



**HAL**  
open science

# Estimation for Stochastic Damping Hamiltonian Systems under Partial Observation. II Drift term

Patrick Cattiaux, José R. León, Clémentine Prieur

► **To cite this version:**

Patrick Cattiaux, José R. León, Clémentine Prieur. Estimation for Stochastic Damping Hamiltonian Systems under Partial Observation. II Drift term. 2013. hal-00877054v1

**HAL Id: hal-00877054**

**<https://hal.science/hal-00877054v1>**

Preprint submitted on 25 Oct 2013 (v1), last revised 29 Jun 2014 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**ESTIMATION FOR STOCHASTIC DAMPING HAMILTONIAN  
SYSTEMS UNDER PARTIAL OBSERVATION.  
II. DRIFT TERM.**

PATRICK CATTIAUX <sup>♠</sup> , JOSÉ R. LEÓN <sup>♣</sup> ,  
AND CLÉMENTINE PRIEUR <sup>♡</sup>

<sup>♠</sup> INSTITUT DE MATHÉMATIQUES DE TOULOUSE

<sup>♣</sup> ESCUELA DE MATEMÁTICA, CARACAS

<sup>♡</sup> LABORATOIRE JEAN KUNTZMANN, GRENOBLE

ABSTRACT. This paper is the second part of our study started with Cattiaux, Leon and Prieur (2012). For some ergodic hamiltonian systems we obtained a central limit theorem for a non-parametric estimator of the invariant density, under partial observation (only the positions are observed). Here we obtain similarly a central limit theorem for a non-parametric estimator of the drift term. This theorem lies on the previous result for the invariant density.

1. INTRODUCTION.

Let  $(Z_t := (X_t, Y_t) \in \mathbb{R}^{2d}, t \geq 0)$  be governed by the following Ito stochastic differential equation:

$$\begin{aligned} dX_t &= Y_t dt \\ dY_t &= \sigma dW_t - (c(X_t, Y_t)Y_t + \nabla V(X_t))dt. \end{aligned} \tag{1.1}$$

Each component  $Y^i$  ( $1 \leq i \leq d$ ) is the velocity of a particle  $i$  with position  $X^i$ . Function  $c$  is called the damping force and  $V$  the potential,  $\sigma$  is some (non-zero) constant and  $W$  a standard brownian motion.

We shall assume that  $c$  and  $V$  are regular enough for the existence and uniqueness of a non explosive solution of (1.1). We shall also assume that the process is ergodic with a unique invariant probability measure  $\mu$ , and that the convergence in the ergodic theorem is quick enough. Some sufficient conditions will be recalled in the next section.

In our previous work Cattiaux, Leon and Prieur (2012) we proposed a non-parametric estimator for the invariant density  $p_s$  of the invariant measure  $\mu$ . We refer to the introduction of Cattiaux et al. (2012) for some references on this problem, as well as short discussion of the physical interest of such models.

In the present paper we attack the problem of estimating the drift term

$$g(x, y) = - (c(x, y)y + \nabla V(x)) . \tag{1.2}$$

---

*Date:* October 25, 2013.

As explained in Cattiaux et al. (2012), if there is an impressive literature on non-parametric estimation for the invariant density or the drift term, most of it deals with elliptic diffusion processes. Here we are looking at a fully degenerate process, but still hypo-elliptic. In addition we intend to propose an estimator based on the observation of the positions  $X$  only, at some discretized observation times.

The main result of Cattiaux et al. (2012) reads as follows: if  $p_s$  denotes the invariant density (see the next section for its existence), then one can find a discretization step  $h_n$ , bandwidths  $b_{1,n}$  and  $b_{2,n}$  and kernels  $K$  such that, defining the estimator

$$\hat{p}_n(x, y) := \frac{1}{nb_{1,n}^d b_{2,n}^d} \sum_{i=1}^n K \left( \frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - \frac{X_{(i+1)h_n} - X_{ih_n}}{h_n}}{b_{2,n}} \right),$$

corresponding to partial observation, it holds

$$\sqrt{nb_{1,n}^d b_{2,n}^d} (\hat{p}_n(x, y) - p_s(x, y)) \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, p_s(x, y) \int K^2(u, v) dudv \right),$$

for all pair  $(x, y)$ . The previous convergence in distribution holds true under the stationary distribution. In the non stationary case we have to shift the summation. See Theorem 3.3 and the comment following the statement of the Theorem in section 3.

Here we introduce another estimator  $\hat{g}_n$  defined by

$$\hat{g}_n(x, y) \hat{p}_n(x, y) := \frac{1}{(n-1)b_{1,n}^d b_{2,n}^d} \sum_{i=1}^{n-1} K \left( \frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - \frac{X_{(i+\frac{1}{3})h_n} - X_{ih_n}}{(h_n/3)}}{b_{2,n}} \right) \frac{\mathfrak{D}_{i,n}}{(h_n/3)^2},$$

where

$$\mathfrak{D}_{i,n} := X_{(i+1)h_n} - 2X_{(i+\frac{2}{3})h_n} + X_{(i+\frac{1}{3})h_n}.$$

The necessity of introducing this kind of increments is hidden in the proof of the main result (see Corollary 4.11) of the present paper which says that, one can find  $h_n$  and  $b_{i,n}$  ( $i = 1, 2$ ) such that in the stationary regime

$$\sqrt{nb_{1,n}^d b_{2,n}^d h_n} (\hat{g}_n(x, y) - g(x, y)) \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, \left( \frac{2}{3} \frac{\sigma^2}{p_s(x, y)} \int K^2(u, v) dudv \right) \mathcal{I} \right),$$

where  $\mathcal{I}$  denotes the identity matrix in  $\mathbb{R}^d$ . As for the invariant density we also prove such a central limit theorem in the non-stationary regime, shifting the summation (see section 5).

It should be very interesting to estimate separately  $c(x, y)$  and  $\nabla V(x)$ , who have different physical interpretations. Actually, there is no explicit relation between the invariant density and the drift term, unless  $c$  is constant, in which case  $p_s(x, y) = \exp \left( -\frac{2c}{\sigma^2} \left( \frac{|y|^2}{2} + V(x) \right) \right)$ , so that we have built estimators for both  $V$  and  $\nabla V$ . In full generality this will require some other ideas.

## 2. THE MODEL AND ITS PROPERTIES.

We are obliged to recall some facts on the model. A more detailed discussion is contained in Cattiaux et al. (2012).

We shall first give some results about non explosion and long time behaviour. In a sense, coercivity can be seen in this context as some exponential decay to equilibrium.

Let us first introduce some sets of assumptions:

**Hypothesis  $\mathcal{H}_1$ :**

- (i) the potential  $V$  is lower bounded, smooth over  $\mathbb{R}^d$ ,  $V$  and  $\nabla V$  have polynomial growth at infinity and

$$+\infty \geq \liminf_{|x| \rightarrow +\infty} \frac{x \cdot \nabla V(x)}{|x|} \geq v > 0,$$

the latter being often called “drift condition”,

- (ii) the damping coefficient  $c(x, y)$  is smooth and bounded, and there exist  $c, L > 0$  so that  $c^s(x, y) \geq cId > 0$ ,  $\forall (|x| > L, y \in \mathbb{R}^d)$ , where  $c^s(x, y)$  is the symmetrization of the matrix  $c(x, y)$ , given by  $\frac{1}{2}(c_{ij}(x, y) + c_{ji}(x, y))_{1 \leq i, j \leq d}$ ,

These conditions ensure that there is no explosion, and that the process is positive recurrent with a unique invariant probability measure  $\mu$ . We will denote by  $P_t f(z) = \mathbb{E}_z(f(Z_t))$  which is well defined for all bounded function  $f$ ,  $P_t$  extends as a contraction semi-group on  $\mathbb{L}^p(\mu)$  for all  $1 \leq p \leq +\infty$ .

Furthermore  $\mu$  admits some exponential moment, hence polynomial moments of any order. Another key feature is that the process is actually  $\alpha$ -mixing, i.e.

**Proposition 2.1.** *There exist some constants  $C > 0$  and  $\rho < 1$  such that:*

$$\begin{aligned} \forall g, f \in \mathbb{L}^\infty(\mu), \quad \forall t \geq 0, \\ |\text{Cov}_\mu(f(Z_t), g(Z_0))| \leq C \rho^{t/2} \left\| g - \int g d\mu \right\|_\infty \left\| f - \int f d\mu \right\|_\infty. \end{aligned} \quad (2.2)$$

*i.e., in the stationary regime,  $(Z_t, t \geq 0)$  is  $\alpha$ -mixing with exponential rate.*

As explained in section 2.2 of Cattiaux et al. (2012), the infinitesimal generator  $L$  is hypo-elliptic, which implies that

$$\mu(dz) = p_s(z) dz$$

with some smooth function  $p_s$ . One can relax the  $C^\infty$  assumption on the coefficients into a  $C^k$  assumption, for a large enough  $k$ , but this is irrelevant.

Furthermore it can be shown that  $p_s$  is everywhere positive, for instance by using an extension of Girsanov theory which is available here.

One can relax some assumptions and still have the same conclusions:

**Hypothesis  $\mathcal{H}_2$ :**

- (a) One can relax the boundedness assumption on  $c$  in  $\mathcal{H}_1$ , assuming that for all  $N > 0$ :  $\sup_{|x| \leq N, y \in \mathbb{R}^d} \|c(x, y)\|_{H.S.} < +\infty$ , where  $H.S.$  denotes the Hilbert-Schmidt norm of matrix; but one has to assume in addition conditions (3.1) and (3.2) in Wu (2001). An interesting example (the Van der Pol model) in this situation is described in Wu (2001) subsection 5.3.
- (b) The most studied situation is the one when  $c$  is a constant matrix. Actually almost all results obtained in Wu (2001) or Bakry, Cattiaux and Guillin (2008) in this situation extend to the general bounded case.

Nevertheless we shall assume now that  $c$  is a constant matrix.

In this case a very general statement replacing  $\mathcal{H}_1$  (i) is given in Theorem 6.5 of Bakry et al. (2008). Tractable examples are discussed in Example 6.6 of the same paper. In particular one can replace the drift condition on  $V$  by

$$\liminf_{|x| \rightarrow +\infty} |\nabla V|^2(x) > 0 \quad \text{and} \quad \|\nabla^2 V\|_{H.S.} \ll |\nabla V|.$$

Notice that one can relax the repelling strength of the potential, and obtain, no more exponential but sub-exponential or polynomial decay (see the discussion in Bakry et al. (2008)).

From now on in the whole paper we will assume that Hypothesis  $\mathcal{H}_1$  (or  $\mathcal{H}_2$ ) is fulfilled. In all the proofs of the paper  $C$  denotes some constant which may vary from line to line.

### 3. ESTIMATION OF THE INVARIANT DENSITY.

In this section we recall the central limit theorem for a non-parametric estimator of the invariant density  $p_s$  proposed in Cattiaux et al. (2012).

First we consider that one can observe the whole process  $Z_t$  at discrete times with discretization step  $h_n$ , i.e we consider

$$\tilde{p}_n(x, y) := \frac{1}{nb_{1,n}^d b_{2,n}^d} \sum_{i=1}^n K \left( \frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - Y_{ih_n}}{b_{2,n}} \right). \quad (3.1)$$

Second we consider the partially observed case, where only the position process  $X_t$  can be observed, and we approximate the velocity, i.e. we consider

$$\hat{p}_n(x, y) := \frac{1}{nb_{1,n}^d b_{2,n}^d} \sum_{i=1}^n K \left( \frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - \frac{X_{(i+1)h_n} - X_{ih_n}}{h_n}}{b_{2,n}} \right). \quad (3.2)$$

In both cases, the kernel  $K$  is some  $C^2$  function with compact support  $A$  such that  $\int_A K(x, y) dx dy = 1$ . We may also assume, without loss of generality that  $A$  is a bounded ball. Moreover, we assume that there exists  $m \in \mathbb{N}^*$  such that for all polynomial  $P(x, y)$

with degree less or equal than  $m$ , and more or equal to 1,  $\int P(u, v)K(u, v)dudv = 0$ . That is, we assume the kernel  $K$  is of order  $m$ .

Let us state the main result in this section (Theorem 3.3 below).

**Theorem 3.3** (Cattiaux et al. (2012)). *Assume Hypothesis  $\mathcal{H}_1$  or  $\mathcal{H}_2$  are fulfilled. Recall that  $p_s$  denotes the density of the invariant measure  $\mu$ . Assume that the bandwidths  $b_{1,n}$ ,  $b_{2,n}$  and the discretization step  $h_n$  tend to zero as  $n$  tends to infinity and satisfy the following assumptions:*

$$(i) \quad n b_{1,n}^d b_{2,n}^d \rightarrow +\infty,$$

$$(ii) \quad \frac{b_{1,n} b_{2,n}}{h_n^2} \rightarrow 0,$$

$$(iii) \quad m \text{ is such that } n b_{1,n}^d b_{2,n}^d \max(b_{1,n}, b_{2,n})^{2(m+1)} \rightarrow 0.$$

Then, in the stationary regime, one gets for any  $(x, y) \in \mathbb{R}^{2d}$

$$\sqrt{n b_{1,n}^d b_{2,n}^d} (\tilde{p}_n(x, y) - p_s(x, y)) \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, p_s(x, y) \int K^2(u, v) dudv \right).$$

If in addition

$$(iv) \quad n h_n \frac{b_{1,n}^d}{b_{2,n}^{2+d}} \rightarrow 0,$$

$$(v) \quad \text{there exists } 1 < p \text{ such that } n h_n^2 \frac{b_{1,n}^{d(2-p)/p}}{b_{2,n}^{2+d}} \rightarrow 0.$$

Then, still in the stationary regime, one gets for any  $(x, y) \in \mathbb{R}^{2d}$

$$\sqrt{n b_{1,n}^d b_{2,n}^d} (\hat{p}_n(x, y) - p_s(x, y)) \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, p_s(x, y) \int K^2(u, v) dudv \right).$$

A similar statement holds true starting from any point  $z_0 \in \mathbb{R}^{2d}$ . In this situation we have to slightly change the definition of our estimators replacing  $\sum_{i=1}^n$  by  $\sum_{i=1+l_n}^{n+l_n}$  for some  $l_n$  such that  $l_n h_n \rightarrow +\infty$  as  $n \rightarrow +\infty$ .

#### 4. ESTIMATION OF THE DRIFT TERM.

In this section we will consider an estimator of the drift function from  $\mathbb{R}^{2d}$  into  $\mathbb{R}^d$ ,  $g(x, y) = -[c(x, y)y + \nabla V(x)]$ .

Let  $K : \mathbb{R}^{2d} \rightarrow \mathbb{R}$  be a  $C^2$   $2d$ -dimensional kernel whose support is compact and such that there exists  $m \in \mathbb{N}^*$  such that for all polynomial  $P(x, y)$  with degree less or equal than  $m$ ,  $\int P(u, v)K(u, v)dudv = 0$ . The estimators of the invariant density,  $\tilde{p}_n(x, y)$  and  $\hat{p}_n(x, y)$  are defined as in Section 3. However, for simplicity we will only use  $\hat{p}_n(x, y)$ .

Define for  $i \in \mathbb{N}^*$ ,  $n \in \mathbb{N}^*$ ,

$$\begin{aligned} \mathfrak{D}_{i,n} &:= X_{(i+1)h_n} - 2X_{(i+\frac{2}{3})h_n} + X_{(i+\frac{1}{3})h_n} \\ &= \int_{(i+\frac{2}{3})h_n}^{(i+1)h_n} (Y_s - Y_{(i+\frac{2}{3})h_n}) ds + \int_{(i+\frac{1}{3})h_n}^{(i+\frac{2}{3})h_n} (Y_{(i+\frac{2}{3})h_n} - Y_s) ds. \end{aligned} \quad (4.1)$$

Because one observes only the position of the particle and not its derivative, we must define our estimator by using only the position. For technical reasons, we first introduce a kernel estimate of  $g(x, y) p_s(x, y)$ , denoted by  $\tilde{H}_n$ , and defined by:

$$\tilde{H}_n(x, y) = \frac{1}{(n-1)b_{1n}^d b_{2n}^d} \sum_{i=1}^{n-1} K\left(\frac{x - X_{ih_n}}{b_{1n}}, \frac{y - Y_{ih_n}}{b_{2n}}\right) \frac{\mathfrak{D}_{i,n}}{(h_n/3)^2}. \quad (4.2)$$

To see why this definition is meaningful, let us first study the asymptotic bias of this estimator in the stationary regime.

Using stationarity we get

$$\begin{aligned} \mathbb{E}[\tilde{H}_n(x, y)] &= \frac{9}{b_{1n}^d b_{2n}^d} \mathbb{E}\left[K\left(\frac{x - X_0}{b_{1n}}, \frac{y - Y_0}{b_{2n}}\right) \frac{\mathfrak{D}_{0,n}}{h_n^2}\right] \\ &= \frac{9}{b_{1n}^d b_{2n}^d} \int_{\mathbb{R}^{2d}} K\left(\frac{x - u}{b_{1n}}, \frac{y - v}{b_{2n}}\right) \mathbb{E}\left[\frac{\mathfrak{D}_{0,n}}{h_n^2} | X_0 = u, Y_0 = v\right] p_s(u, v) dudv. \end{aligned}$$

Using (4.1), we may write

$$\frac{\mathfrak{D}_{0,n}}{h_n^2} = \frac{1}{h_n^2} (\sigma[\int_{\frac{2}{3}h_n}^{h_n} (W_s - W_{\frac{2}{3}h_n}) ds + \int_{\frac{h_n}{3}}^{\frac{2}{3}h_n} (W_{\frac{2}{3}h_n} - W_s) ds] + I(h_n)), \quad (4.3)$$

where

$$I(h_n) = \int_{\frac{2}{3}h_n}^{h_n} \int_{\frac{2}{3}h_n}^t g(X_s, Y_s) ds dt + \int_{\frac{h_n}{3}}^{\frac{2}{3}h_n} \int_t^{\frac{2}{3}h_n} g(X_s, Y_s) ds dt.$$

The independence of the increments of  $W$  and the semigroup properties yield

$$\begin{aligned} \mathbb{E}\left[\frac{\mathfrak{D}_{0,n}}{h_n^2} | X_0 = u, Y_0 = v\right] &= \frac{1}{h_n^2} \left[ \int_{\frac{2}{3}h_n}^{h_n} \int_{\frac{2}{3}h_n}^t P_s g(u, v) ds dt + \int_{\frac{h_n}{3}}^{\frac{2}{3}h_n} \int_t^{\frac{2}{3}h_n} P_s g(u, v) ds dt \right] \\ &:= G(h_n, u, v). \end{aligned}$$

Notice that,  $G(h_n, u, v) \rightarrow \frac{1}{9} g(u, v)$  as  $n \rightarrow +\infty$  since  $h_n \rightarrow 0$ .

A change of variable entails

$$\begin{aligned} \mathbb{E}[\tilde{H}_n(x, y)] &= \frac{9}{b_{1n}^d b_{2n}^d} \int_{\mathbb{R}^{2d}} K\left(\frac{x - u}{b_{1n}}, \frac{y - v}{b_{2n}}\right) G(h_n, u, v) p_s(u, v) dudv \\ &= 9 \int_{\mathbb{R}^{2d}} K(z_1, z_2) G(h_n, x - b_{1n}z_1, y - b_{2n}z_2) p_s(x - b_{1n}z_1, y - b_{2n}z_2) dz_1 dz_2, \end{aligned}$$

which converges to  $g(x, y)p_s(x, y)$  as  $n$  tends to infinity, according to the bounded convergence theorem. Hence  $\tilde{H}_n$  is an asymptotically unbiased estimator of  $gp_s$ .

Starting from this consideration, we first state a central limit theorem for the estimator  $\tilde{H}_n(x, y)$ . Let denote  $\mathcal{I}$  the identity matrix in  $\mathbb{R}^d$ .

During the proof we shall need an additional assumption namely:

**Hypothesis  $\mathcal{H}_3$ :**  $g$  belongs to the domain of the infinitesimal generator  $L$ , in all  $\mathbb{L}^p(\mu)$  for  $1 \leq p < +\infty$ .

According to the properties we recalled before, for  $\mathcal{H}_3$  to be satisfied it is enough that:

**Hypothesis  $\mathcal{H}_4$ :** the function  $c$  (resp.  $V$ ) and its first two derivatives (resp. its first three derivatives) have polynomial growth.

**Theorem 4.4.** *Assume that  $\mathcal{H}_1$ ,  $\mathcal{H}_2$  and  $\mathcal{H}_3$  are satisfied and that the bandwidths  $b_{1,n}$ ,  $b_{2,n}$  and the discretization step  $h_n$  tend to zero as  $n$  tends to  $\infty$ . Assume moreover that the following assumptions are satisfied:*

- i)  $n h_n b_{1,n}^d b_{2,n}^d \rightarrow +\infty$ ,
- ii)  $m \in \mathbb{N}^*$  is such that  $n b_{1,n}^d b_{2,n}^d h_n \max(b_{1,n}, b_{2,n})^{2(m+1)} \xrightarrow[n \rightarrow +\infty]{} 0$ ,
- iii)  $\exists \varepsilon_1 > 0$  such that  $n(b_{1,n}^d b_{2,n}^d)^{1-\varepsilon_1} h_n^3 \xrightarrow[n \rightarrow +\infty]{} 0$ ,
- iv)  $\exists 0 < \varepsilon_2, \varepsilon_3 < 1$  such that  $\frac{h_n^{2(1-\varepsilon_2)}}{(b_{1,n} b_{2,n})^{\varepsilon_3}} \xrightarrow[n \rightarrow +\infty]{} 0$ .

Then, in the stationary regime,

$$\sqrt{n b_{1,n}^d b_{2,n}^d h_n} \left( \tilde{H}_n(x, y) - g(x, y)p_s(x, y) \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, \left( \frac{2}{3} \sigma^2 p_s(x, y) \int K^2(u, v) dudv \right) \mathcal{I} \right).$$

**Proof of Theorem 4.4:**

We may replace  $\sqrt{n}$  by  $\sqrt{n-1}$  without any change. Now decompose

$$\begin{aligned} \mathcal{S}_n &:= \sqrt{(n-1) b_{1,n}^d b_{2,n}^d h_n} \left( \tilde{H}_n(x, y) - g(x, y)p_s(x, y) \right) \\ &= \sqrt{(n-1) b_{1,n}^d b_{2,n}^d h_n} \left( \tilde{H}_n(x, y) - \mathbb{E} \tilde{H}_n(x, y) + \mathbb{E} \tilde{H}_n(x, y) - g(x, y)p_s(x, y) \right) \\ &:= \mathcal{I}_{1n} + \mathcal{I}_{2n}. \end{aligned}$$

To prove Theorem 4.4 we first prove that  $\mathcal{I}_{2n} \rightarrow 0$  and then that

$$\mathcal{I}_{1n} \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, \left( \frac{2}{3} \sigma^2 p_s(x, y) \int K^2(x, y) dx dy \right) \mathcal{I} \right).$$

Define

$$I_i(h_n) := 9 \left( \int_{(i+\frac{2}{3})h_n}^{(i+1)h_n} \int_{(i+\frac{2}{3})h_n}^t g(X_s, Y_s) ds + \int_{(i+\frac{1}{3})h_n}^{(i+\frac{2}{3})h_n} \int_t^{(i+\frac{2}{3})h_n} g(X_s, Y_s) ds \right),$$



and

$$\mathfrak{W}_{i,n} := 9 \left( \int_{(i+\frac{2}{3})h_n}^{(i+1)h_n} (W_s - W_{(i+\frac{2}{3})h_n}) ds + \int_{(i+\frac{1}{3})h_n}^{(i+\frac{2}{3})h_n} (W_{(i+\frac{2}{3})h_n} - W_s) ds \right).$$

The vector  $\sqrt{(n-1)b_{1,n}^d b_{2,n}^d h_n} \tilde{H}_n(x, y)$  can be decomposed in two terms: the one driving the bias in the central limit theorem

$$S_{n,1}(x, y) := \frac{\sqrt{h_n}}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} \sum_{i=1}^{n-1} K \left( \frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - Y_{ih_n}}{b_{2,n}} \right) \frac{1}{h_n^2} I_i(h_n),$$

and the one driving the variance

$$S_{n,2}(x, y) := \frac{\sigma}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} \sum_{i=1}^{n-1} K \left( \frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - Y_{ih_n}}{b_{2,n}} \right) \frac{\mathfrak{W}_{i,n}}{h_n^{3/2}}.$$

Notice that  $\mathbb{E}S_{n,2}(x, y) = 0$ . We thus have

$$\mathcal{I}_{2n} = \mathbb{E}S_{n,1}(x, y) - \sqrt{(n-1)b_{1,n}^d b_{2,n}^d h_n} g(x, y) p_s(x, y), \quad (4.5)$$

while

$$\mathcal{I}_{1n} = (S_{n,1} - \mathbb{E}S_{n,1}(x, y)) + S_{n,2}(x, y). \quad (4.6)$$

### First step: Study of $\mathcal{I}_{2n}$

We define

$$\begin{aligned} \frac{\mathfrak{P}_{i,n}}{9} &:= \int_{(i+\frac{2}{3})h_n}^{(i+1)h_n} \int_{(i+\frac{2}{3})h_n}^t (P_s g(X_{ih_n}, Y_{ih_n}) - g(X_{ih_n}, Y_{ih_n})) ds dt \\ &\quad + \int_{(i+\frac{1}{3})h_n}^{(i+\frac{2}{3})h_n} \int_t^{(i+\frac{2}{3})h_n} (P_s g(X_{ih_n}, Y_{ih_n}) - g(X_{ih_n}, Y_{ih_n})) ds dt. \end{aligned}$$

Thanks to stationarity, it holds

$$\begin{aligned} \mathcal{I}_{2n} &= \sqrt{\frac{(n-1)h_n}{b_{1,n}^d b_{2,n}^d}} \mathbb{E} \left( K \left( \frac{x - X_0}{b_{1,n}}, \frac{y - Y_0}{b_{2,n}} \right) \frac{1}{h_n^2} \mathfrak{P}_{0,n} \right) + \\ &\quad + \sqrt{\frac{(n-1)h_n}{b_{1,n}^d b_{2,n}^d}} \left( \mathbb{E} \left( K \left( \frac{x - X_0}{b_{1,n}}, \frac{y - Y_0}{b_{2,n}} \right) g(X_0, Y_0) \right) - b_{1,n}^d b_{2,n}^d g(x, y) p_s(x, y) \right). \end{aligned}$$

The second summand in the above expression can be treated as in a classical density estimation density problem. More precisely, this term is equal to

$$\sqrt{(n-1)b_{1,n}^d b_{2,n}^d h_n} \int K(u, v) \{g(x - ub_{1,n}, y - vb_{2,n}) p_s(x - ub_{1,n}, y, vb_{2,n}) - g(x, y) p_s(x, y)\} dudv.$$

Thus assuming there exists  $m \in \mathbb{N}^*$  such that for all polynomial  $P(x, y)$  with degree less or equal to  $m$ ,  $\int P(u, v)K(u, v)dudv = 0$ , and performing a Taylor expansion, the above term converges to zero as  $n$  tends to infinity as soon as

$$nb_{1,n}^d b_{2,n}^d h_n \max(b_{1,n}, b_{2,n})^{2(m+1)} \xrightarrow{n \rightarrow +\infty} 0.$$

Let us now study the first summand. As each of the coordinates of the drift function  $g$  belongs to the domain of the infinitesimal generator  $L$  according to  $\mathcal{H}_3$ ,  $\forall 1 \leq p < +\infty$   $(P_t g - g)/t$  is bounded in  $\mathbb{L}^p(\mu)$  uniformly in  $t$  for  $t \in [0, 1]$ , say by  $M_p$ . Now write

$$\sqrt{\frac{(n-1)h_n}{b_{1,n}^d b_{2,n}^d}} \mathbb{E} \left( K \left( \frac{x - X_0}{b_{1,n}}, \frac{y - Y_0}{b_{2,n}} \right) \frac{1}{h_n^2} \mathfrak{P}_{0,n} \right) = 9 \sqrt{\frac{(n-1)h_n}{b_{1,n}^d b_{2,n}^d}} (A_{1n} + A_{2n})$$

with

$$A_{1n} = \frac{1}{h_n^2} \int \int_{2h_n/3}^{h_n} \int_{2h_n/3}^t (P_a g(u, v) - g(u, v)) K \left( \frac{x-u}{b_{1,n}}, \frac{y-v}{b_{2,n}} \right) da dt \mu(du, dv)$$

and  $A_{2n}$  being similar just changing  $\int_{2h_n/3}^{h_n} \int_{2h_n/3}^t$  into  $\int_{h_n/3}^{2h_n/3} \int_t^{2h_n/3}$ . We thus only study  $A_{1n}$ .

Using Fubini's theorem we may first integrate with respect to  $t$  and write

$$A_{1n} = \frac{1}{h_n^2} \int \int_{2h_n/3}^{h_n} (h_n - a) a \left( \frac{P_a g(u, v) - g(u, v)}{a} \right) K \left( \frac{x-u}{b_{1,n}}, \frac{y-v}{b_{2,n}} \right) da \mu(du, dv).$$

Now we integrate with respect to  $\mu$ , use Cauchy-Schwarz inequality and the remark we have made about  $P_a g - g$  (assuming that  $h_n \leq 1$ ). We thus have ( $|\cdot|$  denoting the norm in  $\mathbb{R}^d$ )

$$\begin{aligned} |A_{1n}| &\leq \frac{M_p}{h_n^2} \int_{2h_n/3}^{h_n} (h_n - a) a \left( \int K^r \left( \frac{x-u}{b_{1,n}}, \frac{y-v}{b_{2,n}} \right) \mu(du, dv) \right)^{1/r} da \\ &\leq CM_p h_n \left( \int K^r \left( \frac{x-u}{b_{1,n}}, \frac{y-v}{b_{2,n}} \right) \mu(du, dv) \right)^{1/r} \end{aligned}$$

with  $1 < p, r < +\infty \in \mathbb{N}^*$  such that  $\frac{1}{p} + \frac{1}{r} = 1$ . It follows, using again the change of variables

$$\sqrt{\frac{(n-1)h_n}{b_{1,n}^d b_{2,n}^d}} A_{1n} \leq C_p \sqrt{(n-1)h_n^3} (b_{1,n} b_{2,n})^{d(\frac{1}{r}-\frac{1}{2})}.$$

Thanks to assumption iii), one can choose  $r$  such that this last term tends to zero as  $n$  tends to infinity.

Second step: Study of  $S_{n,2}$

We now consider the term driving the variance. To study the weak convergence of this sequence we adapt the proof of Theorem 3 in Beska et al. (1982) and study the characteristic function of  $S_{n,2}$ . Let us recall that

$$S_{n,2}(x, y) := \frac{\sigma}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} \sum_{i=1}^{n-1} K\left(\frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - Y_{ih_n}}{b_{2,n}}\right) \frac{\mathfrak{W}_{i,n}}{h_n^{3/2}}$$

and that  $\mathbb{E}S_{n,2}(x, y) = 0$ .

The sketch of the proof of this step is the following. We first prove that the sequence of random variables  $\{S_{n,2}(x, y), n \geq 1\}$  is tight. We then prove that if  $S_{n_k,2}(x, y)$  is a subsequence of the original sequence, which converges in distribution, then it converges to  $Y \sim \mathcal{N}(0, 1)$ . We then conclude that the sequence itself converges in distribution to  $Y$ .

We proceed now with the proof.

Define for any  $k \in \mathbb{N}$ ,  $\mathcal{F}_k := \sigma((X_l, Y_l), 0 \leq l \leq k)$ . We now introduce

$$f_n(t) := \prod_{i=1}^{n-1} \mathbb{E}[e^{i < t, \frac{\sigma}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} K(\frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - Y_{ih_n}}{b_{2,n}}) \frac{\mathfrak{W}_{i,n}}{h_n^{3/2}} > | \mathcal{F}_{ih_n}}].$$

Thanks to the independence of the Brownian increments we get

$$f_n(t) = \prod_{i=1}^{n-1} e^{-\frac{t^2 \frac{2}{3} \sigma^2 d}{2(n-1)b_{1,n}^d b_{2,n}^d} K^2(\frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - Y_{ih_n}}{b_{2,n}})} = e^{-\frac{t^2 \frac{2}{3} \sigma^2 d}{2(n-1)b_{1,n}^d b_{2,n}^d} \sum_{i=1}^{n-1} K^2(\frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - Y_{ih_n}}{b_{2,n}})}.$$

Now define

$$Z_n = \frac{1}{(n-1)b_{1,n}^d b_{2,n}^d} \sum_{i=1}^{n-1} K^2\left(\frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - Y_{ih_n}}{b_{2,n}}\right).$$

It satisfies

$$\mathbb{E}Z_n = \frac{1}{b_{1,n}^d b_{2,n}^d} \int_{\mathbb{R}^{2d}} K^2\left(\frac{x - u}{b_{1,n}}, \frac{y - v}{b_{2,n}}\right) p_s(u, v) dudv \rightarrow p_s(x, y) \int K^2(u, v) dudv = A.$$

Furthermore

$$\begin{aligned} \mathbb{E}(|Z_n - A|) &\leq \int \left| \left( \frac{1}{b_{1,n}^d b_{2,n}^d} K^2\left(\frac{x - u}{b_{1,n}}, \frac{y - v}{b_{2,n}}\right) p_s(u, v) \right) - A \right| dudv \\ &\leq \frac{1}{b_{1,n}^d b_{2,n}^d} \int K^2\left(\frac{x - u}{b_{1,n}}, \frac{y - v}{b_{2,n}}\right) |p_s(u, v) - p_s(x, y)| dudv \\ &\leq \int K^2(u, v) |p_s(x - b_{1,n}u, y - b_{2,n}v) - p_s(x, y)| dudv \end{aligned}$$

and the latter goes to 0 by using the bounded convergence theorem and the continuity of  $p_s$ .

Thus  $Z_n \rightarrow p_s(x, y) \int K^2(z_1, z_2) dz_1 dz_2$ , in  $L^1$ . Using the bounded convergence Theorem, we deduce that

$$f_n(t) \xrightarrow{\mathbb{P}} e^{-\frac{t^2(\frac{2}{3})\sigma^2 d}{2}} p_s(x, y) \int K^2(z_1, z_2) dz_1 dz_2 := \phi(t).$$

Passing to subsequences if necessary, we can assume that this convergence holds almost everywhere i.e.

$$f_n(t) \rightarrow \phi(t) \text{ a.e.} \quad (4.7)$$

Let us now define for  $k = 1, \dots, n$  the sets

$$H_{nk} = \left\{ e^{-\frac{t^2 \frac{2}{3} \sigma^2 d}{2(n-1)b_{1,n}^d b_{2,n}^d}} \sum_{i=1}^{k-1} K^2\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right) \geq \frac{1}{2} \phi(t) \right\}.$$

Of course

$$H_{n,n} \subset H_{n,n-1} \subset \dots \subset H_{n,2},$$

and using (4.7)

$$\mathbb{P}\{\limsup_{n \rightarrow \infty} H_{n,n}^c\} = 0. \quad (4.8)$$

Finally, introduce the random variables

$$\zeta_{n,i} = \frac{\sigma}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} K\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right) \frac{\mathfrak{W}_{i,n}}{h_n^{3/2}} \mathbf{1}_{H_{n,i}},$$

and also

$$f_n^*(t) = \prod_{i=1}^{n-1} E[e^{i < t, \zeta_{n,i} >} | \mathcal{F}_{ih_n}] = \prod_{i=1}^{n-1} e^{-\frac{t^2 \frac{2}{3} \sigma^2 d}{2(n-1)b_{1,n}^d b_{2,n}^d} K^2\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right) \mathbf{1}_{H_{n,i}}}.$$

It holds

$$f_n^*(t) \geq \frac{1}{2} \phi(t) \quad \text{by definition and } f_n^*(t) \rightarrow \phi(t) \text{ a.e. by (4.8).}$$

So, we can assume that these two properties hold for the initial variables.

Let us now come back to the study of the weak convergence of  $S_{n,2}(x, y)$ . By using Markov property and the independence of the Brownian increments, we obtain

$$\begin{aligned} \mathbb{E} \left[ \frac{e^{i < t, S_{n,2}(x,y) >}}{f_n(t)} \right] &= E \left[ \frac{e^{i < t, \frac{\sigma}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} K\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right) \frac{\mathfrak{W}_{i,n}}{h_n^{3/2}} >}}{e^{-\frac{t^2 \frac{2}{3} \sigma^2 d}{2(n-1)b_{1,n}^d b_{2,n}^d} K^2\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right)}} \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^{n-2} \frac{e^{i < t, \frac{\sigma}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} K\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right) \frac{\mathfrak{W}_{i,n}}{h_n^{3/2}} >}}{e^{-\frac{t^2 \frac{2}{3} \sigma^2 d}{2(n-1)b_{1,n}^d b_{2,n}^d} K^2\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right)}} \right] \mathbb{E} \left[ \frac{e^{i < t, \frac{\sigma}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} K\left(\frac{x-X_{(n-1)h_n}}{b_{1,n}}, \frac{y-Y_{(n-1)h_n}}{b_{2,n}}\right) \frac{\mathfrak{W}_{(n-1),n}}{h_n^{3/2}} >}}{e^{-\frac{t^2 \frac{2}{3} \sigma^2 d}{2(n-1)b_{1,n}^d b_{2,n}^d} K^2\left(\frac{x-X_{(n-1)h_n}}{b_{1,n}}, \frac{y-Y_{(n-1)h_n}}{b_{2,n}}\right)}} \middle| \mathcal{F}_{(n-1)h} \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^{n-2} \frac{e^{i < t, \frac{\sigma}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} K\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right) \frac{\mathfrak{W}_{i,n}}{h_n^{3/2}} >}}{e^{-\frac{t^2 \frac{2}{3} \sigma^2 d}{2(n-1)b_{1,n}^d b_{2,n}^d} K^2\left(\frac{x-X_{ih_n}}{b_{1,n}}, \frac{y-Y_{ih_n}}{b_{2,n}}\right)}} \right] = 1 \text{ using induction.} \end{aligned}$$

Now we are ready to prove the weak convergence

$$\begin{aligned} \left| \mathbb{E} \left[ e^{i\langle t, S_{n,2}(x,y) \rangle} \right] - \phi(t) \right| &= \left| \mathbb{E} \left[ e^{i\langle t, S_{n,2}(x,y) \rangle} - \mathbb{E} \left[ \phi(t) \frac{e^{i\langle t, S_{n,2}(x,y) \rangle}}{f_n(t)} \right] \right] \right| \\ &= \left| \mathbb{E} \left[ e^{i\langle t, S_{n,2}(x,y) \rangle} \left( 1 - \frac{\phi(t)}{f_n(t)} \right) \right] \right| \leq \frac{2}{\phi(t)} \mathbb{E} |f_n(t) - \phi(t)| \rightarrow 0. \end{aligned}$$

We may conclude the proof of the tightness of  $\{S_{n,2}(x,y), n \geq 1\}$ . To this end, let  $S_{n_k,2}(x,y)$  be a subsequence of the original sequence. We know that  $f_{n_k}(t) \xrightarrow{\mathbb{P}} \phi(t)$ . Whence there exists another subsequence  $f_{n_{k_j}}(t) \xrightarrow{\text{a.e.}} \phi(t)$ . By the above result  $S_{n_{k_j}}$  converges weakly to a r.v.  $Y$ , moreover  $\mathbb{E}[e^{i\langle t, Y \rangle}] = \phi(t)$ . Thus the tightness.

All the limits of the convergent subsequences being the same, we directly conclude that the sequence  $\{S_{n,2}(x,y), n \geq 1\}$  converges weakly and that its limit is  $Y$ . It concludes the proof of the second step.

Third step: study of  $\mathcal{Z}_n := S_{n,1}(x,y) - \mathbb{E}S_{n,1}(x,y)$

Let us denote as before by  $\mathfrak{P}_{i,n}^k$  the  $k^{\text{th}}$  coordinate of the vector  $\mathfrak{P}_{i,n}$ . Defining  $\Gamma_n^k(i, x, y, X, Y) = K\left(\frac{x-X}{b_{1,n}}, \frac{y-Y}{b_{2,n}}\right) \frac{1}{h_n^2} \mathfrak{P}_{i,n}^k$ , we write

$$\mathcal{Z}_n^k = \frac{\sqrt{h_n}}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} \sum_{i=1}^{n-1} \Gamma_n^k(i, x, y, X_{ih_n}, Y_{ih_n}),$$

so that

$$\begin{aligned} &\frac{(n-1)b_{1,n}^d b_{2,n}^d}{h_n} \text{Var}(\mathcal{Z}_n^k) \\ &= \left( \sum_{i=1}^{n-1} \text{Var}(\Gamma_n^k(i, x, y, X_{ih_n}, Y_{ih_n})) + \sum_{i \neq l} \text{Cov}(\Gamma_n^k(i, x, y, X_{ih_n}, Y_{ih_n}), \Gamma_n^k(l, x, y, X_{lh_n}, Y_{lh_n})) \right). \end{aligned}$$

To bound the above expression we first write as we did for the first step

$$\mathbb{E}(\Gamma_n^k(i, x, y, X_{ih_n}, Y_{ih_n}))^2 = U_{1n} + U_{2n}$$

where, using stationarity,

$$U_{1n} = \frac{C}{h_n^4} \mathbb{E} \left[ K^2 \left( \frac{x - X_0}{b_{1,n}}, \frac{y - Y_0}{b_{2,n}} \right) \left( \int_{2h_n/3}^{h_n} (h_n - s) (P_s g(X_0, Y_0) - g(X_0, Y_0)) ds \right)^2 \right],$$

$U_{2n}$  being similar just replacing  $\int_{2h_n/3}^{h_n}(h_n - s)$  by  $\int_{h_n/3}^{2h_n/3}(s - (h_n/3))$ .

Using Cauchy-Schwarz inequality we get

$$\begin{aligned} U_{1n} &\leq \frac{C}{h_n} \int_{2h_n/3}^{h_n} \left( \int K^2 \left( \frac{x-u}{b_{1,n}}, \frac{y-v}{b_{2,n}} \right) (P_s g(u,v) - g(u,v))^2 d\mu \right) ds \\ &\leq C h_n \int_{2h_n/3}^{h_n} \left( \int K^2 \left( \frac{x-u}{b_{1,n}}, \frac{y-v}{b_{2,n}} \right) \left( \frac{P_s g(u,v) - g(u,v)}{s} \right)^2 d\mu \right) ds. \end{aligned}$$

We may argue as in the first step, this time using Hölder inequality for some conjugate pair  $(p, q)$  and  $\mathcal{H}_3$  in  $\mathbb{L}^{2q}$ , to conclude that

$$U_{1n} \leq C h_n^2 (b_{1,n} b_{2,n})^{d/p}.$$

It follows

$$\frac{h_n}{(n-1)b_{1,n}^d b_{2,n}^d} \sum_{i=1}^{n-1} \text{Var}(\Gamma_n^k(i, x, y, X_{ih_n}, Y_{ih_n})) = \mathcal{O} \left( h_n^3 (b_{1,n} b_{2,n})^{d(\frac{1}{p}-1)} \right).$$

One can choose  $p$  such that the right hand term tends to zero thanks to assumption iv).

Let us now compute the covariances.

One has thanks to stationarity and mixing

$$\begin{aligned} &\sum_{i \neq l} \text{Cov}(\Gamma_n^k(i, x, y, X_{ih_n}, Y_{ih_n}), \Gamma_n^k(l, x, y, X_{lh_n}, Y_{lh_n})) \\ &\leq n \sum_{j=1}^{n-2} \min(\rho^{jh_n/2}, \text{Var}(\Gamma_n^k(0, x, y, X_0, Y_0))) \\ &\leq n \sum_{j=1}^{n-2} \min(\rho^{jh_n/2}, (b_{1,n} b_{2,n})^{\frac{d}{p}} h_n^2) \\ &= \mathcal{O} \left( n (b_{1,n} b_{2,n})^{\frac{d}{p}(1-a)} h_n^{1-2a} \right) \quad \text{for any } 0 < a < 1. \end{aligned}$$

We thus get

$$\begin{aligned} &\frac{h_n}{(n-1)b_{1,n}^d b_{2,n}^d} \sum_{i \neq l} \text{Cov}(\Gamma_n^k(i, x, y, X_{ih_n}, Y_{ih_n}), \Gamma_n^k(l, x, y, X_{lh_n}, Y_{lh_n})) \\ &= \mathcal{O} \left( (b_{1,n} b_{2,n})^{\frac{d}{p}(1-a)-d} h_n^{2(1-a)} \right). \end{aligned}$$

One can choose  $p$  and  $a$  such that the right hand term tends to zero as  $n$  tends to infinity thanks to assumption iv). This completes the proof.  $\square$

We now deduce the following theorem:

**Theorem 4.9.** *Define*

$$\hat{H}_n(x, y) = \frac{1}{(n-1)b_{1n}^d b_{2n}^d} \sum_{i=1}^{n-1} K \left( \frac{x - X_{ih_n}}{b_{1n}}, \frac{y - \frac{X_{(i+\frac{1}{3})h_n} - X_{ih_n}}{(h_n/3)}}{b_{2n}} \right) \frac{\mathfrak{D}_{i,n}}{(h_n/3)^2}.$$

Then, under the assumptions of Theorem 4.4 and assuming moreover

v) there exist  $1 < p < +\infty$  and  $\varepsilon > 0$ , such that  $h_n^2 b_{1,n}^{d(\frac{1}{p}-1)} b_{2,n}^{-(2+d)} \xrightarrow{n \rightarrow +\infty} 0$  and

$$h_n \sqrt{nb_{1,n}^{d(\frac{1}{p(1+\varepsilon)} - \frac{1}{2})}} b_{2,n}^{-(\frac{d}{2}+1)} \xrightarrow{n \rightarrow +\infty} 0,$$

one gets

$$\sqrt{nb_{1,n}^d b_{2,n}^d} h_n \left( \hat{H}_n(x, y) - p_s(x, y)g(x, y) \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, \left( \frac{2}{3} \sigma^2 p_s(x, y) \int K^2(u, v) dudv \right) \mathcal{I} \right).$$

**Proof of Theorem 4.9:**

Starting from Theorem 4.4, it remains now to consider  $D_n$  defined by

$$\frac{1}{\sqrt{(n-1)b_{1n}^d b_{2n}^d}} \sum_{j=1}^{n-1} \left[ K \left( \frac{x - X_{jh_n}}{b_{1n}}, \frac{y - Y_{jh_n}}{b_{2n}} \right) - K \left( \frac{x - X_{jh_n}}{b_{1n}}, \frac{y - \frac{X_{(j+\frac{1}{3})h_n} - X_{jh}}{(h_n/3)}}{b_{2n}} \right) \right] \frac{\mathfrak{D}_{j,n}}{h^{3/2}}.$$

Let us define  $A_j = K \left( \frac{x - X_{jh_n}}{b_{1n}}, \frac{y - Y_{jh_n}}{b_{2n}} \right) - K \left( \frac{x - X_{jh_n}}{b_{1n}}, \frac{y - \frac{X_{(j+\frac{1}{3})h_n} - X_{jh}}{(h_n/3)}}{b_{2n}} \right)$ . We then write

$$D_n = \frac{1}{\sqrt{(n-1)b_{1n}^d b_{2n}^d}} \sum_{j=1}^{n-1} A_j \frac{\mathfrak{D}_{j,n}}{h^{3/2}} = \frac{1}{\sqrt{(n-1)b_{1n}^d b_{2n}^d}} \sum_{j=1}^{n-1} A_j \frac{(\sigma \mathfrak{W}_{j,n} + I_j(h_n))}{h^{3/2}}.$$

Using Hölder Inequality, we bound  $\mathbb{E} \left| \frac{1}{\sqrt{(n-1)b_{1n}^d b_{2n}^d}} \sum_{j=1}^{n-1} A_j \frac{I_j(h_n)}{h^{3/2}} \right|$  by

$$\frac{1}{\sqrt{(n-1)b_{1n}^d b_{2n}^d}} h^{-3/2} \sum_{j=1}^{n-1} (\mathbb{E}|A_j|^{1+\varepsilon})^{1/(1+\varepsilon)} (\mathbb{E}|I_j(h_n)|^{(1+\varepsilon)/\varepsilon})^{\varepsilon/(1+\varepsilon)},$$

with  $\varepsilon > 0$ .

We first consider  $\mathbb{E}|I_j(h_n)|^{(1+\varepsilon)/\varepsilon} = \mathbb{E}|I_0(h_n)|^{(1+\varepsilon)/\varepsilon}$  by stationarity. We use (4.3) and we get that  $\mathbb{E} \left[ |I_0(h_n)|^{(1+\varepsilon)/\varepsilon} |X_0 = u, Y_0 = v \right]$  is bounded by

$$\left| \int_{\frac{2}{3}h_n}^{h_n} \int_{\frac{2}{3}h_n}^t P_s g(u, v) ds dt + \int_{\frac{h_n}{3}}^{\frac{2}{3}h_n} \int_t^{\frac{2}{3}h_n} P_s g(u, v) ds dt \right|^{(1+\varepsilon)/\varepsilon}$$

which is itself bounded by

$$C h_n^{\frac{2}{\varepsilon}} \left[ \int_{\frac{2}{3}h_n}^{h_n} \int_{\frac{2}{3}h_n}^t P_s |g|^{\frac{1+\varepsilon}{\varepsilon}}(u, v) ds dt + \int_{\frac{h_n}{3}}^{\frac{2}{3}h_n} \int_t^{\frac{2}{3}h_n} P_s |g|^{\frac{1+\varepsilon}{\varepsilon}}(u, v) ds dt \right]$$

which is equal to

$$C h_n^{\frac{2}{\varepsilon}} h_n^2 G_\varepsilon(h_n, u, v)$$

with  $G_\varepsilon(h_n, u, v) \xrightarrow{n \rightarrow +\infty} |g|^{\frac{1+\varepsilon}{\varepsilon}}(u, v)$ .

Using now similar arguments as in Step 2 in the proof of Theorem 3.3 in Cattiaux et al. (2012) (Theorem 3.3 of the present paper), one bounds  $\mathbb{E}|A_i|^{1+\varepsilon}$  by

$$\mathcal{O} \left( h_n^{1+\varepsilon} b_{1,n}^{\frac{d}{p}} b_{2,n}^{-(1+\varepsilon)} \right). \quad (4.10)$$

Thus  $\mathbb{E} \left| \frac{1}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} \sum_{j=1}^{n-1} A_j \frac{I_j(h_n)}{h_n^{3/2}} \right|$  is bounded by

$$\mathcal{O} \left( h_n \sqrt{n} b_{1,n}^{d(\frac{1}{p(1+\varepsilon)} - \frac{1}{2})} b_{2,n}^{-(\frac{d}{2}+1)} \right)$$

with  $1 < p < +\infty$ . It converges to zero as we assumed vii).

It remains now to bound  $\mathbb{E} \left| \frac{1}{\sqrt{(n-1)b_{1,n}^d b_{2,n}^d}} \sum_{j=1}^{n-1} A_j \frac{\sigma \mathfrak{W}_{j,n}}{h_n^{3/2}} \right|$ . The terms  $A_j \mathfrak{W}_{j,n}$ ,  $1 \leq j \leq n-1$  are centered and uncorrelated. Thus by stationarity we have to bound  $\frac{\sigma^2}{b_{1,n}^d b_{2,n}^d} \mathbb{E} \frac{A_0^2 \mathfrak{W}_{0,n}^2}{h_n^3}$ . First conditioning now on  $Z_{\frac{h_n}{3}}$  one gets  $\mathbb{E} A_0^2 \mathfrak{W}_{0,n}^2 = h_n^3 \mathbb{E} A_0^2$ . We then conclude by using (4.10) for  $\varepsilon = 1$  which yields  $\mathbb{E} A_0^2 \mathfrak{W}_{0,n}^2 = h_n^3 h_n^2 b_{1,n}^{d/p} b_{2,n}^{-2}$  for all  $1 < p < +\infty$ . This completes the proof, using v).  $\square$

To estimate the drift  $g$  we have to divide  $\hat{H}_n$  by some estimator of  $p_s$ . The natural choice would be to define  $\hat{g}_n = \hat{H}_n / \hat{p}_n$ . For technical reasons (which will be explained in the proposition following the corollary), we have to slightly modify this natural choice, introducing two different bandwidths  $b_{i,n}$  and  $c_{i,n}$ , but with the same discretization step.

We state the following corollary, which is the main result in the stationary situation:

**Corollary 4.11.** *Assume that the assumptions in Theorem 4.9 and Theorem 3.3 are satisfied respectively with  $b_{i,n}$  and  $c_{i,n}$ . Define  $\hat{g}_n = \hat{H}_n(b) / \hat{p}_n(c)$  where the indication into braces indicates the bandwidths we are using. Assume in addition that*

$$h_n(b, c) := h_n \frac{b_{1,n}^d b_{2,n}^d}{c_{1,n}^d c_{2,n}^d} \rightarrow 0 \quad \text{as } n \rightarrow +\infty.$$

Then

$$\sqrt{nb_{1,n}^d b_{2,n}^d h_n} (\hat{g}_n(x, y) - g(x, y)) \xrightarrow{n \rightarrow +\infty} \mathcal{N} \left( 0, \left( \frac{2}{3} \frac{\sigma^2}{p_s(x, y)} \int K^2(u, v) dudv \right) \mathcal{I} \right).$$



A similar statement holds defining similarly  $\tilde{g}_n = \tilde{H}_n(b)/\tilde{p}_n(c)$  under the assumptions in Theorem 4.4 and i)-iv) and in Theorem 3.3 respectively.

**Proof of Corollary 4.11:**

We have

$$\hat{g}_n - g = \frac{\hat{H}_n}{\hat{p}_n} - g = \frac{\hat{H}_n - p_s g}{p_s} + \hat{H}_n \left( \frac{1}{\hat{p}_n} - \frac{1}{p_s} \right).$$

For the first term, we have according to Theorem 4.9,

$$\sqrt{nb_{1,n}^d b_{2,n}^d h_n} \frac{\hat{H}_n - p_s g}{p_s}(x, y) \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, \left( \frac{2}{3} \frac{\sigma^2}{p_s(x, y)} \int K^2(u, v) dudv \right) \mathcal{I} \right).$$

We shall show that the second one goes to 0 in probability and then conclude by using Slutsky theorem.

We thus decompose

$$\sqrt{nb_{1,n}^d b_{2,n}^d h_n} \hat{H}_n \left( \frac{1}{\hat{p}_n} - \frac{1}{p_s} \right) = \left( \frac{\sqrt{h_n(b, c)} \hat{H}_n}{\hat{p}_n p_s} \right) \left( \sqrt{nc_{1,n}^d c_{2,n}^d} (p_s - \hat{p}_n) \right).$$

The second term of this product converges in distribution according to Theorem 3.3. We shall show that the first term in the product goes to 0 in probability.

Indeed, according to Cattiaux et al. (2012),  $\hat{p}_n - p_s \rightarrow 0$  in probability as  $n \rightarrow +\infty$ , and as we previously saw,  $\hat{H}_n$  is bounded in  $\mathbb{L}^1$ . Let  $a > 0$ . We have

$$\mathbb{P} \left( \frac{\sqrt{h_n(b, c)} \hat{H}_n}{\hat{p}_n p_s} > a \right) \leq \mathbb{P}(\hat{p}_n \leq (p_s/2)) + \mathbb{P} \left( \hat{H}_n > (a p_s^2 / 2 \sqrt{h_n(b, c)}) \right),$$

and both terms go to 0, using Markov inequality for instance for the second one.

To conclude it remains to recall that if  $U_n$  goes to 0 in probability and  $V_n$  goes to  $V$  in distribution, the product  $U_n V_n$  goes to 0 in probability.  $\square$ .

We conclude this section by giving an explicit class of examples for the parameters  $h_n$ ,  $b_{i,n}$ ,  $c_{i,n}$  to satisfy all the required assumptions in Corollary 4.11.

**Proposition 4.12.** *Choose  $h_n = n^{-\gamma}$ ,  $b_{i,n} = n^{-\alpha_i}$  and  $c_{i,n} = n^{-\beta_i}$  for some  $\gamma, \alpha_i, \beta_i > 0$ . If*

- (1)  $\alpha_2 = \beta_2 - \frac{\varepsilon}{d(4+4d)} = \frac{1-2\varepsilon}{d(4+4d)}$ ,
- (2)  $\beta_1 = \frac{3-2\varepsilon+4d}{d(4+4d)}$ ,  $\alpha_1 = \beta_1 + \frac{\beta_2}{2} - \frac{1}{2d^2}$ ,
- (3)  $\frac{1}{2d} < \gamma < \frac{1}{2d} + \frac{3\varepsilon}{4+4d}$ ,
- (4)  $m > \frac{1+2d}{1-\varepsilon}$ .

for some  $\varepsilon > 0$  small enough. Then Corollary 4.11 holds true.

In addition the rate of convergence is of order strictly smaller than  $n^{\frac{3\varepsilon}{4+4d}}$ .

*Proof.* For Theorem 3.3 to be satisfied in this situation we must have (see Remark 3.4 in Cattiaux et al. (2012))

- (a)  $\beta_1 > \frac{2+\beta_2(3+2d)}{1+2d}$  in particular  $\beta_1 > \beta_2$ ,
- (b)  $1 - \beta_1 d + \beta_2(2 + d) < \gamma < \frac{1}{2}(\beta_1 + \beta_2) < \frac{1}{2d}$ ,
- (c)  $m > \frac{1-d(\beta_1+\beta_2)}{2\beta_2}$ .

Condition (a) ensures that  $1 - \beta_1 d + \beta_2(2 + d) < \frac{1}{2}(\beta_1 + \beta_2)$ , so that we may find some  $\gamma$  sandwiched by both terms. For both (a) and (b) to be satisfied, it is necessary that

$$\beta_2 < \frac{1}{d(4 + 4d)}, \quad (4.13)$$

and then we can choose

$$\frac{1}{d} - \beta_2 > \beta_1 > \frac{2 + \beta_2(3 + 2d)}{1 + 2d}.$$

Remark that our interest is to take  $\beta_2$  as close as possible to its upper bound,  $\beta_1$  as close as possible to  $\frac{1}{d} - \beta_2$  so that  $m$  can be chosen as small as possible (1 is possible).  $\gamma$  can then be chosen smaller than and close to  $1/2d$ .

Now we look at Theorem 4.9 starting with conditions i)-iv) in Theorem 4.4

- (1)  $1 > \gamma + d(\alpha_1 + \alpha_2)$ ,
- (2)  $1 < d(\alpha_1 + \alpha_2) + 2(m + 1)(\alpha_1 \wedge \alpha_2) + \gamma$ ,
- (3)  $1 < 3\gamma + d(\alpha_1 + \alpha_2)$  is enough for getting  $\varepsilon_1 > 0$  in iii),
- (4)  $2\gamma > \varepsilon_3(\alpha_1 + \alpha_2)$  is enough for getting  $\varepsilon_2$  in iv).

We see that the latter will be automatically satisfied for  $\varepsilon_3$  small enough. We have to add

- (5)  $2\gamma > (2 + d)\alpha_2$ ,
- (6)  $\frac{1}{2} < \gamma + \frac{d}{2}(\alpha_1 - \alpha_2) - \alpha_2$ ,

which is enough to furnish both  $p > 1$  and  $\varepsilon > 0$  in v) of Theorem 4.9.

Finally we have to add

$$(7) \quad \gamma > d((\beta_1 + \beta_2) - (\alpha_1 + \alpha_2)).$$

Look at (1). The compatibility with (b) imposes

$$\gamma + d(\alpha_1 + \alpha_2) < 1 < \gamma + d(\beta_1 + \beta_2) - \beta_2(2 + 2d).$$

That is why we cannot take the same bandwidths  $b$  and  $c$ .

We thus have a first necessary condition

$$d(\alpha_1 + \alpha_2) < d(\beta_1 + \beta_2) - \beta_2(2 + 2d),$$

which is satisfied as soon as

$$d(\alpha_1 + \alpha_2) = d(\beta_1 + \beta_2) - \frac{1}{2d}. \quad (4.14)$$

Now we choose for some small  $\varepsilon$  (say less than  $10^{-6}/d^2$ ),

$$\beta_2 = \frac{1 - \varepsilon}{d(4 + 4d)} \quad , \quad \beta_1 = \frac{3 - 2\varepsilon + 4d}{d(4 + 4d)} .$$

(4.13) and the following inequalities are satisfied, as well as the final upper bound in (b). We have thus to choose  $\gamma$  such that, first  $\gamma > \frac{1}{2d}$  for (7) to be satisfied, next for (1) and the first lower bound of (b) to be satisfied. This amounts to

$$\frac{1 - \varepsilon}{2d} + \frac{3\varepsilon}{4 + 4d} < \gamma < \frac{1}{2d} + \frac{3\varepsilon}{4 + 4d} .$$

The left hand side of the previous inequality is less than  $\frac{1}{2d}$  since  $d \geq 1$ . So we only have to choose

$$\frac{1}{2d} < \gamma < \frac{1}{2d} + \frac{3\varepsilon}{4 + 4d} . \quad (4.15)$$

The upper bound in (b) are then satisfied. If we look at (3) it reduces to

$$\frac{1}{2d} + \frac{3\varepsilon}{4 + 4d} < 3\gamma$$

which is satisfied. Now (6) becomes

$$\frac{3\varepsilon}{4 + 4d} + \frac{1}{2d} + (2d + 2)\alpha_2 < 2\gamma$$

which implies (5) and which is satisfied as soon as

$$\frac{3\varepsilon}{4 + 4d} + (2d + 2)\alpha_2 < \frac{1}{2d} .$$

Notice that the latter implies  $\alpha_2 < \alpha_1$ . Notice that for  $\varepsilon > 0$  we may choose  $\alpha_2 = \beta_2 - \frac{\varepsilon}{d(4+4d)}$ . It remains to choose  $m$  for (2) to be satisfied.  $\square$

Remark that contrary to the case of Theorem 3.3, where it is possible to have  $m = 1$  here  $m$  grows linearly with the dimension ( $m > \frac{2+4d+6\varepsilon d+2\varepsilon}{2(1-\varepsilon)}$ ).

## 5. NON-STATIONARY CASE

In Section 3 we stated the central limit theorem for the estimate of the drift  $g(x, y)$  in the case where the process is in the stationary regime. Let us now define the new estimate

$$\bar{g}_n(x, y) = \frac{\frac{1}{(n-1)b_{1n}^d b_{2n}^d} \sum_{i=l_n+1}^{n+l_n-1} K \left( \frac{x - X_{ih_n}}{b_{1n}}, \frac{y - \frac{X_{(i+\frac{1}{3})h_n} - X_{ih_n}}{(h_n/3)}}{b_{2n}} \right) \frac{\mathfrak{D}_{i,n}}{(h_n/3)^2}}{\frac{1}{nb_{1,n}^d b_{2,n}^d} \sum_{i=l_n+1}^{n+l_n} K \left( \frac{x - X_{ih_n}}{b_{1,n}}, \frac{y - \frac{X_{(i+1)h_n} - X_{ih_n}}{h_n}}{b_{2,n}} \right)} \quad (5.1)$$

We remark that given  $Z_0 \sim \mu(dz)$ ,  $\bar{g}_n(x, y) \stackrel{\mathcal{L}}{=} \hat{g}_n(x, y) \quad \forall n \in \mathbb{N}^*$ .

Theorem 5.2 below states that we can estimate  $g(x, y)$  by using  $\bar{g}_n(x, y)$  with  $Z_0 = z_0 = (x_0, y_0)$ .

**Theorem 5.2.** *Under the assumptions in Corollary 4.11, starting from any initial point  $z_0 = (x_0, y_0)$ , it holds*

$$\sqrt{n h_n b_{1,n}^d b_{2,n}^d} (\bar{g}_n(x, y) - g(x, y)) \xrightarrow[n \rightarrow +\infty]{\mathcal{D}} \mathcal{N} \left( 0, \left( \frac{2}{3} \frac{\sigma^2}{p_s(x, y)} \int K^2(u, v) du dv \right) \mathcal{I} \right),$$

provided  $l_n$  appearing in the definition (5.1) of  $\bar{g}_n(x, y)$  satisfies  $l_n h_n \xrightarrow[n \rightarrow +\infty]{} +\infty$ .

**Proof of Theorem 5.2:**

Recall that  $p_s : \mathbb{R}^{2d} \mapsto \mathbb{R}_+$  denotes the invariant density of  $Z$  and  $\mu$  the associated invariant probability measure.

Denote by  $\mathcal{C}_b(\mathbb{R})$  the set of bounded continuous functions  $h : \mathbb{R} \rightarrow \mathbb{R}$ . It is only necessary to prove that, for any  $h \in \mathcal{C}_b(\mathbb{R})$ , the difference

$$\begin{aligned} \Delta_n(h) &= \mathbb{E} \left[ h \left( \sqrt{h_n n b_{1,n}^d b_{2,n}^d} \hat{g}_n(x, y) \mid Z_0 \sim \mu \right) - h \left( \sqrt{h_n n b_{1,n}^d b_{2,n}^d} \bar{g}_n(x, y) \mid Z_0 = z_0 \right) \right] \\ &= \mathbb{E} \left[ h \left( \sqrt{h_n n b_{1,n}^d b_{2,n}^d} \bar{g}_n(x, y) \mid Z_0 \sim \mu \right) - h \left( \sqrt{h_n n b_{1,n}^d b_{2,n}^d} \bar{g}_n(x, y) \mid Z_0 = z_0 \right) \right] \end{aligned}$$

goes to zero as  $n$  tends to infinity. Let  $h \in \mathcal{C}_b(\mathbb{R})$  and denote  $\theta = \|h\|_\infty$ .

To evaluate  $\mathbb{E}(h(\bar{g}_n(x, y)) \mid Z_0 \sim \mu) - \mathbb{E}(h(\bar{g}_n(x, y)) \mid Z_0 = z_0)$ . Let us fix  $n \in \mathbb{N}^*$ . We first make the computations conditionally to  $Z_{jh_n}, j > l_n + 1$ .

One may write

$$\begin{aligned} \bar{g}_n(x, y) &= \bar{g}_n(x, y, Z_{(l_n+1)h_n}, Z_{jh_n}, j > l_n + 1) \text{ and} \\ h_{Z_{jh_n}, j > l_n+1}(z') &= \bar{g}_n(x, y, z', Z_{jh_n}, j > l_n + 1). \end{aligned}$$

Now, conditionally to  $Z_{jh_n}, j > l_n + 1$ , one has:

$$\begin{aligned} & \left| \mathbb{E}(h(\bar{g}_n(x, y)) \mid Z_0 \sim \mu) - \mathbb{E}(h(\bar{g}_n(x, y)) \mid Z_0 = z_0) \right| \\ &= \left| \int h_{Z_{jh_n}, j > l_n+1}(z') (p_s(z') - q_{(l_n+1)h_n}(z_0, z')) dz' \right| \\ &\leq \theta D \rho^{(l_n+1)h_n} \Psi(z_0) \end{aligned} \tag{5.3}$$

using Inequality (2.1) in Cattiaux et al. (2012).

Finally, as  $0 < \rho < 1$ , we can conclude that  $\Delta_n(h)$  goes to zero as  $n$  tends to infinity as soon as  $l_n h_n \xrightarrow[n \rightarrow +\infty]{} +\infty$ , which concludes the proof of Theorem 5.2.  $\square$

## 6. EXAMPLES AND NUMERICAL SIMULATION RESULTS

In this section, we consider examples of stochastic differential equations defined by (1.1) and implement the estimator on simulated data. However, the choice of the optimal bandwidths  $b_{1,n}$ ,  $b_{2,n}$ ,  $c_{1,n}$  and  $c_{2,n}$  as far as the choice of the optimal discretization step  $h_n$ , and the optimal choice of the kernel  $K$ , although interesting, is not the purpose of this section nor this paper. This is a separate study to be addressed in future work.

To simulate sample paths, we use an approximate discrete sampling generated by an explicit or implicit Euler scheme. We consider three specific examples. The first one has been proposed in Pokern et al. (2009). It corresponds to a linear oscillator subject to noise and damping. The second example is one example of generalized Duffing oscillators described in Wu (2001) subsection 5.2. The last example is the Van der Pol oscillator whose damping force depends on both position and velocity coordinates. These three models are of type (1.1) and satisfy assumptions needed to apply our estimation results. Simulations are run with the Epanechnikov kernel.

**6.1. Model I: harmonic oscillator.** We consider an harmonic oscillator that is driven by a white noise forcing:

$$\begin{aligned} dX_t &= Y_t dt \\ dY_t &= \sigma dW_t - (\kappa Y_t + DX_t) dt. \end{aligned} \quad (6.1)$$

with  $\kappa > 0$  and  $D > 0$ . In the following we choose  $D = 2$ ,  $\kappa = 2$  and  $\sigma = 1$ . For this model we know that the stationary distribution is gaussian, with mean zero and an explicit variance matrix given in Gardiner (1985), e.g. With our choice of parameters, the gaussian invariant density is

$$p_s(x, y) = \frac{2\sqrt{2}}{\pi} \exp(-4x^2 - 2y^2).$$

And the drift is defined by  $g(x, y) = -2(y+x)$ . In the following we make use of the explicit Euler scheme to simulate an approximated discrete sampling  $(\tilde{X}_i, \tilde{Y}_i)_{i \in \mathbb{N}}$  of  $(X_t, Y_t)_{t \in \mathbb{R}_+}$ . For a given step  $\delta > 0$ , the scheme is defined as

$$\begin{aligned} \tilde{X}_{i+1} - \tilde{X}_i &= \tilde{Y}_i \delta \\ \tilde{Y}_{i+1} - \tilde{Y}_i &= \sigma (W_{(i+1)\delta} - W_{i\delta}) - (\kappa \tilde{Y}_i + D \tilde{X}_i) \delta \end{aligned} \quad (6.2)$$

$(\tilde{X}_0, \tilde{Y}_0) = (0, 0)$ . We take  $n = 5000$ ,  $h = 0.28$ ,  $b_{1,n} = b_{2,n} = 0.18$ , and the step for the explicit Euler scheme  $\delta = \frac{1}{10}h/3$ .

The drift at some fixed point  $x_0$ ,  $g(x_0, \cdot)$ , is estimated on a grid  $(z_l)_{l=1, \dots, L} = (x_0, y_l)_{l=1, \dots, L}$ .

The drift at some fixed point  $y_0$ ,  $g(\cdot, y_0)$ , is estimated on a grid  $(z_l)_{l=1, \dots, L} = (x_l, y_0)_{l=1, \dots, L}$ .

On Figure 1 below we chose  $L = 40$  and  $y_0 = 0.0597$ .

On Figure 2 below we chose  $L = 40$  and  $x_0 = -0.0183$ .

**6.2. Model II: generalized Duffing oscillator.** We consider the noisy Duffing oscillator known as Kramers oscillator. The system (1.1) writes now

$$\begin{aligned} dX_t &= Y_t dt \\ dY_t &= \sigma dW_t - (\kappa Y_t + \alpha X_t^3 - \beta X_t) dt \end{aligned} \quad (6.3)$$

with  $\sigma, \kappa, \alpha$  and  $\beta > 0$ . The potential is then  $V(x) = \alpha \frac{x^4}{4} - \beta \frac{x^2}{2}$ . The invariant density is in that case

$$p_s(x, y) = \frac{\sqrt{\kappa}}{\sqrt{\pi\sigma C}} \exp\left(\frac{-2\kappa}{\sigma^2} \left(\frac{\alpha x^4}{4} - \frac{\beta x^2}{2} + \frac{y^2}{2}\right)\right),$$

with  $C$  the normalizing constant.

And the drift is defined by  $g(x, y) = -(\kappa y + \alpha x^3 - \beta x)$ .

We make use of the implicit Euler scheme to simulate an approximated discrete sampling. The explicit Euler scheme is indeed known to be unstable for that case (see ?). The choice for the parameters is  $\sigma = 1, \kappa = \alpha = \beta = 1$ .

We take  $n = 10^4, h_n = 0.30, b_{1,n} = b_{2,n} = 0.30$ , and the step for the explicit Euler scheme  $\delta = \frac{1}{10}h/3$ .

The drift is estimated at some fixed point  $y_0, g(\cdot, y_0)$ , on a grid  $(z_l)_{l=1,\dots,L} = (x_l, y_0)_{l=1,\dots,L}$ . It is also estimated at some fixed point  $x_0, g(x_0, \cdot)$ , on a grid  $(z_l)_{l=1,\dots,L} = (x_0, y_l)_{l=1,\dots,L}$ .

On Figure 3 below we chose  $L = 40$  and  $y_0 = 0.0533$ . On Figure 4 below we chose  $L = 40$  and  $x_0 = -0.0498$ .

**6.3. Model III: Van der Pol oscillator.** We consider the Van der Pol oscillator defined by

$$\begin{aligned} dX_t &= Y_t dt \\ dY_t &= \sigma dW_t - ((c_1 X_t^2 - c_2) Y_t + \omega_0^2 X_t) dt \end{aligned} \quad (6.4)$$

with  $\sigma, c_1, c_2$  and  $\omega_0^2 > 0$ . In the following we choose  $\sigma = c_1 = c_2 = \omega_0 = 1$ . The drift is then defined by  $g(x, y) = -((x^2 - 1)y + 4x)$ . The invariant density is unknown in that case. However, we know that it is solution of the following corresponding Fokker-Planck equation

$$\frac{1}{2} \frac{\partial^2 p_s(x, y)}{\partial y^2} - y \frac{\partial p_s(x, y)}{\partial x} + c(x, y) p_s(x, y) + (c(x, y) y + V'(x)) \frac{\partial p_s(x, y)}{\partial y} = 0. \quad (6.5)$$

Thus the invariant density  $p_s(x, y)$  may be approximated by solving Equation (6.5) above, e.g. using a finite difference scheme. Doing so we remark that this density, although positive, is very small in many points. Therefore we plotted  $g(x, y) p_s(x, y)$  to avoid numerical instabilities.

We made use of the explicit Euler scheme to simulate an approximated discrete sampling. We also implemented an implicit scheme, but it slowed too much the simulations, and the results were not better.

On Figures 5 and 6 below we took  $n = 10^5, h_n = 0.18, b_{1,n} = b_{2,n} = 0.10$ , and the step for the explicit Euler scheme  $\delta = \frac{1}{10}h/3$ .

The drift at some fixed point  $y_0$ ,  $g(\cdot, y_0)$ , is estimated on a grid  $(z_l)_{l=1,\dots,L} = (x_l, y_0)_{l=1,\dots,L}$ . On Figure 5 below we chose  $y_0 = 0.0597$  and plotted  $g(\cdot, y_0) * p_s(\cdot, y_0)$ .

The drift at some fixed point  $x_0$ ,  $g(x_0, \cdot)$  is estimated on a grid  $(z_l)_{l=1,\dots,L} = (x_0, y_l)_{l=1,\dots,L}$ .  $L$  was chosen equal to 40.

On Figure 6 below we chose  $x_0 = -0.0183$  and plotted  $g(x_0, \cdot) * p_s(x_0, \cdot)$ .

#### REFERENCES

- D. Bakry, P. Cattiaux, and A. Guillin (2008). Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Func. Anal.* **254**, p. 727-759.
- M. Beska, A. Kłopotowski and L. Ślomiński (1982). Limit Theorems for Random Sums of Dependent  $d$ -Dimensional Random Vectors. *Z. Wahrsch. Verw. Gebiete.* **61**, p. 43-57.
- P. Cattiaux, J. R. León R., and C. Prieur (2012). Estimation for Stochastic Damping Hamiltonian Systems under Partial Observation. I. Invariant density. <http://hal.archives-ouvertes.fr/hal-00721616>
- C. Gardiner (1985). *Handbook of Stochastic Methods*. Springer Verlag, New York.
- Y. Pokern, A. Stuart and P. Wiberg (2009). Parameter estimation for partially observed hypoelliptic diffusions. *J. Roy. Stat. Soc. B* **71**, p. 49-73.
- L. Wu (2001). Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Processes and their Applications* **91**, p. 205-238.

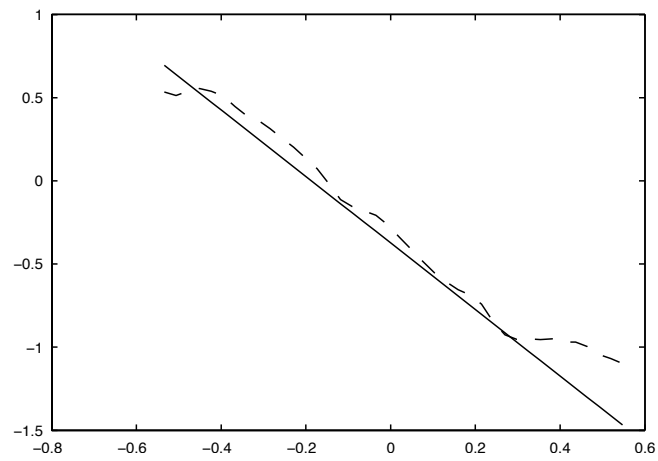


FIGURE 1. Drift's estimation for the harmonic oscillator: theoretical  $g(\cdot, y_0)$  in plain line, estimated in dashed line for  $y_0 = 0.0597$



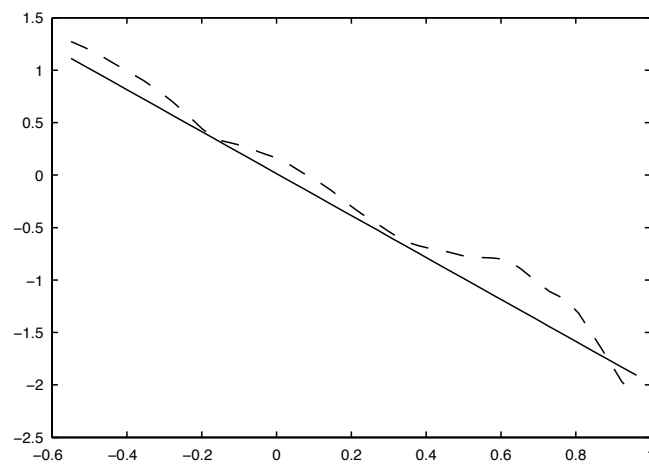


FIGURE 2. Drift's estimation for the harmonic oscillator: theoretical  $g(x_0, \cdot)$  in plain line, estimated in dashed line for  $x_0 = -0.0183$

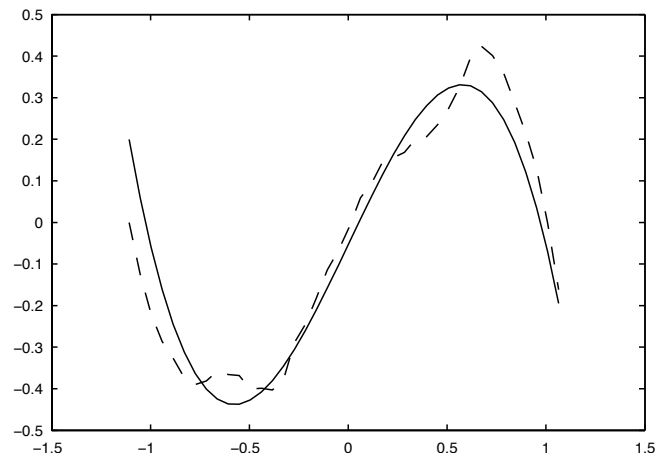


FIGURE 3. Drift's estimation for the Duffing oscillator: theoretical  $g(\cdot, y_0)$  in plain line, estimated in dashed line for  $y_0 = 0.0533$

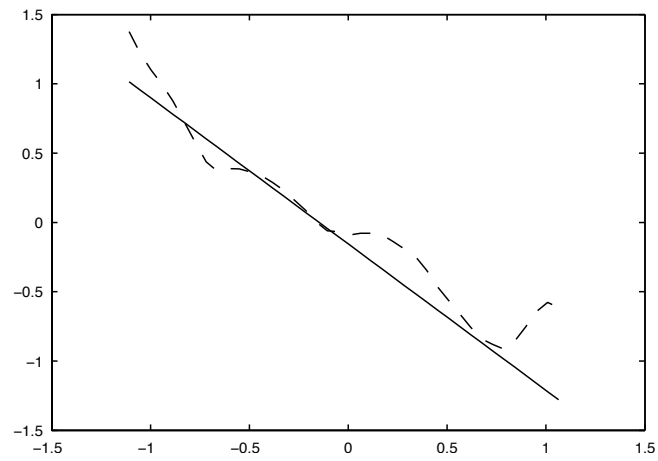


FIGURE 4. Drift's estimation for the Duffing oscillator: theoretical  $g(x_0, \cdot)$  in plain line, estimated in dashed line for  $x_0 = -0.0498$

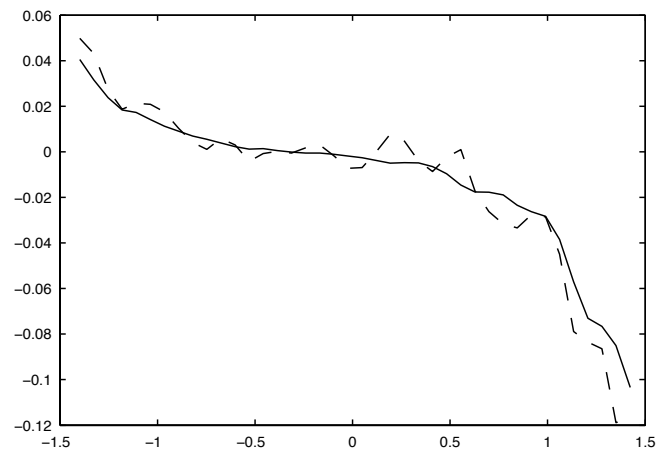


FIGURE 5. Drift's estimation for the Van der Pol oscillator: theoretical  $g(\cdot, y_0) * p_s(\cdot, y_0)$  in plain line, estimated in dashed line for  $y_0 = 0.0597$

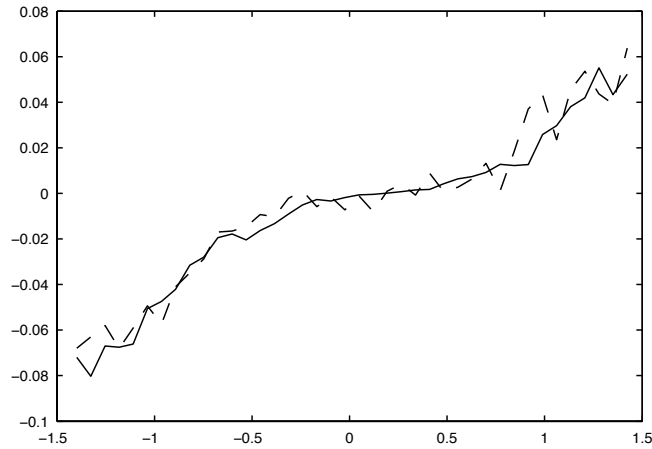


FIGURE 6. Drift's estimation for the Van der Pol oscillator: theoretical  $g(x_0, \cdot) * p_s(x_0, \cdot)$  in plain line, estimated in dashed line for  $x_0 = -0.0183$

**Patrick CATTIAUX**,, INSTITUT DE MATHÉMATIQUES DE TOULOUSE. UNIVERSITÉ DE TOULOUSE., CNRS UMR 5219.,, 118 ROUTE DE NARBONNE, 31062 TOULOUSE CEDEX 09, FRANCE.

*E-mail address:* cattiaux@math.univ-toulouse.fr

**José LEÓN**,, ESCUELA DE MATEMÁTICA, FACULTAD DE CIENCIAS, UNIVERSIDAD CENTRAL DE VENEZUELA., A.P. 47197, LOS CHAGUARAMOS, CARACAS 1041-A, VENEZUELA.

*E-mail address:* jose.leon@ciens.ucv.ve

**Clémentine PRIEUR**,, UNIVERSITÉ JOSEPH FOURIER, LJK/MOISE BP 53., 38041 GRENOBLE CEDEX, FRANCE.

*E-mail address:* clementine.prieur@imag.fr