) - f(b, t)|| + ||\sigma(a, t) - \sigma(b, t)|| \leq K_1||a - b||
\]

**Restriction on growth** there exists a constant \( K_2 > 0 \) such that \( \forall t \in [0, T], \forall a \in \mathbb{R}^n \)

\[
||f(a, t)||^2 + ||\sigma(a, t)||^2 \leq K_2(1 + ||a||^2)
\]
Under these conditions, one can show ([27], p. 105) that equation (II.2) has on \([0, \infty)\) a unique \(\mathbb{R}^n\)-valued solution \(a(t)\), which is continuous with probability one.

In order to investigate the incremental stability properties of system (II.2), consider now two system trajectories \(a(t)\) and \(b(t)\). Our goal will consist of studying the trajectories \(a(t)\) and \(b(t)\) with respect to each other. For this, we consider the augmented system \(x(t) = (a(t), b(t))^T\), which follows the equation

\[
\begin{align*}
\dot{x} &= \begin{pmatrix} f(a, t) \\ f(b, t) \end{pmatrix} dt + \begin{pmatrix} \sigma(a, t) & 0 \\ 0 & \sigma(b, t) \end{pmatrix} \begin{pmatrix} dW_1^d \\ dW_2^d \end{pmatrix} \\
x(0) &= (a(0), b(0)) = (\xi_1, \xi_2)
\end{align*}
\]

(II.3)

**Important remark** As stated in the introduction, the systems \(a\) and \(b\) are driven by distinct and independent Wiener processes \(W_1^d\) and \(W_2^d\). This makes our approach considerably different from [18], where the authors studied two trajectories driven by the same Wiener process.

Our approach enables us to study the stability of the system with respect to differences in initial conditions and to random perturbations: indeed, two trajectories of any real-life system are typically affected by distinct realizations of the noise. In the deterministic domain, incremental stability with respect to different initial conditions and different deterministic inputs (incremental Input-to-State Stability or \(\delta\)ISS) has been studied in [9; 10; 28]. Besides, it should be noted that our approach leads very naturally to nice results on the comparison of noisy and noise-free trajectories (cf. section III-C), which are particularly useful in applications (cf. section IV).

However, because of the very fact that the two trajectories are driven by distinct Wiener processes, one cannot expect the influence of the noise to vanish when the two trajectories get very close to each other. This contrasts with [18], and more generally, with standard stochastic stability approaches, where the noise is assumed to vanish near the origin. The consequences of this will be discussed in detail in section III-B.

**C. Statement and proof of the theorem**

We first recall a Gronwall-type lemma

**Lemma 1:** Let \(g : [0, \infty[ \rightarrow \mathbb{R}\) be a continuous function, \(C\) a real number and \(\lambda\) a strictly positive real number. Assume that

\[
\forall u, t \quad 0 \leq u \leq t \quad g(t) - g(u) \leq \int_u^t -\lambda g(s) + C ds
\]

(II.4)

Then

\[
\forall t \geq 0 \quad g(t) \leq \frac{C}{\lambda} + \left[ g(0) - \frac{C}{\lambda} \right]^+ e^{-\lambda t}
\]

(II.5)

where \([\cdot]^+ = \max(0, \cdot)\).

**Proof** See [29] □

We now introduce two hypotheses

**H1** There exists a state-independent, uniformly positive definite metric \(M(t) = \Theta(t)^T \Theta(t)\), with the lower-bound \(\beta > 0\) (i.e. \(\forall x, t \quad x^T M(t) x \geq \beta \|x\|^2\)) and \(f\) is contracting in that metric, with contraction rate \(\lambda\), i.e. uniformly,

\[
\lambda_{\max} \left( \left( \frac{d}{dt} \Theta(t) + \Theta(t) \frac{\partial f}{\partial a} \right) \Theta^{-1}(t) \right) \leq -\lambda
\]

or equivalently, uniformly,

\[
M(t) \frac{\partial f}{\partial a} + \left( \frac{\partial f}{\partial a} \right)^T M(t) + \frac{d}{dt} M(t) \leq -2\lambda M(t)
\]
(H2) \( \text{tr} \left( (a, t)^T M(t)(a, t) \right) \) is uniformly upper-bounded by a constant \( C \)

**Definition 1:** A system that verifies (H1) and (H2) is said to be **stochastically contracting** in the metric \( M(t) \), with rate \( \lambda \) and bound \( C \).

Consider the Lyapunov-like function \( V(x, t) = (a - b)^T M(t)(a - b) \). Using (H1) and (H2), we derive below an inequality on \( L^V(x, t) \) where \( L \) denotes the differential generator of the Itô process \( x(t) \) ([11], p. 15).

**Lemma 2:** Under (H1) and (H2), one has the inequality

\[
L^V(x, t) \leq -2\lambda V(x, t) + 2C \quad \text{(II.6)}
\]

**Proof** Let us compute first \( L^V \)

\[
L^V(x, t) = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} \bar{f}(x, t) + \frac{1}{2} \text{tr} \left( (a, t)^T \frac{\partial^2 V}{\partial x^2} (a, t) \right)
\]

\[
= (a - b)^T \left( \frac{dT}{dt} M(t) \right) (a - b) + 2(a - b)^T M(t)(f(a, t) - f(b, t)) + \text{tr}(\sigma(a, t)^T M(t)\sigma(a, t)) + \text{tr}(\sigma(b, t)^T M(t)\sigma(b, t))
\]

Fix \( t > 0 \), then, according to [30], there exists \( c \in [a, b] \) such that

\[
(a - b)^T \left[ \frac{dT}{dt} M(t) \right] (a - b) + 2(a - b)^T M(t)(f(a) - f(b))
\]

\[
= (a - b)^T \left[ \frac{dT}{dt} M(t) + M(t) \frac{\partial f}{\partial a}(c, t) + \frac{\partial f}{\partial a}(c, t)^T M(t) \right] (a - b)
\]

\[
\leq -2\lambda(a - b)^T M(t)(a - b) = -2\lambda V(x) \quad \text{(II.7)}
\]

where the inequality is obtained by using (H1).

Finally, combining equation (II.7) with (H2) allows to obtain the desired result \( \Box \)

We can now state the stochastic contraction theorem

**Theorem 2 (Stochastic contraction):** Assume that system (II.2) verifies (H1) and (H2). Let \( a(t) \) and \( b(t) \) be two trajectories whose initial conditions are independent of the noise and given by a probability distribution \( p(\xi_1, \xi_2) \). Then

\[
\forall t \geq 0 \quad \mathbb{E} \left( (a(t) - b(t))^T M(t)(a(t) - b(t)) \right) \leq \frac{C}{\lambda} + e^{-2\lambda t} \int \left[ (a_0 - b_0)^T M(0)(a_0 - b_0) - \frac{C}{\lambda} \right]^+ dp(a_0, b_0) \quad \text{(II.8)}
\]

In particular, \( \forall t \geq 0 \)

\[
\mathbb{E} \left( \|a(t) - b(t)\|^2 \right) \leq \frac{1}{\beta} \left( \frac{C}{\lambda} + \mathbb{E} \left( (\xi_1 - \xi_2)^T M(0)(\xi_1 - \xi_2) \right) e^{-2\lambda t} \right) \quad \text{(II.9)}
\]

**Proof** Let \( x_0 = (a_0, b_0) \in \mathbb{R}^{2n} \). By Dynkin’s formula ([11], p. 10)

\[
\mathbb{E}_{x_0} V(x(t), t) - V(x_0, 0) = \mathbb{E}_{x_0} \int_0^t L^V(x(s), s)ds
\]

Thus one has \( \forall u, t \quad 0 \leq u \leq t < \infty \)

\[
\mathbb{E}_{x_0} V(x(t), t) - \mathbb{E}_{x_0} V(x(u), u) = \mathbb{E}_{x_0} \int_u^t L^V(x(s), s)ds
\]

\[
\leq \mathbb{E}_{x_0} \int_u^t (-2\lambda V(x(s), s) + 2C)ds \quad \text{(II.10)}
\]

\[
= \int_u^t (-2\lambda \mathbb{E}_{x_0} V(x(s), s) + 2C)ds \quad \text{(II.11)}
\]
where inequality (II.10) is obtained by using lemma 2 and equality (II.11) by using Fubini’s theorem (since $s \mapsto \mathbb{E}_{x_0} V(x(s), s)$ is continuous on $[u, t]$, one has $\int_u^t -2\lambda\mathbb{E}_{x_0} V(x(s), s) + 2C|ds < \infty$).

Denote by $g(t)$ the deterministic quantity $\mathbb{E}_{x_0} V(x(t))$. As remarked above, $g(t)$ is a continuous function of $t$. It then satisfies the conditions of the Gronwall-type lemma 1, and as a consequence

$$\forall t \geq 0 \quad \mathbb{E}_{x_0} V(x(t), t) \leq \frac{C}{\lambda} + \left[ V(x_0, 0) - \frac{C}{\lambda} \right] e^{-2\lambda t}$$

which leads to (II.8) by integrating with respect to $(a_0, b_0)$. Next, (II.9) follows from (II.8) by observing that

$$\int \left[ (a_0 - b_0)^T M(0)(a_0 - b_0) - \frac{C}{\lambda} \right]^+ dp(a_0, b_0) \leq \int (a_0 - b_0)^T M(0)(a_0 - b_0) dp(a_0, b_0) = \mathbb{E}((\xi_1 - \xi_2)^T M(0)(\xi_1 - \xi_2))$$

and

$$\|a(t) - b(t)\|^2 \leq \frac{1}{\beta} (a(t) - b(t))^T M(t)(a(t) - b(t)) \square$$

### III. Remarks

#### A. “Optimality” of the mean square bound

Consider the following linear dynamical system, known as the Ornstein-Uhlenbeck (colored noise) process

$$da = -\lambda adt + \sigma dW$$  \hspace{1cm} (III.1)

Clearly, the noise-free system is contracting with rate $\lambda$ and the trace of the noise matrix is upper-bounded by $\sigma^2$. Let $a(t)$ and $b(t)$ be two system trajectories starting respectively at $a_0$ and $b_0$ (deterministic initial conditions). Then by theorem 2, we have

$$\forall t \geq 0 \quad \mathbb{E}((a(t) - b(t))^2) \leq \frac{\sigma^2}{\lambda} + \left[ (a_0 - b_0)^2 - \frac{\sigma^2}{\lambda} \right]^+ e^{-2\lambda t}$$  \hspace{1cm} (III.2)

Let us assess the quality of this bound by solving directly equation (III.1). The solution of equation (III.1) is ([27], p. 134)

$$a(t) = a_0 e^{-\lambda t} + \sigma \int_0^t e^{\lambda(s-t)} dW(s)$$  \hspace{1cm} (III.3)

Compute next the mean square distance between the two trajectories $a(t)$ and $b(t)$

$$\mathbb{E}((a(t) - b(t))^2) = (a_0 - b_0)^2 e^{-2\lambda t} + \sigma^2 \left( \mathbb{E}\left( \left( \int_0^t e^{\lambda(s-t)} dW_1(s) \right)^2 \right) \right) + \mathbb{E}\left( \left( \int_0^t e^{\lambda(u-t)} dW_2(u) \right)^2 \right)$$

$$= (a_0 - b_0)^2 e^{-2\lambda t} + \frac{\sigma^2}{\lambda}(1 - e^{-2\lambda t}) \leq \frac{\sigma^2}{\lambda} + \left[ (a_0 - b_0)^2 - \frac{\sigma^2}{\lambda} \right]^+ e^{-2\lambda t}$$

The last inequality is in fact an equality when $(a_0 - b_0)^2 \geq \frac{\sigma^2}{\lambda}$. Thus, this calculation shows that the upper-bound (III.2) given by theorem 2 is optimal, in the sense that it can be attained.

#### B. No asymptotic almost-sure stability

From the explicit form (III.3) of the solutions, one can deduce that the distributions of $a(t)$ and $b(t)$ converge to the normal distribution $\mathcal{N}(0, \frac{\sigma^2}{2\lambda})$ ([27], p. 135). Since $a(t)$ and $b(t)$ are independent, the
distribution of the difference \( a(t) - b(t) \) will then converge to \( \mathcal{N} \left( 0, \frac{\sigma^2}{\lambda} \right) \). The last observation shows that one cannot – in general – obtain almost-sure stability results.

Indeed, the main difference with the approaches in [16; 17; 18] lies in the term \( 2C \). This extra term comes from the fact that the influence of the noise does not vanish when two trajectories get very close to each other (cf. section II-B). It prevents \( \mathbb{E}V(x(t)) \) from being always non-positive, and as a result, \( V(x(t)) \) is not always non-increasing. Thus, \( V(x(t)) \) is not – in general – a supermartingale, and one cannot then use the supermartingale inequality (or its variations) to obtain asymptotic almost-sure bounds, as in ([11], pp. 47-48) or in [16; 17; 18].

However, if one is interested in finite time bounds then the supermartingale inequality is still applicable, see ([11], p. 86) for details.

C. Noisy and noise-free trajectories

Consider the following augmented system

\[
dx = \begin{pmatrix} f(a, t) \\ f(b, t) \end{pmatrix} dt + \begin{pmatrix} 0 & 0 \\ 0 & \sigma(b, t) \end{pmatrix} \begin{pmatrix} dW_1^1 \\ dW_2^2 \end{pmatrix} = \begin{pmatrix} \tilde{f}(x, t)dt + \tilde{\sigma}(x, t)dW_{2d} \end{pmatrix}
\]

(III.4)

This equation is the same as equation (II.3) except that the a-system is not perturbed by noise. Thus \( V(x) = \|a - b\|^2 \) will represent the distance between a noise-free trajectory and a noisy one. All the calculations will be the same as in the previous development, with \( C \) being replaced by \( C/2 \). One can then derive the following corollary (for simplicity, we consider the case of identity metric; the general case can be easily adapted)

**Corollary 1:** Assume that system (II.2) verifies (H1) and (H2) with \( M = I \). Let \( a(t) \) be a noise-free trajectory starting at \( a_0 \) and \( b(t) \) a noisy trajectory whose initial condition is independent of the noise and given by a probability distribution \( p(\xi_2) \). Then \( \forall t \geq 0 \)

\[
\mathbb{E} \left( \|a(t) - b(t)\|^2 \right) \leq \frac{C}{2\lambda} + \mathbb{E} \left( \|a_0 - \xi_2\|^2 \right) e^{-2Mt}
\]

(III.5)

**Remarks**

- One can note here that the derivation of corollary 1 is only permitted by our initial choice of considering distinct driving Wiener process for the a- and b-systems (cf. section II-B).
- Corollary 1 provides a robustness result for contracting systems, in the sense that any contracting system is automatically protected against noise, as quantified by (III.5). This robustness could be related to the exponential nature of contraction stability.

D. Combination properties

Stochastic contraction inherits naturally from deterministic contraction [8] its convenient combination properties. Because contraction is a state-space concept, such properties can be expressed in more general forms than input-output analogues such as passivity-based combinations [31].

It should be noted that, in the deterministic domain, combination properties have been obtained for \( \delta\text{ISS} \) systems [10; 28] (for the definition of \( \delta\text{ISS} \), see section II-B).

Consider two connected systems

\[
\begin{align*}
\dot{x}_1 &= f_1(x_1, x_2, t)dt + \sigma_1(x_1, t)dW_1 \\
\dot{x}_2 &= f_2(x_1, x_2, t)dt + \sigma_2(x_2, t)dW_2
\end{align*}
\]

where system \( i (i = 1, 2) \) is stochastically contracting with respect to \( M_i = \Theta_i^T \Theta_i \), with rate \( \lambda_i \) and bound \( C_i \) (here, \( M_i \) and \( \Theta_i \) are set to be constant matrices for simplicity; the case of time-varying metrics can be easily adapted).
Assume that these systems are connected by negative feedback \cite{32}, i.e. the Jacobian of their coupling matrices verify \( \Theta_1 J_{12} \Theta_2^{-1} = -k \Theta_2 J_{21}^T \Theta_1^{-1} \), with \( k \) a positive constant. The Jacobian matrix of the augmented noise-free system is given then by

\[
J = \begin{pmatrix}
J_1 & \Theta_2 J_{21}^T \\
J_{21} & -k \Theta_2^{-1} \Theta_1 J_{12} \Theta_1^{-1}
\end{pmatrix}
\]

Consider the coordinate transform \( \Theta = \begin{pmatrix}
\Theta_1 \\
0
\end{pmatrix}
\) associated with the metric \( M = \Theta^T \Theta > 0 \). After some calculations, one has

\[
(\Theta J \Theta^{-1})_s = \begin{pmatrix}
(\Theta_1 J_1 \Theta_1^{-1})_s & 0 \\
0 & (\Theta_2 J_2 \Theta_2^{-1})_s
\end{pmatrix} \leq \max(-\lambda_1, -\lambda_2) I \text{ uniformly}
\]

The augmented system is thus stochastically contracting in the metric \( M \), with rate \( \min(\lambda_1, \lambda_2) \) and bound \( C_1 + kC_2 \).

Similarly, one can show that (with \( \text{sing}(A) \) denoting the largest singular value of \( A \))

- **Hierarchical combination:** If \( J_{12} = 0 \) and \( \text{sing}^2(\Theta_2 J_{21} \Theta_1^{-1}) \leq K \), then the augmented system is stochastically contracting in the metric \( M_\epsilon \), with rate \( \frac{1}{2}(\lambda_1 + \lambda_2 - \sqrt{\lambda_1^2 + \lambda_2^2}) \) and bound \( C_1 + \frac{2C_2 \lambda_1 \lambda_2}{K} \), where \( \epsilon = \sqrt{\frac{2 \lambda_1 \lambda_2}{K}} \).

- **Small gains:** Define \( B_\gamma = \frac{1}{2} \left( \sqrt{\gamma} \Theta_2 J_{21} \Theta_1^{-1} + \frac{1}{\sqrt{\gamma}} (\Theta_1 J_{12} \Theta_2^{-1})^T \right) \). If there exists \( \gamma > 0 \) such that \( \text{sing}^2(B_\gamma) < \lambda_1 \lambda_2 \) then the augmented system is stochastically contracting in the metric \( M_\gamma \), with bound \( C_1 + \gamma C_2 \) and rate \( \lambda \) verifying

\[
\lambda \geq \frac{\lambda_1 + \lambda_2}{2} - \sqrt{\left(\frac{\lambda_1 - \lambda_2}{2}\right)^2 + \text{sing}^2(B_\gamma)}
\]

Taken together, the combination properties presented here allow one to build by recursion stochastically contracting systems of arbitrary size.

**IV. SOME EXAMPLES**

**A. Effect of measurement noise on contracting observers**

Consider a nonlinear dynamical system

\[
\dot{x} = f(x, t)
\]

If a measurement \( y = y(x) \) is available, then it may be possible to choose an output injection matrix \( K(t) \) such that the dynamics

\[
\dot{x} = f(\hat{x}, t) + K(t)(\hat{y} - y)
\]

is contracting, with \( \hat{y} = y(\hat{x}) \). Since the actual state \( x \) is a particular solution of (IV.2), any solution \( \hat{x} \) of (IV.2) will then converge towards \( x \) exponentially.

Assume now that the measurements are corrupted by additive “white noise”. In the case of linear measurement, the measurement equation becomes \( y = \mathbf{H}(t)x + \Sigma(t)\eta(t) \) where \( \eta(t) \) is a multidimensional “white noise” and \( \Sigma(t) \) is the matrix of measurement noise intensities.

The observer equation is now given by the following Itô stochastic differential equation (using the formal rule \( dW = \eta dt)\)

\[
d\hat{x} = (f(\hat{x}, t) + K(t)(\mathbf{H}(t)x - \mathbf{H}(t)\hat{x}))dt + K(t)\Sigma(t)dW
\]

\[
(IV.3)
\]
Next, remark that the solution \( x \) of system (IV.1) is also a solution of the noise-free version of system (IV.3). By corollary 1, one then has, for any solution \( \hat{x} \) of system (IV.3)

\[
\forall t \geq 0 \quad \mathbb{E} \left( \| \hat{x}(t) - x(t) \|^2 \right) \leq \frac{C}{2\lambda} + \| \hat{x}_0 - x_0 \|^2 e^{-2\lambda t}
\]

(IV.4)

where

\[
\lambda = \inf_{x, t} \lambda_{\max} \left| \frac{\partial f(x, t)}{\partial x} - K(t)H(t) \right|
\]

\[
C = \sup_{t \geq 0} \text{tr} \left( \Sigma(t)^T K(t)^T K(t) \Sigma(t) \right)
\]

**Remark** The choice of the injection gain \( K(t) \) is governed by a trade-off between convergence speed \( (\lambda) \) and noise sensitivity \( (C/\lambda) \) as quantified by (IV.4). More generally, the explicit computation of the bound on the expected quadratic estimation error given by (IV.4) may open the possibility of *measurement selection* in a way similar to the linear case. If several possible measurements or sets of measurements can be performed, one may try at each instant (or at each step, in a discrete version) to select the most relevant, i.e., the measurement or set of measurements which will best contribute to improving the state estimate. Similarly to the Kalman filters used in [33] for linear systems, this can be achieved by computing, along with the state estimate itself, the corresponding bounds on the expected quadratic estimation error, and then selecting accordingly the measurement which will minimize it.

### B. Synchronization of noisy FitzHugh-Nagumo oscillators

We analyze in this section the synchronization of two noisy FitzHugh-Nagumo oscillators (see [23] for the references). The interested reader is referred to [34] for a more complete study.

The dynamics of two diffusively-coupled noisy FitzHugh-Nagumo oscillators can be described by

\[
\begin{align*}
\dot{v}_i &= \left( c(v_i + w_i - \frac{1}{3}v_i^3 + I_i) + k(v_0 - v_i) \right) dt + \sigma dW_i \\
\dot{w}_i &= -\frac{1}{c}(v_i - a + bw_i)dt
\end{align*}
\]

where \( i = 1, 2 \). Let next \( x = (v_1, w_1, v_2, w_2)^T \) and \( V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \). The Jacobian matrix of the projected noise-free system is then given by

\[
\begin{pmatrix}
c - \frac{c(v_1^2 + v_2^2)}{2} - k & c \\
-1/c & -b/c
\end{pmatrix}
\]

Thus, if the coupling strength verifies \( k > c \) then the projected system will be stochastically contracting in the diagonal metric \( M = \text{diag}(1, c) \) with rate \( \min(k - c, b/c) \) and bound \( \sigma^2 \). Hence, the average absolute difference between the two membrane potentials \( |v_1 - v_2| \) will be upper-bounded by \( \sigma/\sqrt{\min(1, c) \min(k - c, b/c)} \) after exponential transients (see figure 1 for a numerical simulation).

**ACKNOWLEDGMENTS**

We are grateful to Dr S. Darses, Prof D. Bennequin and Prof M. Yor for stimulating discussions, and to the Associate Editor and the reviewers for their helpful comments.

**REFERENCES**

[1] J.-J. Slotine and W. Li, *Applied Nonlinear Control*. Prentice-Hall, 1991.
[2] D. Lewis, “Metric properties of differential equations,” *American Journal of Mathematics*, vol. 71, pp. 294–312, 1949.
[3] B. Demidovich, “Dissipativity of a nonlinear system of differential equations,” *Ser. Mat. Mekh.*, 1961.
[4] P. Hartmann, *Ordinary differential equations*. Wiley, 1964.
[5] W. Lohmiller and J.-J. Slotine, “Contraction analysis of nonlinear distributed systems,” *International Journal of Control*, vol. 78, 2005.

[6] J. Jouffroy, “Some ancestors of contraction analysis,” in *Proceedings of the IEEE Conference on Decision and Control*, 2005.

[7] E. Sontag and Y. Wang, “Output-to-state stability and detectability of nonlinear systems,” *Systems and Control Letters*, vol. 29, pp. 279–290, 1997.

[8] W. Lohmiller and J.-J. Slotine, “On contraction analysis for nonlinear systems,” *Automatica*, vol. 34, pp. 671–682, 1998.

[9] V. Fromion, G. Scorletti, and G. Ferreres, “Nonlinear performance of a pi controlled missile: an explanation,” *Int. J. Robust Nonlinear Control*, vol. 9, pp. 485–518, 1999.

[10] D. Angeli, “A lyapunov approach to incremental stability properties,” *IEEE Transactions on Automatic Control*, vol. 47, pp. 410–422, 2002.

[11] H. Kushner, *Stochastic Stability and Control*. Academic Press, 1967.

[12] R. Has’minskii, *Stochastic Stability of Differential Equations*. Rockville: Sijthoff and Nordhoff, 1980.

[13] P. Florchinger, “Lyapunov-like techniques for stochastic stability,” *SIAM Journal of Control and Optimization*, vol. 33, pp. 1151–1169, 1995.

[14] X. Mao, *Stability of Stochastic Differential Equations with Respect to Semimartingales*. White Plains, NY: Longman, 1991.

[15] P. Florchinger, “Feedback stabilization of affine in the control stochastic differential systems by the control lyapunov function method,” *SIAM Journal of Control and Optimization*, vol. 35, 1997.

[16] J. Tsinias, “The concept of “exponential iss” for stochastic systems and applications to feedback stabilization,” *Systems and Control Letters*, vol. 36, pp. 221–229, 1999.

[17] H. Deng, M. Krstic, and R. Williams, “Stabilization of stochastic nonlinear systems driven by noise of unknown covariance,” *IEEE Transactions on Automatic Control*, vol. 46, 2001.

[18] T. Caraballo, P. Kloeden, and B. Schmalfuss, “Exponentially stable stationary solutions for stochastic evolution equations and their perturbation,” *Applied Mathematics and Optimization*, vol. 50, pp. 183–207, 2004.

[19] W. Lohmiller and J.-J. Slotine, “Nonlinear process control using contraction theory,” *A.I.Ch.E. Journal*, 2000.

[20] N. Aghannan and P. Rouchon, “An intrinsic observer for a class of lagrangian systems,” *IEEE Transactions on Automatic Control*, vol. 48, 2003.

[21] J. Jouffroy and T. Fossen, “On the combination of nonlinear contracting observers and uges controllers for output feedback,” in *Proceedings of the IEEE Conference on Decision and Control*, 2004.

[22] Y. Zhao and J.-J. Slotine, “Discrete nonlinear observers for inertial navigation,” *Systems and Control Letters*, vol. 54, 2005.

[23] W. Wang and J.-J. E. Slotine, “On partial contraction analysis for coupled nonlinear oscillators.” *Biol*
[24] Q.-C. Pham and J.-J. Slotine, “Stable concurrent synchronization in dynamic system networks.” *Neural Networks*, vol. 20, no. 1, pp. 62–77, Jan. 2007.

[25] B. Girard, N. Tabareau, Q.-C. Pham, A. Berthoz, and J.-J. Slotine, “Where neuroscience and dynamic system theory meet autonomous robotics: a contracting basal ganglia model for action selection,” *Neural Networks*, vol. 21, pp. 628–641, 2008.

[26] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[27] L. Arnold, *Stochastic Differential Equations : Theory and Applications*. Wiley, 1974.

[28] A. Pavlov, N. van de Wouw, and H. Nijmeijer, *Uniform Output Regulation of Nonlinear Systems*. Birkhauser, 2006.

[29] Q.-C. Pham. (2007) A variation of gronwall’s lemma. [Online]. Available: http://arxiv.org/abs/0704.0922

[30] A. Pavlov, A. Pogromsky, N. van de Wouw, and H. Nijmeijer, “Convergent dynamics, a tribute to boris pavlovich demidovich,” *Systems and Control Letters*, vol. 52, pp. 257–261, July 2004.

[31] V. Popov, *Hyperstability of Control Systems*. Springer-Verlag, 1973.

[32] N. Tabareau and J.-J. Slotine, “Notes on contraction theory,” MIT NSL Report, Tech. Rep., 2005.

[33] E. Dickmanns, *Dynamic Vision for Intelligent Vehicles*. Course Notes, MIT EECS dept, 1998.

[34] N. Tabareau, J.-J. Slotine, and Q.-C. Pham. (2008) How synchronization helps reduce the influence of noise in neuronal networks. [Online]. Available: http://arxiv.org/abs/0801.0011
Analysis of Discrete and Hybrid Stochastic Systems by Nonlinear Contraction Theory

Quang-Cuong Pham
Laboratoire de Physiologie de la Perception et de l’Action
Collège de France - CNRS
Paris, France
cuong.pham@normalesup.org

Abstract—We investigate the stability properties of discrete and hybrid stochastic nonlinear dynamical systems. More precisely, we extend the stochastic contraction theorems (which were formulated for continuous systems) to the case of discrete and hybrid resetting systems. In particular, we show that the mean square distance between any two trajectories of a discrete (or hybrid resetting) contracting stochastic system is upper-bounded by a constant after exponential transients. Using these results, we study the synchronization of noisy nonlinear oscillators coupled by discrete noisy interactions.

Index Terms—Discrete systems, hybrid resetting, stochastic systems, nonlinear contraction theory, incremental stability, oscillator synchronization

I. INTRODUCTION

Contraction theory is a set of relatively recent tools that provide a systematic approach to the stability analysis of a large class of nonlinear dynamical systems [1], [2], [3], [4]. A nonlinear nonautonomous system \( \dot{x} = f(x, t) \) is contracting if the symmetric part of the Jacobian matrix of \( f \) is uniformly negative definite in some metric. Using elementary fluid dynamics techniques, it can be shown that contracting systems are incrementally stable, that is, any two system trajectories exponentially converge to each other [1].

From a practical viewpoint, contraction theory has been successfully applied to a number of important problems, such as mechanical observers and controllers design [5], chemical processes control [6], synchronization analysis [2], [7] or biological systems modelling [8].

Recently, contraction analysis has been extended to the case of stochastic dynamical systems governed by Itô differential equations [4]. In parallel, hybrid versions of contraction theory have also been developed [3]. A hybrid system is characterized by a continuous evolution of the system’s state, and intermittent discrete transitions. Such systems are pervasive in both artificial (e.g. analog physical processes controlled by digital devices) and natural (e.g. spiking neurons with subthreshold dynamics) environments.

This paper benefits from these recent developments, and provides an exponential stability result for discrete and hybrid systems governed by stochastic difference and differential equations. More precisely, we prove in section II and III that the mean square distance between any two trajectories of a discrete (respectively hybrid resetting) stochastic contracting system is upper-bounded by a constant after exponential transients. This bound can be expressed as function of the noise intensities and the contraction rates of the noise-free systems. In section IV, we briefly discuss a number of theoretical issues regarding our analysis. In section V, we study, using the previously developed tools, the synchronization of noisy nonlinear oscillators that interact by discrete noisy couplings. Finally, some future directions of research are indicated in section VI.

Notations The symmetric part of a matrix \( A \) is defined as \( A_s = \frac{1}{2} (A + A^T) \). For a symmetric matrix \( A \), \( \lambda_{\min}(A) \) and \( \lambda_{\max}(A) \) denote respectively the smallest and the largest eigenvalue of \( A \). A set of symmetric matrices \( (A_i)_{i \in I} \) is uniformly positive definite if \( \exists \alpha > 0, \forall i \in I, \lambda_{\min}(A_i) \geq \alpha \). Finally, for a process \( x(t) \), we note \( E_x(\cdot) = E(|x(0) = x) \).

II. DISCRETE SYSTEMS

We first prove a lemma that makes explicit the initial “discrete contraction” proof (see section 5 of [1]). Note that a similar proof for continuous systems can be found in [9].

Lemma 1 (and definition): Consider two metrics \( M_i = \Theta_i^T \Theta_i \) defined over \( \mathbb{R}^{n_i} \) \((i = 1, 2) \) and a smooth function \( f : \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2} \). The generalized Jacobian of \( f \) in the metrics \((M_1, M_2)\) is defined by

\[
F = \Theta_2 \frac{\partial f}{\partial x} \Theta_1^{-1}
\]

Assume now that \( f \) is contracting in the metrics \((M_1, M_2)\) with rate \( \beta (0 < \beta < 1) \), i.e.

\[
\forall x \in \mathbb{R}^{n_1}, \lambda_{\max}(F(x)^T F(x)) \leq \beta
\]

Then for all \( u, v \in \mathbb{R}^n \), one has

\[
d_{M_2}(f(u), f(v))^2 \leq \beta d_{M_1}(u, v)^2
\]

where \( d_M \) denotes the distance associated with the metric \( M \) (the distance between two points is defined by the infimum of the lengths in the metric \( M \) of all continuously differentiable curves connecting these points).

Proof Consider a \( C^1 \) curve \( \gamma : [0, 1] \rightarrow \mathbb{R}^{n_1} \) that connects \( u \) and \( v \) (i.e. \( \gamma(0) = u \) and \( \gamma(1) = v \)). The \( M_1 \)-length of such a curve is given by

\[
L_{M_1}(\gamma) = \int_0^1 \sqrt{\left( \frac{\partial \gamma}{\partial u}(u) \right)^T M_1 \left( \frac{\partial \gamma}{\partial u}(u) \right)} \, du
\]
Since \( f \) is a smooth function, \( f(\gamma) \) is also a \( C^1 \) curve, with
\[
L_{M_2}(f(\gamma)) = \int_0^1 \sqrt{\left( \frac{\partial f(\gamma)}{\partial u} (u) \right)^T M_2 \left( \frac{\partial f(\gamma)}{\partial u} (u) \right)} \, du
\]
The chain rule next implies that
\[
\frac{\partial f(\gamma)}{\partial u} (u) = \frac{\partial \gamma}{\partial x} \frac{\partial f(\gamma)}{\partial u} (u)
\]
which leads to
\[
L_{M_2}(f(\gamma)) = \int_0^1 \sqrt{\left( \frac{\partial f(\gamma)}{\partial x} \right)^T \Theta \frac{\partial f(\gamma)}{\partial x} } \, du
\]
\[
= \sqrt{\beta} L_{M_1}(\gamma)
\]
Choose now a sequence of curves \( (\gamma_n)_{n \in \mathbb{N}} \) such that
\[
\lim_{n \to \infty} L_{M_1}(\gamma_n) = d_{M_1}(u, v).
\]
From (1), one has \( \forall n \in \mathbb{N}, L_{M_2}(f(\gamma_n)) \leq \sqrt{\beta} L_{M_1}(\gamma_n) \). By definition of distance, one then has \( \forall n \in \mathbb{N}, d_{M_2}(f(u), f(v)) \leq \sqrt{\beta} L_{M_1}(\gamma_n) \). Finally, by letting \( n \) go to infinity in the last inequality, one obtains the desired result. □

**Theorem 1 (Discrete stochastic contraction):** Consider the stochastic difference equation
\[
a_{k+1} = f(a_k, k) + \sigma(a_k, k)w_{k+1}
\]
where \( f \) is a \( \mathbb{R}^n \times \mathbb{N} \to \mathbb{R}^n \) function, \( \sigma \) is a \( \mathbb{R}^n \times \mathbb{N} \to \mathbb{R}^{nd} \) matrix-valued function, and \( \{w_k, k = 1, 2, \ldots \} \) is a sequence of independent \( d \)-dimensional Gaussian noise vectors, with \( w_k \sim \mathcal{N}(0, Q_k) \).

Assume that the system verifies the following two hypotheses

\begin{align*}
\text{(H1)} & \quad \text{the dynamics } f(a, k) \text{ is contracting in the metrics } (M_k, M_{k+1}), \text{ with contraction rate } \beta \ (0 < \beta < 1), \text{ and the metrics } (M_k)_{k \in \mathbb{N}} \text{ are uniformly positive definite.} \\
\text{(H2)} & \quad \text{the impact of noise is uniformly upper-bounded by a constant } \sqrt{C} \text{ in the metrics } M_k
\end{align*}

\[\forall a, k \quad d_{M_k}(f(a, k), f(a, k) + \sigma(a, k)w_k) \leq \sqrt{C}\]

Let \( a_k \) and \( b_k \) be two trajectories whose initial conditions are given by a probability distribution \( p(x_0) = p(a_0, b_0) \). Then for all \( k \geq 0 \)
\[
\mathbb{E}(d_{M_k}(a_k, b_k)) \leq \frac{2 \sqrt{C}}{1 - \beta} + \sqrt{\beta} \left[ \int_{x \in \mathbb{R}^n} d_{M_0}(a, b) - \frac{2 \sqrt{C}}{1 - \beta} \right]_{[]}^+ \, dp(a, b)
\]
where \( [\cdot]_+ = \max(0, \cdot) \).

This implies in particular that for all \( k \geq 0 \)
\[
\mathbb{E}(d_{M_k}(a_k, b_k)) \leq \frac{2 \sqrt{C}}{1 - \beta} + \sqrt{\beta} \mathbb{E}(d_{M_0}(a_0, b_0))
\]

**Proof** Let \( x = (a, b) \in \mathbb{R}^{2n} \). We have by the triangle inequality (to avoid long formulas, we drop the second argument of \( f \) and \( \sigma \) in the following calculations)
\[
d_{M_{k+1}}(a_{k+1}, b_{k+1}) \leq d_{M_{k+1}}(f(a_k), f(b_k)) + d_{M_{k+1}}(f(a_k), f(a_k) + \sigma(a_k)w_{k+1}) + d_{M_{k+1}}(f(b_k), f(b_k) + \sigma(b_k)w_{k+1})
\]

Let us examine the conditional expectations of the three terms of the right hand side

\[\text{• From (H1) and lemma 1 one has:}\]
\[
\mathbb{E}_x(d_{M_{k+1}}(f(a_k), f(b_k)))) \leq \sqrt{\beta} \mathbb{E}_x(d_{M_k}(a_k, b_k))
\]

\[\text{• Next, from (H2) } \]
\[
\mathbb{E}_x(d_{M_{k+1}}(f(a_k), f(a_k) + \sigma(a_k)w_{k+1})) \leq \sqrt{C}
\]

and similarly for \( d_{M_{k+1}}(f(b_k), f(b_k) + \sigma(b_k)w_{k+1}) \).

If we now set \( u_k = \mathbb{E}_x(d_{M_k}(a_k, b_k)) \) then the above implies
\[
\mathbb{E}_x(u_k) \leq \sqrt{\beta} \mathbb{E}_x(u_k)
\]

Define next \( v_k = \frac{u_k}{\sqrt{\beta}} \leq \mathbb{E}_x(u_k) \). Then replacing \( u_k \) by \( v_k \) and \( \sqrt{C}/(1 - \sqrt{\beta}) \) in (5) yields
\[
v_{k+1} \leq \sqrt{\beta} v_k
\]

This implies that \( \forall k \geq 0, v_k \leq v_0 \sqrt{\beta}^k \leq |v_0|^+ \sqrt{\beta}^k \).

Replacing \( v_k \) by its expression in terms of \( u_k \) then yields
\[
\forall k \geq 0, u_k \leq \frac{2 \sqrt{C}}{1 - \sqrt{\beta}} + \sqrt{\beta} \left( u_0 - \frac{2 \sqrt{C}}{1 - \sqrt{\beta}} \right)_{[+]^+}
\]
which is the desired result.

Next, integrating the last inequality with respect to \( x \) leads to (3). Finally, (4) follows from (3) by remarking that
\[
\mathbb{E}_x(d_{M_0}(a, b) - \frac{\sqrt{C}}{1 - \sqrt{\beta}})_+ \, dp(a, b) \leq \mathbb{E}_x(d_{M_0}(a, b)) \]

**Remark** In the particular context of state-independent metrics, hypothesis (H2) is equivalent to the following simpler condition
\[
\forall a, k \quad \sigma(a, k)^T M_{k+1} \sigma(a, k) \leq C
\]

Also, for state-independent metrics, one has
\[
d_{M_k}(a_k, b_k)^2 = ||a_k - b_k||^2_{M_k} = (a_k - b_k)^T M_k (a_k - b_k)
\]
which leads to the following stronger result instead of (4)
\[
\mathbb{E}(||a_k - b_k||^2_{M_k}) \leq \frac{2C}{1 - \beta} + \beta^k \mathbb{E}(||a_0 - b_0||^2_{M_0})
\]

### III. HYBRID SYSTEMS

We have derived above the discrete stochastic contraction theorem for time- and state-dependent metrics, contrary to the context of continuous systems, where the state-dependent-metrics version of the contraction theorem is still unproved [4].
We now address the case of hybrid systems, but due to the current limitations of continuous stochastic contraction, only state-independent metrics will be considered.

For clarity, we assume in this paper constant dwell-times, although more elaborate conditions regarding dwell-times can be adapted from [3].

Consider the hybrid resetting stochastic dynamical system

\[ \forall k \geq 0 \quad a(k\tau^+ -) = f_d(a(k\tau^-), k) + \sigma_d(a(k\tau^-), k)w_k \]

\[ \forall k \geq 0, \forall t \in ]k\tau, (k + 1)\tau[ \quad da = f(a, t)dt + \sigma(a, t)dW \]

All the contraction properties below will be stated with respect to a uniformly positive definite time-varying metric \(M(t) = \Theta(t)^T\Theta(t)\). Furthermore, it will be assumed that for all \(k \geq 0\), \(M\) is continuously differentiable in \(]k\tau, (k + 1)\tau[\). Finally, \(M(k\tau^-)\) and \(M(k\tau^+)\) will respectively denote the left and right limits of \(M(t)\) at \(t = k\tau\) (and similarly for \(\Theta\)).

A. The discrete and continuous parts are both contracting

**Theorem 2 (Hybrid stochastic contraction):** Assume the following conditions

(i) For all \(k\), the discrete part is stochastically contracting at \(k\tau\) with rate \(\beta < 1\) and bound \(C_d\), i.e.

\[ \forall a \in \mathbb{R}^n \quad \lambda_{\max}(F(k\tau)^T F(k\tau)) \leq \beta \]

where \(F(k\tau) = \Theta(k\tau)^T \Theta(k\tau^-)\), and

\[ \forall a \in \mathbb{R}^n \quad \operatorname{tr}(\sigma_d(a, k)^T M(k\tau^+)\sigma_d(a, k)Q_k) \leq C_d \]

(ii) For all \(k\), the continuous part is stochastically contracting in \([k\tau, (k + 1)\tau[\) with rate \(\lambda > 0\) and bound \(C_c\), i.e.

\[ \forall a \in \mathbb{R}^n \quad \lambda_{\max}\left(\frac{d}{dt} \Theta(t)^T \Theta(t) + \Theta(t)^T \frac{\partial f}{\partial a}(t) \Theta^{-1}(t)\right) \leq -\lambda \]

\[ \operatorname{tr}(\sigma_c(a, t)^T M(t)\sigma_c(a, t)) \leq C_c \]

Let \(a(t)\) and \(b(t)\) be two trajectories whose initial conditions are given by a probability distribution \(p(x(0)) = p(a(0), b(0))\). Then for all \(t \geq 0\)

\[ \mathbb{E}\left(\|a(t) - b(t)\|^2_{M(t)}\right) \leq C_1 + \mathbb{E}\left(\|a(0) - b(0)\|^2_{M(0)}\right)\beta^{(t/\tau)}e^{-2\lambda t} \]

where \(C_1 = \frac{2C_d(1-\beta)(1+\beta-\tau r_1)}{1-\beta (1-\tau r_1)}\) and \(r_1 = \beta e^{-2\lambda t}\).

**Proof** For all \(t \geq 0\), let \(u(t) = \mathbb{E}\left(\|a(t) - b(t)\|^2_{M(t)}\right)\) and let us study the evolution of \(u(t)\) between \(k\tau^+\) and \((k + 1)\tau^+\).

Condition (ii) and theorem 2 of [4] yield

\[ u((k + 1)\tau^-) \leq \frac{C_c}{\lambda} + u(k\tau^+)e^{-2\lambda(\tau^-)} \]

Next, condition (i) and theorem 1 above yield

\[ u((k + 1)\tau^+) \leq \frac{2C_d}{1-\beta} + \beta u((k + 1)\tau^-) \]

Substituting (9) into (10) leads to

\[ u((k + 1)\tau^+) \leq \frac{2C_d}{1-\beta} + \beta \left(\frac{C_c}{\lambda} + u(k\tau^+)e^{-2\lambda(\tau^-)}\right) \]

\[ = \frac{2C_d}{1-\beta} + \frac{2C_c}{\lambda} + \beta e^{-2\lambda r_1} u(k\tau^+) \]

Define \(D_1 = \frac{2C_d}{1-\beta} + \frac{2C_c}{\lambda}\) and \(v_k = u(k\tau^+) - D_1/(1-r_1)\). Then, similarly to the proof of theorem 1, we have \(v_{k+1} \leq r_1v_k\), and then \(v_k \leq r_1^{k-1}v_0\), which implies

\[ u(k\tau^+) \leq \frac{D_1}{1-r_1} + \left[u(0^+) - \frac{D_1}{1-r_1}\right]r_1^k \]

\[ \leq \frac{D_1}{1-r_1} + u(0^+)r_1^k \]

Now, for any \(t \geq 0\), choose \(k = \lfloor t/\tau \rfloor\). Then

\[ u(t) \leq \frac{C_c}{\lambda} + u(k\tau^+)e^{-2\lambda(t-\tau^+)} \]

\[ \leq \frac{C_c}{\lambda} + \frac{D_1 e^{-2\lambda(t-\tau^+)} }{1-r_1} + u(0^+)\beta^ke^{-2\lambda t} \]

which leads to the desired result after some algebraic manipulations. □

B. Only the discrete part is contracting

Let us examine now the more interesting case when the continuous part is not contracting, more precisely when \(\lambda \leq 0\) in (8). For this, we shall need to revisit the proof of theorem 2 in [4].

**Theorem 3 (Case \(\lambda = 0\)):** Assume all the hypotheses of theorem 2 except that \(\lambda = 0\) in (8). Then for all \(t \geq 0\)

\[ \mathbb{E}\left(\|a(t) - b(t)\|^2_{M(t)}\right) \leq C_2 + \mathbb{E}\left(\|a(0) - b(0)\|^2_{M(0)}\right)\beta^{(t/\tau)} \]

where \(C_2 = \frac{2C_d+3(1-\beta)C_c}{1-\beta}\).

**Proof** As in the proof of theorem 2 in [4], let

\[ V(x(t), t) = V((a, b)^T, t) = (a - b)^T M(t)(a - b) \]

Lemma 1 of [4] is unchanged, yielding (see [4] for more details)

\[ \forall t \in [k\tau, (k + 1)\tau[ \quad \tilde{A}V(x(t), t) \leq 2C_c \]

where \(\tilde{A}\) is the infinitesimal operator associated with the process \(x(t)\) (see section 2.1.2 of [4] or p. 15 of [10] for more details).

By Dynkin’s formula [10], one then obtains for all \(x \in \mathbb{R}^{2n}\)

\[ \mathbb{E}_x V(x(t), t) - V(x, k\tau^+) = \mathbb{E}_x \int_{k\tau^+}^t \tilde{A}V(x(s), s)ds \leq \mathbb{E}_x \int_{k\tau^+}^t 2C_c ds = 2C_c(t - k\tau^+) \]

Integrating the above inequality with respect to \(x\) then yields

\[ \forall t \in [k\tau, (k + 1)\tau[ \quad u(t) \leq 2C_c(t - k\tau) + u(k\tau^+) \]

In particular, (9) becomes

\[ u((k + 1)\tau^-) \leq 2C_c\tau + u(k\tau^+) \]


which leads to, after substitution into (10),
\[ u((k + 1)\tau^+) \leq \frac{2C_d}{1 - \beta} + 2\beta C_c \tau + \beta u(k\tau^+) \]

This finally implies
\[ u(k\tau^+) \leq \frac{2C_d}{1 - \beta} + 2\beta C_c \tau + \beta (u(0^+))\beta^k \]

The remainder of the proof can be adapted from that of theorem 2. □

**Theorem 4 (Case \( \lambda < 0 \)):** Assume all the hypotheses of theorem 2 except that \( \lambda < 0 \) in (8). Let \( k = [t/\tau] \). There are two cases:

- If \( \beta < e^{-2|\lambda|\tau} \), then let \( r_2 = \beta e^{2|\lambda|\tau} < 1 \). For all \( t \geq 0 \)
  \[ \mathbb{E} \left( \|a(t) - b(t)\|_M(t) \right)^2 \leq C_3 + \mathbb{E} \left( \|a(0) - b(0)\|_M(0) \right)^2 e^{2|\lambda|\tau^k} \]

  where \( C_3 = \frac{2|\lambda|C_d + (1 - \beta)\beta e^{2|\lambda|\tau^k} C_c}{1 - \beta - \beta^2|\lambda|^2} \).
- If \( \beta \geq e^{-2|\lambda|\tau} \), then there is – in general – no finite bound

  on \( \mathbb{E} \left( \|a(t) - b(t)\|_M(t) \right)^2 \) as \( t \to +\infty \).

**Proof** One has now for all \( t \in ]k\tau, (k + 1)\tau[ \)

\[ \bar{A}V(x(t), t) \leq 2|\lambda|V(x(t), t) + 2C_c \]

with \( |\lambda| > 0 \). By Dynkin’s formula, one has, for all \( x \in \mathbb{R}^{2n} \)

\[ \mathbb{E}_x V(x(t), t) - V(x, k\tau^+) \leq \mathbb{E}_x \int_{k\tau}^{t} (2|\lambda|V(x(s), s) + 2C_c)ds \]

Let now \( g(t) = \mathbb{E}_x V(x(t), t) \). The above equation then yields

\[ g(t) = V(x, k\tau^+) + 2C_c (t - k\tau) + 2|\lambda| \int_{k\tau}^{t} g(s)ds \]

Applying the classical Gronwall’s lemma [11] to \( g(t) \) leads to

\[ g(t) \leq V(x, k\tau^+) + 2C_c (t - k\tau) + \frac{2|\lambda|}{\lambda} \int_{k\tau}^{t} V(x, k\tau^+) + 2C_c s \exp \left( \int_{k\tau}^{t} 2|\lambda|du \right)ds \]

= \( C_3 e^{2|\lambda|(t-k\tau)} - 1 + V(x, k\tau^+) e^{2|\lambda|(t-k\tau)} \)

Integrating the above inequality with respect to \( x \) then yields \( \forall t \in ]k\tau, (k + 1)\tau[ \)

\[ u(t) \leq \frac{C_3}{|\lambda|} \left( e^{2|\lambda|(t-k\tau)} - 1 \right) + u(k\tau^+) e^{2|\lambda|(t-k\tau)} \]

which implies

\[ u((k + 1)\tau^+) \leq D_2 + \beta e^{2|\lambda|\tau} u(k\tau^+) \] \hspace{1cm} (11)

where \( D_2 = \frac{2C_d}{1 - \beta} + \frac{2\beta C_c}{|\lambda|} (e^{2|\lambda|\tau - 1}) \).

There are three cases:

- If \( \beta < e^{-2|\lambda|\tau} \), then \( r_2 = \beta e^{2|\lambda|\tau} < 1 \).

  By the same reasoning as in theorem 1, one obtains

  \[ u(k\tau^+) \leq \frac{D_2}{1 - r_2} + u(0^+) r_2^k \]

  The remainder of the proof can be adapted from that of theorem 2.

- If \( \beta = e^{-2|\lambda|\tau} \), then (11) reads

  \[ u((k + 1)\tau^+) \leq D_2 + u(k\tau^+) \]

  which implies \( \forall k \geq 0 \), \( u(k\tau^+) \leq kD_2 + u(0^+) \). From this, it is clear that there is – in general – no finite bound

  for \( u(k\tau^+) \).

- If \( \beta > e^{-2|\lambda|\tau} \), then \( r_2 = \beta e^{2|\lambda|\tau} > 1 \).

  By the same reasoning as in theorem 1, one obtains

  \[ u(k\tau^+) \leq \left( u(0^+) + \frac{D_2}{2r_2 - 1} \right) r_2^k - \frac{D_2}{2r_2 - 1} \]

Since \( r_2 > 1 \) in this case, it is clear that there is – in general – no finite bound for \( u(k\tau^+) \). □

**Remarks** Theorems 3 and 4 show that it is possible to stabilize an unstable system by discrete resettings. If the continuous system is indifferent (\( \lambda = 0 \)), then any sequence of uniformly contracting resettings is stabilizing. However, it should be noted that the asymptotic bound \( C_2 \to \infty \) when \( \beta \to 1 \). In contrast, if the continuous system is strictly unstable (\( \lambda < 0 \)), then specific contraction rates (depending on the dwell-time and the “expansion” rate of the continuous system) of the resettings are required. Finally, note that in both cases, the asymptotic bounds \( C_2 \) and \( C_3 \) are increasing functions of the dwell-time \( \tau \).

**IV. Comments**

**A. Modelling issue: distinct driving noise**

In the same spirit as [4], and contrary to previous works on the stability of stochastic systems [12], the \( a \) and \( b \) systems considered in sections II and III are driven by distinct noise processes. This approach enables us to study the stability of the system with respect to variations in initial conditions and to random perturbations: indeed, two trajectories of any real-life system are typically affected by distinct realizations of the noise. In addition, this approach leads very naturally to nice results regarding the comparison of noisy and noise-free trajectories (see section IV-B), which are particularly useful in applications (see e.g. section V).

However, because of the very fact that the two trajectories are driven by distinct noise processes, we cannot expect the influence of noise to vanish when the two trajectories get very close to each other. As a consequence, the asymptotic bounds \( 2C/(1 - \beta) \) (for discrete systems) and \( C_1, C_2, C_3 \) (for hybrid systems) are strictly positive. These bounds are nevertheless optimal, in the sense that they can be attained (adapt the Ornstein-Uhlenbeck example in section 2.3.1 of [4]).

**B. Noisy and noise-free trajectories**

Instead of considering two noisy trajectories \( a \) and \( b \) as in theorem 1, we assume now that \( a \) is noisy, while \( b \) is noise-free. More precisely, for all \( k \in \mathbb{N} \)

\[ a_{k+1} = f(a_k, k) + \sigma(a_k, k)w_{k+1} \]

\[ b_{k+1} = f(b_k, k) \]
To show the exponential convergence of $a$ and $b$ to each other, one can follow the same reasoning as in the proof of theorem 1, with $C$ is replaced by $C/2$. This leads to the following result

Corollary 1: Assume all the hypothesis of theorem 1 and consider a noise-free trajectory $b_0$, and a noisy trajectory $a_k$ whose initial conditions are given by a probability distribution $p(a_0)$. Then, for all $k \in \mathbb{N}$

$$
E \left( \|a_k - b_k\|_{\tilde{M}_k} \right) \leq \frac{C}{1-\beta} + \beta^k \int \left[ \|a - b_0\|_{\tilde{M}_0} - \frac{C}{1-\beta} \right]^+ dp(a)
$$

(12)

Remarks

- The above derivation of corollary 1 is only permitted by our choice of considering distinct driving noise processes for systems $a$ and $b$ (see section IV-A).
- Based on theorems 2, 3 and 4, similar corollaries can be obtained for hybrid systems.
- These corollaries provide a robustness result for contracting discrete and hybrid systems, in the sense that any contracting system is automatically protected against noise, as quantified by (12). This robustness could be related to the exponential nature of contraction stability.

V. APPLICATION: OSCILLATOR SYNCHRONIZATION BY DISCRETE COUPLINGS

Using the above developed tools, we study in this section the synchronization of nonlinear oscillators in presence of random perturbations. The novelty here is that the interactions between the oscillators occur at discrete time instants, contrary to many previous works devoted to synchronization in the state-space$^1$ [14], [7].

Specifically, consider the Central Pattern Generator (CPG) delivering $2\pi/3$-phase-locked signals of section 5.3 in [7]. This CPG consists of a network of three Andronov-Hopf oscillators $x_i = (x_{i1}, x_{i2}, x_{i3})^T$, $i = 1, 2, 3$. We construct below a discrete-couplings version of this CPG.

At instants $t = k\tau$, $k \in \mathbb{N}$, the three oscillators are coupled in the following way (assuming noisy measurements)

$$
x_i(k\tau^+) = x_i(k\tau^-) + \gamma \left( R \left( x_{i+1}(k\tau^-) + \frac{\sigma_d}{\sqrt{2}} w_k \right) - x_i(k\tau^-) \right)
$$

with $x_4 = x_1$ and

$$
R = \begin{pmatrix}
-\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}}
\end{pmatrix}
$$

Between two interaction instants, the oscillators follow the uncoupled, noisy, dynamics

$$
dx_i = f(x_i)dt + \frac{\sigma_c}{\sqrt{2}} dW
$$

where

$$
f(x_i) = f \left( \begin{array}{l}
 x_{i1} \\
 x_{i2} \\
 x_{i3}
\end{array} \right) = \left( \begin{array}{l}
 x_{i1} - y_i - x_{i1}^3 - x_{i1} y_i^2 \\
 x_{i2} + y_i - x_{i2}^3 - y_i x_{i2}^2 \\
 x_{i3} - x_{i3}^3 - x_{i2} y_{i3}^2
\end{array} \right)
$$

We apply now the projection technique developed in [7], [4]. We recommend the reader to refer to these papers for more details about the following calculations.

Consider first the (linear) subspace $M$ of the global state space (the global state is defined by $\tilde{x} = (x_1, x_2, x_3)^T$) where the oscillators are $2\pi/3$-phase-locked

$$
M = \left\{ (R^2(x), R(x), x)^T : x \in \mathbb{R}^2 \right\}
$$

Let $V$ and $U$ be two orthonormal projections on $M^\perp$ and $M$ respectively and consider $\tilde{y} = V \tilde{x}$. Since the mapping is linear, using Itô differentiation rule yields the following dynamics for $\tilde{y}$

$$
\forall k \in \mathbb{N} \quad \tilde{y}(k\tau^+) = g_d(\tilde{y}(k\tau^-)) + \gamma \frac{\sigma_d}{\sqrt{2}} w_k
$$

(13)

$$
\forall t \in [k\tau, (k+1)\tau] \quad d\tilde{y} = g_c(\tilde{y})dt + \frac{\sigma_c}{\sqrt{2}} dW
$$

(14)

with

$$
g_d(\tilde{y}) = VL\tilde{x} = VL(V^T\tilde{y} + U^T U\tilde{x}) = VLV^T \tilde{y}
$$

$$
g_c(\tilde{y}) = \tilde{f}(V^T \tilde{y} + U^T U\tilde{x})
$$

where

$$
L = \begin{pmatrix}
(1-\gamma)I_2 & \gamma R & 0 \\
0 & (1-\gamma)I_2 & \gamma R \\
\gamma R & 0 & (1-\gamma)I_2
\end{pmatrix}
$$

$$
\tilde{f}(\tilde{x}) = (f(x_1), f(x_2), f(x_3))^T
$$

Remark that $g_d(0) = 0$ and $g_c(0) = 0$ (the last equality holds because of the symmetry of $f$: $Vx$, $f(Rx) = R(f(x))$). Thus, 0 is a particular solution to the noise-free version of the hybrid stochastic system (13,14).

Let us now examine the contraction properties of equations (13) and (14).

We have first

$$
\frac{\partial g_d}{\partial \tilde{y}} T \frac{\partial g_d}{\partial \tilde{y}} = VL^T V^T V L = (3\gamma^2 - 3\gamma + 1)I_4
$$

so that $\lambda_{\text{max}}(\frac{\partial g_d}{\partial \tilde{y}} T \frac{\partial g_d}{\partial \tilde{y}}) = 3\gamma^2 - 3\gamma + 1 < 1$ (for $0 < \gamma < 1$).

Second,

$$
\frac{\partial g_c}{\partial \tilde{y}} = V \frac{\partial f}{\partial x} V^T = V \begin{pmatrix}
\frac{\partial f}{\partial x}(x_1) & 0 & 0 \\
0 & \frac{\partial f}{\partial x}(x_2) & 0 \\
0 & 0 & \frac{\partial f}{\partial x}(x_3)
\end{pmatrix} V^T
$$

Now observe that $\lambda_{\text{max}}(\frac{\partial f}{\partial x}) = 1 - x^2 - y^2 \leq 1$. Since $V$ is an orthonormal projection, one then has $\lambda_{\text{max}}(\frac{\partial g_c}{\partial \tilde{y}}) \leq 1$.

Therefore, if

$$
3\gamma^2 - 3\gamma + 1 < e^{-2\tau}
$$

(15)

then theorem 4 together with the corollaries of section IV-B
imply that, after exponential transients,
\[ E(\|\tilde{y}\|^2) \leq \frac{2\gamma^2 \sigma^2_{\epsilon} + (1 - \beta)(1 + \beta) - \beta e^{2\tau})e^{2\tau} \sigma^2_{\epsilon}}{2(1 - \beta)(1 - \beta e^{2\tau})} \]
where \( \beta = 3\gamma^2 - 3\gamma + 1 \).

To conclude, observe that
\[ \|\tilde{y}\|^2 = \|Vx\|^2 = \frac{1}{3} \sum_{i=1}^{3} \|Rx_{i+1} - x_i\|^2 \]
Define the phase-locking quality \( \delta \) by
\[ \delta = \sum_{i=1}^{3} \|Rx_{i+1} - x_i\|^2 \]
then one finally obtains
\[ E(\delta) \leq \frac{6\gamma^2 \sigma^2_{\epsilon} + 3(1 - \beta)(1 + \beta) - \beta e^{2\tau})e^{2\tau} \sigma^2_{\epsilon}}{2(1 - \beta)(1 - \beta e^{2\tau})} \tag{16} \]
after exponential transients.

A numerical simulation is provided in Fig. 1.

![Numerical simulation using the Euler-Maruyama algorithm](image)

Fig. 1. Numerical simulation using the Euler-Maruyama algorithm [15]. The following set of parameters was used: \( \sigma_{\epsilon} = 0.1, \sigma_d = 0.05, \tau = 0.1 \). Two coupling strengths were tested: \( \gamma_{\text{weak}} = 0.01 \) for plots (a), (b), (c), and \( \gamma_{\text{strong}} = 0.2 \) for plots (d), (e), (f). Note that \( \gamma_{\text{weak}} \) does not satisfy condition (15), while \( \gamma_{\text{strong}} \) does, and yields the theoretical bound \( \gamma_{\text{bound}} \approx 0.446 \) (as provided by (16)) on the phase-locking quality \( \delta \). Plots (a) and (d) show the 2D trace of sample trajectories of the three oscillators for \( t \in [0, 1] \). Plots (b) and (e) show sample trajectories of the first coordinates of \( x_1, R(x_2) \) and \( R^2(x_3) \) as functions of time. Plot (c) and (f) show three sample trajectories of \( \delta \).

VI. PERSPECTIVES

We are now focusing on the following directions of research:
- proving the state-dependent-metrics version of the continuous and hybrid stochastic contraction theorems,
- developing more elaborate conditions on dwell-times, and also hybrid *switched* versions of the theorems,
- applying the synchronization-by-discrete-couplings analysis to other types of coupled dynamical systems,
- studying the robustness of hybrid controllers and observers against random perturbations (for instance, the discrete observer for inertial navigation developed in [16]).

ACKNOWLEDGMENT

The author is grateful to Prof J.-J. Slotine and N. Tabareau for stimulating discussions, and to Dr H. Hicheur for the careful reading of the manuscript. This work has been supported by EC - contract number FP6-IST-027140, action line: Cognitive Systems. This publication reflects only the author’s views. The European Community is not liable for any use that may be made of the information contained therein.

REFERENCES

[1] W. Lohmiller and J.-J. Slotine, “On contraction analysis for nonlinear systems,” *Automatica*, vol. 34, pp. 671–682, 1998.
[2] W. Wang and J.-J. E. Slotine, “On partial contraction analysis for coupled nonlinear oscillators.” *Biol Cybern*, vol. 92, no. 1, pp. 38–53, Jan. 2005.
[3] K. El Rifai and J.-J. Slotine, “Compositional contraction analysis of resetting hybrid systems.” *IEEE Transactions on Automatic Control*, vol. 51, no. 9, pp. 1536–1541, 2006.
[4] Q.-C. Pham, N. Tabareau, and J.-J. Slotine, “Nonlinear process control using contraction theory,” *A.I.Ch.E. Journal*, 2000.
[5] Q.-C. Pham and J.-J. Slotine, “Stable concurrent synchronization in dynamic system networks.” *Neural Netw*, vol. 20, no. 1, pp. 62–77, Jan. 2007.
[6] B. Girard, N. Tabareau, Q.-C. Pham, A. Berthoz, and J.-J. Slotine, “Where neuroscience and dynamic system theory meet autonomous robotics: a contracting basal ganglia model for action selection.” *Neural Networks*, vol. 21, pp. 628–641, 2008.
[7] N. Aghannan and P. Rouchon, “An intrinsic observer for a class of lagrangian systems.” *IEEE Transactions on Automatic Control*, vol. 48, 2003.
[8] H. Kushner, *Stochastic Stability and Control*. Academic Press, 1967.
[9] J. Robbin. (2006) Gronwall’s inequality. [Online]. Available: www.math.wisc.edu/~robbin/angelic/grimwall.pdf
[10] P. Florchinger, “Lyapunov-like techniques for stochastic stability.” *SIAM Journal of Control and Optimization*, vol. 33, pp. 1151–1169, 1995.
[11] E. Izhikevich, “Weakly pulse-coupled oscillators, FM interactions, synchronization, and oscillatory associative memory.” *Neural Networks, IEEE Transactions on*, vol. 10, no. 3, pp. 508–526, 1999.
[12] A. Pogromsky, G. Santoboni, and H. Nijmeijer, “Partial synchronization: from symmetry towards stability.” *Physica D: Nonlinear Phenomena*, vol. 172, no. 1-4, pp. 65–87, 2002.
[13] D. Higham, “An algorithmic introduction to numerical simulation of stochastic differential equations.” *SIAM Review*, vol. 43, pp. 525–546, 2001.
[14] Y. Zhao and J.-J. Slotine, “Discrete nonlinear observers for inertial navigation.” *Systems and Control Letters*, vol. 54, 2005.