A STABILISATION ALGORITHM FOR LINEAR CONTROLLED SDE'S

F. CAMPILLO, A. TRAORÉ
INRIA, BP 93
06902 Sophia–Antipolis Cédex, France

1 Introduction

The stability properties of linear deterministic systems $\dot{X}_t = A\,X_t$, $X_0 = x$ is well known. We can now suppose that the matrix A is subject to a Gaussian white noise perturbation η, i.e. the previous systems reads $\dot{X}_t = [A_0 + \sum_{l=1}^{L} A_l\,\eta_t^l]\,X_t$, $X_0 = x \neq 0$, or as a Stratonovich stochastic differential equation (SDE)[1] :

$$dX_t = A_0\,X_t\,dt + \sum_{l=1}^{L} A_l\,X_t \circ dW_t^l\ , \quad X_0 = x \tag{1}$$

where W is a standard Brownian motion. Point 0 is an equilibrium state of Equation (1). A natural idea to "stabilise" system (1) is to introduce a linear feedback control, hence we get :

$$dX_t = (A_0 + B_0\,\rho)\,X_t\,dt + \sum_{l=1}^{L}(A_l + B_l\,\rho)\,X_t \circ dW_t^l\ , \quad X_0 = x\ . \tag{2}$$

We note : $A_l^\rho \stackrel{\triangle}{=} A_l + \rho\,B_l$ $(l = 0, \dots, L)$. The Lyapunov exponent of (2) starting at x_0 with parameter ρ is defined by

$$\lambda_\rho(x_0) \stackrel{\triangle}{=} \lim_{t\to\infty} t^{-1} \log \|X_t\|\ . \tag{3}$$

[1] We can also consider non linear SED's : $dX_t = f(X_t)\,dt + g(X_t) \circ dW_t$, $X_0 = x$ (where $f(0)$ and $g(0)$ are zero) for which 0 is also an equilibrium state. The stability of non linear sytem has been well studied. The first results concerning the stability of non linear systems were given by Lyapunov. R.S. Bucy [7] stated that the "stochastic Lyapunov function" should be a supermartingale (see also Wonham [20]). Different books present this problem, but the main reference is Khasminskii [13] (which summarises all his past works). There are alternative approaches like Haussmann [12], Curtain [9].

One of the first problem is to define the notion of stability in a stochastic context. This notion is quite natural in the deterministic case, in the stochastic case many possibilities come to mind : pathwise stability, stability on the moments, weak stability (i.e. the probability for the process to be in a centred ball of a given diameter goes to zero as the times converges to infinity) etc.

Here, we focus on linear SED's (1). A huge amount of results are available in this context : proceedings Arnold [3], Arnold-Crauel-Eckman [1] and the review paper Arnold [2] summarise them.

Oseledec's multiplicative ergodic theorem states that the limit (3) exists with probability one and that there are d fixed numbers $\lambda_1 \geq \cdots \geq \lambda_d$ – called the Lyapunov exponents of (2) – such that the random variable $\lambda_\rho(x_0)$ takes on only these values (see [1] for a review). Moreover, (2) is exponentially stable with probability one if and only if $\lambda_1 < 0$.

We define the projection of X_t onto the sphere by

$$U_t \triangleq \|X_t\|^{-1}\, X_t \in S^{d-1} \triangleq \{x \in \mathbb{R}^d ; \|x\| = 1\} .$$

U_t is the solution of the following SDE on S^{d-1}

$$dU_t = h(A_0^\rho, U_t)\, dt + \sum_{l=1}^{L} h(A_l^\rho, U_t) \circ dW_l(t) , \quad U_0 = u_0 \triangleq \|x_0\|^{-1}\, x_0 \tag{4}$$

where $h(C,u) \triangleq C\,u - [C\,u]^*\,u\,u$, and $\|x\|^2 = x^*\,x$. Moreover

$$\|X_t\| = \|x_0\| \times e^{\int_0^t [q(\rho,U_s)\,ds + \sum_{l=1}^{L} p_l(\rho,U_s)\,dW_s^l]} \tag{5}$$

with

$$\begin{aligned} p_0'(\rho,u) &\triangleq \tfrac{1}{2}\sum_{l=1}^{L}[((A_l^\rho)^2\,u,u) + \|A_l^\rho\,u\|^2 - 2(A_l^\rho\,u,u)^2] , \\ p_l(\rho,u) &\triangleq (A_l^\rho\,u,u) , \quad l = 0,\ldots,L , \qquad q(\rho,u) \triangleq p_0(\rho,u) + p_0'(\rho,u) . \end{aligned}$$

For each matrix M, $h(M,-u) = -h(M,u)$, so that $h(M,\cdot)$ can be viewed as a vector field on the projective space P^{d-1} (obtained from S^{d-1} by identifying u and $-u$). Therefore (4) can be considered as a stochastic differential equation on P^{d-1} and (5) is still valid with this definition.

Hypothesis 1.1 *For all $u \in P^{d-1}$ and $\rho \in \mathcal{R}$*

$$\dim \mathsf{Lie\ Algebra}\{h(A_l^\rho,\cdot)\,;\, l = 1,\ldots,L\}(u) = d-1 .$$

Theorem 1.2 (Proof in [3]) *Under Hypothesis 1.1, for all $\rho \in \mathcal{R}$:*

The diffusion process U_t admits a unique invariant probability measure $\mu_U(\rho,du)$ which has a C^∞, strictly positive density $p_U(\rho,u)$ with respect to the Lebesgue measure on P^{d-1} solution of the following Fokker–Planck equation $[\mathcal{L}_\rho^*\, p_U(\rho,\cdot)](u) = 0$, $\forall u \in P^{d-1}$, where $\mathcal{L}_\rho$ is the infinitesimal generator associated with Equation (4).

The number $\lambda_\rho \triangleq \int_{P^{d-1}} q(\rho,u)\,\mu_U(\rho,du)$ is equal to the top Lyapunov exponent λ_1.

For all $x_0 \in \mathbb{R}^d$, $x_0 \neq 0$, $\lambda_\rho(x_0) = \lambda_\rho$ a.s. When $\lambda_\rho < 0$, the system (2) is exponentially stable with probability one.

Proposition 1.3 (Proof in [8]) *Under Hypothesis 1.1, the application $\mathcal{R} \ni \rho \mapsto \lambda_\rho \in \mathbb{R}$ is continuous.*

We deduce from this proposition that $\mathcal{D} \triangleq \{\lambda_\rho ; \rho \in \mathcal{R}\}$ is a connected interval of $\mathbb{R}$. By definition, we will say that the system (2) is said to be *stabilisable* (resp. *non stabilisable*) if $\mathcal{D} \cap [-\infty, 0[\neq \emptyset$ (resp. $\mathcal{D} \cap [-\infty, 0[= \emptyset$).

2 Gradient of the Lyapunov exponent

In this section we restrict ourself to the case of one standard real Wiener input process (i.e. $L = 1$). We use the notation of the previous section. Throughout this section, ρ is fixed. Equation (2) reads

$$dX_t = A_0^\rho\, X_t\, dt + A_1^\rho\, X_t\, \circ dW_t\ , \quad X_0 = x_0 \neq 0\ , \tag{6}$$

This system is exponentially stable with probability one if and only if $\lambda_\rho < 0$. In order to find ρ which minimise λ_ρ we can use a gradient algorithm.

The next step is to derive the derivative of λ_ρ with respect to ρ :

$$G(\rho) \stackrel{\triangle}{=} \partial\lambda_\rho / \partial\rho\ .$$

We know that under Hypothesis 1.1 : $\lambda_\rho = \text{lim–a.s.}_{T\to\infty}\, T^{-1} \int_0^T q(\rho, U_t)\, dt$. Because all coefficients of Equation (4) are smooth enough, we can define :

$$V_t \stackrel{\triangle}{=} \partial U_t / \partial\rho \quad \text{and} \quad Z_t \stackrel{\triangle}{=} [U_t, V_t]^*\ .$$

In this section we want show that, under certain conditions, the process Z_t admits a unique invariant measure $\mu_Z(\rho, du, dv)$. Which will imply that

$$G(\rho) = \underset{T\to\infty}{\text{lim–a.s.}}\, T^{-1} \int_0^T Q(\rho, U_t, V_t)\, dt\ ,\ \text{with}\ Q(\rho, u, v) \stackrel{\triangle}{=} \frac{\partial q(\rho, u)}{\partial\rho} + \frac{\partial q(\rho, u)}{\partial u}.v\ .$$

In the two–dimensional case (i.e. $d = 2$), let $\theta_t = (\cos(U_t), \sin(U_t))^*$ and

$$\xi_t \stackrel{\triangle}{=} \partial\theta_t / \partial\rho\ . \tag{7}$$

Process θ_t is solution of the following SDE

$$d\theta_t = h(A_0^\rho, \theta_t)\, dt + h(A_1^\rho, \theta_t)\, \circ dW_t\ . \tag{8}$$

Let $F_i^\rho(\theta) \stackrel{\triangle}{=} \partial h(A_i^\rho, \theta)/\partial\theta$ and $f_i^\rho(\theta) \stackrel{\triangle}{=} \partial h(A_i^\rho, \theta)/\partial\rho$, for $i = 0, 1$, so we get

$$d\xi_t = [F_0^\rho(\theta_t)\, \xi_t + f_0^\rho(\theta_t)]\ dt + [F_1^\rho(\theta_t)\, \xi_t + f_1^\rho(\theta_t)]\ \circ dW_t\ . \tag{9}$$

Let $\varphi_{t,s}$ be the fundamental solution of Equation (9), and $\varphi_t \stackrel{\triangle}{=} \varphi_{t,0}$.

Hypothesis 2.1 *There exist $\alpha < 0$ such that $t^{-1} \log \|\varphi_t\| \to \alpha$ a.s. when $t \to \infty$.*

Proposition 2.2 *Under hypotheses 1.1 and 2.1, Equation (9) admits a unique stationary solution.*

The proof of this result is technical and will be presented elsewhere.

This result imply that under Hypotheses 1.1 and 2.1, the process $Z_t = (U_t, V_t)$ admits a unique invariant measure $\mu_Z(\rho, du, dv)$. So we can have the following representation for $G(\rho)$

$$G(\rho) = \underset{T\to\infty}{\text{lim–a.s.}}\, T^{-1} \int_0^T Q(\rho, U_t, V_t)\, dt = \int Q(\rho, u, v)\, \mu_Z(\rho, du, dv)\ .$$

3 Stabilisation algorithm

Suppose that System (2) is stabilisable in the sense given in section 1. We can try to find a parameter ρ such that $\lambda_\rho < 0$ and the more this term is negative, the more System (2) will be stable. So a natural idea is to use a gradient algorithm which minimises λ_ρ : $\rho_{n+1} \leftarrow \rho_n - \gamma_n\, G(\rho_n)$ where $\gamma_n > 0$ is a decreasing gain parameter. Of course, this algorithm is not useful in practice because $G(\rho_n)$ cannot be computed explicitly. We present now a way to approximate such an expression.

Stabilisation techniques derived from Lyapunov's approach are also avalaible[2].

Let Δ be a time step. Given ρ_n, we simulate the processes U_t and V_t with fixed parameter ρ_n for $t \in [n\,\Delta, (n+1)\,\Delta)$, then we use the following update rule

$$\rho_{n+1} = \rho_n - \gamma_n\, \tilde{G}(\rho_n)\ , \quad n \geq 0, \dots\ , \quad \rho_0 \text{ given,} \tag{10}$$

where $\tilde{G}(\rho_n) \triangleq \Delta^{-1} \int_{n\Delta}^{(n+1)\Delta} Q(\rho_n, \tilde{U}_t^n, \tilde{V}_t^n)\, dt$.

Where $(\tilde{U}^n, \tilde{V}^n)$ is given by "freezing" parameter ρ at the value ρ_n over the interval $t \in [n\,\Delta, (n+1)\,\Delta]$, i.e. it is the solution of (we explicitly denote the dependency with respect to parameter ρ)

$$\begin{aligned} dU_t^\rho &= h(A_0^\rho, U_t^\rho)\, dt + h(A_1^\rho, U_t^\rho) \circ dW_t\ , && (11)\\ dV_t^\rho &= [\ell_0^\rho(U_t^\rho) + L_0^\rho(U_t^\rho, V_t^\rho)]\, dt + [\ell_1^\rho(U_t^\rho) + L_1^\rho(U_t^\rho, V_t^\rho)] \circ dW_t\ , && (12)\end{aligned}$$

for $n\,\Delta < t \leq (n+1)\,\Delta$, $\rho = \rho_n$ and initial conditions $U_{n\Delta}^\rho = \tilde{U}_{n\Delta}^n$, $V_{n\Delta}^\rho = \tilde{V}_{n\Delta}^n$, where,for $i = 0, 1$, $L_i^\rho(U_t^\rho, V_t^\rho) \triangleq \partial h(A_i^\rho, U_t^\rho)/\partial u\, V_t^\rho$ and $\ell_i^\rho(U_t^\rho) \triangleq \partial h(A_i^\rho, U_t^\rho)/\partial \rho$.

The convergence properties of such algorithms were studied by many authors[3]. Given a certain time-space scaling defined with the gain sequence γ_n, the trajectories of $\{\rho_n\}_{n\in\mathbb{N}}$ converge to the trajectories of the following ODE :

$$\dot{\rho}(t) = -h(\rho(t))\ , \quad t \in \mathbb{R}\ . \tag{13}$$

For simulation, we discretise Equations (11,12) in time using a Milshtein scheme :

1st step We simulate (X_t^ρ, Y_t^ρ) for $t \in [n\Delta, (n+1)\Delta[$ with $\rho = \rho_n$ by :

$$\begin{aligned} dX_t^\rho &= [A_0^\rho + \tfrac{1}{2}(A_1^\rho)^2]\, X_t^\rho\, dt + A_1^\rho\, X_t^\rho\, dW_t\ ,\\ dY_t^\rho &= S_t^\rho(X_t^\rho, Y_t^\rho)\, dt + [A_1^\rho\, Y_t^\rho + B_1\, X_t^\rho]\, dW_t\ ,\\ S_t^\rho(X_t^\rho, Y_t^\rho) &\triangleq [A_0^\rho + \tfrac{1}{2}(A_1^\rho)^2]\ Y_t^\rho + \tfrac{1}{2}(A_1\, B_1 + B_1\, A_1)\, X_t^\rho + (B_0 + \rho\, B_1^2)\, X_t^\rho\ .\end{aligned}$$

Let $\bar{X}_n$ and $\bar{Y}_n$ be the approximations of $\tilde{X}_{n\Delta}^{\rho_n}$ and $\tilde{Y}_{n\Delta}^{\rho_n}$ respectively. We get :

$$\begin{aligned} \bar{X}_{n+1} &\leftarrow \bar{X}_n + A_1^{\rho_n}\, \bar{X}_n\, w_n + A_0^{\rho_n}\, \bar{X}_n\, \Delta + \tfrac{1}{2}\, (A_1^{\rho_n})^2\, \bar{X}_n\, w_n^2\\ \bar{Y}_{n+1} &\leftarrow \bar{Y}_n + [A_1^{\rho_n}\, \bar{Y}_n^{\rho_n} + B_1\, \bar{X}_n^{\rho_n}]\, w_n + [A_0^{\rho_n}\, \bar{Y}_n + \tfrac{1}{2}\, B_1\, A_1^{\rho_n}\, \bar{X}_n + B_0\, \bar{X}_n]\, \Delta\\ &\qquad + \tfrac{1}{2}\ [(A_1^{\rho_n})^2\, \bar{Y}_n + A_1^{\rho_n}\, B_1\, \bar{X}_n]\ w_n^2\end{aligned}$$

where $w_n \triangleq W_{t_{n+1}} - W_{t_n} \sim \mathcal{N}(0, \Delta)$ is a white Gaussian noise.

[2]See Florchinger [10]. Other approaches by Mao [16], Kozin-Zhang [14], Willems-Willems [19], Gao-Ahmed [11], Baxendale [4, 5] treat the moment stability properties.

[3]Like Kushner-Clark [15], Benveniste-Métivier-Priouret [6] etc.

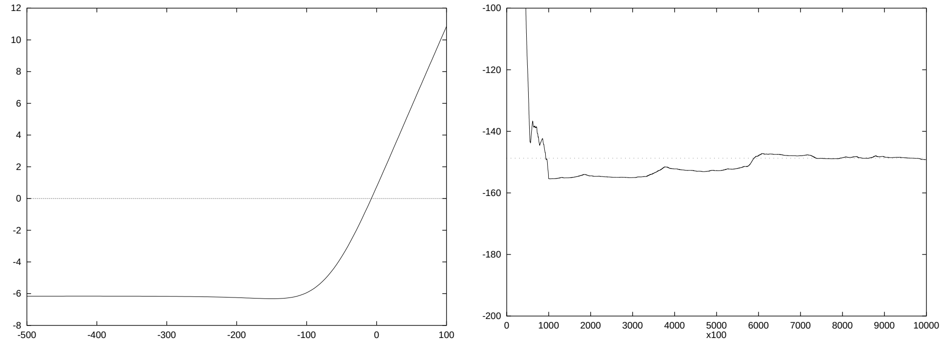

Figure 1: Example 1 : Graph $\rho \mapsto \lambda(\rho)$ (left). Graph $n \mapsto \rho_n$ (right).

2nd step We compute $(\bar{U}_{n+1}, \bar{V}_{n+1})$, the approximation of $(\tilde{U}^{\rho_n}_{(n+1)\Delta}, \tilde{V}^{\rho_n}_{(n+1)\Delta})$, from $(X_(n+1)^{\rho_n}, Y_(n+1)^{\rho_n})$:

$$\bar{U}_{n+1} \leftarrow \left\|\bar{X}_{n+1}\right\|^{-1} \bar{X}_{n+1} \ , \quad \bar{V}_{n+1} \leftarrow \left\|\bar{X}_{n+1}\right\|^{-1} [\bar{Y}_{n+1} - (\bar{U}_{n+1}, \bar{Y}_{n+1})\, \bar{U}_{n+1}] \ .$$

We end up with the following algorithm :

$$\rho_{n+1} = \rho_n - \gamma_n\, Q(\rho_n, \bar{U}_{n+1}, \bar{V}_{n+1}) \ , \quad n \geq 0 \ ,$$

ρ_0 given. For practical implementation we use the standard following gain coefficient : $\gamma_n = c + b/\max(1, n - n_0)$ where $c \geq 0$, $b > 0$, et $n_0 \geq 1$ are given.

For this algorithm we also use a averaging technique proposed by Polyak-Juditsky [18].

4 Numerical simulation

We present two examples where the initial value of ρ the system is unstable (i.e. $\lambda_{\rho_0} > 0$), but using the proposed algorithm we can tune parameter ρ up in such way that process X stabilises.

Example 1

$$\begin{aligned} dX^1_t &= [2 + 0.1\,\rho]\, X^1_t\, dt + [-2\, X^2_t + 0.01\,\rho\, X^1_t] \circ dW_t \ , \\ dX^2_t &= [X^1_t - 5\, X^2_t]\, dt + [X^1_t + 0.01\,\rho\, X^2_t] \circ dW_t \ . \end{aligned}$$

The stochastic gradient algorithm parameters are : $n_0 = 10^4$, $c = 10^{-2}$, $b = 1$, $\rho_0 = 200$.

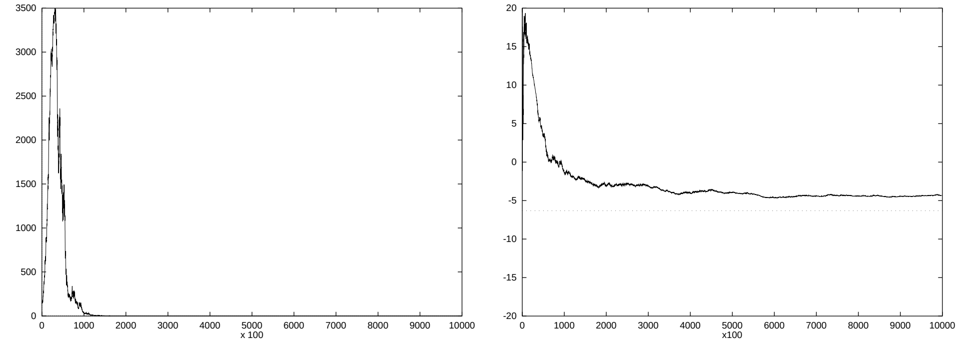

Figure 2: Example 1 : Graph $n \mapsto |\bar{X}_n|$ (left). Graph $n \mapsto \lambda_n \simeq \lambda(\rho_n)$ (right).

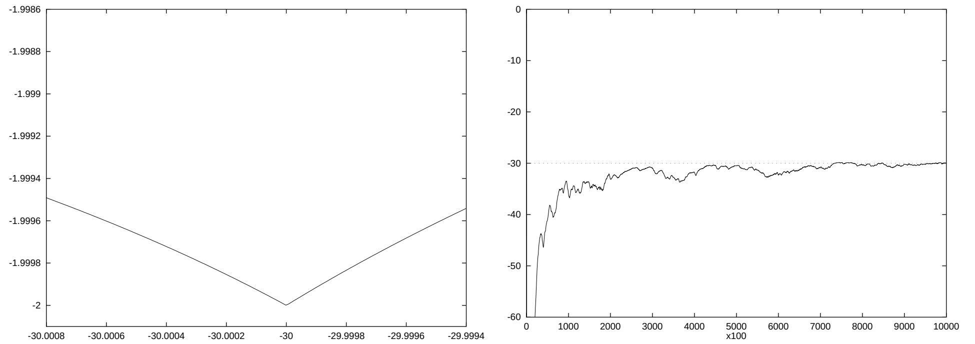

Figure 3: Example 2 : Graph $\rho \mapsto \lambda(\rho)$ (left). Graph $n \mapsto \rho_n$ (right).

Example 2

$$
\begin{aligned}
dX_t^1 &= [1 + 0.1\,\rho]\,X_t^1\,dt + [0.1\,X_t^2 + 0.01\,\rho\,X_t^1] \circ dW_t \ , \\
dX_t^2 &= -2\,X_t^2\,dt + [0.1\,X_t^1 + 0.01\,\rho\,X_t^2] \circ dW_t \ .
\end{aligned}
$$

This system does not satisfy Conditions 1.1 for $\rho = \rho^* = -30$ and $\theta = \pi/4$. The stochastic gradient algorithm parameters are : $n_0 = 2.10^4$, $c = 1$, $b = 0.7$, $\rho_0 = 100$.

References

[1] L. ARNOLD. Stabilization by noise revisited. *Zeitschrift für angewandte Mathematik und Mechanik*, 70(7):235–246, 1990.

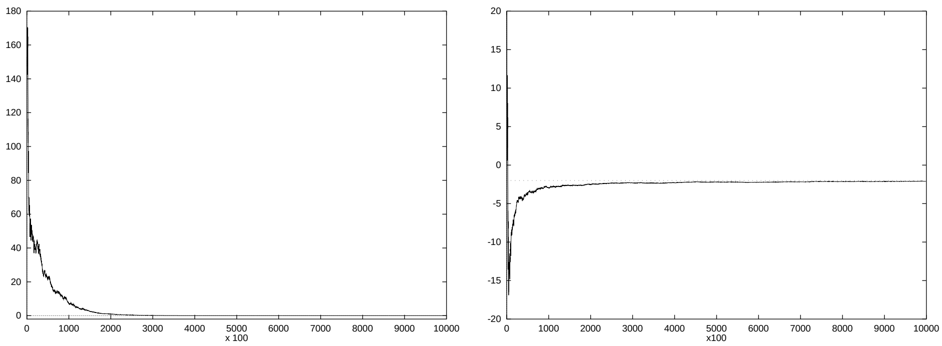

Figure 4: Example 2 : Graph $n \mapsto |\bar{X}_n|$ (left). Graph $n \mapsto \lambda_n \simeq \lambda(\rho_n)$ (right).

[2] L. ARNOLD, H. CRAUEL, and J.P. ECKMAN, editors. *Lyapunov Exponent, Oberwolfach 1990*, volume 1486 of *Lecture Notes in Mathematics*, Berlin, 1991. Springer Verlag.

[3] L. ARNOLD and V. WIHSTUTZ, editors. *Lyapunov Exponents, Bremen–1984*, volume 1186 of *Lecture Notes in Mathematics.* Springer Verlag, Berlin, 1986.

[4] P.H. BAXENDALE. Moment stability and large deviations for linear stochastic differential equations. In N.Ikeda, editor, *Probabilistic Methods in Mathematical Physics, Katata and Kyoto 1985*, pages 31–54, Tokyo, 1987. Kinokuniya.

[5] P.H. BAXENDALE and E.M. HENNING. Stabilization of a linear system via rotaional control. *Random & Computational Dynamics*, 1(4):395–421, 1992–93.

[6] A. BENVENISTE, M. METIVIER, and P. PRIOURET. *Algorithmes Adaptatifs et Approximations Stochastiques.* Techniques Stochastiques. Masson, Paris, 1987.

[7] R.S. BUCY. Stability and positive supermartingales. *Journal of Differential Equations*, 1:198, 1965.

[8] F. CAMPILLO and A. TRAORÉ. Lyapunov exponents of controlled SDE's and stabilizability property : Some examples. Rapport de Recherche 2397, INRIA, November 1994.

[9] R.F. CURTAIN, editor. *Stability of Stochastic Dynamical Systems, Warwick 1972*, volume 294 of *Lecture Notes in Mathematics*, Berlin, 1972. Springer Verlag.

[10] P. FLORCHINGER. Lyapunov–like techniques for stochastic stability. *SIAM Journal on Control and Optimization*, 33(4):1151–1169, July 1995.

[11] Z.Y. GAO and N.U. AHMED. Feedback stabilizability of non–linear stochastic systems with state dependent noise. *Internation Journal of Control*, 45(2):729–737, 1987.

[12] U.G. HAUSSMANN. Stability of linear systems with control dependent noise. *SIAM Journal on Control and Optimization*, 11(2):382–394, 1973.

[13] R.Z. KHASMINSKII. *Stochastic Stability of Differential Equations.* Sijthoff and Noordhoff, Alphen aan den Rijn, 1980.

[14] F. KOZIN and Z.Y. ZHANG. The exact almost sure stability for a specific class of non–linear Ito differential equations. In W. Schiehlen, editor, *Nonlinear Dynamics in Engineering Systems.* UTIAM Symposium, Stuttgart, Germany 1989, Springer Verlag, 1990.

[15] H. J. KUSHNER and D. S. CLARK. *Stochastic approximation methods for constrained and unconstrained systems*, volume 26 of *Applied mathematical sciences.* Springer Verlag, 1978.

[16] X. MAO. *Stability of stochastic differential equations with respect to semi–martingales*, volume 251 of *Pitman Research Notes in Mathematics.* Longman, London, 1991.

[17] E. PARDOUX and V. WIHSTUTZ. Lyapunov exponent of linear stochastic systems with large diffusion term. *Stochastic Processes and their Applications*, 40(2):289–308, 1992.

[18] B.T. POLYAK and A.B. JUDITSKY. Acceleration of stochastic approximation by averaging. *SIAM Journal on Control and Optimization*, 30(4):838–854, July 1992.

[19] J.L. WILLEMS and J.C. WILLEMS. Feedback stabilizability for stochastic system with state and control dependent noise. *Automatica*, 12:277–283, 1976.

[20] W.M. WONHAM. Liapunov criteria for weak stochastic stability. *Journal of Differential Equations*, 2:195–207, 1966.