Combining Gamma and Brownian processes for degradation modeling in presence of explanatory variables

Laurent Bordes, Christian Paroissin, Ali Salami

Combining Gamma and Brownian processes for degradation modeling in presence of explanatory variables

L. Bordes[a,*], C. Paroissin[a], A. Salami[a]

[a] *Université de Pau et des Pays de l'Adour, Laboratoire de Mathématiques et de leurs Applications - UMR CNRS 5142, Avenue de l'Université, 64013 Pau cedex, France.*

Abstract

A new approach has been proposed recently to describe a system deterioration. In this new approach the authors have considered the degradation as the sum of a gamma process and a Brownian motion, independent of the gamma process. The Brownian motion may describe, for example, degradation measurement error. In this paper covariates are introduced in order to take into account environmental effects or systems heterogeneity. From the observation of n independent items at regular instants over a finite time interval, the model parameters are estimated by a two-stage least-square method. Asymptotic properties of the estimators are provided. Finally both numerical simulations and real data applications are supplied to illustrate our method.

Keywords: gamma process, Wiener process, covariates, least-square estimators, consistency, asymptotic normality
AMS Classification: 62F10, 62F12, 62N05

1. Introduction

Modelling the degradation of a system can be done using a stochastic process. Degradation is usually measured at several times and may be influenced by the system environment. Thus it is necessary to account covariates describing this environment. Degradation levels can also reflect measurement errors or, for example, minor repairs of the system. As a consequence, a good statistical model should take into account all these sources of variation.

For certain types of degradation a process involving independent non-negative increments is appropriate, as for example the gamma process (van Noortwijk (2009)). These kind of processes imply that the system state cannot be improved over time, and then this system cannot return to its original state without external maintenance actions. The gamma process can be regarded as a compound Poisson process of gamma-distributed increments in which the Poisson rate tends to infinity and increment size tend to zero in proportion (Lawless and Crowder (2004)). It was originally proposed by Abdel-Hameed (1975) in order to describe a degradation phenomenon. This process is frequently used in the literature since it is preferable from the physics point of view (monotonic deterioration). Moreover the independent increments property makes the subsequent mathematical treatment quite tractable. In several papers, covariates or random effects have been incorporated to a gamma process in order to take into account the environment or the individual heterogeneity. Bagdonavičius and Nikulin (2001) propose an accelerated life test (ALT) model where covariates modify the time scale of the gamma process. Alternatively Lawless and Crowder (2004) and Crowder and Lawless (2007) assume that the scale parameter depends on covariates and is also proportional to a random effect. Let us recall the definition of this process in order to fix our notations. Let $\xi \in \mathbb{R}_+$ and $\eta = (\eta_t)_{t\geqslant 0}$ a real-valued non-decreasing function such that $\eta_0 = 0$. The stochastic process Y is a gamma process with parameters (ξ, η) if:

1. $Y(0) = 0$;
2. Y has independent increments;

*http://lma-umr5142.univ-pau.fr, Tél. 05 59 40 75 38, Fax 05 59 40 75 55
Email addresses: `laurent.bordes@univ-pau.fr` (L. Bordes), `cparoiss@univ-pau.fr` (C. Paroissin), `ali.salami@univ-pau.fr` (A. Salami)

3. For $0 \leq s < t$, $Y(t) - Y(s)$ is gamma distributed with parameters $(\xi, \eta_t - \eta_s)$.

The probability distribution function of $Y(t)$ is defined by

$$f_{Y(t)}(y) = \frac{\xi^{\eta_t}}{\Gamma(\eta_t)} y^{\eta_t - 1} e^{-\xi y} 1_{\{y \geq 0\}},$$

where $\Gamma(\cdot)$ is the gamma function and $1_{\{\cdot\}}$ is the set indicator function. A special case is mainly considered in the literature when the function η is linear, say for instance $\eta_t = \alpha t$. Indeed, in such case, Y is a stationary process: for any $t \geq 0$ and $\delta > 0$, $Y(t + \delta) - Y(t)$ and $Y(\delta)$ have the same distribution. Thus Y turns to be a Lévy process. This special case allows to do many explicit computations. Some non-linear cases have been also studied. The most common non-linear model is obtained for $\eta_t = \alpha t^{\beta}$. When dealing with statistical purpose the parameter β is generally assumed to be known. Lawless and Crowder (2004) considered another case where η is defined by the Paris-Erdogan curve. For more details on gamma processes see van Noortwijk (2009). Hereafter we only consider the linear case.

Another interesting Markov process used in the literature as a degradation model is the Wiener diffusion process (Barker (2006); Doksum and Hóyland (1992); Wang (2010); Whitmore (1995) and Whitmore et al. (1998)). The reason why such Markov processes (the gamma process and the Brownian motion) have been and are still extensively used is that they belong to a class of time-dependent stochastic processes known as Lévy processes. Let us recall some basic facts about the Wiener process. A process W is said to be a Brownian motion with drift μ and variance σ^2 if $W(t) = \sigma B(t) + \mu t$ where B is a standard Brownian motion with variance t. The process W has the following properties:

1. $W(0) = 0$;
2. W has continuous sample paths;
3. W has independent increments;
4. For any $0 \leq s < t$, the random variable $W(t) - W(s)$ has normal distribution with mean $\mu(t - s)$ and variance $\sigma^2(t - s)$.

The drawback of the Gaussian assumption is that it leads to a process with non monotone sample paths. As claimed by Park and Padgett (2005), this is the reason why this process can be inadequate in modelling monotone deterioration.

The aim of this paper is to propose a model joining the two previously mentioned approaches (the gamma process and the Brownian motion). Then we assume that covariates only act on the gamma process part involved in our degradation process. Covariates are incorporated as in Bagdonavičius and Nikulin (2001). From the observation of n independent items at regular instants over a finite time interval, the model parameters are estimated by a two-stage least-square method. Asymptotic properties of the estimators are then provided and, finally, both numerical simulations and real data applications are supplied to illustrate our method.

2. Degradation model

As claimed above a new approach has been proposed recently (see Bordes et al. (2010)) to describe a system degradation, considering the following degradation process:

$$\forall t \geq 0\,,\; D(t) = Y(t) + \tau B(t)\,,$$

where Y is a gamma process with parameters α and ξ (as defined previously) and where B is a Brownian motion. This model is defined for $\tau \in \mathbb{R}$ and the processes Y and B are assumed to be independent. Without loss of generality, we can assume that $\tau \geq 0$ since $\tau B(t)$ and $-\tau B(t)$ have the same distribution for all $t \geq 0$.

The motivations for considering such a model are the following ones. First, this model joins the gamma process and the Brownian motion into a single model. Indeed, it is clear that when $\tau = 0$, this model turns to be a gamma process. Moreover, if α/ξ tends to $b > 0$ and α/ξ^2 tends to 0, then this model converges weakly to a Brownian motion with positive drift b. Second, measurements of degradation levels reflect measurement errors. Hence, the role of Brownian motion in this model can be interpreted as measurement errors. Finally, our model can take into account minor repairs carried out on the system over time that can be responsible of non-monotone degradation. Using this model Bordes et al. (2010) assume that n independent items are observed at irregular instants and provide estimators of the model parameters using the moment method approach.

In this paper, we assume that covariates act on the gamma process part involved in our degradation process. Covariates are integrated as in Bagdonavičius and Nikulin (2001). Thus if $x^T = \left(x^{(1)}, \ldots, x^{(p)}\right)$ is a vector of p covariates, conditionally on x our model is defined by

$$\forall t \geq 0, \; D_x(t) = Y\left(te^{\beta^T x}\right) + \tau B(t),$$

where $\beta = \left(\beta_1, \ldots, \beta_p\right)^T$ is a vector of unknown parameters. It follows that Y is a stationary gamma process with scale parameter ξ and shape parameter $\alpha e^{\beta^T x}$. Moreover, we assume that the covariates vector x is an observed value of a random vector X having density function f_X with respect to a σ-finite measure μ_p on $\mathbb{R}^p$ and we denote by F_X its distribution function.

To end this section, we introduce some simulations of the proposed degradation process. Here we consider that items are subject to a one dimensional covariate. This covariate can take one of three possible values corresponding to three different solicitations (low, medium and high stress levels). For instance, it can be the temperature as in NIST/SEMATECH (2010). On Figure 1, simulated trajectories are plotted (dotted lines), along with the three averages of degradations (solid lines). Parameters are set to $\xi = 1$, $\alpha = 2, \beta = (-0.5, 0, 0.4)$ and $\tau^2 = 1$.

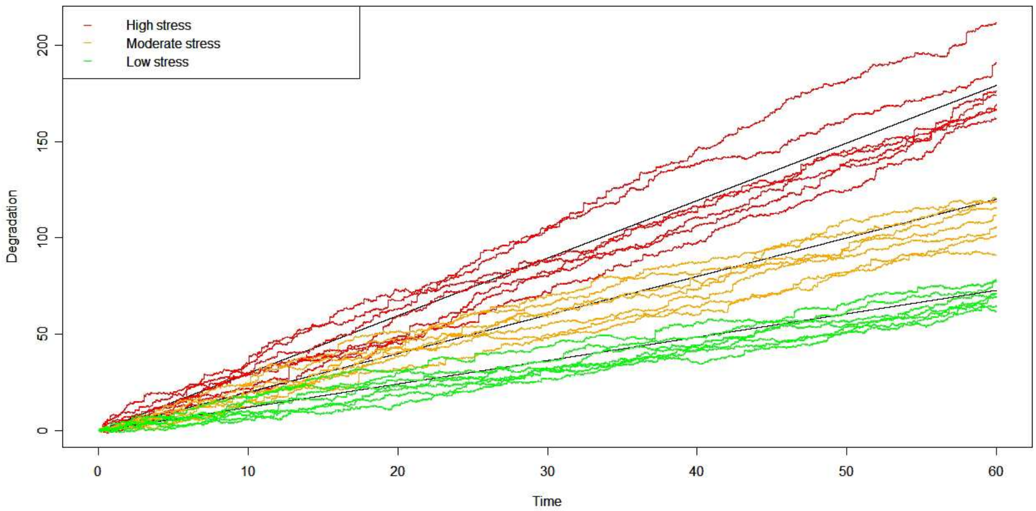

Figure 1: Simulation example

One can observe that items have to be monitored over a sufficient long time interval in order to make appear the three distinct groups. Of course, this "optimal" time interval depends on parameters values which are unknown when considering real applications.

3. Parameter estimation

We denote by $D_{x_1}^{(1)}, \ldots, D_{x_n}^{(n)}$ n independent degradation processes such that for any $i \in \{1, \ldots, n\}$ and for any $t \geq 0$,

$$D_{x_i}^{(i)}(t) = Y^{(i)}\left(te^{\beta^T x_i}\right) + \tau B^{(i)}(t),$$

where $D_{x_1}^{(1)}, \ldots, D_{x_n}^{(n)}$ have the same distribution as D_x in Section 2 and $x_1, \ldots, x_n$ are n independent and identically distributed (i.i.d.) realizations of X. We suppose that we observe the process $D_{x_i}^{(i)}$ at regular instants $t_j = jT/N$ for

$j \in \{0, \ldots, N\}$. For any $1 \leq j \leq N$ and $k > 0$, we denote by $m_\theta^{(k)}(x_i)$ the k-th moment of increments

$$\Delta D_{x_i}^{(i)}\left(t_j\right) = D_{x_i}^{(i)}\left(t_j\right) - D_{x_i}^{(i)}\left(t_{j-1}\right),$$

where $\theta = \left(\theta^{(1)}, \theta^{(2)}\right) = \left((\gamma, \beta), \left(\alpha, \tau^2\right)\right)$ with $\gamma = \alpha/\xi$. We assume that θ belongs to $\Theta = \Theta_1 \times \Theta_2 \subset (\mathbb{R}_+^* \times \mathbb{R}^p) \times (\mathbb{R}_+^* \times \mathbb{R}_+)$ the parameter space of the model. In the sequel we denote by $\theta_0 = \left(\theta_0^{(1)}, \theta_0^{(2)}\right) = \left((\gamma_0, \beta_0), \left(\alpha_0, \tau_0^2\right)\right)$ the true value of the model parameter. Hereafter we denote by $\mathbb{E}_\theta\left[\varphi\left(D_x(t)\right)\right]$ the conditional mean $\mathbb{E}_\theta\left[\varphi\left(D_X(t)\right)|X = x\right]$ for any real-valued measurable function φ.

The first moment is equal to

$$\begin{aligned} m_\theta^{(1)}(x_i) &= \mathbb{E}_\theta\left[\Delta D_{x_i}^{(i)}\left(t_j\right)\right] = \mathbb{E}_\theta\left[\Delta D_{X_i}^{(i)}\left(t_j\right)|X_i = x_i\right] \\ &= \frac{\gamma T}{N} e^{\beta^T x_i} = m_{\theta^{(1)}}^{(1)}(x_i). \end{aligned}$$

The second moment is equal to

$$\begin{aligned} m_\theta^{(2)}(x_i) &= \mathbb{E}_\theta\left[\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2\right] = \mathbb{E}_\theta\left[\left(\Delta D_{X_i}^{(i)}\left(t_j\right)\right)^2 |X_i = x_i, \theta^{(1)}\right] \\ &= m_{\theta^{(1)}}^{(1)}(x_i)\frac{\gamma}{\alpha} + \left(m_{\theta^{(1)}}^{(1)}(x_i)\right)^2 + \tau^2 T/N = m_{\theta^{(2)}}^{(2)}\left(x_i, \theta^{(1)}\right). \end{aligned}$$

To estimate $\theta^{(1)} = (\gamma, \beta)$, we minimize the following first regression function

$$d_1^{(n)}\left(\theta^{(1)}\right) = \sum_{i=1}^{n}\sum_{j=1}^{N}\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta^{(1)}}^{(1)}(x_i)\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{N}\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - \frac{\gamma T}{N}e^{\beta^T x_i}\right)^2.$$

Then we set $\hat{\theta}_n^{(1)} = \underset{\theta^{(1)} \in \Theta_1}{\arg\min}\, d_1^{(n)}\left(\theta^{(1)}\right)$. Once this parameter is estimated, we estimate $\theta^{(2)} = \left(\alpha, \tau^2\right)$ by minimizing the following second regression function

$$\begin{aligned} d_2^{(n)}\left(\theta^{(2)}, \hat{\theta}_n^{(1)}\right) &= \sum_{i=1}^{n}\sum_{j=1}^{N}\left(\left[\Delta D_{x_i}^{(i)}\left(t_j\right)\right]^2 - m_{\theta^{(2)}}^{(2)}\left(x_i, \hat{\theta}_n^{(1)}\right)\right)^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{N}\left(\left[\Delta D_{x_i}^{(i)}\left(t_j\right)\right]^2 - m_{\hat{\theta}_n^{(1)}}^{(1)}(x_i)\frac{\hat{\gamma}}{\alpha} - \left(m_{\hat{\theta}_n^{(1)}}^{(1)}(x_i)\right)^2 - \tau^2 T/N\right)^2. \end{aligned}$$

As a consequence $\theta^{(2)}$ is estimated by $\hat{\theta}_n^{(2)} = \underset{\theta^{(2)} \in \Theta_2}{\arg\min}\, d_2^{(n)}\left(\theta^{(2)}, \hat{\theta}_n^{(1)}\right)$. The final estimator of θ is therefore $\hat{\theta}_n = \left(\hat{\theta}_n^{(1)}, \hat{\theta}_n^{(2)}\right)$. At each stage of our approach, estimators are obtained numerically by using a differentiable optimization method.

4. Theoretical results

First we denote by $d^{(n)}(\theta)$ and $d(\theta)$ the two following matrices below:

$$d^{(n)}(\theta) = \frac{1}{nN}\begin{pmatrix} d_1^{(n)}\left(\theta^{(1)}\right) \\ d_2^{(n)}\left(\theta^{(2)}, \theta^{(1)}\right)\end{pmatrix} \text{ and } d(\theta) = \begin{pmatrix} d_1\left(\theta^{(1)}\right) \\ d_2\left(\theta^{(2)}, \theta^{(1)}\right)\end{pmatrix},$$

where it is established in the proof of Theorem 2 that d_1 and d_2 are respectively the almost sure limits of $d_1^{(n)}$ and $d_2^{(n)}$, and are respectively defined by:

$$d_1\left(\theta^{(1)}\right) = \int_{\mathbb{R}^p} \mathbb{E}_{\theta_0}\left[D_x\left(\frac{T}{N}\right) - m_{\theta^{(1)}}^{(1)}(x)\right]^2 f_X(x)\,\mathrm{d}\mu_p(x),$$

and

$$d_2(\theta) = d_2\left(\theta^{(2)}, \theta^{(1)}\right) = \int_{\mathbb{R}^p} \mathbb{E}_{\theta_0}\left[D_x^2\left(\frac{T}{N}\right) - m_{\theta^{(2)}}^{(2)}\left(x, \theta^{(1)}\right)\right]^2 f_X(x)\,\mathrm{d}\mu_p(x).$$

4.1. Consistency

Let us prove the consistency of $\left(\hat{\theta}_n\right)_{n\geq 1}$. First we need to show the following lemma. In the sequel $\|\cdot\|$ denotes the Euclidean norm.

Lemma 1. *Assume that the following conditions are fulfilled:*

1. Θ *is a compact set;*
2. *For* $i \in \{1, 2\}$*, the application* $\theta \longmapsto d_i(\theta)$ *is continuous and satisfies:*
   - $d_i(\theta) \geq 0$ *for all* $\theta \in \Theta$*,*
   - d_1 *and* d_2 *admit unique minima at* $\theta_0^{(1)} \in \mathring{\Theta}_1$ *and* $\theta_0 \in \mathring{\Theta}$ *respectively.*
3. $\sup_{\theta\in\Theta} \|d^{(n)}(\theta) - d(\theta)\| \xrightarrow[n\to\infty]{Pr} 0.$

Then, as n tends to infinity, $(\hat{\theta}_n)_{n\geq 0}$ *converges in probability to* θ_0*.*

Proof. To ensure that $\hat{\theta}_n$ converges in probability to θ_0, we have to prove that $\hat{\theta}_n^{(1)}$ (respectively $\hat{\theta}_n^{(2)}$) converges in probability to $\theta_0^{(1)}$(respectively $\theta_0^{(2)}$). First let us check that $\hat{\theta}_n^{(1)} \xrightarrow[n\to\infty]{Pr} \theta_0^{(1)}$.

Without loss of generality, we assume that $d_1\left(\theta_0^{(1)}\right) = 0$. Θ_1 is a compact set and $B\left(\theta_0^{(1)}, \epsilon\right) \subset \Theta_1$ is the open ball with center $\theta_0^{(1)}$ and radius $\epsilon > 0$. By Assumption 2 there exists $\eta > 0$ such that

$$\forall\, \theta^{(1)} \in B^c\left(\theta_0^{(1)}, \epsilon\right), \quad d_1\left(\theta^{(1)}\right) \geq \eta.$$

Then, if $\hat{\theta}_n^{(1)} \in B^c\left(\theta_0^{(1)}, \epsilon\right)$, we have

$$d_1^{(n)}\left(\hat{\theta}_n^{(1)}\right) \;\geq\; \eta - \sup_{\theta^{(1)}\in\Theta_1} \left|d_1^{(n)}\left(\theta^{(1)}\right) - d_1\left(\theta^{(1)}\right)\right|. \tag{1}$$

Since $\hat{\theta}_n^{(1)} \;=\; \arg\min_{\theta^{(1)}\in\,\Theta_1} d_1^{(n)}\left(\theta^{(1)}\right)$ and $d_1\left(\theta_0^{(1)}\right) = 0$, we have

$$d_1^{(n)}\left(\hat{\theta}_n^{(1)}\right) \;\leq\; \sup_{\theta^{(1)}\in\Theta_1} \left|d_1^{(n)}\left(\theta^{(1)}\right) - d_1\left(\theta^{(1)}\right)\right|. \tag{2}$$

Inequalities (1) and (2) implies that

$$\left\{\|\hat{\theta}_n^{(1)} - \theta_0^{(1)}\| > \epsilon\right\} \subseteq \left\{\eta < 2 \sup_{\theta^{(1)}\in\Theta_1} \left|d_1^{(n)}\left(\theta^{(1)}\right) - d_1\left(\theta^{(1)}\right)\right|\right\}.$$

and then

$$\mathbb{P}_{\theta^{(1)}}\left[\|\hat{\theta}_n^{(1)} - \theta_0^{(1)}\| > \epsilon\right] \leq \mathbb{P}_{\theta^{(1)}}\left[\sup_{\theta^{(1)}\in\Theta_1} \left|d_1^{(n)}\left(\theta^{(1)}\right) - d_1\left(\theta^{(1)}\right)\right| > \eta/2\right].$$

Hence we conclude by Assumption 3 that

$$\lim_{n\longrightarrow\infty} \mathbb{P}_{\theta^{(1)}}\left[\|\hat{\theta}_n^{(1)} - \theta_0^{(1)}\| > \epsilon\right] = 0,$$

which means that $\hat{\theta}_n^{(1)} \xrightarrow{Pr} \theta_0^{(1)}$. Next, let us check that $\hat{\theta}_n^{(2)} \xrightarrow{Pr} \theta_0^{(2)}$. Similarly as above Θ_2 is a compact set and $B\left(\theta_0^{(2)}, \epsilon\right) \subset \Theta_2$ is an open ball with center $\theta_0^{(2)}$ and radius $\epsilon > 0$. By Assumptions 1 and 2 there exists $\tilde{\eta}, \tilde{\epsilon} > 0$ such that

$$\forall\, \theta = \left(\theta^{(1)}, \theta^{(2)}\right) \in B^c\left(\theta_0, \tilde{\epsilon}\right), \quad d_2(\theta) \geq \tilde{\eta}.$$

Without loss of generality, we assume that $d_2\left(\theta_0^{(1)}, \theta_0^{(2)}\right) = 0$. Then, for any $\hat{\theta}_n \in B^c(\theta_0, \tilde{\epsilon})$, we have

$$d_2^{(n)}\left(\hat{\theta}_n^{(1)}, \hat{\theta}_n^{(2)}\right) \;\geq\; \tilde{\eta} - \sup_{\theta\in\Theta}\left|d_2^{(n)}(\theta) - d_2(\theta)\right|. \tag{3}$$

Because $\hat{\theta}_n^{(2)} = \underset{\theta^{(2)}\in\Theta_2}{\arg\min}\, d_2^{(n)}\left(\hat{\theta}^{(1)}, \theta^{(2)}\right)$ we have

$$d_2^{(n)}\left(\hat{\theta}_n^{(1)}, \hat{\theta}_n^{(2)}\right) \leq \sup_{\theta\in\Theta}\left|d_2^{(n)}(\theta) - d_2(\theta)\right| + d_2\left(\hat{\theta}_n^{(1)}, \theta_0^{(2)}\right). \tag{4}$$

Inequalities (3) and (4) implies that

$$\left\{\|\hat{\theta}_n^{(2)} - \theta_0^{(2)}\| > \epsilon\right\} \subseteq \left\{\tilde{\eta} < 2\sup_{\theta\in\Theta}\left|d_2^{(n)}(\theta) - d_2(\theta)\right| + d_2\left(\hat{\theta}_n^{(1)}, \theta_0^{(2)}\right)\right\},$$

and then

$$\mathbb{P}_\theta\left[\|\hat{\theta}_n^{(2)} - \theta_0^{(2)}\| > \epsilon\right] \leq \mathbb{P}_\theta\left[\tilde{\eta}/2 < \sup_{\theta\in\Theta}\left|d_2^{(n)}(\theta) - d_2(\theta)\right| + d_2\left(\hat{\theta}_n^{(1)}, \theta_0^{(2)}\right)\right].$$

By Assumption 3 we have

$$\sup_{\theta\in\Theta}\left|d_2^{(n)}(\theta) - d_2(\theta)\right| \xrightarrow[n\to\infty]{Pr} 0.$$

Since $\hat{\theta}_n^{(1)} \xrightarrow[n\to\infty]{Pr} \theta_0^{(1)}$ and d_2 is continuous at θ_0, it follows that $d_2\left(\hat{\theta}_n^{(1)}, \theta_0^{(2)}\right) \xrightarrow[n\to\infty]{Pr} 0$. Thus we obtain that $\hat{\theta}_n^{(2)} \xrightarrow{Pr} \theta_0^{(2)}$. □

Next, we prove that Assumptions 2 and 3 of Lemma 1 are satisfied in our set-up, leading to consistency.

Theorem 2. *Under the following assumptions:*

(A_1) *Θ is a compact set such that $\theta_0 \in \mathring{\Theta}$ and $\beta_0 \neq 0$;*

(A_2) *X is a bounded random vector;*

(A_3) *Let $\mathcal{A} \subset \mathbb{R}^p$ such that $\mu_p\left(\overline{\mathcal{A}}\right) = 0$. There exist $x_1, \ldots, x_{p+1} \in \mathcal{A}$ such that*

- *a) $\forall i \in \{1, \ldots, p+1\}$, $f_X(x_i) > 0$*
- *b) For any $i \in \{1, \ldots, p+1\}$ we set $\tilde{x}_i^T = \begin{pmatrix}1 & x_i^T\end{pmatrix}$. Then $\tilde{x}_1, \ldots, \tilde{x}_{p+1}$ are linearly independent.*

we have $\hat{\theta}_n \xrightarrow[n\to\infty]{Pr} \theta_0$.

Proof. To show that $\hat{\theta}_n$ converges in probability to θ_0 we must check Assumptions 2 and 3 of Lemma 1 since Θ is a compact set by Assumption (A_1). We begin by showing that for $i \in \{1, 2\}$, $\theta \longmapsto d_i(\theta)$ is a continuous map.

$\theta^{(1)} \longmapsto d_1\left(\theta^{(1)}\right)$ **is continuous**. We have

$$\begin{aligned}
d_1\left(\theta^{(1)}\right) &= \int_{\mathbb{R}^p} \mathbb{E}_{\theta_0}\left(D_x\left(\frac{T}{N}\right) - m^{(1)}_{\theta^{(1)}}(x)\right)^2 dF_X(x) \\
&= \int_{\mathbb{R}^p} \mathbb{E}_{\theta_0}\left(\left(D_x\left(\frac{T}{N}\right) - m^{(1)}_{\theta_0^{(1)}}(x)\right) + \left(m^{(1)}_{\theta_0^{(1)}}(x) - m^{(1)}_{\theta^{(1)}}(x)\right)\right)^2 dF_X(x) \\
&= d_1\left(\theta_0^{(1)}\right) + \int_{\mathbb{R}^p}\left(m^{(1)}_{\theta_0^{(1)}}(x) - m^{(1)}_{\theta^{(1)}}(x)\right)^2 dF_X(x).
\end{aligned}$$

As $d_1\left(\theta_0^{(1)}\right)$ is a constant, let us check the continuity of $\theta^{(1)} \longmapsto \int_{\mathbb{R}^p}\left(m^{(1)}_{\theta_0^{(1)}}(x) - m^{(1)}_{\theta^{(1)}}(x)\right)^2 dF_X(x)$. We denote by

$$f^{(1)}_{\theta^{(1)}}(x) = \left(\frac{\gamma_0 T}{N} e^{\beta_0^T x} - \frac{\gamma T}{N} e^{\beta^T x}\right)^2 f_X(x).$$

For any $x \in \mathbb{R}^p, \theta^{(1)} \longmapsto f^{(1)}_{\theta^{(1)}}(x)$ is a continuous function and can be bounded as follows:

$$\left|f^{(1)}_{\theta^{(1)}}(x)\right| \leq \left(\frac{\gamma_0 T}{N} e^{\|\beta_0\|.\|x\|} + \frac{\gamma T}{N} e^{\|\beta\|.\|x\|}\right)^2 f_X(x).$$

As $\theta^{(1)} \in \Theta_1$ a compact set, then there exist two constants $k_1, k_2 > 0$ such that

$$\left|f^{(1)}_{\theta^{(1)}}(x)\right| \leq k_1 e^{k_2\|x\|} f_X(x).$$

Using Assumption (A_2) we deduce that

$$\left|f^{(1)}_{\theta^{(1)}}(x)\right| \leq C_1 f_X(x) \in L_1\left(\mu_p\right)$$

where C_1 is a constant. Hence, applying the dominated convergence theorem, we obtain the continuity of

$$\theta^{(1)} \longmapsto \int_{\mathbb{R}^p} f^{(1)}_{\theta^{(1)}}(x)\, dF_X(x).$$

Finally we deduce that $\theta^{(1)} \longmapsto d_1\left(\theta^{(1)}\right)$ is a continuous map. We obtain similarly the continuity of $\theta = \left(\theta_0^{(1)}, \theta^{(2)}\right) \longmapsto d_2(\theta)$. Next we prove that d_1 admits a unique minimum at $\theta_0^{(1)}$.

d_1 **admits a minimum at** $\theta_0^{(1)} \in \mathring{\Theta}_1$. We have

$$d_1\left(\theta^{(1)}\right) = d_1\left(\theta_0^{(1)}\right) + \int_{\mathbb{R}^p} \left(m^{(1)}_{\theta_0^{(1)}}(x) - m^{(1)}_{\theta^{(1)}}(x)\right)^2 dF_X(x).$$

Thus we deduce that for any $\theta^{(1)} \in \Theta_1$, $d_1\left(\theta_0^{(1)}\right) \leq d_1\left(\theta^{(1)}\right)$ since $\int_{\mathbb{R}^p} \left(m^{(1)}_{\theta_0^{(1)}}(x) - m^{(1)}_{\theta^{(1)}}(x)\right)^2 dF_X(x) \geq 0$.

Uniqueness of $\theta_0^{(1)}$. We see that $d_1\left(\theta_0^{(1)}\right) = d_1\left(\theta^{(1)}\right)$ if $\int_{\mathbb{R}^p} \left(m^{(1)}_{\theta_0^{(1)}}(x) - m^{(1)}_{\theta^{(1)}}(x)\right)^2 dF_X(x) = 0$. Thus to check the uniqueness of $\theta_0^{(1)}$, first note that:

$$\int_{\mathbb{R}^p} \left(m^{(1)}_{\theta_0^{(1)}}(x) - m^{(1)}_{\theta^{(1)}}(x)\right)^2 dF_X(x) = \int_{\mathbb{R}^p} \left(\gamma_0 e^{\beta_0^T x} - \gamma e^{\beta^T x}\right)^2 f_X(x) d\mu_p(x) = 0.$$

From assumption (A_3), there exist at least $x_1, \ldots, x_{p+1} \in \mathcal{A} \subset \mathbb{R}^p$ such that $f_X(x_i) > 0$ and $\mu_p\left(\overline{\mathcal{A}}\right) = 0$, it implies that for any $i \in \{1, \ldots, p+1\}$

$$\gamma_0 e^{\beta_0^T x_i} = \gamma e^{\beta^T x_i} \tag{5}$$

then, for $i \in \{1, \ldots, p+1\}$

$$\ln \gamma_0 + \beta_0^T x_i = \ln \gamma + \beta^T x_i.$$

However, the last equality can be written as follows:

$$\begin{pmatrix} 1 & x_1^T \\ \vdots & \vdots \\ 1 & x_{p+1}^T \end{pmatrix} \begin{pmatrix} \ln \gamma \\ \beta \end{pmatrix} = \begin{pmatrix} \ln \gamma_0 + \beta_0^T x_1 \\ \vdots \\ \ln \gamma_0 + \beta_0^T x_{p+1} \end{pmatrix} \Leftrightarrow \begin{pmatrix} \ln \gamma \\ \beta \end{pmatrix} = \begin{pmatrix} 1 & x_1^T \\ \vdots & \vdots \\ 1 & x_{p+1}^T \end{pmatrix}^{-1} \begin{pmatrix} 1 & x_1^T \\ \vdots & \vdots \\ 1 & x_{p+1}^T \end{pmatrix} \begin{pmatrix} \ln \gamma_0 \\ \beta_0 \end{pmatrix},$$

because $\tilde{x}_1, \ldots, \tilde{x}_{p+1}$ are linearly independent, we deduce that $\gamma_0 = \gamma$ and $\beta_0 = \beta$.

$d_2(\theta)$ **admits a minimum at** $\theta_0 \in \mathring{\Theta}$**.** We have

$$d_2\left(\theta_0^{(1)}, \theta^{(2)}\right) = d_2(\theta_0) + \int_{\mathbb{R}^p} \left(m^{(2)}_{\theta_0^{(2)}}\left(x, \theta_0^{(1)}\right) - m^{(2)}_{\theta^{(2)}}\left(x, \theta_0^{(1)}\right)\right)^2 dF_X(x).$$

Thus we deduce that $\forall \theta^{(2)} \in \Theta_2,\ d_2(\theta_0) \leq d_2\left(\theta_0^{(1)}, \theta^{(2)}\right)$ since

$$\int_{\mathbb{R}^p} \left(m^{(2)}_{\theta_0^{(2)}}\left(x, \theta_0^{(1)}\right) - m^{(2)}_{\theta^{(2)}}\left(x, \theta_0^{(1)}\right)\right)^2 dF_X(x) \geq 0.$$

Uniqueness of $\theta_0^{(2)}$**.** We have

$$\begin{aligned}
&\int_{\mathbb{R}^p} \left(m^{(2)}_{\theta_0^{(2)}}\left(x, \theta_0^{(1)}\right) - m^{(2)}_{\theta^{(2)}}\left(x, \theta_0^{(1)}\right)\right)^2 dF_X(x) \\
&= \int_{\mathbb{R}^p} \left(\frac{\gamma_0^2 T}{N\alpha_0} e^{\beta_0^T x} + \left(\frac{\gamma_0 T}{N} e^{\beta_0^T x}\right)^2 + \tau_0^2 \frac{T}{N} - \frac{\gamma_0^2 T}{N\alpha} e^{\beta_0^T x} - \left(\frac{\gamma_0 T}{N} e^{\beta_0^T x}\right)^2 - \tau^2 \frac{T}{N}\right)^2 dF_X(x) \\
&= \int_{\mathbb{R}^p} \left(\frac{\gamma_0^2 T}{N\alpha_0} e^{\beta_0^T x} + \tau_0^2 \frac{T}{N} - \frac{\gamma_0^2 T}{N\alpha} e^{\beta_0^T x} - \tau^2 \frac{T}{N}\right)^2 dF_X(x) = 0.
\end{aligned}$$

From assumption (A_3), there exist at least $x_1, \ldots, x_{p+1} \in \mathcal{A}$ such that $f_X(x_i) > 0$ and $\mu_p\left(\overline{\mathcal{A}}\right) = 0$, it implies that for any $i \in \{1, \ldots, p+1\}$

$$\left(\frac{\gamma_0^2}{\alpha_0} - \frac{\gamma_0^2}{\alpha}\right) e^{\beta_0^T x_i} = \tau^2 - \tau_0^2. \tag{6}$$

In this case, if $\alpha \neq \alpha_0$ then $e^{\beta_0^T x_i} = e^{\beta_0^T x_j}$ for any i and j. It follows that $e^{\beta_0^T (x_i - x_j)} = 1$ for any $i \neq j$. This implies that β_0 is orthogonal to Span$\left\{x_i - x_j; 1 \leq i < j \leq p+1\right\}$. Thus $\beta_0 = 0$ which cannot hold since $\beta_0 \neq 0$ by Assumption (A_1). Then we deduce that $\alpha = \alpha_0$ and $\tau^2 = \tau_0^2$. Finally we conclude that $d_2(\theta)$ admits a unique minimum at $\theta_0 \in \mathring{\Theta}$.

In the sequel we prove a little bit more than Assumption 3 of Lemma 1 since almost sure uniform convergence results are obtained.

Almost sure convergence of $\sup_{\theta_1 \in \Theta_1} \left| d_1^{(n)}\left(\theta^{(1)}\right) - d_1\left(\theta^{(1)}\right)\right|$ **to zero.** Taking into account that covariates X_i are uniformly bounded with respect to i, we have to verify that

$$\begin{aligned}
&\sup_{\theta^{(1)} \in \Theta_1} \left| d_1^{(n)}\left(\theta^{(1)}\right) - d_1\left(\theta^{(1)}\right)\right| \\
= &\sup_{\theta^{(1)} \in \Theta_1} \left| \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{N} \sum_{j=1}^N \left(\Delta D^{(i)}_{X_i}\left(t_j\right) - m^{(1)}_{\theta^{(1)}}(X_i)\right)^2\right] - \mathbb{E}\left[\frac{1}{N} \sum_{j=1}^N \mathbb{E}_{\theta_0}\left(\Delta D^{(i)}_{X_i}\left(t_j\right) - m^{(1)}_{\theta^{(1)}}(X_i)\right)^2\right]\right| \xrightarrow[n \to \infty]{a.s.} 0.
\end{aligned}$$

Let $\mathbb{G}_1 = \left\{ w \in \mathbb{R}^{p+N} \longmapsto g^{(1)}_{\theta^{(1)}}(w); \theta^{(1)} \in \Theta_1 \subset \mathbb{R}^{p+1}\right\}$ be a collection of measurable functions indexed by a bounded set Θ_1 such that $g^{(1)}_{\theta^{(1)}}$ is defined as follows:

$$\begin{aligned}
g^{(1)}_{\theta^{(1)}}(W_i) &= g^{(1)}_{\theta^{(1)}}\left(X_i^{(1)}, \ldots, X_i^{(p)}, D^{(i)}_{X_i}(t_1), \ldots, D^{(i)}_{X_i}(t_N)\right) \\
&= \frac{1}{N} \sum_{j=1}^N \left(\Delta D^{(i)}_{X_i}\left(t_j\right) - m^{(1)}_{\theta^{(1)}}(X_i)\right)^2 \\
&= \frac{1}{N} \sum_{j=1}^N \left(D^{(i)}_{X_i}\left(t_j\right) - D^{(i)}_{X_i}\left(t_{j-1}\right) - \frac{\gamma T}{N} e^{\beta^T X_i}\right)^2,
\end{aligned}$$

where $W_i = \left(X_i^{(1)}, \ldots, X_i^{(p)}, D_{X_i}^{(i)}(t_1), \ldots, D_{X_i}^{(i)}(t_N)\right)$, for any $i \in \{1, \ldots, n\}$, are i.i.d. random vectors in $\mathbb{R}^{p+N}$. Thus our goal is to check that

$$\sup_{\theta^{(1)} \in \Theta_1} \left| \frac{1}{n} \sum_{i=1}^{n} g_{\theta^{(1)}}^{(1)}(W_i) - \mathbb{E}\left[g_{\theta^{(1)}}^{(1)}(W_i)\right] \right| \xrightarrow[n\to\infty]{a.s.} 0, \tag{7}$$

where

$$\begin{aligned}
\mathbb{E}\left[g_{\theta^{(1)}}^{(1)}(W_i)\right] &= \int_{\mathbb{R}^p} \frac{1}{N} \sum_{j=1}^{N} \mathbb{E}_{\theta_0} \left(\Delta D_x^{(i)}\left(t_j\right) - m_{\theta^{(1)}}^{(1)}(x)\right)^2 f_X(x)\, \mathrm{d}\mu_p(x) \\
&= \int_{\mathbb{R}^p} \mathbb{E}_{\theta_0} \left(\Delta D_x\left(\frac{T}{N}\right) - m_{\theta^{(1)}}^{(1)}(x)\right)^2 f_X(x)\, \mathrm{d}\mu_p(x).
\end{aligned}$$

We note that the convergence in (7) is true whenever $\mathbb{G}_1$ is Glivenko-Cantelli (van der Vaart and Wellner (1996) or van der Vaart (1998)). Note that we have the following decomposition:

$$\begin{aligned}
\left|g_{\theta^{(1)}}^{(1)}(W_i)\right| &= \frac{1}{N} \sum_{j=1}^{N} \left(\Delta D_{X_i}^{(i)}\left(t_j\right) - \frac{\gamma T}{N} e^{\beta^T X_i}\right)^2 \\
&= \left| \frac{1}{N} \sum_{j=1}^{N} \left(\Delta D_{X_i}^{(i)}\left(t_j\right)\right)^2 - \frac{2}{N} \sum_{j=1}^{N} \left(\Delta D_{X_i}^{(i)}\left(t_j\right)\right) \frac{\gamma T}{N} e^{\beta^T X_i} + \left(\frac{\gamma T}{N} e^{\beta^T X_i}\right)^2 \right| \\
&\leq C_1(W_i) + C_2(W_i)\, \gamma e^{\beta^T X_i} + C_3 \left(\gamma e^{\beta^T X_i}\right)^2,
\end{aligned}$$

where $C_1(W_i)$ and $C_2(W_i)$ only depend on the last N components of W_i and $C_3 = T/N$. Moreover we have that for any $w \in \mathbb{R}^{p+N}$, $\theta^{(1)} \longmapsto g_{\theta^{(1)}}^{(1)}(w)$ is a continuous function. To prove that $\mathbb{G}_1$ is Glivenko-Cantelli it is sufficient to show, using Example 19.8 in van der Vaart (1998), that $\mathbb{G}_1$ has an integrable envelope function G_1, that is:

- $\sup_{\theta^{(1)} \in \Theta_1} |g_{\theta^{(1)}}(w)| \leq G_1(w)$ for all $w \in \mathbb{R}^{p+N}$,
- $\mathbb{E}[G_1(W_1)] < +\infty$.

In this case, for $\gamma \leq \Upsilon$, we have

$$\forall w \in \mathbb{R}^{p+N}, G_1(w) = C_1(w) + C_2(w)\, \Upsilon e^{\|\beta\| \times \|x\|} + C_3 \Upsilon^2 e^{2\|\beta\| \times \|x\|},$$

where x is the p-component sub-vector of w. Because by (A_1) $\theta^{(1)} = (\gamma, \beta)$ belongs to the compact set Θ_1 and since by (A_2) we have $\| X_i \| \leq \kappa$ almost surely, we obtain

$$\mathbb{E}[G_1(W_1)] \leq \left(\frac{\gamma^2 T}{N\alpha} e^{\kappa\|\beta\|} + \left(\frac{\gamma T}{N} e^{\kappa\|\beta\|}\right)^2 + \tau^2 \frac{T}{N}\right) + \Upsilon \left(\frac{\gamma T}{N} e^{\kappa\|\beta\|}\right) e^{\kappa\|\beta\|} + C_3 \Upsilon^2 e^{2\kappa\|\beta\|} < +\infty.$$

Thus it follows that $\mathbb{G}_1$ is Glivenko-Cantelli. By similar arguments we prove also the almost sure convergence of $\sup_{\theta \in \Theta} \left|d_2^{(n)}(\theta) - d_2(\theta)\right|$ to zero. Finally, applying Lemma 1 we obtain that $\hat{\theta}_n \xrightarrow{Pr} \theta_0$ as $n \to +\infty$. □

Remark 1. From Equation (6), the parameter β_0 must be different from zero. Indeed if $\beta_0 = 0$ both γ and β can be identified by (5), however Equation (6) becomes

$$\frac{\gamma_0^2}{\alpha_0} - \frac{\gamma_0^2}{\alpha} = \tau^2 - \tau_0^2,$$

which is satisfied by many couples $\left(\alpha, \tau^2\right)$. Hence if $\beta_0 = 0$ we lose the identifiability of α and τ^2.

4.2. Asymptotic normality

We now provide the result about asymptotic normality of our estimator. First let us remark that the matrix of partial derivatives of $d_1^{(n)}$ at $\theta_0^{(1)}$ and of $d_2^{(n)}$ at θ_0 can be viewed as a sum of independent random vectors

$$\begin{pmatrix} \frac{\partial d_1^{(n)}}{\partial \theta^{(1)}}\left(\theta_0^{(1)}\right) \\ \frac{\partial d_2^{(n)}}{\partial \theta^{(2)}}\left(\theta_0\right) \end{pmatrix} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\begin{pmatrix} -2TN^{-2}e^{\beta_0^T x_i}\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta^{(1)}}^{(1)}\left(x_i\right)\right) \\ -2TN^{-2}\gamma_0 x_i e^{\beta_0^T x_i}\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta^{(1)}}^{(1)}\left(x_i\right)\right) \\ 2TN^{-2}\gamma_0^2\alpha_0^{-2}e^{\beta_0^T x_i}\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta^{(2)}}^{(2)}\left(x_i,\theta^{(1)}\right)\right) \\ -2TN^{-2}\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta^{(2)}}^{(2)}\left(x_i,\theta^{(1)}\right)\right) \end{pmatrix} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}V_{x_i,j}.$$

For any $n \geq 1$, we denote by:

$$I_n^{(1)}\left(\theta\right) = \frac{\partial^2 d_1^{(n)}}{\partial\theta^{(1)}\partial\theta^{(1)T}}\left(\theta^{(1)}\right), \quad I_n^{(2)}\left(\theta\right) = \frac{\partial^2 d_2^{(n)}}{\partial\theta^{(2)}\partial\theta^{(2)T}}\left(\theta^{(1)},\theta^{(2)}\right)$$

and

$$I_n^{(3)}\left(\theta\right) = \frac{\partial^2 d_2^{(n)}}{\partial\theta^{(1)}\partial\theta^{(2)T}}\left(\theta^{(1)},\theta^{(2)}\right),$$

three matrices. For $1 \leq k \leq 3$, we set $I_\infty^{(k)}\left(\theta\right) = \lim_{n\to\infty} I_n^{(k)}\left(\theta\right)$ (provided it exists) and we denote by

$$I_\infty\left(\theta_0\right) = \begin{pmatrix} I_\infty^{(1)}\left(\theta_0\right) & 0 \\ I_\infty^{(3)}\left(\theta_0\right) & I_\infty^{(2)}\left(\theta_0\right) \end{pmatrix},$$

which is defined in Appendix A.

Theorem 3. *If $(A_1 - A_3)$ and the following assumptions are fulfilled*

- (A_4) *for any $\theta \in \Theta$, $I_\infty\left(\theta\right)$ is invertible,*
- (A_5) *for $\rho_1 \in \{0,1,2,4\}$ and $\rho_2 \in \{1,\ldots,7\}$, there exist $\epsilon > 0$ and deterministic functions E_{ρ_1,ρ_2} such that:*

$$\sup_{\beta\in B(\beta_0,\epsilon)}\left|\frac{1}{n}\sum_{i=1}^{n} x_i^{\otimes\rho_1} e^{\rho_2\beta^T x_i} - E_{\rho_1,\rho_2}\left(\beta\right)\right| \xrightarrow[n\longrightarrow\infty]{} 0,$$

where

$$x_i^{\otimes\rho_1} = \begin{cases} 1 & si \quad \rho_1 = 0, \\ x_i & si \quad \rho_1 = 1, \\ x_i x_i^T & si \quad \rho_1 = 2, \\ x_i x_i^T x_i^{\otimes 2} & si \quad \rho_1 = 4. \end{cases}$$

then we have

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow[n\longrightarrow\infty]{d} N\left(0,\mathcal{M}\right),$$

where the variance-covariance matrix $\mathcal{M}$ is defined by

$$\mathcal{M} = \left(I_\infty^{-1}\left(\theta_0\right)\right)^T \Sigma^{(\infty)} I_\infty^{-1}\left(\theta_0\right),$$

and where $\Sigma^{(\infty)}$, defined in Appendix A, is the limit of $\Sigma^{(n)} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N} Cov\left(V_{x_i,j}\right)$ when n tends to infinity.

Proof. Applying a first order Taylor's expansion to $\partial d_1^{(n)}/\partial\theta^{(1)}$ at $\hat{\theta}_n^{(1)}$, we obtain that

$$\frac{\partial d_1^{(n)}}{\partial\theta^{(1)}}\left(\hat{\theta}_n^{(1)}\right) = 0_{\mathbb{R}^{p+1}} = \frac{\partial d_1^{(n)}}{\partial\theta^{(1)}}\left(\theta_0^{(1)}\right) + \frac{\partial^2 d_1^{(n)}}{\partial\theta^{(1)}\partial\theta^{(1)\,T}}\left(\overline{\theta}_n^{(1)}\right)\left(\hat{\theta}_n^{(1)} - \theta_0^{(1)}\right),$$

where $\overline{\theta}_n^{(1)}$ belongs to the line segment with extremities $\hat{\theta}_n^{(1)}$ and $\theta_0^{(1)}$. Similarly, applying Taylor's expansion to $\partial d_2^{(n)} / \partial \theta^{(2)}$ at $\hat{\theta}_n$, the vector of partial derivatives of $d_2^{(n)}$ with respect to $\theta^{(2)}$ at $\hat{\theta}_n$ is given by:

$$\frac{\partial d_2^{(n)}}{\partial \theta^{(2)}}\left(\hat{\theta}_n\right) = 0_{\mathbb{R}^2} = \frac{\partial d_2^{(n)}}{\partial \theta^{(2)}}\left(\theta_0\right) + \frac{\partial^2 d_2^{(n)}}{\partial \theta^{(2)} \partial \theta^{(2)\,T}}\left(\tilde{\theta}_n\right)\left(\hat{\theta}_n^{(2)} - \theta_0^{(2)}\right) + \frac{\partial^2 d_2^{(n)}}{\partial \theta^{(1)} \partial \theta^{(2)\,T}}\left(\tilde{\theta}_n\right)\left(\hat{\theta}_n^{(1)} - \theta_0^{(1)}\right),$$

where $\tilde{\theta}_n$ belongs to the line segment with extremities $\hat{\theta}_n$ and θ_0.

Thus one gets that

$$\sqrt{n}\begin{pmatrix} I_n^{(1)}\left(\overline{\theta}_n^{(1)}\right) & 0 \\ I_n^{(3)}\left(\tilde{\theta}_n\right) & I_n^{(2)}\left(\tilde{\theta}_n\right) \end{pmatrix}\begin{pmatrix} \hat{\theta}_n^{(1)} - \theta_0^{(1)} \\ \hat{\theta}_n^{(2)} - \theta_0^{(2)} \end{pmatrix} = -\sqrt{n}\begin{pmatrix} \dfrac{\partial d_1^{(n)}}{\partial \theta^{(1)}}\left(\theta_0^{(1)}\right) \\ \dfrac{\partial d_2^{(n)}}{\partial \theta^{(2)}}\left(\theta_0\right) \end{pmatrix}.$$

Next, using Lindeberg-Feller theorem, we prove that $n^{1/2}\partial d^{(n)}\left(\theta_0\right)/\partial\theta$ satisfies:

$$\sqrt{n}\begin{pmatrix} \dfrac{\partial d_1^{(n)}}{\partial \theta^{(1)}}\left(\theta_0^{(1)}\right) \\ \dfrac{\partial d_2^{(n)}}{\partial \theta^{(2)}}\left(\theta_0\right) \end{pmatrix} \xrightarrow[n\longrightarrow\infty]{d} N\left(0, \Sigma^{(\infty)}\right). \tag{8}$$

To prove this asymptotic normality, we must check the two conditions of the Lindeberg-Feller theorem (see e.g. van der Vaart (1998)). Let us recall the first condition of this theorem:

$$\forall \epsilon > 0,\ \frac{n}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(V_{x_i,j}^{(1)}\right)^2 + \| V_{x_i,j}^{(2)}\|_2^2 + \left(V_{x_i,j}^{(3)}\right)^2 + \left(V_{x_i,j}^{(4)}\right)^2\right)1_{\left\{\sqrt{\left(V_{x_i,j}^{(1)}\right)^2 + \|V_{x_i,j}^{(2)}\|_2^2 + \left(V_{x_i,j}^{(3)}\right)^2 + \left(V_{x_i,j}^{(4)}\right)^2} > \epsilon\sqrt{n}\right\}}\right] \xrightarrow[n\longrightarrow\infty]{} 0.$$

We have:

$$\begin{aligned}
&\left(V_{x_i,j}^{(1)}\right)^2 + \| V_{x_i,j}^{(2)}\|_2^2 + \left(V_{x_i,j}^{(3)}\right)^2 + \left(V_{x_i,j}^{(4)}\right)^2 \\
&\qquad = \frac{4T^2}{N^4}\left(1 + \gamma_0^2\|x_i\|_2^2\right)e^{2\beta_0^T x_i}\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}\left(x_i\right)\right)^2 \\
&\qquad\quad + \frac{4T^2}{N^4}\left(1 + \frac{\gamma_0^4}{\alpha_0^4}e^{2\beta_0^T x_i}\right)\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i, \theta^{(1)}\right)\right)^2,
\end{aligned}$$

Then it follows that for any $\epsilon > 0$

$$\begin{aligned}
&\frac{n}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(V_{x_i,j}^{(1)}\right)^2 + \| V_{x_i,j}^{(2)}\|_2^2 + \left(V_{x_i,j}^{(3)}\right)^2 + \left(V_{x_i,j}^{(4)}\right)^2\right)1_{\left\{\sqrt{\left(V_{x_i,j}^{(1)}\right)^2 + \|V_{x_i,j}^{(2)}\|_2^2 + \left(V_{x_i,j}^{(3)}\right)^2 + \left(V_{x_i,j}^{(4)}\right)^2} > \epsilon\sqrt{n}\right\}}\right] \\
\leq\ &\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left[c_1\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}\left(x_i\right)\right)^2 + c_2\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i, \theta^{(1)}\right)\right)^2\right]\right. \\
&\qquad \left. \times\, 1_{\left\{\left|\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}\left(x_i\right)\right| > \epsilon\sqrt{\frac{n}{2c_1}}\right\}}\right] \\
&+\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left[c_1\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}\left(x_i\right)\right)^2 + c_2\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i, \theta^{(1)}\right)\right)^2\right]\right. \\
&\qquad \left. \times\, 1_{\left\{\left|\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i, \theta^{(1)}\right)\right| > \epsilon\sqrt{\frac{n}{2c_2}}\right\}}\right] \\
=\ &B_1^{(n)} + B_2^{(n)}
\end{aligned}$$

where, using Assumption (A_2), there exist two constants $C_{0,2}(\beta_0)$ and $C_{2,2}(\beta_0)$ such that $e^{2\beta_0^T x_i} \leq C_{0,2}(\beta_0)$ and $x_i^{\otimes 2} e^{2\beta_0^T x_i} \leq C_{2,2}(\beta_0)$, then c_1 and c_2 are equal to:

$$c_1 = \frac{4T^2}{N^4}\left(C_{0,2}(\beta_0) + \gamma_0^2 C_{2,2}(\beta_0)\right) \text{ and } c_2 = \frac{4T^2}{N^4}\left(1 + \frac{\gamma_0^4}{\alpha_0^4} C_{0,2}(\beta_0)\right).$$

Since the Lyapunov condition implies the Lindeberg condition we prove in the sequel that $B_1^{(n)} + B_2^{(n)}$ tends to 0 as n tends to infinity. First it is shown in the proof of Lemma 4 in Bordes et al. (2010) that $\mathbb{E}\left[\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^q\right] = \mathrm{Pol}_q\left(\Delta t_j\right)$ where Pol_q denotes a polynomial of order q with respect to Δt_j, the coefficients of which depend only on θ. Here the process $D_{x_i}^{(i)}$ is observed at regular instants. Then

$$\mathbb{E}_\theta\left[\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^q\right] = \mathrm{Pol}_q\left(\frac{T}{N} e^{\beta^T x_i}\right).$$

Using Assumption (A_5) it follows that

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^q\right] = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathrm{Pol}_q\left(\frac{T}{N} e^{\beta^T x_i}\right) \tag{9}$$

tends to a constant independent of i. Next, for any $\epsilon > 0$, we have:

$$\begin{aligned}
B_1^{(n)} &\leq \frac{c_1\sqrt{2c_1}}{\epsilon\, n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left|\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}(x_i)\right|^3\right] \\
&\quad + \frac{c_2\sqrt{2c_1}}{\epsilon\, n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right)^2\left|\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}(x_i)\right|\right] \\
&\leq \frac{c_1\sqrt{2c_1}}{\epsilon\, n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}(x_i)\right)^2\left(\Delta D_{x_i}^{(i)}\left(t_j\right) + m_{\theta_1}^{(1)}(x_i)\right)\right] \\
&\quad + \frac{c_2\sqrt{2c_1}}{\epsilon\, n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right)^2\left(\Delta D_{x_i}^{(i)}\left(t_j\right) + m_{\theta_1}^{(1)}(x_i)\right)\right]
\end{aligned}$$

which tends to 0 as n tends to infinity. Indeed, by Assumption (A_5) and Equality (9), we obtain that:

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}(x_i)\right)^2\left(\Delta D_{x_i}^{(i)}\left(t_j\right) + m_{\theta_1}^{(1)}(x_i)\right)\right]$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right)^2\left(\Delta D_{x_i}^{(i)}\left(t_j\right) + m_{\theta_1}^{(1)}(x_i)\right)\right]$$

tend to constants independent of i. Thus, one deduces that $B_1^{(n)} \xrightarrow[n\longrightarrow\infty]{} 0$. Similarly, for any $\epsilon > 0$, we have

$$\begin{aligned}
B_2^{(n)} &\leq \frac{c_1\sqrt{2c_2}}{\epsilon\, n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left|\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right|\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}(x_i)\right)^2\right] \\
&\quad + \frac{c_2\sqrt{2c_2}}{\epsilon\, n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left|\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right|^3\right] \\
&\leq \frac{c_1\sqrt{2c_2}}{\epsilon\, n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 + m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right)\left(\Delta D_{x_i}^{(i)}\left(t_j\right) - m_{\theta_1}^{(1)}(x_i)\right)^2\right] \\
&\quad + \frac{c_2\sqrt{2c_2}}{\epsilon\, n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 - m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right)^2\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2 + m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right]
\end{aligned}$$

which tends to 0 as n tends to infinity. Indeed, by Assumption (A_5) and Equality (9), it follows that:

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2+m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right)\left(\Delta D_{x_i}^{(i)}\left(t_j\right)-m_{\theta_1}^{(1)}\left(x_i\right)\right)^2\right]$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\mathbb{E}_\theta\left[\left(\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2-m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right)^2\left(\Delta D_{x_i}^{(i)}\left(t_j\right)\right)^2+m_{\theta_2}^{(2)}\left(x_i,\theta^{(1)}\right)\right]$$

tend to constants independent of i. Thus, one deduces that $B_2^{(n)}\xrightarrow[n\longrightarrow\infty]{}0$.

Next by painful and straightforward calculations, using Assumption (A_5), we check the second condition of Lindeberg-feller Theorem:

$$\Sigma^{(n)}=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N}\operatorname{Cov}\left(V_{x_i,j}\right)\xrightarrow[n\longrightarrow\infty]{}\Sigma^{(\infty)}$$

where the entries $\sigma_{uv}^{(\infty)}$ $(1\le u\le v\le 4)$ of the variance-covariance matrix $\Sigma^{(\infty)}$ are given in Appendix A.

Finally, we have:

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_n^{(1)}-\theta_0^{(1)}\\ \hat{\theta}_n^{(2)}-\theta_0^{(2)}\end{pmatrix} = -\begin{pmatrix}I_n^{(1)}\left(\overline{\theta}_n^{(1)}\right) & 0\\ I_n^{(3)}\left(\tilde{\theta}_n\right) & I_n^{(2)}\left(\tilde{\theta}_n\right)\end{pmatrix}^{-1}\sqrt{n}\begin{pmatrix}\dfrac{\partial d_1^{(n)}}{\partial\theta^{(1)}}\left(\theta_0^{(1)}\right)\\ \dfrac{\partial d_2^{(n)}}{\partial\theta^{(2)}}\left(\theta_0\right)\end{pmatrix}. \tag{10}$$

Because of Theorem 2 we have $\overline{\theta}_n^{(1)}\xrightarrow{Pr}\theta_0^{(1)}$ and $\tilde{\theta}_n\xrightarrow{Pr}\theta_0$, then by (A_5) we have:

$$\lim_{n\longrightarrow\infty}\begin{pmatrix}I_n^{(1)}\left(\overline{\theta}_n^{(1)}\right) & 0\\ I_n^{(3)}\left(\tilde{\theta}_n\right) & I_n^{(2)}\left(\tilde{\theta}_n\right)\end{pmatrix}=\begin{pmatrix}I_\infty^{(1)}\left(\theta_0\right) & 0\\ I_\infty^{(3)}\left(\theta_0\right) & I_\infty^{(2)}\left(\theta_0\right)\end{pmatrix}=I_\infty\left(\theta_0\right),$$

which is invertible by (A_4). Thus, using Equality (10), Slutsky lemma, Assumption (A_5) and (8), it holds that

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_n^{(1)}-\theta_0^{(1)}\\ \hat{\theta}_n^{(2)}-\theta_0^{(2)}\end{pmatrix}\xrightarrow[n\longrightarrow\infty]{d}N\left(0,\mathcal{M}\right),$$

where $\mathcal{M}=\left(I_\infty^{-1}\left(\theta_0\right)\right)^T\Sigma^{(\infty)}I_\infty^{-1}\left(\theta_0\right)$ (entries of the matrices $\Sigma^{(\infty)}$ and $I_\infty\left(\theta_0\right)$ are given in Appendix A). □

Remark 2. The matrix $\mathcal{M}$ can be estimated by the matrix

$$\widehat{\mathcal{M}}=\left(\widehat{I_\infty}^{-1}\right)^T\widehat{\Sigma}^{(\infty)}\widehat{I_\infty}^{-1},$$

where the coefficients of matrices $\widehat{I_\infty}$ and $\widehat{\Sigma}^{(\infty)}$ are calculated by estimating $E_{\rho_1,\rho_2}\left(\beta_0\right)$ by $n^{-1}\sum_{i=1}^{n}x_i^{\otimes\rho_1}e^{\rho_2\hat{\beta}_n^T x_i}$ and $\theta_0^T=\left(\gamma_0,\beta_0^T,\alpha_0,\tau_0^2\right)$ by $\hat{\theta}_n^T=\left(\hat{\gamma}_n,\hat{\beta}_n^T,\hat{\alpha}_n,\hat{\tau}_n^2\right)$. Another possibility is to estimate $\Sigma^{(\infty)}$ by the empirical variance-covariance matrix of vectors $V_{x_i,j}$ (where θ is replaced by $\hat{\theta}_n$).

Remark 3. In the proof of the above theorem, we obtained asymptotic normality of $n^{1/2}\partial d^{(n)}\left(\theta_0\right)/\partial\theta$ conditional on covariates X_i, by applying the Lindeberg-Feller theorem. It is a little bit more than necessary since Theorem 3 remains valid in case we proceed unconditionally on the X_i's. Indeed, assuming the X_i's i.i.d., Assumption (A_5) holds since by classical empirical process approach we get:

$$\sup_{\beta\in B(\beta_0,\epsilon)}\left|\frac{1}{n}\sum_{i=1}^{n}x_i^{\otimes\rho_1}e^{\rho_2\beta^T x_i}-E_{\rho_1,\rho_2}\left(\beta\right)\right|\xrightarrow[n\longrightarrow\infty]{}0.$$

Thus using Assumptions $(A_1) - (A_4)$ mentioned in Theorem 3, the two conditions of the Lindeberg-Feller theorem are fulfilled, leading to:

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_n^{(1)} - \theta_0^{(1)} \\ \hat{\theta}_n^{(2)} - \theta_0^{(2)}\end{pmatrix} \xrightarrow[n\longrightarrow\infty]{d} N(0, \mathcal{M}).$$

5. Numerical illustration and concluding remarks

Bias and MSE.

Here we illustrate our theoretical results through Monte Carlo simulations. The model under consideration is the same as the one we used in Section 2 to simulate a few trajectories (see Figure 1). We recall that the model parameters were setted to: $\xi = 1$, $\alpha = 2$, $\beta = (-0.5, 0, 0.4)$ and $\tau^2 = 1$. We notice that one coordinate of β has to be fixed to a given value in order to ensure identifiability. A classical choice here is to set $\beta_2 = 0$ since it corresponds to the medium stress level. The number of observations for each item was set to $N = 10$ instants between 10 and $T = 100$. We have computed the empirical bias and the empirical mean squared error (MSE) for 1000 repetitions. Table 1 and Table 2 report respectively the empirical bias and the empirical MSE for several sample sizes n. During simulations we have reported a few cases where the convergence of the estimation method failed (see NCC: Non Converging Cases). We have then calculated the empirical bias and the empirical MSE deleting these cases. Based on results given in the Tables 1 and 2 and on other results not shown here, we note that the percentage of the NCC decreases when n increases. However, the situation is better when considering the estimation of γ. It means that the average of degradation is well estimated whatever the sample size. It is also the case for the parameters β (covariates) and τ (Brownian motion part). Finally, based on our simulation studies, the larger is n, the better are the estimation results towards the bias and the MSE.

Table 1: Empirical bias

n	20	50	100	200
γ	4.02e-4	-6.34e-4	8.55e-4	4.51e-4
β_1	-1.13e-3	-2.47e-4	1.12e-3	-6.78e-4
β_3	-3.56e-4	-5.61e-4	-7.44e-4	-2.69e-4
α	4.81e-1	3.08e-1	1.63e-1	8.66e-2
τ^2	-5.92e-2	3.31e-2	1.88e-3	-2.47e-3
ξ	2.41e-1	1.54e-1	8.11e-2	4.32e-2
NCC	4.2%	1.3%	1.1%	0.8%

Table 2: Empirical MSE

n	20	50	100	200
γ	4.91e-3	2.17e-3	9.48e-4	5.91e-4
β_1	3.58e-3	1.55e-3	7.02e-4	4.38e-4
β_3	2.09e-3	8.55e-4	3.95e-4	2.44e-4
α	7.62e-1	6.42e-1	3.24e-1	1.76e-1
τ^2	4.35e-1	2.99e-1	1.71e-1	1.08e-1
ξ	1.94e-1	1.61e-1	8.06e-2	4.41e-2
NCC	10.2%	3.2%	2.2%	1.2%

Asymptotic confidence intervals.

Table 3 summarizes an example of results obtained with 1000 simulations for $n = 100, N = 10, T = 10, \xi = 1, \alpha = 2, \tau^2 = 1, \beta_1 = -0,5, \beta_2 = 0$ and $\beta_3 = 0,4$. We estimated the model parameters and we constructed a 95% confidence intervals for each model parameter. We observe that the true values of each model parameters belong to the limit of the confidence regions.

Table 3: Parameter estimation and 95% confidence interval

Parameters	γ	β_1	β_3	α	τ^2
Estimation	$2,006$	$-0,494$	$0,395$	$2,034$	$0,991$
Confidence interval	$[0; 4,36]$	$[-1,16; 0,17]$	$[0,03; 0,75]$	$[1.81; 2,26]$	$[0,91; 1,06]$

Non equidistant observation times.
For the model we introduced, we proposed a methodology to estimate the parameters. The procedure consists in a two-stage least-square method, and to avoid to much complexity items were assumed to be observed at regular instants. Of course, our methodology based on the two-stage least-square method can also be applied to items observed at irregular instants t_{ij} where $0 \le j \le N_i$ and $1 \le i \le n$ with i indexes items and N_i is the number of observation times for the ith item. Indeed, to estimate $\theta^{(1)} = (\gamma, \beta)$, we can minimize the following first regression function

$$d_1^{(n)}\left(\theta^{(1)}\right) = \sum_{i=1}^{n}\sum_{j=1}^{N_i}\left(\Delta D_{x_i}^{(i)}\left(t_{ij}\right) - m_{ij}^{(1)}\left(x_i, \theta^{(1)}\right)\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{N_i}\left(\Delta D_{x_i}^{(i)}\left(t_{ij}\right) - \gamma \Delta t_{ij} e^{\beta^T x_i}\right)^2,$$

where $\Delta D_{x_i}^{(i)}\left(t_{ij}\right) = D_{x_i}^{(i)}\left(t_{ij}\right) - D_{x_i}^{(i)}\left(t_{ij-1}\right)$. Then we set $\hat{\theta}_n^{(1)} = \underset{\theta^{(1)} \in \Theta_1}{\arg\min}\, d_1^{(n)}\left(\theta^{(1)}\right)$. Once this parameter is estimated, we can estimate $\theta^{(2)} = \left(\alpha, \tau^2\right)$ by minimizing the following second regression function

$$\begin{aligned} d_2^{(n)}\left(\theta^{(2)}, \hat{\theta}_n^{(1)}\right) &= \sum_{i=1}^{n}\sum_{j=1}^{N_i}\left(\left[\Delta D_{x_i}^{(i)}\left(t_{ij}\right)\right]^2 - m_{ij}^{(2)}\left(x_i, \hat{\theta}_n^{(1)}, \theta^{(2)}\right)\right)^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{N_i}\left(\left[\Delta D_{x_i}^{(i)}\left(t_{ij}\right)\right]^2 - m_{ij}^{(1)}\left(x_i, \hat{\theta}_n^{(1)}\right)\frac{\hat{\gamma}}{\alpha} - \left(m_{ij}^{(1)}\left(x_i, \hat{\theta}_n^{(1)}\right)\right)^2 - \tau^2 \Delta t_{ij}\right)^2. \end{aligned}$$

As a consequence $\theta^{(2)}$ is estimated by $\hat{\theta}_n^{(2)} = \underset{\theta^{(2)} \in \Theta_2}{\arg\min}\, d_2^{(n)}\left(\theta^{(2)}, \hat{\theta}_n^{(1)}\right)$. The final estimator of θ is $\hat{\theta}_n = \left(\hat{\theta}_n^{(1)}, \hat{\theta}_n^{(2)}\right)$.

Illustrative application.
An example of degradation data can be found on-line, on the website of the National Institute of Standards and Technology (NIST/SEMATECH (2010)). In this dataset, fifteen components were tested under three different temperatures 65 °C, 85 °C and 105 °C corresponding respectively to regression parameters $\beta_1, \beta_2 = 0$ and β_3. Degradation percent values were read out at 200, 500 and 1000 hours. We first estimate the model parameters without integrating the covariate, using the method of moments proposed in Bordes et al. (2010). Then parameters of the model with covariate are estimated using the methodology we proposed (two-stage least-square method). Table 4 summarizes results.

Table 4: Pointwise estimation for the NIST dataset

Parameters	without covariate	with covariate
α	2.403e-3	2.098e-3
ξ	7.599e-2	1.057e-1
τ^2	7.671e-2	6.191e-2
β_1		-1.251
β_3		0.905

Table 5: Average degradation per unit of time

low stress	medium stress	high stress
5.678e-3	1.985e-2	4.907e-2

As one can expect, we obtained that $\hat{\beta}_1 < 0 < \hat{\beta}_3$. This is reasonable, since the interpretation of levels of the covariate. The estimation of τ has decreased from the model without covariate to the one including covariate. Indeed variability of the observations has been taken into account also in the covariate, implying a less important role of the Brownian motion. Let us remark also, Table 5, that $\hat{\alpha}e^{\hat{\beta}_1}/\hat{\xi} < \hat{\alpha}/\hat{\xi} < \hat{\alpha}e^{\hat{\beta}_3}/\hat{\xi}$ which imply that the degradation mean increases with respect to covariates.

However we must be careful when dealing with asymptotic properties (consistency and asymptotic normality). These asymptotic properties require stability conditions of the same type as those given in Bordes et al. (2010). Finally we have to mention that we can construct asymptotic confidence intervals of any regular function of the parameters using the δ-method and the estimated variance-covariance matrix mentioned in Remark 2. For example, confidence intervals may be calculated for α/ξ^2, τ^2 allowing to test the gamma process with covariates versus the Brownian motion with positive drift driven by covariates as in Bordes et al. (2010). Classical chi-square type statistics can also be used in order to test significance (of subset) of covariates.

Appendix A. Calculations results

The variance-covariance matrix $\Sigma^{(\infty)}$ is defined by:

$$
\begin{aligned}
\sigma_{11}^{(\infty)} &= \frac{4T^3\gamma_0^2}{N^4\alpha_0}E_{0,3}(\beta_0) + \frac{4T^3\tau_0^2}{N^4}E_{0,2}(\beta_0),\\
\sigma_{12}^{(\infty)} &= \frac{4T^3\gamma_0^3}{N^4\alpha_0}E_{1,3}(\beta_0) + \frac{4T^3\gamma_0\tau_0^2}{N^4}E_{1,2}(\beta_0),\\
\sigma_{13}^{(\infty)} &= \left(\frac{8T^3\gamma_0^7}{N^4\alpha_0^6} + \frac{8T^4\gamma_0^5\tau_0^2}{N^5\alpha_0^4}\right)E_{0,3}(\beta_0) + \frac{8T^4\gamma_0^7}{N^5\alpha_0^5}E_{0,4}(\beta_0),\\
\sigma_{14}^{(\infty)} &= \left(\frac{8T^3\gamma_0^3}{N^4\alpha_0^2} + \frac{8T^4\gamma_0\tau_0^2}{N^5}\right)E_{0,2}(\beta_0) + \frac{8T^4\gamma_0^3}{N^5\alpha_0}E_{0,3}(\beta_0),\\
\sigma_{22}^{(\infty)} &= \frac{4T^3\gamma_0^4}{N^4\alpha_0}E_{2,3}(\beta_0) + \frac{4T^3\gamma_0^2\tau_0^2}{N^4}E_{2,2}(\beta_0),\\
\sigma_{23}^{(\infty)} &= \left(-\frac{8T^3\gamma_0^6}{N^4\alpha_0^4} - \frac{8T^4\gamma_0^4\tau_0^2}{N^5\alpha_0^2}\right)E_{1,3}(\beta_0) - \frac{8T^4\gamma_0^6}{N^5\alpha_0^3}E_{1,4}(\beta_0),\\
\sigma_{24}^{(\infty)} &= \left(\frac{8T^3\gamma_0^4}{N^4\alpha_0^2} + \frac{8T^4\gamma_0^2\tau_0^2}{N^5}\right)E_{1,2}(\beta_0) + \frac{8T^4\gamma_0^4}{N^5\alpha_0}E_{1,3}(\beta_0),\\
\sigma_{33}^{(\infty)} &= \frac{8T^4\gamma_0^4\tau_0^4}{N^5\alpha_0^4}E_{0,2}(\beta_0) + \left(\frac{16T^4\gamma_0^6\tau_0^2}{N^5\alpha_0^5} + \frac{24T^3\gamma_0^8}{N^4\alpha_0^7}\right)E_{0,3}(\beta_0) + \left(\frac{16T^5\gamma_0^6\tau_0^2}{N^6\alpha_0^4} + \frac{40T^4\gamma_0^8}{N^5\alpha_0^6}\right)E_{0,4}(\beta_0) + \frac{16T^5\gamma_0^8}{N^6\alpha_0^5}E_{0,5}(\beta_0),\\
\sigma_{34}^{(\infty)} &= -\frac{8T^4\gamma_0^2\tau_0^4}{N^5\alpha_0^2}E_{0,1}(\beta_0) + \left(-\frac{16T^4\gamma_0^4\tau_0^2}{N^5\alpha_0^3} - \frac{24T^3\gamma_0^6}{N^4\alpha_0^5}\right)E_{0,2}(\beta_0) + \left(-\frac{16T^5\gamma_0^4\tau_0^2}{N^6\alpha_0^2} - \frac{40T^4\gamma_0^6}{N^5\alpha_0^4}\right)E_{0,3}(\beta_0) - \frac{16T^5\gamma_0^6}{N^6\alpha_0^3}E_{0,4}(\beta_0),\\
\sigma_{44}^{(\infty)} &= \frac{8\tau_0^4T^4}{N^5} + \left(\frac{24T^3\gamma_0^4}{N^4\alpha_0^3} + \frac{16T^4\gamma_0^2\tau_0^2}{N^5\alpha_0}\right)E_{0,1}(\beta_0) + \left(\frac{16T^5\gamma_0^2\tau_0^2}{N^6} + \frac{40T^4\gamma_0^4}{N^5\alpha_0^2}\right)E_{0,2}(\beta_0) + \frac{16T^5\gamma_0^4}{N^6\alpha_0}E_{0,3}(\beta_0).
\end{aligned}
$$

The matrix $I_\infty(\theta_0)$ is defined by:

$$
\begin{pmatrix}
\begin{pmatrix}
\frac{2T^2}{N^2}E_{0,2}(\beta_0) & \frac{2T^2\gamma_0}{N^2}E_{1,2}(\beta_0)\\
\frac{2T^2\gamma_0}{N^2}E_{1,2}(\beta_0) & \frac{2T^2\gamma_0^2}{N^2}E_{2,2}(\beta_0)
\end{pmatrix}_{((p+1)\times(p+1))} & \begin{pmatrix} 0 & 0\\ 0 & 0\end{pmatrix}_{(2\times2)}\\
\begin{pmatrix}
-\frac{4\gamma_0^3T^2}{N^2\alpha_0^3}E_{0,2}(\beta_0) - \frac{4\gamma_0^3T^3}{N^3\alpha_0^2}E_{0,3}(\beta_0) & -\frac{2T^2\gamma_0^4}{N^2\alpha_0^3}E_{1,2}(\beta_0) - \frac{4T^3\gamma_0^4}{N^3\alpha_0^2}E_{1,3}(\beta_0)\\
\frac{4T^2\gamma_0}{N^2\alpha_0}E_{0,1}(\beta_0) + \frac{4T^3\gamma_0}{N^3}E_{0,2}(\beta_0) & \frac{2T^2\gamma_0^2}{N^2\alpha_0}E_{1,1}(\beta_0) + \frac{4T^3\gamma_0^2}{N^3}E_{1,2}(\beta_0)
\end{pmatrix}_{(2\times(p+1))} & \begin{pmatrix}
\frac{2\gamma_0^4T^2}{N^2\alpha_0^4}E_{0,2} & -\frac{2T^2\gamma_0^2}{N^2\alpha_0^2}E_{0,1}(\beta_0)\\
-\frac{2T^2\gamma_0^2}{N^2\alpha_0^2}E_{0,1}(\beta_0) & 0
\end{pmatrix}_{(2\times2)}
\end{pmatrix}.
$$

References

Abdel-Hameed, M., 1975. A gamma wear process. IEEE Transactions on Reliability 24 (2), 152–153.

Bagdonavičius, V., Nikulin, M. S., 2001. Estimation in degradation models with explanatory variables. Lifetime Data Analysis 7 (1), 85–103.

Barker, C., 2006. Maintenance policies to guarantee optimal performance of stochastically deteriorating multi-component systems. Ph.D. thesis, School of Enginerring and Mathematical Sciences.

Bordes, L., Paroissin, C., Salami, A., 2010. Parametric inference in a perturbed gamma degradation process. arXiv:1005.1214v1.

Crowder, M., Lawless, J., 2007. On a scheme for a predictive maintenance. European Journal of Operational Research 176 (3), 1713–1722.

Doksum, K., Hóyland, A., 1992. Models for variable-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution. Technometrics 34 (1), 74–82.

Lawless, J., Crowder, M., 2004. Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Analysis 10 (3), 213–227.

NIST/SEMATECH, 2010. e-Handbook of Statistical Methods. http://www.itl.nist.gov/div898/handbook/apr/section4/apr423.htm.

Park, C., Padgett, W. J., 2005. Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. Lifetime Data Analysis 11 (4), 511–527.

van der Vaart, A. W., 1998. Asymptotic Statistics. Cambridge University Press.

van der Vaart, A. W., Wellner, J., 1996. Weak Convergence and Empirical Processes: With Applications to Statistics. Springer Series in Statistics.

van Noortwijk, J. M., 2009. A survey of the application of gamma processes in maintenance. Reliability Engineering and System Safety 94 (1), 2–21.
Wang, X., 2010. Wiener processes with random effects for degradation data. Journal of Multivariate Analysis 101 (2), 340–351.
Whitmore, G., 1995. Estimating degradation by a Wiener diffusion process subject to measurement error. Lifetime Data Analysis 1 (3), 307–319.
Whitmore, G., Crowder, M., Lawless, J., 1998. Failure inference from a marker process based on a bivariate Wiener model. Lifetime Data Analysis 4 (3), 229–251.