

## Approximation of some discrete-time stochastic processes by differential equations

Raúl Fierro \*

### Abstract

This work deals with solutions of ordinary differential equations as approximations of some discrete-time stochastic processes. Similarly, these stochastic processes may be seen as schemes of approximation for this solution. Indeed, these stochastic schemes are defined and their convergence to the solution of a differential equation is proven. Moreover, the asymptotic distribution of the fluctuations about the limit solution is studied. This fact gives the asymptotic distribution of a random global error of approximation. Main results are illustrated by means of the so called SIS epidemic model and numerical simulations are carried out.

*2000 AMS Classification:* 60F17, 34K28.

**Keywords:** Law of Large Numbers; Central Limit Theorem; Numerical schemes; SIS epidemic model.

Received 26 : 06 : 2013 : Accepted 04 : 11 : 2013

### 1. Introduction

Often processes associated to population dynamics are mathematically modeled by differential equations and/or stochastic processes, which are of continuous or discrete time. Because the analysis of a model based on differential equations is less cumbersome and more efficient, both from a mathematical point of view as computational, by introducing a stochastic model for a given process is desirable that it can be approximated by the solution of an Ordinary Differential Equation (ODE), as is also the case studied in this work. Some authors such as Kurtz [9, 10, 11] and Darling and Norris [3] have studied the approximation of continuous-time Markov processes with pure jump by solving an ODE. The convergence shown by these authors is almost surely and based on the Markov property of these processes. Our interest is to analyze such an approach for a class of discrete

---

\*Instituto de Matemáticas, Pontificia Universidad Católica de Valparaíso, Avda. Brasil 2950, Valparaíso, Chile. Instituto de Matemática, Universidad de Valparaíso.  
Email: rfierro@ucv.cl

time stochastic processes which are not necessarily Markovian, but including discrete-time Markov chains. Conversely, given an ordinary differential equation, it is possible to approximate its solution through this kind of processes. Indeed, our class of stochastic processes can be seen as a stochastic variant of the Euler scheme to approximate the solution to an ODE. These schemes of approximation are presented as discrete-time stochastic processes, which includes but are not limited to Markov chains. Recently some authors, such as Abbasbandy and Bervillier in [1], Eslahchi *et al.* in [4] Parand *et al.* in [5], among others, have studied the problem of approximation for ordinary differential equations by different deterministic methods. Also, stochastic schemes of approximation have been developed by Fierro and Torres in [6] and Kloeden and Platen in [8]. This latter reference deals with schemes of approximation for stochastic differential equations (SDE). In [12], Kushner and Dupuis present a stochastic scheme of approximation for SDE based on a Markov chain, which, in particular, can be applied for approximating solutions to ODEs. Even though, in general, our schemes need not be Markovian, this model can be included in our setting, whenever the noise part of the equation is zero.

The main results presented in this work are the convergence of the mentioned schemes to the solution to the ODE and a central limit theorem, which allows to know the asymptotic distribution of the global error of the approximation. These results are applied to an example coming from the biomathematical literature. Indeed, the differential equation modeling the well-known SIS epidemic model is analyzed under our framework by means of two natural schemes of approximation.

In order to quantify the probability of error in the approximation, the central limit theorem presented allows to know the asymptotic distribution of the global error, i.e. of the fluctuations of the process around the solution to the ODE. By this result it is possible to establish confidence bands around this solution, which are determined by a preassigned probability. Therefore, when a particular heuristic model is defined by a solution to an ODE, it is possible to perform an asymptotical statistical test to validate the model. Indeed, by considering the stochastic model as the observed process, these observations should be close to the solution to the ODE insofar this solution to be a good model for the heuristic situation. Hence, the asymptotic distribution of the global error allows to carry out a goodness of fit test for both the random and deterministic model.

The plan of this paper is as follows. In Section 2, by means of a recursive condition, we define a family of discrete-time stochastic processes, which approximate the solution of the ODE. Main results of this work, along with their proofs, are stated in Section 3. Since our schemes are stochastic, the global error is so. The asymptotic distribution of it is analyzed in Section 4. Moreover, some dispersion measures and their estimators are defined in this section. An example is included in Section 5. Indeed, the solution to the differential equation defining the so called SIS epidemic model is approximated through two schemes included in our framework. Both schemes are compared and numerical simulations are carried out.

## 2. Preliminaries

Let  $x_0 \in \mathbb{R}^d$  and  $b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a continuous function satisfying the following Lipschitz condition:

$$(L) \quad \|b(t, x) - b(t, y)\| \leq K \|x - y\|, \text{ for all } t \in \mathbb{R}_+,$$

where  $K$  is a positive constant and  $\|\cdot\|$  stands for the usual norm in  $\mathbb{R}^d$ . Hence, the initial value problem

$$(2.1) \quad \dot{x}(t) = b(t, x(t)) \quad x(0) = x_0,$$

has one and only one solution.

In this work, a stochastic scheme of approximation for the solution to (2.1) is stated.

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and  $\{t_k^n\}_{k \in \mathbb{N}}$  the sequence of non-negative real numbers defined as  $t_k^n = Ck/n$ , ( $C > 0, n \in \mathbb{N} \setminus \{0\}$ ). In what follows and without loss of generality, we assume  $C = 1$ . An approximation of the solution to (2.1) is obtained by means of a sequence  $\{Z^n\}_{n \in \mathbb{N}}$  of stochastic processes defined on  $\mathbb{R}_+ \times \Omega$ . Such an approximation is obtained by defining  $\mathcal{F}_k^n = \sigma(Z^n(t_1^n), \dots, Z^n(t_k^n))$  as the sigma algebra generated by  $Z^n(t_1^n), \dots, Z^n(t_k^n)$ ,  $x^n = x^n(0) + \frac{1}{n}Z^n$  and assuming the following condition:

$$(C) \quad \mathbb{E}(\Delta Z^n(t_k^n) | \mathcal{F}_{k-1}^n) = b(t_{k-1}^n, x^n(t_{k-1}^n)), \quad (k \geq 1),$$

where for any stochastic process  $Z$ ,  $\Delta Z(t_k^n) = Z(t_k^n) - Z(t_{k-1}^n)$ .

For a real number  $x$ ,  $[x]$  stands for the integer part of  $x$  and

$$L^n(t) = \frac{1}{n} \sum_{k=1}^{[nt]} \xi_k^n, \quad (t \geq 0),$$

where  $\xi_k^n = \Delta Z^n(t_k^n) - b(t_{k-1}^n, x^n(t_{k-1}^n))$ . By defining  $\mathcal{F}_t^n = \mathcal{F}_{[nt]}^n$ , ( $t \geq 0$ ), we have  $L^n$  is a  $d$ -dimensional  $\mathcal{F}_t^n$ -martingale and

$$(2.2) \quad Z^n(t) = Z^n(0) + n \sum_{k=1}^{[nt]} b(t_{k-1}^n, x^n(t_{k-1}^n)) \Delta t^n + nL^n(t), \quad (t > 0).$$

Given any  $d$ -dimensional martingale  $L$ , its predictable quadratic variation, at time  $t$ , is denoted by  $\langle L \rangle(t)$ . Thus,  $\langle L \rangle(t)$  is a  $d \times d$ -matrix and it directly follows that

$$(2.3) \quad \langle L^n \rangle(t) = \frac{1}{n^2} \sum_{k=1}^{[nt]} \mathbb{E}(\xi_k^n \xi_k^{n\top} | \mathcal{F}_{k-1}^n), \quad (t \geq 0)$$

From (2.1) and (2.2), we have

$$(2.4) \quad x^n(t) - x(t) = x^n(0) - x(0) + \int_0^{[nt]/n} \{b([nu]/n, x^n(u)) - b(u, x(u))\} du + L^n(t) + \epsilon^n(t),$$

where,  $\epsilon^n(t) = x([nt]/n) - x(t)$ . Note that  $\sup_{0 \leq u \leq t} \|\epsilon^n(s)\| \leq S_t/n$ , where  $S_t = \sup_{0 \leq u \leq t} \|b(u, x(u))\|$ .

### 3. Main results

In the sequel,  $x$  stands for the solution to (2.1). In this section, the convergence of  $x^n$  to  $x$  is stated, which means  $\{x^n\}_{n \in \mathbb{N}}$  converges uniformly in probability, on compact subsets of  $\mathbb{R}_+$ , to  $x$  as  $n$  goes to  $\infty$ .

**3.1. Theorem.** *Assume conditions (C) and (L) are satisfied. Moreover, suppose the following two conditions hold:*

$$(3.1.1): x^n(0) \xrightarrow{\mathbb{P}} x_0.$$

$$(3.1.2): \text{For each } t \geq 0,$$

$$\frac{1}{n^2} \sum_{k=1}^{[nt]} \mathbb{E}(\|\xi_k^n\|^2 | \mathcal{F}_{k-1}^n) \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \rightarrow \infty.$$

Then, for each  $T > 0$ ,  $\sup_{0 \leq t \leq T} \|x^n(t) - x(t)\| \xrightarrow{\mathbb{P}} 0$ , as  $n \rightarrow \infty$ .

*Proof.* Fix  $T > 0$  and let  $g^n(t) = \sup_{0 \leq s \leq t} \|x^n(s) - x(s)\|$ , ( $t \in [0, T]$ ). From (2.4) and (L), we obtain

$$g^n(t) \leq \alpha_n + K \int_0^t g^n(u) du,$$

where  $\alpha_n = g^n(0) + \sup_{0 \leq t \leq T} \|L^n(t)\| + S_T/n$ . Since, by (3.1.1),  $\{g^n(0)\}_{n \in \mathbb{N}}$  converges in probability to zero, by Gronwall's inequality, it suffices to verify that  $\{\sup_{0 \leq t \leq T} \|L^n(t)\|\}_{n \in \mathbb{N}}$  converges in probability to zero.

From Theorem 1 by Lenglart [13], for any  $\epsilon, \eta > 0$ , we have

$$\begin{aligned} \mathbb{P}(\sup_{0 \leq t \leq T} \|L^n(t)\|^2 > \epsilon) &\leq \frac{1}{\epsilon} \mathbb{E}(\text{tr}\langle L^n \rangle(T) \wedge \eta) + \mathbb{P}(\text{tr}\langle L^n \rangle(T) > \eta) \\ &< \frac{\eta}{\epsilon} + \mathbb{P}\left(\frac{1}{n^2} \sum_{k=1}^{[nT]} \text{tr} \mathbb{E}(\xi_k^n \xi_k^{n\top} | \mathcal{F}_{k-1}^n) > \eta\right) \\ &= \frac{\eta}{\epsilon} + \mathbb{P}\left(\frac{1}{n^2} \sum_{k=1}^{[nT]} \mathbb{E}(\|\xi_k^n\|^2 | \mathcal{F}_{k-1}^n) > \eta\right) \end{aligned}$$

and hence, by (3.1.2),

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\sup_{0 \leq t \leq T} \|L^n(t)\|^2 > \epsilon\right) = 0,$$

which concludes the proof.  $\blacksquare$

For each  $t \in \mathbb{R}_+$ , let  $\tilde{b}_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be such that  $\tilde{b}_t(x) = b(t, x)$  and suppose for each  $t \in \mathbb{R}_+$ ,  $\tilde{b}_t$  has continuous partial derivatives. The following result aims to the problem of finding confident bands for the approximate solution to (2.1). Before stating it, for each  $(t, a) \in \mathbb{R}_+ \times \mathbb{R}^d$ , let  $D(b)(t, a)$  denote the Jacobian matrix of  $\tilde{b}_t$  at  $a$ .

A function  $v : \mathbb{R}_+ \rightarrow \mathbb{R}^{d \times d}$  is said to be a positive-definite, if for each  $\theta \in \mathbb{R}^m$ ,  $f^\theta : \mathbb{R}_+ \rightarrow \mathbb{R}$  defined as  $f^\theta(t) = \theta^\top v(t) \theta$  is continuous, increasing and  $f^\theta(0) = 0$ .

**3.2. Theorem.** *Let  $v$  be a positive-definite function and  $y^n = \sqrt{n}(x^n - x)$ . Suppose the following conditions hold:*

**(3.2.1):** *The partial derivatives of  $\tilde{b}_t$  exist and are continuous in  $\mathbb{R}^d$ .*

**(3.2.2):** *For each  $\epsilon > 0$  and  $t \geq 0$ ,  $\frac{1}{n} \sum_{k=1}^{[nt]} \mathbb{E}(\|\xi_k^n\|^2 \mathbb{I}_{\{\|\xi_k^n\| > \epsilon\sqrt{n}\}} | \mathcal{F}_{k-1}^n) \xrightarrow{\mathbb{P}} 0$ , as  $n$  goes to  $\infty$ .*

**(3.2.3):**  *$\{y^n(0)\}_{n \in \mathbb{N}}$  converges in distribution to a random variable  $\eta$ .*

**(3.2.4):** *For each  $t \geq 0$ ,  $\sup_{0 \leq s \leq t} \left\| \frac{1}{n} \sum_{k=1}^{[ns]} \mathbb{E}(\xi_k^n \xi_k^{n\top} | \mathcal{F}_{k-1}^n) - v(s) \right\| \xrightarrow{\mathbb{P}} 0$ , as  $n$  goes to  $\infty$ .*

*Then, the sequence  $\{y^n\}_{n \in \mathbb{N}}$  converges in law to the solution  $y$  satisfying the following stochastic differential equation:*

$$(3.1) \quad dy(t) = D(b)(t, x(t))y(t)dt + dm(t), \quad y(0) = \eta,$$

*where  $m$  is a  $d$ -dimensional continuous martingale starting at zero with predictable quadratic variation, at  $t \geq 0$ , given by  $\langle m \rangle(t) = v(t)$ .*

*Proof.* Condition (3.2.2) implies the jump asymptotic rarefaction condition in [14, Theorem 8, Chapter II.5] by Rebolledo, for the sequence of martingales  $\{m^n\}_{n \in \mathbb{N}}$ , where  $m^n = \sqrt{n}L^n$ . This fact along with condition (3.2.4) imply  $\{m^n\}_{n \in \mathbb{N}}$  converges in law to a continuous martingale  $m$  starting at zero and having predictable quadratic variation  $\langle m \rangle$  given by  $\langle m \rangle(t) = v(t)$ .

Let  $b_i$  be the  $i$ -th coordinate of  $b$ , ( $i = 1, \dots, d$ ). By the Value Mean Theorem, there exists  $\theta_i^n(t) \in \mathbb{R}^d$  between  $x(t)$  and  $x^n(t)$  such that  $b_i(t, x^n(t)) - b_i(t, x(t)) = D(b_i)(t, \theta_i^n(t))(x^n(t) - x(t))$ , where  $D(b_i)(t, a)$  is the Jacobian matrix of  $b_i(t, \cdot)$  at  $a \in \mathbb{R}^d$ .

From (2.4), it is derived

$$(3.2) \quad y^n(t) = y^n(0) + \int_0^t D_n(u)y^n(s) ds + m^n(t) + \sqrt{n}\epsilon^n(t),$$

where  $D_n(u) = (D(b_1)(u, \theta_1^n(u)), \dots, D(b_d)(u, \theta_d^n(u)))^\top$ .

Consequently,

$$\sup_{0 \leq u \leq t} \|y^n(u)\| \leq \|y^n(0)\| + C(t) \int_0^t \sup_{0 \leq u \leq s} \|y^n(u)\| ds + \sup_{0 \leq u \leq t} \|m^n(u)\| + \frac{S_t}{\sqrt{n}},$$

where  $C_n(t) = \sup_{0 \leq u \leq t} \sup_{\|y\|=1} \|D_n(u)y\|$  and  $C(t) = \sup_{n \in \mathbb{N}} C_n(t)$ . From (3.2.1) and Theorem 3.1,  $\{\sup_{0 \leq u \leq t} \|D_n(u)\|\}_{n \in \mathbb{N}}$  converges in probability to  $\sup_{0 \leq u \leq t} \|D(b)(u, x(u))\|$  and thus,  $C(t) < \infty$ .

Hence, from a standard application of the Gronwall inequality, we obtain

$$(3.3) \quad \sup_{0 \leq u \leq t} \|y^n(u)\| \leq (\|y^n(0)\| + \sup_{0 \leq u \leq t} \|m^n(u)\| + S_t/\sqrt{n}) e^{tC(t)}.$$

In order to prove the convergence in law of  $\{y^n\}_{n \in \mathbb{N}}$  and that its limit has continuous trajectories, Theorem 15.5 by Billingsley (1968) is used. Since  $\{y^n(0)\}_{n \in \mathbb{N}}$  converges in distribution, Theorem 6.2 in Billingsley (1968) implies this sequence is tight, which means for each  $\epsilon > 0$ , there exists  $a > 0$  such that  $\sup_{n \in \mathbb{N}} \mathbb{P}(\|y^n(0)\| > a) < \epsilon$ . Hence

$$(3.4) \quad \lim_{a \rightarrow \infty} \sup_{n \in \mathbb{N}} \mathbb{P}(\|y^n(0)\| > a) = 0.$$

Fix  $T > 0$  and let us define the modulus of continuity  $\omega_T$  as

$$\omega_T(z, \delta) = \sup_{|s-t| < \delta} \|z(s) - z(t)\|,$$

where  $\delta > 0$  and  $z : [0, T] \rightarrow \mathbb{R}^d$  is right continuous and left-hand limited.

From (3.2) we have

$$(3.5) \quad \omega_T(y^n, \delta) \leq \delta C(T) \sup_{0 \leq t \leq T} \|y^n(t)\| + \omega_T(m^n, \delta) + 2S_T/\sqrt{n}.$$

Since  $\{m^n\}_{n \in \mathbb{N}}$  converges in distribution to  $m$ , it follows from Theorem 15.2 in Billingsley [2] that for each  $\epsilon > 0$ ,  $\lim_{\delta \rightarrow 0} \sup_{n \in \mathbb{N}} \mathbb{P}(\omega_T(m^n, \delta) > \epsilon) = 0$ . Hence, from (3.3), for each  $\epsilon > 0$ ,

$$(3.6) \quad \lim_{\delta \rightarrow 0} \sup_{n \in \mathbb{N}} \mathbb{P}(\omega_T(y^n, \delta) > \epsilon) = 0.$$

Conditions (3.4) and (3.6) imply the sequence  $\{P_n\}_{n \in \mathbb{N}}$  of probabilities measures, where  $P_n$  is the law of  $y^n$ , satisfies the hypotheses of Theorem 15.5 in Billingsley [2] and hence,  $\{P_n\}_{n \in \mathbb{N}}$  is tight and every limit point  $P$  of this sequence satisfies  $P(C) = 1$ , where  $C$  is the space of continuous functions from  $\mathbb{R}_+$  to  $\mathbb{R}^d$ . This fact, along Theorem 6.1 in Billingsley [2], imply that  $\{P_n\}_{n \in \mathbb{N}}$  is relatively compact. Let  $\{y^{n_k}\}_{k \in \mathbb{N}}$  a subsequence converging in distribution to a process  $y$ . Since, by Theorem 3.1,  $\{D_n\}_{n \in \mathbb{N}}$  converges uniformly in probability to  $D(b)(\cdot, x(\cdot))$ ,  $\{m^n\}_{n \in \mathbb{N}}$  converges in law to  $m$  and  $\{\sqrt{n}\epsilon^n\}_{n \in \mathbb{N}}$  converges uniformly to 0, it follows from (3.2) that  $y$  is a solution to (3.1). Finally, uniqueness of solutions to (3.1) implies  $\{y^n\}_{n \in \mathbb{N}}$  converges in distribution to this solution  $y$ , which concludes the proof.  $\blacksquare$

### Remarks

**R1:** By Itô's rule, the unique solution to (3.1) is given by

$$y(t) = \Psi(t) \left( \eta + \int_0^t \Psi(s)^{-1} dm(s) \right), \quad 0 \leq t \leq 1,$$

where  $\Psi$  is the unique solution to the matrix differential equation

$$\Psi'(t) = D(b)(t, x(t))\Psi(t), \quad \Psi(0) = \text{identity matrix.}$$

**R2:** Condition (3.2.4) holds whenever for each  $t \geq 0$  and  $\epsilon > 0$ ,

$$\frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}(\|\xi_k^n\|^2 \mathbf{1}_{\{\|\xi_k^n\| > \epsilon\sqrt{n}\}}) \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \text{ goes to } \infty. \quad (\text{Lindeberg condition}).$$

## 4. Random global discretization error

In this section, we assume the partial derivatives of  $\tilde{b}_t$  exist and are continuous in  $\mathbb{R}^d$  and for each  $n \in \mathbb{N}$ ,  $x^n(0) = x(0)$ .

**4.1. Some definitions.** In order to analyze the error produced by the discretization scheme introduced here, for a fixed  $T > 0$ , we define the random global error to be  $\widehat{e}_T^n = \|x^n(T) - x(T)\|$  and, for  $p \geq 1$ , the  $p$ -mean global error to be  $e_T^n(p) = \mathbb{E}(\|x^n(T) - x(T)\|^p)^{1/p}$ , whenever  $\mathbb{E}(\|x^n(T)\|^p) < \infty$ , i.e.  $e_T^n(p)$  is the usual norm of  $\widehat{e}_T^n$  defined on  $L^p(\Omega, \mathcal{F}, \mathbb{P})$ , the space of random variables  $x$  such that  $\mathbb{E}(|x|^p) < \infty$ . We refer to  $e_T^n(2)$  as the square mean global error. Since, even in simple cases, it is not possible to know or calculate  $e_T^n(p)$ , an estimator of this one is obtained by defining

$$\widehat{e}_T^n(p, m) = \left( \frac{1}{m} \sum_{i=1}^m \widehat{e}_T^n(i)^p \right)^{1/p},$$

where  $\widehat{e}_T^n(1), \dots, \widehat{e}_T^n(m)$  are independent random variables with the same distribution than  $\widehat{e}_T^n$ . Anyway the distribution of  $\widehat{e}_T^n$  needs to be known. Theorem 3.2 allows to obtain an approximation of this distribution.

It follows from the Strong Law of Large Numbers by Kolmogorov that the estimator  $\widehat{e}_T^n(p, m)$  is strongly consistent, i.e.

$$\lim_{m \rightarrow \infty} \widehat{e}_T^n(p, m) = e_T^n(p).$$

Consequently, by carrying out simulations of  $\widehat{e}_T^n$ , an approximation of  $e_T^n(p)$  can be obtained. In particular, the sample variance of  $\widehat{e}_T^n$  can be consistently estimated by means of

$$S_m^{n,2} = \frac{1}{m} \sum_{i=1}^m \left( \widehat{e}_T^n(i) - \frac{1}{m} \sum_{i=1}^m \widehat{e}_T^n(i) \right)^2 = \widehat{e}_T^n(2, m)^2 - \widehat{e}_T^n(1, m)^2.$$

Since  $(S_m^{n,2}; m \in \mathbb{N})$  converges  $\mathbb{P}$ -a.s. to  $\text{Var}(\widehat{e}_T^n) = e_T^n(2)^2 - e_T^n(1)^2$  and  $\text{Var}(\widehat{e}_T^n)$  is a measure of dispersion, small values of  $S_m^{n,2}$  suggest no much simulations of  $\widehat{e}_T^n$  are necessary to carry out a suitable estimation of the square mean global error.

**4.2. Asymptotic distribution of the global error.** In this subsection we examine the asymptotical distribution of  $\widehat{e}_T^n$ . Indeed, let  $y^n$  be as in Theorem 3.2 and suppose the hypotheses of this theorem hold. Thus,  $\widehat{e}_T^n = (\Delta t^n)^{1/2} \|y^n(T)\|$  and from Theorem 3.2,  $\widehat{e}_T^n$  is asymptotically distributed as  $\|y(T)\|/\sqrt{n}$ , where  $y$  is the solution to (3.1) with  $y(0) = 0$ . From Remark R1,

$$y(T) = \int_0^T B(T, s) dm(s)$$

where  $B(t, s) = \Psi(t)\Psi(s)^{-1}$ . Hence, by taking into account that, for almost sure  $s \geq 0$ , there exists  $v'(s)$ , the derivative of  $v$  at  $s$ , we have

$$(4.1) \quad \mathbb{E}(\|y(T)\|^2) = \int_0^T B(T, s)v'(s)B(T, s)^\top ds$$

and for large values of  $n$ ,  $e_T^n(2)$ , the square mean global error can be approximated by  $\sqrt{\mathbb{E}(\|y(T)\|^2)/n}$ , whenever  $\{\|y^n(T)\|^2\}_{n \in \mathbb{N}}$  is uniformly integrable (see Theorem 5.4 in Billingsley [2]).

**4.3. Hypothesis testing.** Fix  $T > 0$ . The global error could be used to develop an asymptotic hypothesis testing to reject or not the validity of the model. This procedure is performed in the following natural manner: given a significance level  $\alpha \in (0, 1)$ , we choose  $t_\alpha > 0$  such that  $\mathbb{P}(\|y(T)\| > t_\alpha) = \alpha$ . Then, we compare the statistic  $\sqrt{n}\widehat{e}_T^n$  with  $t_\alpha$ . If  $\sqrt{n}\widehat{e}_T^n > t_\alpha$  we reject the hypothesis as false, while if  $\sqrt{n}\widehat{e}_T^n \leq t_\alpha$ , we conclude that there is no sufficient evidence that the model is incorrect. In this case, although the null hypothesis need not to be true, no change in the model is recommended.

## 5. An example

In this section, we apply the results of this work to a known differential equation coming from the biomathematical literature. A brief description of this model is given in the first subsection and two probabilistic schemes, which are approximated by the solution to this equation, are presented. Results of this work are illustrated by means numerical simulations in Subsection 2.

**5.1. The SIS epidemic model.** One of the most commonly used differential equations in the biomathematical literature is that correspondig to the SIS epidemic model. In this model it is assumed that at time  $t \geq 0$ ,  $x(t)$  and  $y(t)$  represent the densities of infective and susceptible individuals, respectively, and they satisfy the following system of ordinary differential equations:

$$\begin{aligned}\frac{dx}{dt}(t) &= \beta x(t)y(t) - \gamma y(t) \\ \frac{dy}{dt}(t) &= -\beta x(t)y(t) + \gamma y(t).\end{aligned}$$

Since for each  $t \geq 0$ ,  $x(t) + y(t) = 1$ , this model is completely determined by the ordinary differential equation:

$$(5.1) \quad \frac{dx}{dt}(t) = \beta(1 - x(t))x(t) - \gamma x(t).$$

This test equation, given  $x(0) = x_0 \in ]0, 1[$ , has the unique solution

$$x(t) = \begin{cases} \frac{x_0}{x_0\beta t + 1} & \text{if } \beta = \gamma \\ \frac{x_0(\beta - \gamma)e^{(\beta - \gamma)t}}{\beta - \gamma + x_0\beta(e^{(\beta - \gamma)t} - 1)} & \text{if } \beta \neq \gamma. \end{cases}$$

Let  $x^n(0) = [nx_0]/n$ . In accordance with our setting, let  $Z^n = A^n - B^n$  and  $x^n = x^n(0) + \frac{1}{n}Z^n$ , where  $A^n$  and  $B^n$  are independent. Two probability distributions are defined below, in order to  $Z^n$  takes values in the nonnegative integers. The first one, which is labeled by distribution D1 is recursively defined as follows. Let  $I^n(t_{k-1}^n) = [nx_0] + Z^n(t_{k-1}^n)$  and  $S^n(t_{k-1}^n) = n - I^n(t_{k-1}^n)$ , conditional on  $\mathcal{F}_{k-1}^n$ ,  $\Delta A^n(t_k^n)$  and  $\Delta B^n(t_k^n)$  have Binomial distribution with parameters  $(S^n(t_{k-1}^n), \beta x^n(t_{k-1}^n)\Delta t^n)$  and  $(I^n(t_{k-1}^n), \gamma\Delta t^n)$ . I.e., for each  $a \in \{0, \dots, S^n(t_{k-1}^n)\}$  and  $b \in \{0, \dots, I^n(t_{k-1}^n)\}$ ,

$$\mathbb{P}(\Delta A^n(t_k^n) = a | \mathcal{F}_{k-1}^n) = \binom{S^n(t_{k-1}^n)}{a} p_{n,k-1}^a (1 - p_{n,k-1})^{S^n(t_{k-1}^n) - a}$$

and

$$\mathbb{P}(\Delta B^n(t_k^n) = b | \mathcal{F}_{k-1}^n) = \binom{I^n(t_{k-1}^n)}{b} q_n^b (1 - q_n)^{I^n(t_{k-1}^n) - b},$$

where  $p_{n,k-1} = \beta x^n(t_{k-1}^n) \Delta t^n$  and  $q_n = \gamma \Delta t^n$ . Here, it is suppose  $n$  is large enough to  $\beta/n \leq 1$  and  $\gamma/n \leq 1$ .

Since for each  $k \in \mathbb{N}$ ,  $0 \leq \Delta A^n(t_k^n), \Delta B^n(t_k^n) \leq n$ ,  $\mathbb{P}$ -a.s., on a time interval  $[0, T]$ , ( $T > 0$ ), the state space of  $x^n$  is a finite subset of  $\frac{1}{n}\mathbb{Z}_+ = \{k/n : k \in \mathbb{Z}_+\}$ .

By using notations before, we have

$$\xi_k^n = \Delta A^n(t_k^n) - \Delta B^n(t_k^n) - \{\beta S^n(t_{k-1}^n) x^n(t_{k-1}^n) \Delta t^n - \gamma I^n(t_{k-1}^n) \Delta t^n\}$$

and consequently,

$$\mathbb{E}(|\xi_k^n|^2 | \mathcal{F}_{k-1}^n) = \beta(1 - x^n(t_{k-1}^n)) x^n(t_{k-1}^n) (1 - \beta x^n(t_{k-1}^n) \Delta t^n) + \gamma x^n(t_{k-1}^n) (1 - \gamma \Delta t^n).$$

Since

$$\mathbb{E}(|\xi_k^n|^3 | \mathcal{F}_{k-1}^n) = S^n(t_{k-1}^n) p_{n,k-1} (1 - p_{n,k-1})^2 + I^n(t_{k-1}^n) q_n (1 - q_n)^2 \leq \beta + \gamma,$$

$\{|\xi_k^n|^2; n, k \geq 1\}$  is uniformly integrable and Lindeberg condition stated in R2 holds. In addition,  $\mathbb{E}(|\xi_k^n|^2 | \mathcal{F}_{k-1}^n) \leq 1$  and hence conditions (3.1.2) and (3.2.2) of Theorem 3.1 and 3.2, respectively, are satisfied. Consequently, Theorem 3.1 implies for each  $t \geq 0$ ,

$$\sup_{0 \leq s \leq t} \left| \frac{1}{n} \sum_{k=1}^{[ns]} \mathbb{E}(|\xi_k^n|^2 | \mathcal{F}_{k-1}^n) - v_1(s) \right| \xrightarrow{\mathbb{P}} 0, \text{ as } n \text{ goes to } \infty,$$

where

$$(5.2) \quad v_1(s) = \int_0^s \{\beta(1 - x(u)) + \gamma\} x(u) \, du.$$

As shown in [7], the distribution of  $x^n$  has a biomathematical sense, where  $n$  denotes the population size. However, other distributions allow  $x^n$  is approximated by the solution to (5.1). Indeed, a second distribution for  $x^n$ , which has less variability, and we label by D2, is defined as follows. Assumed that, conditional to  $\mathcal{F}_{k-1}^n$ ,  $\Delta A^n(t_k^n)$  and  $\Delta B^n(t_k^n)$  have Bernoulli distribution with parameters  $\beta S^n(t_{k-1}^n) x^n(t_{k-1}^n) \Delta t^n$  and  $\gamma I^n(t_{k-1}^n) \Delta t^n$ , respectively. For large enough values of  $n$ , these conditional parameters are equal or less than one and it is easy to see the hypotheses of Theorems 3.1 and 3.2 hold with a positive-definite function  $v_2$  defined by

$$(5.3) \quad v_2(s) = \int_0^s \{\beta(1 - x(u))(1 - \beta(1 - x(u))x(u)) + \gamma(1 - \gamma x(u))\} x(u) \, du$$

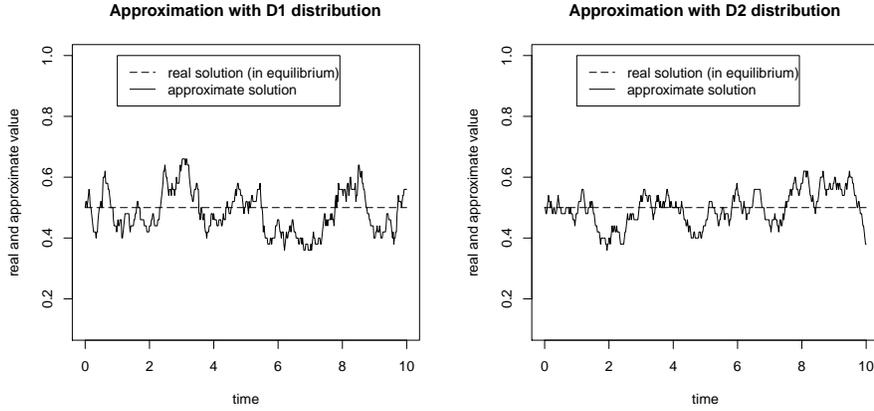
and satisfying, for each  $t \geq 0$ ,

$$\sup_{0 \leq s \leq t} \left| \frac{1}{n} \sum_{k=1}^{[ns]} \mathbb{E}(|\xi_k^n|^2 | \mathcal{F}_{k-1}^n) - v_2(s) \right| \xrightarrow{\mathbb{P}} 0, \text{ as } n \text{ goes to } \infty,$$

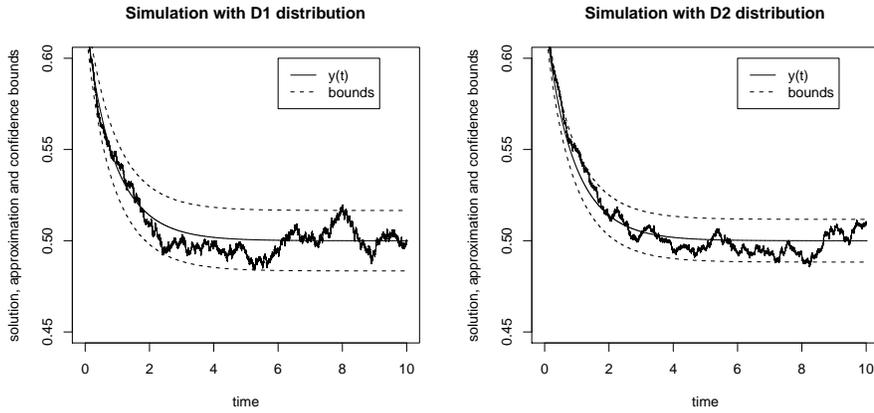
This latter probabilistic scheme has some advantages regarding the conditional Binomial jumps case, labeled by D1. One of them is that, according to D2 distribution,  $\Delta Z^n$  takes values in the set  $\{-1, 0, 1\}$  instead of  $\{-n, \dots, 0, \dots, n\}$  as in the D1 distribution. In addition, from (5.2) and (5.3), for each  $s \geq 0$ ,

$$(5.4) \quad v_1'(s) - v_2'(s) = \{\beta(1 - x(s))x(s)\}^2 + \{\gamma x(s)\}^2 \geq 0.$$

This inequality and (4.1) imply the square mean global error is lesser, for the D2 distribution, than the corresponding error for the D1 distribution.



**Figure 1.** Approximations of the equilibrium solution with  $n = 50$ ,  $\beta = 2$  and  $\gamma = 1$ .



**Figure 2.** Approximations for the solution and confidence bounds with  $n = 7,000$ ,  $\beta = 2$  and  $\gamma = 1$ .

**5.2. Numerical simulations.** In the sequel, some of the concepts presented before are applied to the solution to (5.1) with the scheme of approximations labeled by D1 and D2. First, in order to appreciate the difference in variability of the schemes D1 and D2, the equilibrium solution to (5.1) is considered, i.e.  $x_0 = 1 - \gamma/\beta$ . We simulated both approximations for  $T = 10$ ,  $\beta = 2$ ,  $\gamma = 1$  and  $n = 50$ ; see Figure 1.

Let  $a(t) = \beta - \gamma - 2\beta x(t)$  and,  $v_1$  and  $v_2$  defined by (5.2) and (5.3), respectively. From (4.1),  $Var_1(y(t))$  and  $Var_2(y(t))$ , the variances of  $y(t)$  according to the D1 and D2 distribution, respectively, are given by

$$Var_i(y(t)) = \int_0^t v'_i(u) e^{2 \int_u^t a(s) ds} du,$$

where  $v'_1(u) = \{\beta(1 - x(u)) + \gamma\}x(u)$  and  $v'_2(u) = \{\beta(1 - x(u))(1 - \beta(1 - x(u))x(u)) + \gamma(1 - \gamma x(u))\}x(u)$ .

Let  $0 < \alpha < 1$  and  $\Phi$  be the cumulative function of a standard normal distribution. Since for the D1 approximation scheme and for large  $n$ ,  $\sqrt{n}(x^n(t) - x(t))$  has approximately normal distribution with mean zero and variance  $Var_1(y(t))$ , by defining  $u_\alpha^\pm(t) = x(t) \pm w_{\alpha/2}\sqrt{Var_1(y(t))/n}$ , we have  $x^n(t) \in [u_\alpha^-(t), u_\alpha^+(t)]$  with an approximate probability  $1 - \alpha$  for large values of  $n$ . Analogously,  $x(t) \pm w_{\alpha/2}\sqrt{Var_2(y(t))/n}$  allow to obtain confidence bands for the scheme of approximation based upon the D2 distribution. In Figure 2, simulations of  $x^n$ , starting at  $x_0 = 5/8$ , are carried out according to the D1 and D2 distributions with  $n = 7,000$ ,  $T = 10$ ,  $\beta = 2$  and  $\gamma = 1$ . In both cases, the bounds  $u_\alpha^-$  and  $u_\alpha^+$  are pictured with dash lines for  $\alpha = .05$ , which gives  $w_{\alpha/2} = 1.96$ .

**5.3. About the appropriate value of  $n$ .** In order to choose an appropriate value of  $n$  that provides a good approximation for the global error to a normal distribution, a goodness-of-fit test is developed for each of the both distributions we are considering.

Let  $CHI_i^2(n) = n(\hat{e}_T^n)^2 / Var_i(y(T))$ , ( $i = 1, 2$ ). The values of the variances are given by  $Var_1(y(T)) = 0.5016656$  and  $Var_2(y(T)) = 0.2508328$ . For large enough values of  $n$ , it is expected  $CHI_i^2(n)$ , ( $i = 1, 2$ ), has an approximate  $\chi^2$ -distribution with one degree of freedom, whether is D1 or D2, respectively, the assumed distribution for the model. Let  $F$  be the accumulative distribution function corresponding to a  $\chi^2$ -distribution with one degree of freedom. Consequently, we expect  $F(CHI_i^2(n))$  has an approximately uniform distribution for large values of  $n$ . We use the goodness-of-fit  $\chi^2$ -test to evaluate this concordance. For this purpose, we partition the positive part of the real straight line by  $m$  subintervals determined by  $0 = t_0 < t_1 < \dots < t_{m-1} < t_m = \infty$ , where  $t_0, \dots, t_m$  have been chosen in such a way that  $F(t_v) - F(t_{v-1}) = 1/m$ . Then,  $CHI_i^2(n)$  is simulated repeatedly, recording the number of times that  $CHI_i^2(n)$  fall into each subinterval  $[t_{v-1}, t_v[$ , for each  $v = 1, \dots, m$ . By choosing  $m = 10$ , we have  $t_1 = 0.016$ ,  $t_2 = 0.064$ ,  $t_3 = 0.148$ ,  $t_4 = 0.275$ ,  $t_5 = 0.455$ ,  $t_6 = 0.708$ ,  $t_7 = 1.074$ ,  $t_8 = 1.642$  and  $t_9 = 2.706$ . In addition, for  $m = 10$ , the expected percentage falling into each subinterval is 10%. A  $\chi^2$  test is performed for different values of  $n$ .

First, we analyzed the approximate normality of  $CHI_1^2(n)$ . To this end,  $CHI_1^2(n)$  is simulated  $10^3$  times and the percentages of  $CHI_1^2(n)$  falling into these subintervals are determined by the values in Table 1.

**Table 1.** Percentages of observations of  $CHI_2^2(n)$  for the indicated value of  $n$  and p-values of the corresponding  $\chi^2$  test ( $\beta = 2$ ,  $\gamma = 1$  and  $T = 10$ ).

$n$	$[t_0, t_1[$	$[t_1, t_2[$	$[t_2, t_3[$	$[t_3, t_4[$	$[t_4, t_5[$	$[t_5, t_6[$	$[t_6, t_7[$	$[t_7, t_8[$	$[t_8, t_9[$	$[t_9, t_{10}[$	p-value
40	7.3	0.0	15.9	14.7	13.6	0.0	14.4	7.5	12.4	14.2	0.00015575
50	8.3	15.7	0.0	14.5	13.9	10.1	7.5	8.5	10.8	10.7	0.03813725
60	6.1	13.8	13.1	0.0	11.3	10.8	9.4	15.0	9.9	10.6	0.05308202
70	6.8	11.9	11.8	12.9	0.0	9.0	15.8	9.0	11.1	11.7	0.05671326
80	5.8	15.0	11.0	10.4	9.4	9.2	9.3	11.6	8.8	9.5	0.83829960
90	5.6	12.3	10.1	11.2	8.7	9.1	9.7	13.0	11.2	9.1	0.91180700
100	6.7	12.1	9.0	8.8	9.5	10.5	12.8	13.1	7.6	9.9	0.90154290

Observed percentages of the values of  $CHI_2^2(n)$  falling in the corresponding time intervals, for the seven values of  $n$  given in Table 1, are organized in the matrix

$$\mathbf{A} = (A_{uv}) = \begin{pmatrix} 7.3 & 0.0 & 15.9 & 14.7 & 13.6 & 0.0 & 14.4 & 7.5 & 12.4 & 14.2 \\ 8.3 & 15.7 & 0.0 & 14.5 & 13.9 & 10.1 & 7.5 & 8.5 & 10.8 & 10.7 \\ 6.1 & 13.8 & 13.1 & 0.0 & 11.3 & 10.8 & 9.4 & 15.0 & 9.9 & 10.6 \\ 6.8 & 11.9 & 11.8 & 12.9 & 0.0 & 9.0 & 15.8 & 9.0 & 11.1 & 11.7 \\ 5.8 & 15.0 & 11.0 & 10.4 & 9.4 & 9.2 & 9.3 & 11.6 & 8.8 & 9.5 \\ 5.6 & 12.3 & 10.1 & 11.2 & 8.7 & 9.1 & 9.7 & 13.0 & 11.2 & 9.1 \\ 6.7 & 12.1 & 9.0 & 8.8 & 9.5 & 10.5 & 12.8 & 13.1 & 7.6 & 9.9 \end{pmatrix}.$$

For the purpose of carrying out the test, the statistics

$$\chi_u^2 = \sum_{v=1}^{10} \frac{(O_{uv} - E_{uv})^2}{E_{uv}} \sim \chi^2(9), \quad u = 1, \dots, 7,$$

have been defined, where  $O_{uv} = 10 \times A_{uv}$  and  $E_{uv} = 100$ , for  $u = 1, \dots, 7$  and  $v = 1, \dots, 10$ . For each  $u = 1, \dots, 7$ , the rejection region is defined as  $\{\chi_u^2 > c\}$ , where  $c$  is chosen in such a way that  $\mathbb{P}(\chi_u^2 > c) = .05$ .

We compute the p-values associated with the  $\chi^2$  test statistic to evaluate the goodness-of-fit of  $CHI_1^2(n)$ ; see Table 1. Since for  $n = 60$  the p-value is approximately the significance level .05, we think, the distribution of  $CHI_1^2(60)$  is well approximated by the  $\chi^2$  distribution with one degree of freedom.

Next, the former test is performed for D2 distribution and 9 values of  $n$  are considered. The simulated values of  $CHI_2^2(n)$ , along with the corresponding p-values for each  $n$ , are shown in Table 2.

**Table 2.** Percentages of observations of  $CHI_2^2(n)$  for the indicated value of  $n$  and p-values of the corresponding  $\chi^2$  test ( $\beta = 2$ ,  $\gamma = 1$  and  $T = 10$ ).

$n$	$[t_0, t_1[$	$[t_1, t_2[$	$[t_2, t_3[$	$[t_3, t_4[$	$[t_4, t_5[$	$[t_5, t_6[$	$[t_6, t_7[$	$[t_7, t_8[$	$[t_8, t_9[$	$[t_9, t_{10}[$	p-value
50	9.7	0.0	20.5	0.0	21.3	0.0	14.9	10.3	8.9	14.4	0.000
80	9.0	17.5	0.0	13.8	15.3	0.0	13.9	9.1	12.7	8.7	0.000
100	7.0	14.5	0.0	15.7	15.3	10.8	9.8	6.7	11.0	9.2	0.01612624
110	6.6	14.0	14.5	0.0	13.0	12.1	9.1	7.6	12.4	10.7	0.04275268
115	15.9	0.0	13.8	11.5	10.7	11.3	9.4	7.1	9.8	10.5	0.06137558
120	8.4	14.8	11.7	0.0	12.5	11.5	8.0	13.9	10.1	9.1	0.07337148
150	6.0	13.4	10.5	11.4	11.6	10.9	8.2	5.6	13.1	9.3	0.68034140
200	5.0	9.5	10.2	11.9	10.8	10.8	13.6	9.7	7.4	11.1	0.82372460
500	10.6	7.9	13.7	6.6	11.3	11.2	8.1	11.3	9.1	10.2	0.91595900

It is obtained for  $n$  between 110 and 115 the p-value of the corresponding  $\chi^2$  test is approximately the significance level .05. Hence, for these values of  $n$ , it is reasonable to assume  $CHI_2^2(60)$  has an approximated  $\chi^2$  distribution with one degree of freedom.

Although, under D1,  $x^n$  has more variability than under D2, the conducted simulation shows that the global error, under D1, attains approximate normality for lower values on  $n$  than under D2.

## 6. Conclusions

A family of discrete-time stochastic processes is presented and it is proven that these processes can be approximate by means of the solution to an ODE. Conversely, these processes may be seen as schemes of approximation for this solution. For this reason,

a stochastic version of the global error associated to these schemes are defined and its asymptotic distribution is studied. The uniform convergence in probability, on compact subsets of the positive real numbers, is proven and a central limit theorem for the fluctuations of the stochastic processes is derived. This fact allows us to find confidence bands, where with a preassigned probability the trajectories of the stochastic processes are bounded by these bands. Our results are illustrated by an emblematic model coming from the mathematical literature. Indeed, two discrete time stochastic processes are approximated by the solution of the differential equation corresponding to the SIS epidemic model. Simulations of their trajectories are carried out and compared with the solution of the SIS deterministic model. Moreover,  $\chi^2$  tests are carried out to evaluate the goodness of the discretization, in order to obtain approximate normality for the global error.

### Acknowledgments

The author wish to thank two anonymous referees for their helpful comments that aided for improving this article. This work was partially supported by project number 124.733/2012 of Dirección de Investigación e Innovación de la Pontificia Universidad Católica de Valparaíso and by FONDECYT project number 1120879.

### References

- [1] S. Abbasbandy, C. Bervillier, Analytic continuation of Taylor series and the boundary value problems of some nonlinear ordinary differential equations, *Applied Mathematics and Computation* vol. 218 (no. 5) (2011) 2178–2199.
- [2] P. Billingsley, *Convergence of Probability Measures*, Wiley & Sons, New York, 1968.
- [3] R.W.R Darling, J.R. Norris, Differential equation approximations for Markov Chain, *Probability Surveys* Vol. 5 (2008) 37-79.
- [4] M. D. M.R. Eslahchi, S. Ahmadi-Asl, The general Jacobi matrix method for solving some nonlinear ordinary differential equations, *Applied Mathematical Modelling* (In Press) (2012) 1–9.
- [5] S. K. K. Parand, S. Abbasbandy, J. Rad, A novel application of radial basis functions for solving a model of first-order integro-ordinary differential equation, *Communications in Nonlinear Science and Numerical Simulation* Vol. 16 (No. 11) (2012) 4250–4258.
- [6] R. Fierro, S. Torres, A stochastic scheme of approximation for ordinary differential equations, *Electronic Communications in Probability* Vol. 13 (2008) 1–9.
- [7] R. Fierro, A class of stochastic epidemic models and its deterministic counterpart, *Journal of the Korean Statistical Society* Vol. 39 (No. 4) (2010) 397–407.
- [8] P. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, New York, 1995.
- [9] T.G. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes, *Journal of Applied Probability*, Vol. 7 (No.1) (1970) 49-58.
- [10] T.G. Kurtz, Limit theorems for sequences of jump Markov processes approximating ordinary differential processes, *Journal of Applied Probability*, Vol. 8 (No.2) (1971) 344-356.
- [11] T.G. Kurtz, *Approximation of Population Processes*, Series in Applied Mathematics 36. Philadelphia: SIAM, 1981.
- [12] H. Kushner, P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer, New York, 1991.
- [13] E. Lenglart, Relation de domination entre deux processus, *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques* Vol. 13 (No. 2) (1977) 171–179.
- [14] R. Rebolledo, La méthode des martingales appliquée à l'étude de la convergence en loi de processus, *Mémoires de la Société Mathématique de France* Vol. 62 (No. 4) (1979) 1–125.