Estimating Stochastic Dynamical Systems Driven by a Continuous-Time Jump Markov Process

Julien Chiquet, Nikolaos Limnios

Methodol Comput Appl Probab (2006) 8: 431–447
DOI 10.1007/s11009-006-0423-z

Estimating Stochastic Dynamical Systems Driven by a Continuous-Time Jump Markov Process

Julien Chiquet · Nikolaos Limnios

Received: 25 July 2006 / Revised: 25 July 2006 / Accepted: 4 August 2006

Abstract We discuss the use of a continuous-time jump Markov process as the driving process in stochastic differential systems. Results are given on the estimation of the infinitesimal generator of the jump Markov process, when considering sample paths on random time intervals. These results are then applied within the framework of stochastic dynamical systems modeling and estimation. Numerical examples are given to illustrate both consistency and asymptotic normality of the estimator of the infinitesimal generator of the driving process. We apply these results to fatigue crack growth modeling as an example of a complex dynamical system, with applications to reliability analysis.

Keywords Stochastic dynamical system · Markov process · Estimation · Fatigue crack growth

AMS 2000 Subject Classification 60H10 · 60K40 · 62M05

1 Introduction

A general stochastic differential system may be written as follows:

$$\frac{\mathrm{d}Z_t}{\mathrm{d}t} = f(Z_t, X_t), \tag{1}$$

J. Chiquet (✉) · N. Limnios
Université de Technologie de Compiègne, Centre de Recherche de Royallieu, LMAC, BP 20529, 60205 Compiègne cedex, France
e-mail: julien.chiquet@utc.fr

N. Limnios
e-mail: nikolaos.limnios@utc.fr

J. Chiquet
Commissariat à l'Énergie Atomique, Centre de Recherche de Saclay, DM2S/SERMA/LCA, 91191 Gif sur Yvette cedex, France

where Z_t is a process describing the dynamical system to be modeled, X_t a stochastic process and f a given function with the appropriate properties, assuring both existence and uniqueness of a solution. We refer to the stochastic process X_t as the *driving process*. Such an interpretation of Eq. 1 comes from control theory, when observing an evolution Z_t to be controlled by a random process X_t. Yet whereas control theory looks toward optimizing a functional of X_t, the study of stochastic differential systems such as Eq. 1 aims at modeling dynamical systems with random components.

This formulation is quite general, thus an appropriate function f must be given to be consistent with a specific physical phenomenon. This function may be built as follows: first, a deterministic part describes the general behavior of the system; second, because the system is not completely known or has an intrinsically stochastic behavior, a stochastic part is needed to take the random component into account. This is the aim of the process X_t, which disturbs the deterministic description of the phenomenon. The nature of the driving process is consequently of major interest, because it conditions the general behavior of the system (1).

When X_t is a very irregular noisy process, it is approximated by the Wiener process. Thus, we are in the framework of Stochastic Differential Equations which constitutes the basic methodology for modeling dynamical systems subject to parametric random noises. Hence, with the appropriate initial conditions, the system (1) can be reformulated as an Itô stochastic differential equation, whose solution is a diffusion process (see, e.g., Øksendal, 2003; Sobczyk, 1991, for theory and applications). Estimation problems in stochastic differential equations are discussed by, e.g., Jankunas and Khasminskii (1997), Kutoyants (2004).

In this paper, we use a quite different approach, modeling the driving process by a continuous-time, finite state space Markov process. The first contribution goes back to Goldstein (1951) : he introduced a two-state Markov chain as a driving process, in an effort to give a stochastic representation of the telegraph equation. Then, the more general case of a continuous time finite state space Markov process as the driving process has been studied. In this way, the system (1) could be interpreted as a model for the movement of particles, with X_t being the velocity of the system. Hence, the couple markovian process (Z_t, X_t) is called a "transport process" and has some interesting probabilistic properties which can be found in the literature (see, e.g., Dautray, 1989; Lapeyre and Pardoux, 1998). In control theory, these systems are known as "Piecewise deterministic Markov processes" (see Davis, 1993). We suggest here to use this kind of approach to model a wide class of dynamical systems.

Once the stochastic formulation of a dynamical system has been given through a random evolution such as Eq. 1, the fundamental key point of the estimation remains.

As a matter of fact, the observed data which will be exploited to achieve the estimation of the different parameters of the system only concerns the process Z_t. Consequently, some methods are needed to estimate both the function f and the Markov driving process X_t from sample paths of Z_t. Concerning the function f, it is common to make an assumption on its general form, depending on the physical phenomenon to be modeled. Then, the parameters of f are estimated with results and methodologies from data analysis (for an example of data analysis of a stochastic dynamical system in structural mechanics, see Spencer et al., 1989). In this paper, the main interest is the estimation of the Markov driving process X_t from trajectories of the observed process Z_t.

We try to be as general as possible about the estimation, in order to obtain results available for a wide class of applications. Thus, we consider that we observe K sample paths of Z_t defined on random time intervals $[0, T_k]_{k=1\dots K}$, assuming that the random variables (**r.v.**) T_k are independent and identically distributed (**i.i.d.**). This is an extension of the case when sample paths are defined on the same fixed-time interval $[0, T]$, with constant $T > 0$. Such data could come from a stochastic system whose sample paths are observed until failure. In this case, the trajectories are i.i.d., yet not of the same size.

Provided that sample paths of Z_t and X_t have the same duration, one of the major task of the paper is the estimation of the Markov driving process X_t on randomly censored data. In fact, we observe only the trajectories of the process Z_t, while the process X_t remains unobservable. Statistical inferences for a Markov process have been discussed by Billingsley (1961) and Albert (1962). More precisely, Albert gave fundamental results on the estimation of the infinitesimal generator of X_t and the associated asymptotic properties when observing sample paths on the same fixed-time interval $[0, T]$. We will look for the same results on random time intervals. Once the estimation of the driving process is available, we can return to the estimation of the whole stochastic dynamical system.

Our main motivation for the present study is stochastic fatigue crack growth modeling, an engineering problem widely studied in the last three decades. To the authors' knowledge, there is no work modeling fatigue crack growth with dynamical system driven by a jump Markov process.

The paper is organized as follows: in Section 2, we briefly present the well-known maximum-likelihood estimator (**MLE**) of the infinitesimal generator of a continuous-time finite state space Markov process and its asymptotic properties, when sample paths are of the same length. In Section 3, we extend these results for random censored data. In Section 4, we use the results of Section 3 within the framework of dynamical systems modeling: we study a stochastic differential system driven by a jump Markov process, with given initial conditions. In Section 5, we present some numerical applications to the results of Sections 3 and 4: first, the good behavior of the estimator with random censored data is illustrated; next, we introduced a numerical example in the framework of fatigue crack growth modeling. We give applications to reliability calculation and estimation.

2 Preliminaries

In this section, we introduce general notation and we remind some results on the MLE of the infinitesimal generator of a Markov process.

Let $(X_t, t \geq 0)$ be a time-homogeneous, continuous-time Markov process with finite state space $E = \{1, 2, \dots, d\}$. The initial distribution of X_t is denoted by $\alpha = (\alpha(i))_{i \in E}$, where $\alpha(i) = \mathbb{P}[X_0 = i]$. We denote by $\mathbf{A} = (a_{ij})_{i,j \in E}$ its infinitesimal generator, which is stable, that is

$$\begin{aligned} a_{ij} &\geq 0, && \text{for all } i \neq j, \\ \text{and } a_{ii} &= -a_i && = -\textstyle\sum_{k \in E, i \neq k} a_{ik}. \end{aligned}$$

Suppose we observe K sample paths of X_t, censored by the i.i.d. r.v. $(T_k)_{k=1\ldots K}$. The T_k are independent of X_t with common distribution function F on $\mathbb{R}_+$. We denote $\left(X_t^k, 0 \le t \le T_k\right)_{k=1,\ldots,K}$ the censored sample paths of X_t, which are independent.

In the particular case where $T_k = T$ for all $k = 1, \ldots, K$, with T a strictly positive constant, we are in the framework of Albert (1962). It means the distribution function F of T_k is $F(t) = \mathbb{1}_{\{T \le t\}}$. We introduce the following notation:

- $N_{ij}(K)$ is the total number of transitions from state i to state j on the K trials;
- $V_i(K)$ is the total length of time that state i is occupied on the K trials.

Albert (1962) performed the maximum-likelihood estimation of $\mathbf{A}$ and proved both consistency and asymptotic normality when considering sample paths defined on the same fixed-time interval $[0, T]$:

Theorem 1 (Albert, 1962) *Suppose we have K independent sample paths of a jump Markov process $(X_t, t \ge 0)$ on an interval $[0, T]$ with $T > 0$.*

(a) *The maximum-likelihood estimator $\widehat{a}_{ij}(K)$ of the infinitesimal generator of the jump Markov process estimated on the K sample paths is given by*

$$\widehat{a}_{ij}(K) = \begin{cases} \dfrac{N_{ij}(K)}{V_i(K)} & \text{if } i \ne j \text{ and } V_i(K) \ne 0, \\ -\dfrac{\sum_{l \in E \setminus \{i\}} N_{il}(K)}{V_i(K)} & \text{if } i = j \text{ and } V_i(K) \ne 0, \\ 0 & \text{if } V_i(K) = 0. \end{cases} \tag{2}$$

(b) *The estimator $\widehat{a}_{ij}(K)$ is strongly consistent as $K \to \infty$, i.e. $\widehat{a}_{ij}(K) \xrightarrow[K\to\infty]{a.s.} a_{ij}$.*

(c) *If $\mathbb{P}[X_t = i] > 0$, for all $i \in E$, the set of r.v. $\left\{\sqrt{K}\left(\widehat{a}_{ij}(K) - a_{ij}\right)\right\}_{i \ne j}$ are asymptotically normal as $K \to \infty$ with zero mean and covariances given by*

$$\sigma^2(i, j; p, q) = D(i, j; p, q) a_{ij} \Big/ \int_0^T \mathbb{P}\left[X_t = i\right] \mathrm{d}t\,, \tag{3}$$

where

$$D(i, j; p, q) = \begin{cases} 1 & \text{if} \quad (i, j) = (p, q), \quad \text{for all } (i, j), (p, q) \in E^2 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

3 Estimating the Markov Driving Process

In this section, we give results on the estimation of the generator when considering random censored data, that is when the T_k are positive r.v. Let us introduce the following additional notation:

- $N_{ij}^k(T_k)$, the total number of transitions from state i to j on the sample path $\left(X_t^k, 0 \le t \le T_k\right)$;
- $V_i^k(T_k)$, the total length of time that state i is occupied on the sample path $\left(X_t^k, 0 \le t \le T_k\right)$.

We naturally have

$$\sum_{k=1}^{K} N_{ij}^{k}(T_k) = N_{ij}(K),$$

$$\text{and} \sum_{k=1}^{K} V_{i}^{k}(T_k) = V_{i}(K). \tag{5}$$

We consider $\widehat{a}_{ij}(K)$ given by Eq. 2 from Theorem 1, where $N_{ij}(K)$ and $V_i(K)$ are defined in expression (5), that is:

$$\widehat{a}_{ij}(K) := \frac{N_{ij}(K)}{V_i(K)} = \frac{\sum_{k=1}^{K} N_{ij}^{k}(T_k)}{\sum_{k=1}^{K} V_{i}^{k}(T_k)}. \tag{6}$$

For any censored time distribution F, part (b) and (c) of Theorem 1 do not apply since the K sample paths are defined on random time intervals. Thence, we will prove that both consistency and asymptotic normality remain true for the extended definition (6) of $\widehat{a}_{ij}(K)$.

Proposition 1 (Consistency) Estimator (6) is strongly consistent when considering K sample paths on random-time intervals $[0, T_k]$, $k = 1, \ldots, K$, that is $\widehat{a}_{ij}(K) \xrightarrow[K\to\infty]{a.s.} a_{ij}$.

Proof Expression (6) of $\widehat{a}_{ij}$ can be written

$$\widehat{a}_{ij}(K) = \frac{1}{K}\sum_{k=1}^{K} N_{ij}^{k}(T_k) \Big/ \frac{1}{K}\sum_{k=1}^{K} V_{i}^{k}(T_k).$$

Since $N_{ij}^{k}(T_k)$, respectively $V_{i}^{k}(T_k)$, are i.i.d., the Strong Law of Large Numbers gives

$$\widehat{a}_{ij}(K) \xrightarrow[K\to\infty]{a.s.} \frac{\mathbb{E}N_{ij}^{1}(T_1)}{\mathbb{E}V_{i}^{1}(T_1)},$$

where $\mathbb{E}$ is the expectation relative to the sample paths of X_t defined on random time intervals. Hence, conditioning on the r.v. T, we may write

$$\frac{\mathbb{E}N_{ij}^{1}(T_1)}{\mathbb{E}V_{i}^{1}(T_1)} = \int_0^\infty \mathbb{E}N_{ij}^{1}(s)\mathrm{d}F(s) \Big/ \int_0^\infty \mathbb{E}V_{i}^{1}(s)\mathrm{d}F(s),$$

where F is the distribution function of T_1. According to Theorem 5.1 from Albert (1962), the following relations holds for all $s > 0$:

$$\mathbb{E}N_{ij}^{1}(s) = a_{ij}\int_0^s \mathbb{P}\big[X(t) = i\big]\mathrm{d}t,$$

$$\mathbb{E}V_{i}^{1}(s) = \int_0^s \mathbb{P}\big[X(t) = i\big]\mathrm{d}t.$$

Finally

$$\frac{\mathbb{E}N_{ij}^{1}(T_1)}{\mathbb{E}V_{i}^{1}(T_1)} = a_{ij}$$

which proves the proposition. □

Proposition 2 (Asymptotic normality) Let estimator (6) be given. The set of r. v. $\left\{\sqrt{K}\left(\widehat{a}_{ij}(K)-a_{ij}\right)\right\}_{i\neq j}$ are asymptotically normal as $K\to\infty$, provided that $\mathbb{P}[X_t=i]>0$, with zero mean and covariances

$$\sigma^2(i,\,j;\,p,q)=D(i,\,j;\,p,q)a_{ij}\Big/\int_0^\infty\int_0^s\mathbb{P}\big[X_t=i\big]\mathrm{d}t\mathrm{d}F(s)\,, \tag{7}$$

with $D(i,\,j;\,p,q)$ defined in Eq. 4.

Proof By definition of the estimator (6), we have for any $i,\,j\in E$, with $i\neq j$,

$$\begin{aligned}\sqrt{K}\left(\widehat{a}_{ij}(K)-a_{ij}\right)&=\sqrt{K}\left(\frac{N_{ij}(K)}{V_i(K)}-a_{ij}\right)\\&=\frac{K}{\sum_{k=1}^K V_i^k(T_k)}\left(\frac{N_{ij}(K)-a_{ij}V_i(K)}{\sqrt{K}}\right).\end{aligned}$$

Applying Slutsky's theorem, and since V_i^k are i.i.d., the r.v.'s $\left\{\sqrt{K}\left(\widehat{a}_{ij}(K)-a_{ij}\right)\right\}_{i\neq j}$ have the same asymptotic distribution as

$$\left\{\frac{1}{\mathbb{E}V_i^1(T_1)}\left(\frac{N_{ij}(K)-a_{ij}V_i(K)}{\sqrt{K}}\right)\right\}_{i\neq j}.$$

Then, from the Central Limit Theorem, this random vector asymptotically follows a normal distribution with zero mean and the elements of its covariance matrix given by

$$\sigma^2(i,\,j;\,p,q)=\frac{\mathbb{E}\left[\left(N_{ij}^1(T_1)-a_{ij}V_i^1(T_1)\right)\left(N_{pq}^1(T_1)-a_{pq}V_p^1(T_1)\right)\right]}{\mathbb{E}V_i^1(T_1)\mathbb{E}V_p^1(T_1)}. \tag{8}$$

Again, from Theorem 5.1 of Albert (1962), we have for all $s>0$,

$$\mathbb{E}\left[\left(N_{ij}^1(s)-a_{ij}V_i^1(s)\right)\left(N_{pq}^1(s)-a_{pq}V_p^1(s)\right)\right]=D(i,\,j;\,p,q)a_{ij}\int_0^s\mathbb{P}[X_t=i]\mathrm{d}t.$$

We should now write the numerator of Eq. 8 conditioning on T:

$$\begin{aligned}&\mathbb{E}\left[\left(N_{ij}^1(T_1)-a_{ij}V_i^1(T_1)\right)\left(N_{pq}^1(T_1)-a_{pq}V_p^1(T_1)\right)\right]\\&\quad=\mathbb{E}\int_0^\infty\left[\left(N_{ij}^1(s)-a_{ij}V_i^1(s)\right)\left(N_{pq}^1(s)-a_{pq}V_p^1(s)\right)\right]\mathrm{d}F(s)\\&\quad=\int_0^\infty\mathbb{E}\left[\left(N_{ij}^1(s)-a_{ij}V_i^1(s)\right)\left(N_{pq}^1(s)-a_{pq}V_p^1(s)\right)\right]\mathrm{d}F(s)\\&\quad=D(i,\,j;\,p,q)a_{ij}\int_0^\infty\int_0^s\mathbb{P}\big[X_t=i\big]\mathrm{d}t\mathrm{d}F(s).\end{aligned}$$

where we used the independence between T and X_t to exchange the order of integration. Moreover, remembering

$$\mathbb{E}V_i^1(T_1)=\int_0^\infty\int_0^s\mathbb{P}\big[X_t=i\big]\mathrm{d}t\mathrm{d}F(s),\ \text{ for all }\ i\in E,$$

we get expression (7) of the covariances. □

Remark 1 When all the T_k equal the same constant $T > 0$, then $F(t) = \mathbb{1}_{\{T \le t\}}$ and we have

$$\sigma^2(i, j; p, q) = D(i, j; p, q) a_{ij} \Big/ \int_0^\infty \int_0^s \mathbb{P}\big[X_t = i\big] \mathrm{d}t \mathrm{d}F(s) ,$$
$$= D(i, j; p, q) a_{ij} \Big/ \int_0^T \mathbb{P}\big[X_t = i\big] \mathrm{d}t ,$$

which is the case of Albert's estimation on fixed-time interval $[0, T]$.

4 Dynamical System Estimation

We present here the general formulation of a stochastic dynamical system driven by a jump Markov process. First, we manage to give an analytical form of the system's reliability function. Next, a method for the estimation of the system is proposed.

4.1 Preliminaries

Let us consider a process $(Z_t, t \ge 0)$ describing a stochastic dynamical system on $[0, L]$, with L a strictly positive constant. The driving process is a jump Markov process $(X_t, t \ge 0)$ with finite state space E, infinitesimal generator $\mathbf{A}$ and initial distribution α. The process Z_t is governed by

$$\begin{cases} \dfrac{\mathrm{d}Z_t}{\mathrm{d}t} &= f(Z_t, X_t), \;\; f : \big[0, L\big] \times E \longrightarrow \mathbb{R}_+, \\ Z_0 &= z_0, \end{cases} \tag{9}$$

where the function f is assumed to be known with the following assumptions:

A 1. *The function $f : [0, L] \times E \longrightarrow \mathbb{R}_+$ is positive and measurable;*
A 2. *The process $(X_t, t \ge 0)$ is a finite state space irreducible Markov process with stationary distribution $\pi = (\pi_i, i \in E)$;*
A 3. *There is a function $k : E \longrightarrow \mathbb{R}_+$ such that, for $z, z' \in [0, L]$ and $x \in E$,*

$$\left| f(z, x) - f(z', x) \right| \le k(x) \left| z - z' \right| .$$

We are interested in studying system (9) when the process Z_t evolves over $[0, L]$ and finally reaches a certain point denoted $\{\Delta\}$, where it stays indefinitely. Point $\{\Delta\}$ is an *absorbing point*. Using classical notation of reliability analysis, we employ the following two-set partition for the state space of Z_t: a set of working states $U = [0, \Delta[$ and a set of down states $D = [\Delta, L]$, with $0 < \Delta \le L$.

The initial random condition Z_0 belongs to a subset $[0, C] \subset [0, \Delta[$, with $0 < C < \Delta$. The subset $[0, C]$ is introduced to avoid having $\sup_K Z_0^k = \{\Delta\}$, otherwise, it would mean the system could start in the subset D. The distribution of the r.v. Z_0 is denoted by β, that is $\beta(B) = \mathbb{P}[Z_0 \in B]$, with B a Borel subset of $[0, C]$.

We also define the hitting time τ of D by $\tau = \inf\{t \ge 0 : Z_t \in D\}$. Because $D = [\Delta, L]$ with $\{\Delta\}$ an absorbing point, the time τ when the system reaches point $\{\Delta\}$ is also the failure time.

In the following, we suppose we have K independent copies of Z_t, denoted by $\left(Z_t^k, 0 \leq t \leq T_k\right)$, for $k = 1, \ldots, K$, defined on random time intervals $[0, T_k]$. The T_k are K independent copies of τ.

The coupled process (Z_t, X_t) is markovian with well-known theoretical framework (see, e.g., Davis, 1993; Lapeyre and Pardoux, 1998). It allows us to give some analytical results for (Z_t, X_t) and for the reliability analysis of the system (9). The generator $\mathbf{B}$ of the Markov process $(Z_t, X_t, t \geq 0)$ is given by

$$\mathbf{B}g(z,i) = f(z,i)\frac{\partial}{\partial z}g(z,i) + \sum_{j \in E} \mathbf{A}(i,j)g(z,j), \tag{10}$$

with $i, j \in E$, $z \in [0, L]$ and g a continuous and differentiable function on the first argument.

The initial distribution of the couple (Z_t, X_t) is denoted

$$\mu(B,i) = \mathbb{P}\left[Z_0 \in B, X_0 = i\right], \text{ with } B \text{ a Borel subset of } [0,C] \text{ and } i \in E. \tag{11}$$

An alternative for the definition of the failure time τ is the following:

$$\tau = \inf\left\{t \geq 0 : (Z_t, X_t) \in D \times E\right\}. \tag{12}$$

It is much more convenient to describe the failure time τ as a function of the couple process (Z_t, X_t) rather than of Z_t only. As matter of fact, we can note the direct link between the distribution function F of τ and the reliability function R of the couple (Z_t, X_t), defined by

$$R(t) = \mathbb{P}\left[(Z_s, X_s) \in U \times E, \forall s \leq t\right]. \tag{13}$$

Introducing $P_{ij}(t; z, B)$ for the transition probability function of the Markov process (Z_t, X_t), that is

$$P_{ij}(t; z; B) = \mathbb{P}\left(Z_t \in B, X_t = j | Z_0 = z, X_0 = i\right), \tag{14}$$

the expression of F becomes

$$F(t) = 1 - R(t) = 1 - \sum_{i,j \in E} \int_{[0,C]} \mu(\mathrm{d}z, x) P_{ij}(t; z, U). \tag{15}$$

When the infinitesimal generator $\mathbf{B}$ is bounded, the transition probability function $P(t)$ is given through the Kolmogorov equations, and $P(t) = \exp t\mathbf{B}$.

4.2 Estimating the Generator of the Markov Driving Process

We remind that system (9) aims at modeling a dynamical system described by a stochastic process Z_t. Consequently, it is justified to assume that we have K independent sample paths $\left(Z_t^k, 0 \leq t \leq T_k\right)$ of the process Z_t. These sample paths could be obtained from laboratory measurements on the real dynamical system to be modeled.

Nevertheless, we also need some sample paths corresponding to the driving process X_t. As a matter of fact, we need some trajectories $\left(X_t^k, 0 \leq t \leq T_k\right)$ to the estimate generator $\mathbf{A}$ using results from Section 2. It is obviously not possible to observe directly trajectories of the driving process from measurements. Hence, we need to "extract" an estimation of X_t^k from the available measurements, that is from the Z_t^k. We propose the following way.

We denote the derivative by $\dot{Z}_t = \mathrm{d}Z_t/\mathrm{d}t$ and $\left(\dot{Z}_t^k\right)_{k=1\ldots K}$ the associated samples. From system (9), the following relationship holds between Z_t^k, $\dot{Z}_t^k$ and X_t^k:

$$\dot{Z}_t^k = f\left(Z_t^k, X_t^k\right). \tag{16}$$

We only observe the Z_t^k. Yet, there are a lot of methods which give an estimation of the derivatives $\dot{Z}_t^k$ from sample paths Z_t^k, e.g., the simple secant method (see Virkler et al., 1979, for a complete analysis of the different methods available in the framework of fatigue crack growth). If we denote by Δt the time discretization step of the data set, the estimator of $\dot{Z}_t^k$ is

$$\widehat{\dot{Z}}_t^k = \frac{Z_{t+\Delta t/2}^k - Z_{t-\Delta t/2}^k}{\Delta t}. \tag{17}$$

Moreover, we must assume there exists a function h from expression (16) giving X_t^k as a function of Z_t^k and $\widehat{\dot{Z}}_t^k$. Such a function not always exists, yet when the stochastic process X_t is a linear additive or multiplicative term in the function f, we may easily find the corresponding function h. Hence, we may estimate the X_t^k with the following relationship:

$$\widehat{X}_t^k = h\left(Z_t^k, \widehat{\dot{Z}}_t^k\right). \tag{18}$$

In the present context, it is assumed that the states of the process X_t are known, thus the paths $\widehat{X}_t^k$ are filtered according to this known state space. If the state space is not known, which is the most common case, some clustering methods can be used to fit as well as possible with the estimated sample paths of X_t. In Chiquet et al. (2006), we apply the classical K-means algorithm to estimate the state space of X_t.

With the estimated paths $\widehat{X}_t^k$, which are defined on the same random time intervals as Z_t^k, that is $[0, T_k]$, $k = 1, \ldots, K$, we may estimate the generator $\mathbf{A}$ of the jump Markov process X_t with expression (6). Note that it is not the optimal estimator for $\mathbf{A}$ in the particular context defined by the dynamical system (1), because the censoring time of the paths is also the failure time τ, which obviously depends on X_t from Eq. 15. Yet, we will see that it still give some good results in the following section, by investigating some numerical validations.

5 Numerical Applications

In this section, we first aim at giving a numerical application for Markov processes estimation, strictly corresponding to Propositions 1 and 2 of Section 2. We illustrate both the estimator consistency and its asymptotic normality. Next, a complete numerical example for fatigue crack growth is given, when modeling crack propagation by a stochastic system driven by a jump Markov process, thus exploiting the first part of the numerical applications.

5.1 Markov Process with Random Censoring

We shall consider a given state space E, a generator $\mathbf{A}$ and an initial distribution α for the Markov jump process X_t, in order to simulate sample paths. We have chosen a uniform initial distribution, the 3-state space $E = \{1, 2, 3\}$ and the following generator:

$$\mathbf{A} = \begin{bmatrix} -0.02 & 0.02 & 0 \\ 0.027 & -0.03 & 0.003 \\ 0.01 & 0 & -0.01 \end{bmatrix}. \tag{19}$$

The sample paths are censored at times $(T_k)_{k=1,\ldots,K}$, which follow an exponential law with parameter λ, i.e. $T_k \sim \mathcal{E}(\lambda)$ for all $k = 1, \ldots, K$ (e.g. $\lambda = 0.001$). It is then easy to simulate K sample paths of X_t to estimate $\mathbf{A}$ and look at the consistency and asymptotic normality of estimator $\widehat{\mathbf{A}}$ as a function of K, the number of trajectories.

From Proposition 2, the r.v.'s $\sqrt{K}\left(\widehat{a}_{ij}(K) - a_{ij}\right)$ for $i \neq j$ are asymptotically normal as $K \to \infty$, with zero mean and covariances given by Eq. 7. Then we have

$$\mathbb{P}\left[-u_{1-\gamma/2} \leq \frac{\sqrt{K}\left(\widehat{a}_{ij}(K) - a_{ij}\right)}{\sigma(i, j; p, q)} \leq u_{1-\gamma/2}\right] = 1 - \gamma,$$

where u_γ is the $\gamma -$ *percentile* of the centered and reduced normal r.v. We denote by $\widehat{\sigma}^2$ a consistent estimator of the covariance (see, e.g. Sadek and Limnios, 2005), that is verifying $\widehat{\sigma}^2(K) \xrightarrow[K\to\infty]{a.s.} \sigma^2$. It means that the following confidence interval holds asymptotically for a_{ij}:

$$\left[\widehat{a}_{ij}(K) - u_{1-\gamma/2}\frac{\widehat{\sigma}(i, j; p, q; K)}{\sqrt{K}}; \widehat{a}_{ij}(K) + u_{1-\gamma/2}\frac{\widehat{\sigma}(i, j; p, q; K)}{\sqrt{K}}\right]. \tag{20}$$

We can calculate the covariance $\sigma^2(i, j; p, q)$ using

$$\begin{aligned}\sigma^2(i, j; p, q) &= D(i, j; p, q)a_{ij} \int_0^\infty \int_0^s \mathbb{P}\left[X_t = i\right] \mathrm{d}t \mathrm{d}F(s) \\ &= D(i, j; p, q)a_{ij} \int_0^\infty \int_0^s \alpha \exp(t\mathbf{A})\mathbf{e}_i \lambda \exp(-\lambda s) \mathrm{d}t\mathrm{d}s,\end{aligned}$$

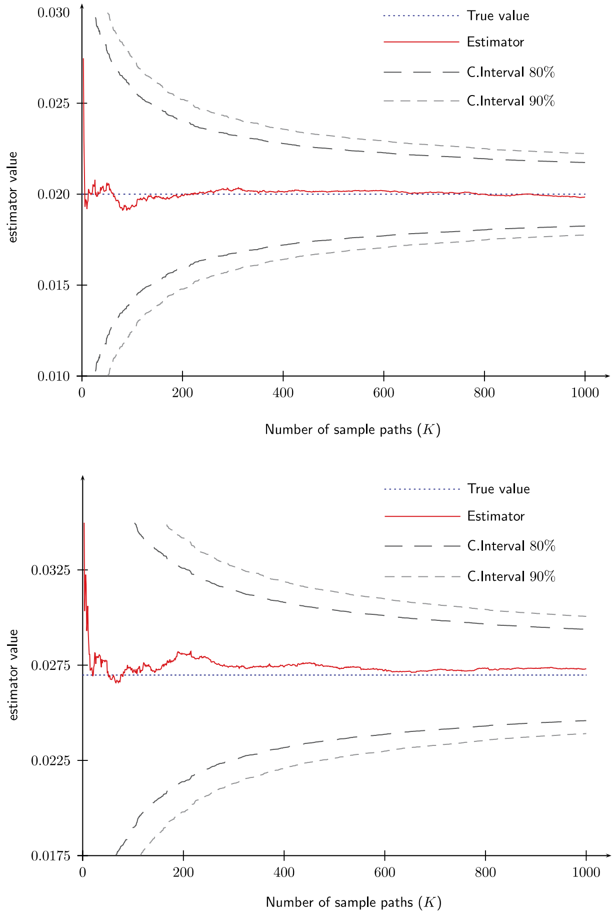
where $\mathbf{e}_i$ is the i^{th} vector of the canonical basis of $\mathbb{R}^d$, with d the number of states of E, that is $d = 3$ in the present case. Sadek and Limnios (2005) gave results on the estimation of reliability measurements for continuous-time Markov processes. Henceforth, they allow us to give the following estimator for the covariance:

$$\widehat{\sigma}^2(i, j; p, q; K) = D(i, j; p, q)\widehat{a}_{ij}(K) \int_0^\infty \int_0^s \alpha \exp\left(t\widehat{\mathbf{A}}\right)\mathbf{e}_i \widehat{\lambda}(K) \exp\left(-\widehat{\lambda}(K)s\right) \mathrm{d}t\mathrm{d}s, \tag{21}$$

where $\widehat{\lambda}(K)$ is the MLE of λ, given by

$$\widehat{\lambda}(K) = \frac{1}{\bar{T}}, \text{ with } \bar{T} = \frac{1}{K}\sum_{k=1}^{K} T_k. \tag{22}$$

Fig. 1 Asymptotic normality and consistency for elements a_{12} and a_{21} of the matrix generator



With the above statistical results, we can represent the r.v. $\left(\widehat{a}_{ij}(K)\right)_{i\neq j}$ as a function of the number K of sample paths used for the estimation and the corresponding confidence interval (20). Figure 1 represents this r.v. for, respectively, a_{12} and a_{21}. Results for a_{23} and a_{31} are of the same kind.

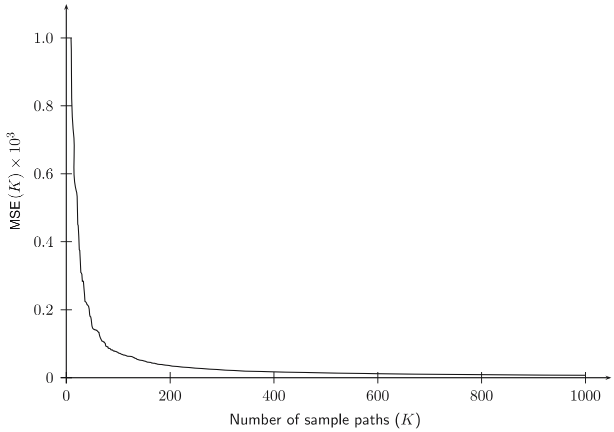
Another way to evaluate the consistency of the estimator $\widehat{a}_{ij}(K)$ is to observe a given error function between $\widehat{a}_{ij}(K)$ and a_{ij}. We suggest to estimate the classical Mean Squared Error (**MSE**: see, e.g. Härdle, 1990), adapted to the present context, that is

$$\mathrm{MSE}(K) = \sum_{\substack{i,j\in E\\ i\neq j}} \mathbb{E}\left[\left(\widehat{a}_{ij}(K) - a_{ij}\right)^2\right]. \tag{23}$$

We can again take advantage of the asymptotic normality of the $\left\{\sqrt{K}\left(\widehat{a}_{ij}(K) - a_{ij}\right)\right\}_{i\neq j}$. We denote $Y(i,j) = \sqrt{K}\left(\widehat{a}_{ij}(K) - a_{ij}\right)$ a r.v. with a cumulative distribution function G_{ij}. Thus,

$$\mathrm{MSE}(K) = \sum_{\substack{i,j\in E\\ i\neq j}} \frac{1}{K}\int_{\mathbb{R}_+} y^2 \mathrm{d}G_{ij}(y). \tag{24}$$

Fig. 2 Estimated Mean Squared Error



The distribution functions G_{ij} can be estimated by their asymptotic value, i.e. by the normal distribution with zero mean and covariance σ^2 whose estimator is given by Eq. 21. Moreover, noticing that $\int_{\mathbb{R}_+} y^2 \mathrm{d}G_{ij}$ is nothing but the second order moment of the normal r. v. Y, we suggest the following estimator for the MSE:

$$\widehat{\mathrm{MSE}}(K) = \frac{1}{K} \sum_{\substack{i,j \in E \\ i \neq j}} \widehat{\sigma}^2(i, j; p, q; K). \tag{25}$$

The results are illustrated on Fig. 2.

From Figs. 1 and 2, not only can we see that the estimator is consistent and asymptotically normal, but also it quickly reaches an acceptable value, for a mean censored-time which is reasonable ($1/\lambda = 1000$). This property is interesting for applications where laboratory measurements are expensive, i.e. when the number of sample paths is small.

5.2 Estimating a Dynamical System: Fatigue Crack Growth Analysis

Let us now give a numerical example in the framework of a process Z_t describing a dynamical system such as Eq. 9.

This application is motivated by stochastic fatigue crack growth modeling, which is an engineering problem widely studied during the last three decades in structural analysis and system reliability (see, e.g, Lin and Yang, 1985; Sobczyk, 1993; Tanaka, 1999; Tsurui and Ishikawa, 1986; Yang and Manning, 1996). Here the size of a crack is described by the process Z_t and the dynamical system models its random growth. The crack size Z_t takes its values on $[0, L]$, where L is the wall thickness of a structure subject to some fatigue crack growth degradation. The absorbing point $\{\Delta\}$ could represent a critical crack size for which the structure collapses once reached.

We choose the function $f : (z, x) \longrightarrow a z^b v(x)$, with $a, b \in \mathbb{R}_+$, as suggested by Lin and Yang (1985) and later by Yang and Manning (1996) to model the crack growth law. We also consider the one-to-one function $v : E \longrightarrow \mathbb{R}_+$ which allows us to modulate easily the effect of the process X_t on the stochastic component of the system. The corresponding random evolution is the following:

$$\begin{cases} \dfrac{\mathrm{d}Z_t}{\mathrm{d}t} &= aZ_t^b \times v(X_t),\ a, b \in \mathbb{R}_+, \\ Z_0 &= z_0. \end{cases} \tag{26}$$

The driving process X_t is the jump Markov process considered in the previous numerical application. Let $v : (1, 2, 3) \to (1, 0.9, 1.2)$. Hence, the term $v(X_t)$ in system (26) models the variabilities of the growth rate of Z_t. Depending on the state occupied by X_t, which modulates the growth rate of Z_t, the system will evolve faster or more slowly: when X_t is in state 1, $v(X_t) = 1$ and the growth rate of Z_t is considered as "moderate" ; when X_t is in state 2, $v(X_t) = 0.9$ and the system evolves more slowly; when X_t is in state 2, $v(X_t) = 1.2$ and the system goes faster. The process Z_t starts almost surely from an initial point $z_0 \in]0, \Delta[$ and finally reaches the absorbing point $\{\Delta\}$ where it stops its evolution. Hence, the process Z_t takes its values on $[z_0, \Delta]$. Provided that $\Delta < L < \infty$, assumptions A1 and A3 are satisfied for the function f.

The data set is generated from system (26), with the true generator $\mathbf{A}$. Initial distributions α and β, constants a and b and absorbing point $\{\Delta\}$ are

$$\begin{cases} \alpha(1) = 1, \alpha(2) = 0, \alpha(3) = 0; \\ \beta(z) = \delta_{z_0},\ \text{with } z_0 = 9; \\ a = 5.10^{-4}, b = 1.5; \\ \Delta = 50. \end{cases} \tag{27}$$

Figure 3 represents 20 trajectories generated from system (26) with parameters (27), as an illustration. The whole generated data set is composed of 10,000 sample paths. In the following, these simulated data will be used to estimate parameters and reliability measurements of the system.

Concerning the generator, we use the method described in paragraph 4.2 to perform the estimation, that is:

1. Estimate the $\dot{Z}_t^k$ with the secant method, which is appropriate to system (26);
2. Estimate the X_t^k with function h defined in the general case by Eq. 18, which links Z_t, X_t and $\dot{Z}_t$. In the case of system (26), it is easy to see that

$$\widehat{X}_t^k = v^{-1}\left(\frac{1}{a}\left(Z_t^k\right)^{-b}\hat{\dot{Z}}_t^k\right);$$

3. Estimate the generator $\mathbf{A}$ from the estimated sample paths $\widehat{X}_t^k$ of the jump Markov process, with results from Section 2.

When comparing the initial true generator $\mathbf{A}$ and the estimator $\widehat{\mathbf{A}}$ obtained as described above, we obtain results similar to those obtained in the simpler case of a Markov process with random censoring, particularly a fast consistency. Thus, in this example, the fact that X_t and τ are dependent has no significant impact on the estimation of $\mathbf{A}$.

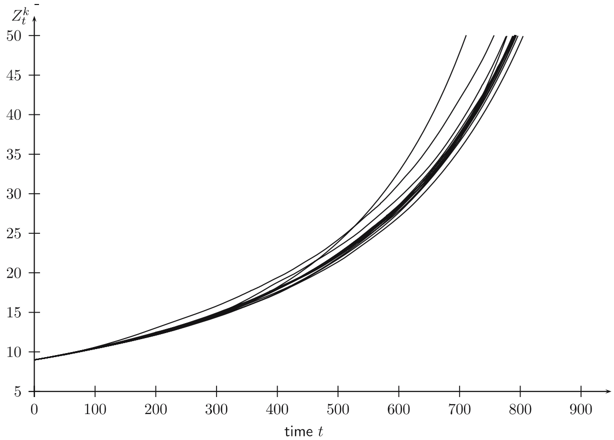


Fig. 3 Data set generated from system (26)

We can now take advantage of the classical Monte-Carlo method to obtain some reliability measurements of system (26). As an example, we propose to estimate the reliability function R defined by Eq. 13.

The true value of R could be calculated from result (15), yet we need a method to discretize the continuous generator $\mathbf{B}$. Another simpler approach is the estimation of the true value by the Monte-Carlo method. Estimating the distribution of the failure time τ is quite easy, and here, we have $R(t) = 1 - F(t)$. Once the estimation of R is achieved on the whole data generated with the *true generator* $\mathbf{A}$, we compare the estimation of R achieved on sample paths simulated with the *estimated generator* $\widehat{\mathbf{A}}$, using Monte-Carlo technique each time.

Figure 4 presents some results of reliability estimation: the "true" reliability has been obtained on the whole data set described above. We compare it with the estimation of the system's reliability obtained with 10, 100 and 2, 500 sample paths generated with $\widehat{\mathbf{A}}$.

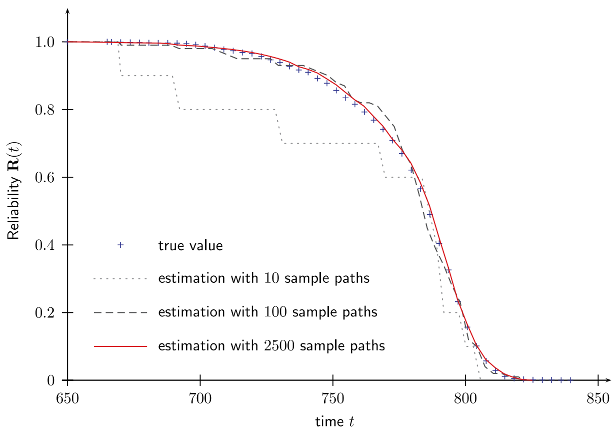


Fig. 4 Reliability estimated for different numbers of sample paths

We quickly reach the true value of the reliability: we may reasonably assume that the method we proposed to estimate the generator of the Markov driving process within the framework of stochastic dynamical systems modeling is good.

Until now, the parameters a and b of the function f were supposed to be known, which allowed us to study the behavior of the estimator of the generator, because this is the main interest of our work. Nevertheless, in applications, we don't necessary know parameters a and b. In the best case, we might suppose we know the general form of the function f. Consequently, we propose to estimate not only the generator but also the parameters a and b from the data set. With the estimated values $\widehat{a}$, $\widehat{b}$ and $\widehat{\mathbf{A}}$, we can make some simulations in order to estimate the reliability and compare it with the one obtained with the true values of a, b and $\mathbf{A}$.

As a first approximation, we propose to estimate a and b through application of the ordinary least-squares method. Taking logarithm on both sides of Eq. 26, then

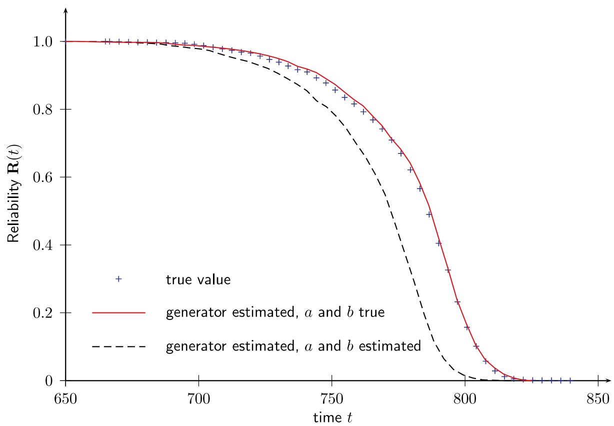
$$\ln \dot{Z}_t = \ln a + b \ln Z_t + \ln v(X_t). \tag{28}$$

From the data set, we have N sample points $\{(t_i, Z_{t_i})\}$. We transform this N-sample to obtain the N-sample $\left\{\left(\ln Z_{t_i}, \ln \widehat{\dot{Z}}_{ti}\right)\right\}$, much more adapted for regression analysis according to Eq. 28. For notational convenience, we denote $x_i = \ln Z_{t_i}$, $y_i = \ln \widehat{\dot{Z}}_{ti}$, $\varepsilon_i = \ln v(X_{t_i})$ and then we have $y_i = \ln a + b x_i + \varepsilon_i$. Hence, we perform a classical least-squares regression on the N-sample $\{(x_i, y_i)\}_{i=1\ldots N}$, by minimizing the square of the so-called residuals ε_i that is, $\sum_{i=1}^{N} (y_i - \ln a - b x_i)^2$. The following estimators are well-known:

$$\begin{cases} \widehat{a} = \exp\left(\bar{y} - \widehat{b}\bar{x}\right) \\ \widehat{b} = \dfrac{\sum x_i y_i - N\bar{x}\bar{y}}{\sum x_i^2 - N\bar{x}^2} \end{cases}, \text{ where } \bar{x} = \frac{1}{N}\sum x_i, \quad \bar{y} = \frac{1}{N}\sum y_i.$$

In this method, the $\varepsilon_i = \ln X_{t_i}$ describes the individual behavior of the sample i. When performing linear regression analysis, it is assumed that the ε_i are independent with zero mean, which is approximately true for this particular numerical example, and more generally in stochastic fatigue crack growth modeling; then, this approach could be a first approximation in order to estimate parameters a and b in our model.

Fig. 5 Reliability estimated for a and b true and estimated



The algorithm for the whole system estimation is the following:

1. Estimate the $\dot{Z}_t^k$ with the secant from the data set;
2. Estimate a and b with the least-squares method on the N-sample $\left\{\left(\ln Z_{t_i}, \ln \widehat{\dot{Z}}_{ti}\right)\right\}_{i=1\ldots N}$ as described above;
3. Estimate the sample paths of X_t with function h given by

$$\widehat{X}_t^k = v^{-1}\left(\frac{1}{\widehat{a}}\left(Z_t^k\right)^{-\widehat{b}}\,\widehat{\dot{Z}}_t^k\right);$$

4. Estimate generator $\mathbf{A}$ from the $\widehat{X}_t^k$;
5. Estimate the reliability function from simulated paths of Z_t through system (26) with parameters $\widehat{a}$, $\widehat{b}$ and $\widehat{\mathbf{A}}$.

Figure 5 illustrates the results, comparing the estimator of the reliability when parameters a and b are assumed to be known and when they are estimated.

We can see that the estimation of R is rather conservative when estimating the parameters of the function. It does not mean that the estimation of the infinitesimal generator is not good, yet the system we have chosen is sensitive to parameters a and b: we find $\widehat{a} = 5.003 \times 10^{-4}$ and $\widehat{b} = 1.506$ with the ordinary least-square method, which are close to the true values given by Eq. 27. However, we can observe a slight discrepancy in the estimation of the reliability, which proves the sensitiveness of this system.

6 Conclusions

In this paper, we first gave some results on the estimation of continuous-time Markov process when sample paths are observed on randomly censored intervals, where the censoring time is independent to the Markov process. We proved both consistency and asymptotic normality of the estimator of the generator. These results can be exploited in a wide class of engineering applications, as soon as Markov processes are used to model a system whose observations are defined on random time intervals. We gave a numerical application to illustrate the theoretical results.

Next, we studied the framework of stochastic dynamical systems driven by continuous-time jump Markov processes. We were able to estimate the generator of the driving process only from sample paths of the stochastic process to be modeled, using results from the first part of the paper. We gave the theoretical reliability of the system and its estimation with Monte-Carlo method. We gave a complete numerical application to the problem of fatigue crack growth, which is a widely studied complex dynamical system. We approached the reliability with the generator estimated and with the parameters of the function f successively known and unknown, illustrating the good behavior of the estimator of the generator of the jump process.

As a further work, we think that asymptotic results on the whole stochastic differential system (9) could be of great interest. Lapeyre and Pardoux (1998) studied asymptotic approximation of transport processes to a diffusion process, and Korolyuk and Limnios (2005) gave asymptotic results for general stochastic evolutionary systems which may help us.

Another key point is to try to take into account the censoring time in the estimation procedure of the generator of the driving process when they are not independent, that is in the context of a dynamical system observed until failure. The authors are working on this issue at the present time.

References

A. Albert, "Estimating the infinitesimal generator of a continuous time, finite state Markov process," *Annals of Mathematical Statistics* vol. 38 pp. 727–753, 1962.

P. Billingsley, *Statistical Inference for Markov Processes*, University of Chicago Press: Chicago, 1961.

J. Chiquet, N. Limnios, and M. Eid, "Modeling and estimating the reliability of stochastic dynamical systems with markovian switching." In *ESREL 2006 - Safety and Reliability for managing Risks*, 2006.

R. Dautray (ed.), *Méthodes probabilistes pour les équations de la physique*, Synthèse. Eyrolles: Paris, France, 1989.

M. H. A. Davis, *Markov Models and Optimization*, Monographs On statistics and Applied Probability 49, Chapman & Hall, UK, 1993.

S. Goldstein, "On diffusion by discontinuous movements, and on the telegraph equation," *Quarterly Journal of Mechanics and Applied Mathematics* vol. 4 pp. 129–156, 1951.

W. Härdle, *Applied Nonparametric Regression*, Econometric Society Monographs, Cambridge University Press, Cambridge, UK, 1990.

A. Jankunas and R. Z. Khasminskii, "Estimation of parameters of linear homogeneous stochastic differential equation," *Stochastic Processes and their Applications* vol. 72 pp. 205–219, 1997.

V. S. Korolyuk and N. Limnios, *Stochastic Systems in Merging Phase Space*, World Scientific: Singapore, 2005.

Y. A. Kutoyants, *Statistical Inference for Ergodic Diffusion Processes*, Springer: Berlin Heidelberg New York, 2004.

B. Lapeyre and E. Pardoux, *Méthodes de Monte-Carlo pour les équations de transport et de diffusion*, Springer: Berlin Heidelberg New York, 1998.

Y. K. Lin and J. N. Yang, "A stochastic theory of fatigue crack propagation," *AIAA Journal* vol. 23 pp. 117–124, 1985.

B. Øksendal, *Stochastic Differential Equations*, 6th edn., Springer: Berlin Heidelberg New York, 2003.

A. Sadek and N. Limnios, "Nonparametric estimation of reliability and survival function for continuous time finite Markov process," *Journal of Statistical Planning and Inference* vol. 133 pp. 1–21, 2005.

K. Sobczyk, *Stochastic Differential Equations with Applications to Physics and Engineering*, Kluwer: Dordrecht, The Netherlands, 1991.

K. Sobczyk (ed.), *Stochastic Approach to Fatigue*, Springer: Berlin Heidelberg New York, 1993.

B. F. Spencer, J. Tang, and M. E. Artley, "A stochastic approach to modeling fatigue crack growth," *AIAA Journal* vol. 114 pp. 1628–1635, 1989.

H. Tanaka, "Importance sampling simulation for a stochastic fatigue crack growth model," In R. E. Melchers and M. G. Stewart (eds.), *Proceeding of ICASP8* vol. 2, pp. 907–914, 1999.

A. Tsurui and H. Ishikawa, "Application of the Fokker-Plank equation to a stochastic fatigue crack growth model," *Structural Safety* vol. 4 pp. 15–29, 1986.

D. A. Virkler, B. M. Hillberry, and P. K. Goel, "The statistical nature of fatigue crack propagation," *Journal of Engineering Material Technology ASME* vol. 101 pp. 148–153, 1979.

J. N. Yang and S. D. Manning, "A simple second order approximation for stochastic crack growth analysis," *Engineering Fracture Mechanics* vol. 53 pp. 677–686, 1996.