# Estimation non-paramétrique dans des modèles d'équations différentielles stochastiques à effets aléatoires

Charlotte Dion [1,2] & Valentine Genon-Catalot [2]

[1] *LJK, UMR CNRS 5224, Université Joseph Fourier Grenoble, charlotte.dion@imag.fr*
[2] *MAP5, UMR CNRS 8145, Université Paris Descartes Paris*

**Résumé.** Nous étudions des modèles d'équations différentielles stochastiques à effets aléatoires dans le coefficient de dérive. En observant de manière continue $N$ trajectoires d'un processus donné, nous proposons une procédure pour estimer la densité commune des effets aléatoires, supposés indépendants et identiquement distribués. Les différences entre les observations sont alors dues au mouvement Brownien et à la réalisation des effets aléatoires, c'est pourquoi ces derniers font l'objet d'une étude approfondie. Dans un premier temps nous nous intéressons au modèle d'Ornstein-Uhlenbeck à un effet aléatoire puis nous étudions un modèle de diffusion général à deux effets aléatoires linéaires dans la dérive. Nous proposons dans les deux cas une procédure d'estimation des effets aléatoires et de leur densité : univariée ou bivariée.
**Mots-clés.** Équation différentielle stochastique, effet aléatoire, estimation non-paramétrique.

**Abstract.** We investigate stochastic differential models with random effects in the drift. Observing continuously $N$ trajectories with dynamics ruled by a given process, we aim at recovering the density of the random effects which are assumed to be independent and identically distributed. Differences between observations are due to the realization of a Brownian motion and the random effects, and this is why it is a matter of interest . First we deal with the special case of the Ornstein-Uhlenbeck model with one random effect, then we study a more general diffusion with two linear random effects in the drift. We propose in both cases an estimation procedure of the random effects and of their density : univariate or bivariate.
**Keywords.** Stochastic differential equation, random effects, nonparametric estimation.

## 1 Introduction

Lorsque l'on observe des potentiels membranaires de neurones entre deux spikes, on reconnaît une forme commune. Cependant il existe une variabilité claire entre observations qui peut être modélisée par des effets aléatoires. Le modèle d'équations différentielles stochastiques (EDS) d'Ornstein-Uhlenbeck avec un effet aléatoire additif est le modèle le plus simple utilisé pour décrire ces données. Nous nous y intéressons dans une première partie. Nous proposons un estimateur de la densité des effets aléatoires adaptatif obtenu à l'aide d'une méthode de Goldenshluger et Lepski [5] à partir d'un estimateur de Comte *et al.* [1]. Cet estimateur est implémenté sur données réelles en particulier, dans l'article Dion [2] et fournit une estimation non-paramétrique intéressante pour la modélisation de l'activité neuronale.

Puis dans Comte *et al.* [1] il est proposé une méthode d'estimation dans le cas d'un effet aléatoire multiplicatif dans la dérive. Nous proposons maintenant une méthode dans le cas d'une dérive dépendant de deux effets aléatoires, modèle comprenant les effets additifs et multiplicatifs. En particulier le modèle de Cox-Ingersoll-Ross qui est plus adapté pour décrire les données neuronales (par exemple) peut être étudié. Autoriser un coefficient de diffusion non constant ou non borné est une difficulté supplémentaire. La procédure d'estimation que nous proposons permet de traiter ce modèle en particulier. En effet en nous appuyant sur les travaux de Genon-Catalot and Larédo [4] nous estimons d'abord les effets aléatoires d'une nouvelle manière et nous obtenons un résultat sur le risque $\mathbb{L}^2$ de cet estimateur. Puis nous estimons leur densité commune, bivariée cette fois. La méthodologie complète est présentée dans Dion and Genon-Catalot [3] ainsi qu'une illustration numérique des résultats.

# 2 Modèle d'Ornstein-Uhlenbeck à un effet aléatoire

On observe $N$ trajectoires du processus d'Ornstein-Uhlenbeck à un effet aléatoire sur l'intervalle $[0, T]$ ($T$ connu) :

$$\begin{cases} dX_j(t) & = \left(\phi_j - \frac{X_j(t)}{\alpha}\right) dt + \sigma dW_j(t) \\ X_j(0) & = x_j \end{cases} \tag{1}$$

où $(W_j)_{1 \leq j \leq N}$ sont $N$ processus de Wiener indépendants, et $(\phi_j)_{1 \leq j \leq N}$ sont $N$ variables aléatoires réelles, non observées, et *i.i.d.* de densité commune $f$. Les suites $(\phi_j)_{1 \leq j \leq N}$ et $(W_j)_{1 \leq j \leq N}$ sont indépendantes et donc les $X_j(t)$, $j = 1, \ldots, N$ sont *i.i.d.* pour $t$ fixé. Les valeurs $(x_1, \ldots, x_N)$ sont connues. Les constantes strictement positives $\sigma$ et $\alpha$ sont supposées connues (et sont estimées en pratique).

## 2.1 Estimateur à noyau de la densité de l'effet aléatoire

Regardons les variables aléatoires suivantes, pour $j = 1, \ldots, N$ et $\tau \in ]0, T]$,

$$Z_{j,\tau} := \frac{X_j(\tau) - X_j(0) - \int_0^\tau (-\frac{X_j(s)}{\alpha} ds)}{\tau} = \phi_j + \frac{\sigma}{\tau} W_j(\tau). \tag{2}$$

Les $Z_{j,\tau}$ peuvent être considérés comme estimateur de $\phi_j$ et les $Z_{j,\tau} - \phi_j$ sont des variables centrées. Pour $\tau$ fixé, les $(Z_{j,\tau})_{j=1,\ldots,N}$ sont *i.i.d.* Pour réduire la variance de l'estimateur on fixe $\tau = T$. L'estimateur de la densité $f$ des $\phi_j$ résultant est, pour tout $x \in \mathbb{R}$,

$$\widehat{f}_h(x) = \frac{1}{N}\sum_{j=1}^{N} K_h(x - Z_{j,T}) \tag{3}$$

où $h > 0$ est la fenêtre et $K : \mathbb{R} \to \mathbb{R}$ un noyau $\mathcal{C}^2$ tel que $\int K(u)du = 1$, $\|K\|^2 = \int K^2(u)du < +\infty$, $\int (K''(u))^2 du < +\infty$, $K_h(x) = K(x/h)/h$. On définit $f_h(x) := K_h \star f(x) = \int K_h(x-y)f(y)dy$ et on note $\|f\|_p = (\int |f(x)|^p dx)^{1/p}$ et pour $p = 2$ $\|f\|_2 = \|f\|$. On rappelle le résultat suivant.

**Proposition 1.** *(Comte et al.[1]) L'estimateur $\widehat{f}_h$ donné par (3), vérifie*

$$\mathbb{E}[\|\widehat{f}_h - f\|^2] \leq 2\|f - f_h\|^2 + \frac{\|K\|^2}{Nh} + \frac{\sigma^4\|K''\|^2}{3T^2h^5}. \tag{4}$$

Le premier terme de la borne (4) est le terme de biais décroissant avec $h$. Le deuxième est un terme de variance qui croit quand $h$ décroit. Et le dernier terme est un terme d'erreur, due à l'approximation de $\phi_j$ par $Z_{j,T}$ qui décroit avec $h$. Nous allons voir maintenant comment réaliser un compromis entre ces termes pour choisir la meilleure fenêtre.

## 2.2 Sélection de la fenêtre

Le meilleur choix pour $h$ parmi $\mathcal{H}_{N,T}$, l'ensemble des fenêtres, est celui qui minimise la borne du risque (4), réalisant ainsi le compromis biais-variance. Cette borne est inconnue en pratique, et doit donc être estimée. On note $V$ un estimateur des deux termes de variance :

$$V(h) = \kappa_1 \frac{\|K\|_1^2\|K\|^2}{Nh} + \kappa_2 \frac{\sigma^4\|K\|_1^2\|K"\|^2}{3T^2h^5}$$

avec $\kappa_1$, $\kappa_2$ deux constantes numériques positives. On utilise le critère de Goldenshluger et Lepski [5] pour estimer le biais :

$$A(h) = \sup_{h' \in \mathcal{H}_{N,T}} \left(\|\widehat{f}_{h,h'} - \widehat{f}_{h'}\|^2 - V(h')\right)_+, \quad \widehat{f}_{h,h'}(x) := K_{h'} \star \widehat{f}_h(x) = \frac{1}{N}\sum_{j=1}^{N} K_{h'} \star K_h(x - Z_{j,T}).$$

La fenêtre sélectionnée est alors donnée par :

$$\widehat{h} = \underset{h \in \mathcal{H}_{N,T}}{\operatorname{argmin}} \left(A(h) + V(h)\right)$$

avec $\mathcal{H}_{N,T} = \{h > 0,\ \frac{1}{Nh} \leq 1, \frac{1}{h^5T^2} \leq 1,\ \forall c > 0, \exists\, \Sigma(c) < \infty,\ \sum_{h \in \mathcal{H}_{N,T}} h^{-1/2} e^{-c/\sqrt{h}} \leq \Sigma(c)\}$. On prouve alors le théorème suivant :

**Theorem 2.** *L'estimateur $\widehat{f}_{\hat{h}}$ donné par (3) vérifie*

$$\mathbb{E}[\|\widehat{f}_{\hat{h}} - f\|^2] \leq C_1 \inf_{h \in \mathcal{H}_{N,T}} \left\{\|f - f_h\|^2 + V(h)\right\} + \frac{C_2}{N}$$

*avec $C_1, C_2$ deux constantes positives telles que $C_1$ dépend on $\|K\|_1, \kappa_1, \kappa_2$ et $C_2$ dépend de $\|f\|, \|K\|, \|K\|_1,\ \|K\|_{4/3}$.*

Ce théorème, non asymptotique montre que l'estimateur réalise automatiquement le compromis biais-variance. La fenêtre est sélectionnée seulement en fonction des observations. Un ensemble de fenêtres convenable est par exemple $\mathcal{H}_{N,T} = \{1/k^2, k = 1, \dots, \sqrt{N}\}$. La méthodologie est illustrée sur simulations et sur données réelles dans [2].

# 3 Modèle d'EDS à deux effets aléatoires.

On observe maintenant $N$ trajectoires sur l'intervalle $[0, T]$ du processus suivant :

$$\begin{cases} dX_j(t) & = {}^t b(X_j(t))\phi_j dt + \sigma(X_j(t))dW_j(t) \\ X_j(0) & = \gamma_j \end{cases} \tag{5}$$

où $\phi_j = {}^t(\phi_{j,1}, \phi_{j,2}) \in \mathbb{R}^2$ est un effet aléatoire bidimensionnel, $b(.) = {}^t(b_1(.), b_2(.))$, $\sigma(.)$ sont des fonctions connues définies sur $\mathbb{R}$, $(W_j)_{1 \leq j \leq N}$ sont $N$ processus de Wiener indépendants et $\gamma_j$ est une variable aléatoire réelle. Les variables aléatoires $((\phi_1, \gamma_1), \dots, (\phi_N, \gamma_N))$ sont *i.i.d.* et les suites $((\phi_1, \gamma_1), \dots, (\phi_N, \gamma_N))$ et $(W_1, \dots, W_N)$ sont indépendantes. On suppose que les $\phi_j = {}^t(\phi_{j,1}, \phi_{j,2})$ ont une densité commune bivariée notée $f$.

On suppose que $(W_j)_{j=1\dots,N}$ et $(\phi_j, \gamma_j)_{j=1,\dots,N}$ sont définis sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. On suppose qu'il existe une unique solution forte $X_j(.)$ de (5) pour tous $(\phi_j, \gamma_j) \in \mathbb{R}^3$, adaptée à la filtration $\mathcal{F}_t = \sigma((\phi_j, \gamma_j)_{j=1,\dots,N},\ W_j(s), s \leq t,\ j = 1\dots, N), t \geq 0$. On note $\varphi = {}^t(\varphi_1, \varphi_2)$ une valeur fixée de $\Phi$ et $X^\varphi$ le processus solution de l'équation différentielle stochastique pour $\varphi$ fixé. On suppose qu'il existe un ensemble ouvert $\Phi$ de $\mathbb{R}^2$ et un intervalle $(l, r) \subset \mathbb{R}$ tel que $\sigma^2(x) > 0$, tel que $\forall \varphi \in \Phi$, le processus $X^\varphi$ soit récurrent positif et admette une unique distribution stationnaire $\pi_\varphi(x)dx$. On suppose que $(\phi_j, \gamma_j)$ a pour distribution sur $\Phi \times (l, r)$ : $\pi(d\varphi, dx) = f(\varphi)d\varphi \otimes \pi_\varphi(x)dx$, ainsi $((\phi_j, X_j(t)), t \geq 0)$ est strictement stationnaire.

L'estimation est basée sur les vecteurs $U_j(T)$ et les matrices symétrique $2 \times 2$ $V_j(T)$ :

$$U_j(T) := \int_0^T \frac{b}{\sigma^2}(X_j(s))dX_j(s), \quad V_j(T) := \int_0^T \frac{b\,{}^t b}{\sigma^2}(X_j(s))ds$$

On peut réécrire $U_j(T) = V_j(T)\phi_j + M_j(T)$ avec $M_j(T) := \int_0^T (b/\sigma)(X_j(s))dW_j(s)$ et $V_j(T) =< M_j >_T$ où $< . >$ est la variation quadratique d'une martingale locale continue. Pour toute fonction mesurable $h : (l, r) \to \mathbb{R}$ $\pi_\varphi$-intégrable pour tout $\varphi \in \Phi$, on définit la

variable aléatoire $\pi_{\phi_j}(h) := \int_l^r h(x)\pi_{\phi_j}(x)dx$.
**(H1)** Pour $\varphi \in \Phi$ et $i \in \{1,2\}$, on suppose $\pi_\varphi(b_i^2\sigma^2) = \int_l^r (b_i^2/\sigma^2)(x)\pi_\varphi(x)dx < +\infty$. On définit la matrice limite

$$L_j := \pi_{\phi_j}\left(\frac{b\,{}^t b}{\sigma^2}\right).$$

**(H2)** Pour $j = 1, \ldots, N$, $\mathbb{P}\left(L_j \text{ invertible}\right) = 1$.
Notons que cette hypothèse implique que $V_j(T)$ est inversible pour $T$ assez grand. Sous **(H1)-(H2)**, Genon-Catalot and Larédo donnent dans [4] le résultat de convergence :

$$\frac{1}{T}\int_0^T \frac{b\,{}^t b}{\sigma^2}(X_j(s))ds = \frac{V_j(T)}{T} \underset{T\to+\infty}{\to} L_j,\ p.s.$$

On note dans la suite $\|.\|$ la norme $L^2(\mathbb{R})$, $\|\|_2$ la norme euclidienne de $\mathbb{R}^2$ et $\|\|_F$ la norme de Frobenius définie pour $A \in M_2(\mathbb{R})$ par $\|A\|_F^2 = \sum_{i,j} A_{i,j}^2 = \mathrm{Tr}({}^t AA)$ et $S_2(\mathbb{R})$ le sous espace des matrices symétriques de $M_2(\mathbb{R})$.

Notre but est double : d'abord estimer les effets aléatoires $\phi_j$ et ensuite leur densité $f$, à partir des observations $(X_j(t), 0 \le t \le T)_{1\le j\le N}$, pour $T$ grand.

## 3.1 Estimation des effets alétaoires

Pour chaque $j$ on définit un estimateur de la variable aléatoire $\phi_j$ basé sur la trajectoire $X_j(t),\ t \in [0,T]$ :

$$A_j(T) := {}^t(A_{j,1}(T), A_{j,2}(T)) = V_j(T)^{-1}U_j(T),$$

(qui correspond à l'estimateur du maximum de vraisemblance conditionnelle sachant $\phi_j = \varphi$ en $\varphi$, voir [4]). De plus on a : $A_j(T) = \phi_j + V_j(T)^{-1}M_j(T)$, donc $A_j(T)$ est un estimateur consistant de $\phi_j$ quand $T$ tend vers l'infini. La présence de la matrice inverse $V_j(T)^{-1}$ rend le calcul des moments de $A_j(T)$ difficile. On considère alors un estimateur tronqué :

$$\widehat{A_j}(T) := A_j(T)\mathbf{1}_{B_j(T)},\quad B_j(T) := \{V_j(T) \ge \kappa\sqrt{T} I_2\}$$

où $I_2$ est la matrice identité de $M_2(\mathbb{R})$. L'inégalité dans la définition de $B_j(T)$ a un sens matriciel : $(A,B) \in S_2(\mathbb{R})$, $A \le B \Leftrightarrow A - B$ est une matrice positive. On énonce le résultat non asymptotique qui est la clé de l'estimation non-paramétrique qui suit.

**Proposition 3.** *Soient $(X_j(t), j = 1, \ldots, N)$ donnés par (5) alors sous les hypothèses **(H1)-(H2)** et si pour $j = 1, \ldots, N$,*

$$\mathbb{E}\left[\frac{\|L_j\|_F^2}{\det(L_j)^2}\mathrm{Tr}\left(L_j\right)\right] < \infty,\quad \mathbb{E}\left[\frac{\|L_j\|_F^4}{\det(L_j)^4}\mathrm{Tr}\left(\pi_{\phi_j}\left(\left(\frac{b\,{}^t b}{\sigma^2}\right)^2\right)\right)\right] < \infty,$$

$$\mathbb{E}\left[\left\|\sqrt{T}\left(L_j - \frac{V_j(T)}{T}\right)\right\|_F^4\right] < \infty, \quad \mathbb{E}\left[\|\phi_j\|_2^2\left(\frac{1}{L_{j,1,1}^2} + \frac{1}{L_{j,2,2}^2}\right)\left\|\sqrt{T}\left(L_j - \frac{V_j(T)}{T}\right)\right\|_F^2\right] < \infty$$

*il existe une constante $C > 0$ telle que* $\mathbb{E}\left[\|\widehat{A}_j(T) - \phi_j\|_2^2\right] \leq C/T$.

### 3.2 Estimateur à noyau de la densité des effets aléatoires

On suppose que $f \in \mathbb{L}^2(\mathbb{R}^2)$. Soit $K$ un noyau $\mathcal{C}^1(\mathbb{R}^2)$ tel que ses dérivées partielles $\frac{\partial K}{\partial u}$ and $\frac{\partial K}{\partial v}$ sont dans $\in L^2(\mathbb{R}^2)$, tel que $K$ est intégrable, $\iint K(u,v)dudv = 1$ et $\|K\|^2 = \iint K^2(u,v)dudv < +\infty$. Pour tout $h = (h_1, h_2), h_1 > 0, h_2 > 0$, pour tout $(u,v) \in \mathbb{R}^2$, on note $K_h(u,v) = K\left(u/h_1, v/h_2\right)/(h_1h_2)$. Alors pour $x \in \mathbb{R}^2$ on propose l'estimateur de $f$ :

$$\widehat{f}_h(x) = \frac{1}{N}\sum_{j=1}^{N} K_h(x - \widehat{A}_j(T)). \tag{6}$$

On note $f_h(x) := K_h \star f(x) = \iint f(y_1,y_2)K_h(x_1 - y_1, x_2 - y_2)dy_1dy_2$. Alors il vient :

**Proposition 4.** *Sous (H1)-(H2) et les hypothèses de la Proposition 3,*

$$\mathbb{E}[\|\widehat{f}_h - f\|^2] \leq 2\|f - f_h\|^2 + 2\max\left(\frac{1}{h_1^3h_2}, \frac{1}{h_1h_2^3}\right)\left(\left\|\frac{\partial K}{\partial u}\right\|^2 + \left\|\frac{\partial K}{\partial v}\right\|^2\right)\frac{C}{T} + \frac{\|K\|^2}{Nh_1h_2}.$$

Les commentaires sur la borne de la Proposition 1 restent valables ici. Les ordres obtenus généralisent en dimension 2 ceux trouvés par Comte *et al.* [1] dans le cadre d'un seul effet multiplicatif : $1/(Th^3) + 1/(Nh)$. Il est possible de mener la sélection adaptative de la fenêtre bidimensionnelle en adaptant un critère de Goldenshluger et Lepski. On renvoie à l'article [3] pour l'application de cette procédure sur des exemples.

# Bibliographie


[1] Comte, F. Genon-Catalot, V. Samson, A (2013), Nonparametric estimation for stochastic differential equation with random effects, Stochastic Processes and their Applications, **7**, 2522–2551.

[2] Dion, C (2014), Nonparametric estimation in a mixed-effect Ornstein-Uhlenbeck model, preprint MAP5 2014-24

[3] Dion, C. Genon-Catalot, V. (2015), Bidimensional Random Effect Estimation in Mixed Stochastic Differential Model, preprint MAP5 2015-02

[4] Genon-Catalot, V. Larédo, C. (2013), Estimation for stochastic differential equations with mixed effects, preprint MAP5 2013-09

[5] Goldenshluger, A. Lepski, O. (2011), Bandwidth selection in kernel density estimation : oracle inequalities and adaptive minimax optimality, The Annals of Statistics, **39**, 1608–1632.

[6] Kutoyants, Y. (2004), Statistical Inference for Ergodic Diffusion Processes, Springer.