

# Asymptotic lower bounds in estimating jumps

Emmanuelle Clément\*      Sylvain Delattre†      Arnaud Gloter‡

Revised version, September, 12 2012

## Abstract

We study the problem of the efficient estimation of the jumps for stochastic processes. We assume that the stochastic jump process  $(X_t)_{t \in [0,1]}$  is observed discretely, with a sampling step of size  $1/n$ . In the spirit of Hajek's convolution theorem, we show some lower bounds for the estimation error of the sequence of the jumps  $(\Delta X_{T_k})_k$ . As an intermediate result, we prove a LAMN property, with rate  $\sqrt{n}$ , when the marks of the underlying jump component are deterministic. We deduce then a convolution theorem, with an explicit asymptotic minimal variance, in the case where the marks of the jump component are random. To prove that this lower bound is optimal, we show that a threshold estimator of the sequence of jumps  $(\Delta X_{T_k})_k$  based on the discrete observations, reaches the minimal variance of the previous convolution theorem.

**MSC 2010.** Primary: 62G20, Secondary: 60F05, 60H05.

**Key words:** Itô process, LAMN property, convolution theorem.

## 1 Introduction

The statistical study of stochastic processes with jumps, from high frequency data, has been the subject of many recent works. A major issue is to determine if the jump part is relevant to model the observed phenomenon. Especially, for modelling of asset prices, the assessment of the part due to the jumps in the price is an important question. This has been addressed in several

---

\*Université Paris-Est, LAMA (UMR 8050), UPEMLV, UPEC, CNRS, F-77454, Marne-la-Vallée, France

†Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, Université Denis Diderot, 77251 Paris Cedex 05, France

‡Université d'Évry Val d'Essonne, Département de Mathématiques, 91025 Évry Cedex, France. This research benefited by the support of the 'Chaire Risque de crédit', Fédération Bancaire Française.

works, either by considering multi-power variations ([6], [7], [11]) or by truncation methods (see [19], [20]). Another issue is to test statistically if the stochastic process has continuous paths. The question has been addressed in many works (see [1], [2], [4]) and is crucial to the hedging of options. A clearly related question is to determine the degree of activity of the jump component of the process. Estimators of the Blumenthal-Gettoor index of the Lévy measure of the process are proposed in several papers ([3], [8], [23]).

In that context, the main statistical difficulty comes from the fact that one observes a discrete sampling of the process, and consequently, the exact values of the jumps are unobserved. As a matter of fact, a lot of statistical procedures rely on the estimation of a functional of the jumps. In [13], Jacod considers the estimation, from a high frequency sampling, of the functional of the jumps

$$\sum_{0 \leq s \leq 1, \Delta X_s \neq 0} f(\Delta X_s) = \sum_k f(\Delta X_{T_k}),$$

for a smooth function  $f$  vanishing at zero (see Theorems 2.11 and 2.12 in [13] for precise assumptions). In particular, he studies the difference between the unobserved quantity  $\sum_{0 \leq s \leq 1} f(\Delta X_s)$  and the observed one  $\sum_{i=0}^{n-1} f(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})$ . When  $X$  is a semi-martingale, it is shown that the difference between the two quantities goes to zero with rate  $\sqrt{n}$ . Rescaled by this rate, the difference is asymptotically distributed as

$$\sum_k f'(\Delta X_{T_k}) [\sigma_{T_k-} \sqrt{U_k} N_k^- + \sigma_{T_k} \sqrt{1 - U_k} N_k^+], \quad (1)$$

where the variables  $U_k$  are uniform variables on  $[0, 1]$  and  $N_k^-$ ,  $N_k^+$  are standard Gaussian variables. The quantity  $\sigma_{T_k-}$  (resp.  $\sigma_{T_k}$ ) is the local volatility of the semi martingale  $X$  before (resp. after) the jump at time  $T_k$ . This result serves as the basis for studying the statistical procedures developed in [4, 15].

However, the problem of the efficiency of these methods seems to have never been addressed. Motivated by these facts, we discuss, in this paper, the notion of efficiency to estimate the jumps from the discrete sampling  $(X_{\frac{i}{n}})_{0 \leq i \leq n}$ .

Let us stress, that the meaning of efficiency is not straightforward here. Indeed, we are not dealing with a standard parametric statistical problem, and it is not clear which quantity can stand for the Fisher's information. In this paper, we restrict ourself to processes  $X$  solutions of

$$X_t = x_0 + \int_0^t b(s, X_s) ds + \int_0^t a(s, X_s) dW_s + \sum_{T_k \leq t} c(X_{T_k-}, \Lambda_k),$$

where we assume that the number of jumps on  $[0, 1]$ , denoted by  $K$ , is finite. We note  $J = (\Delta X_{T_1}, \dots, \Delta X_{T_K})$  the vector of jumps, and  $\Lambda = (\Lambda_1, \dots, \Lambda_K)$  the random marks. The notion of efficiency will be stated in this context as a convolution result in Theorem 1. More precisely, we prove that for any estimator  $\tilde{J}^n$  such that the error  $\sqrt{n}(\tilde{J}^n - J)$  converges in law to some variable  $Z$ , the law of  $Z$  is a convolution between the law of the vector

$$[a(T_k, X_{T_k-})\sqrt{U_k}N_k^- + a(T_k, X_{T_k})\sqrt{1 - U_k}N_k^+]_{k=1, \dots, K}$$

and some other law. Contrarily to the standard convolution theorem, we do not need the usual regularity assumption on the estimator. The explanation is that we are not estimating a deterministic (unknown) parameter, but we estimate some random (unobserved) variable  $J$ .

The proof of this convolution result relies on the study of a preliminary parametric model: we consider the parametric model where the values of the marks  $\Lambda$  are considered as an unknown deterministic parameter  $\lambda \in \mathbb{R}^K$ . The resulting model is a stochastic differential equation with jumps, whose coefficients depend on this parameter  $\lambda$ . We establish then in Theorem 2, that this statistical experiment satisfies the LAMN property, with rate  $\sqrt{n}$  and some explicit Fisher's information matrix  $I(\lambda)$ .

By Hajek's theorem, it is well known that the LAMN property implies a convolution theorem for any regular estimator of the parameter  $\lambda$  (see [17], [22]). However our context differs from the usual Hajek's convolution theorem on at least two points. First, the parameter  $\lambda$  is randomized and second the target of the estimator  $J = (c(X_{T_k-}, \Lambda_k))_k$  depends both on the randomized parameter and on some unobserved quantities  $X_{T_k-}$ . As a result, the connection between the minimal law of the convolution theorem and the Fisher's information of the parametric model is not straightforward. The proof of the convolution theorem, when  $c(X_{T_k-}, \Lambda_k) = c(\Lambda_k)$  does not depend on  $X_{T_k-}$ , is simpler and is given in Theorem 4.

Remark that it is certainly possible to state a general result about the connection between the LAMN property and convolution theorems for the estimation of unobserved random quantities. The proof of the Proposition 4 is a step in this direction. However giving such general results is beyond the scope of the paper.

The outline of the paper is as follows. In Section 2, we state a convolution Theorem, which establishes an asymptotic lower bound for the asymptotic error of any estimator of the jumps. The LAMN property is enounced in Section 3. In Section 4, we show that the threshold estimator, introduced by Mancini (see [19], [20]), reaches the lower bound of Theorem 1. This proves

that this lower bound is optimal. The proofs of these results are postponed to the Section 5.

## 2 Convolution Theorem

### 2.1 Notation

Consider  $(X_t)_{t \in [0,1]}$  an adapted c.à.d.l.à.g., one dimensional, stochastic process defined on some filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,1]}, \mathbb{P})$ . We assume that the sample paths of  $X$  almost surely admit a finite number of jumps. We denote by  $K$  the random number of jumps on  $[0, 1]$  and  $0 < T_1 < \dots < T_K < 1$  the instants of these jumps. We assume that the process  $X$  is a solution of the stochastic differential equation with jumps

$$X_t = x_0 + \int_0^t b(s, X_s) ds + \int_0^t a(s, X_s) dW_s + \sum_{T_k \leq t} c(X_{T_k-}, \Lambda_k), \quad (2)$$

where  $W$  is a standard  $(\mathcal{F}_t)_t$  Brownian motion. The vector of marks  $(\Lambda_k)_k$  is random. The Brownian motion, the jump times and the marks are independent.

We will note  $J = (J_k)_{k \geq 1}$  the sequence of the jumps of the process, defined by  $J_k = c(X_{T_k-}, \Lambda_k) = \Delta X_{T_k}$ , for  $1 \leq k \leq K$  and  $J_k = 0$ , for  $k > K$ .

Remark that if  $T_k - T_{k-1}$  is exponentially distributed, the jumps times are arrival times of a Poisson process. Then, if the marks  $(\Lambda_k)_k$  are i.i.d. variables, the process  $\sum_{T_k \leq t} \Lambda_k$  is a compound Poisson process. In this particular case, the equation (2) becomes a standard S.D.E. with jumps based on a random Poisson measure with finite intensity.

It is convenient to assume that the process is realized on the canonical product space of the Brownian part and the jumps parts  $\Omega = \Omega^1 \times \Omega^2$ ,  $\mathbb{P} = \mathbb{P}^1 \otimes \mathbb{P}^2$ . More precisely, we note  $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1) = (\mathcal{C}([0,1]), \mathcal{B}, \mathbb{W})$ , the space of continuous functions endowed with the Wiener measure on the Borelian sigma-field and  $(\mathcal{F}_t^1)_{t \in [0,1]}$  the filtration generated by the canonical process. We introduce  $(\Omega^2, \mathcal{F}^2, \mathbb{P}^2) = (\mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}} \otimes \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}, \mathbb{P}^2)$ , where  $\mathbb{P}^2$  is the law of two independent sequences of random variables  $(T_k)_{k \geq 1}$ ,  $(\Lambda_k)_{k \geq 1}$ . We assume that,  $\mathbb{P}^2$ -almost surely, the sequence  $(T_k)_{k \geq 1}$  is non decreasing and such  $K = \text{Card}\{k, T_k \leq 1\}$  is finite. Then,  $((W_t)_{t \in [0,1]}, (T_k)_{k \geq 1}, (\Lambda_k)_{k \geq 1})$  are the canonical variables on  $\Omega$ . We assume that  $(\mathcal{F}_t)_t$  is the right continuous, completed, filtration based on  $(\mathcal{F}_t^1 \times \mathcal{F}^2)_t$  and  $\mathcal{F} = \mathcal{F}_1$ .

In order to describe the asymptotic law of any estimator of the jumps, we need some additional notation. Following [13], we introduce an extension of our initial probability space. We

consider an auxiliary probability space  $(\Omega', \mathcal{F}', \mathbb{P}')$  which contains  $U = (U_k)_{k \geq 1}$  a sequence of independent variables with uniform law on  $[0, 1]$ , and  $N^- = (N_k^-)_{k \geq 1}$ ,  $N^+ = (N_k^+)_{k \geq 1}$  two sequences of independent variables with standard Gaussian law. All these variables are mutually independent. We extend the initial probability space by setting  $\tilde{\Omega} = \Omega \times \Omega'$ ,  $\tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$ ,  $\tilde{\mathbb{P}} = \mathbb{P} \otimes \mathbb{P}'$ ,  $\tilde{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{F}'$ .

## 2.2 Main result

We need some more assumptions on the process. Especially, to avoid cumbersome notation we will first assume in the next subsection that the number of jumps is deterministic. We will show in Section 2.2.2 that this is not a real restriction, since we can reformulate our result by conditioning on the number of jumps  $K$ .

### 2.2.1 Deterministic number of jumps

Since the number of jumps  $K$  is deterministic, the probability space  $\Omega$  introduced in Section 2.1 is simplified accordingly :  $\Omega = \Omega^1 \times \Omega^2$ ,  $\Omega^1 = \mathcal{C}([0, 1])$  and  $\Omega^2 = \mathbb{R}^K \times \mathbb{R}^K$ . The space  $\tilde{\Omega} = \Omega \times \Omega'$  with  $\Omega' = \mathbb{R}^{3K}$  extends the initial space with the sequences  $N^- = (N_k^-)_{1 \leq k \leq K}$ ,  $N^+ = (N_k^+)_{1 \leq k \leq K}$ ,  $U = (U_k)_{1 \leq k \leq K}$ .

**H0 (law of the jump times).** The number of jumps  $K$  is deterministic and the law of  $T = (T_1, \dots, T_K)$  is absolutely continuous with respect to the Lebesgue measure. We note  $f_T$  its density.

**H1 (smoothness assumption).** The functions  $(t, x) \mapsto a(t, x)$  and  $(t, x) \mapsto b(t, x)$  are  $\mathcal{C}^{1,2}$  on  $[0, 1] \times \mathbb{R}$ . We note  $a'$  and  $b'$  their derivatives with respect to  $x$  and we assume that  $a'$  and  $b'$  are  $\mathcal{C}^{1,2}$  on  $[0, 1] \times \mathbb{R}$ . Moreover the functions  $a$ ,  $b$ , and their derivatives are uniformly bounded. The function  $(x, \theta) \mapsto c(x, \theta)$  is  $\mathcal{C}^{2,1}$  on  $\mathbb{R} \times \mathbb{R}$ , with bounded derivatives. We note  $c'$  its derivative with respect to  $x$  and  $\dot{c}$  its derivative with respect to  $\theta$ . We assume moreover that  $\dot{c}$  is  $\mathcal{C}^{1,1}$  with bounded derivatives.

**H2 (non-degeneracy assumption).** We assume that there exist two constants  $\underline{a}$  and  $\bar{a}$  such that

$$\forall (t, x) \in [0, 1] \times \mathbb{R} \quad 0 < \underline{a} \leq a(t, x) \leq \bar{a};$$

$$\forall (x, \theta) \in \mathbb{R} \times \mathbb{R} \quad |1 + c'(x, \theta)| \geq \underline{a}.$$

**H3 (“randomness” of the jump sizes).** The law of  $\Lambda = (\Lambda_1, \dots, \Lambda_K)$  is absolutely continuous with respect to the Lebesgue measure and we note  $f_\Lambda$  its density. We assume also

$$\forall(x, \theta) \in \mathbb{R} \times \mathbb{R} \quad \dot{c}(x, \theta) \neq 0.$$

Let us comment on these assumptions. First, the assumption that the vector of jump times admits a density, hypothesis H0, is crucial to prove the convergence in law of the fractional part of  $(nT_k)_k$  to the vector of uniform laws  $(U_k)_k$ . In order to find a lower bound, we need to deal with a kind of regular model, this explains the assumption H1. Moreover, it is clear that if the diffusion coefficient  $a$  is equal to zero, one will expect a rate of convergence for the estimation of the jumps faster than  $\sqrt{n}$ . In that case, the LAMN property will not be satisfied with rate  $\sqrt{n}$ . This clarifies why we assume a strictly positive lower bound on  $a$ .

Remark that the non degeneracy of  $|1 + c'(x, \theta)|$  is a standard assumption which implies that the equation (2) admits a flow. The assumption H3 is more specifically related to our statistical problem. We want to prove a lower bound for the estimation of the random jump sizes. Indeed, if these quantities do not exhibit enough randomness, it could be possible to estimate them with a rate faster than  $\sqrt{n}$ . For instance, the condition H3 excludes that the jump sizes do not depend on the underlying random marks.

We can now state our main result. We recall that  $J = (J_k)_{1 \leq k \leq K} = (c(X_{T_k-}, \Lambda_k))_{1 \leq k \leq K} = (\Delta X_{T_k})_{1 \leq k \leq K} \in \mathbb{R}^K$  is the sequence of the jumps of the process.

We will call  $(\tilde{J}^n)_{n \geq 1}$  a sequence of estimators if for each  $n$ ,  $\tilde{J}^n \in \mathbb{R}^K$  is a measurable function of the observations  $(X_{i/n})_{i=0, \dots, n}$ .

**Theorem 1** *Assume H0–H3. Let  $\tilde{J}^n$  be any sequence of estimators such that*

$$\sqrt{n}(\tilde{J}^n - J) \xrightarrow[\text{law}]{n \rightarrow \infty} \tilde{Z} \quad (3)$$

*for some variable  $\tilde{Z}$ . Then, the law of  $\tilde{Z}$  is necessarily a convolution:*

$$\tilde{Z} \stackrel{\text{law}}{=} \left( \sqrt{U_k} a(T_k, X_{T_k-}) N_k^- + \sqrt{1 - U_k} a(T_k, X_{T_k}) N_k^+ \right)_k + \tilde{R}, \quad (4)$$

*where the random vector  $\tilde{R}$  is independent of  $(N_k^-, N_k^+)_k$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$ .*

We will say that an estimator  $\tilde{J}^n$  of the jumps is efficient if the asymptotic distribution of  $\sqrt{n}(\tilde{J}^n - J)$  is equal in law to  $(\sqrt{U_k} a(T_k, X_{T_k-}) N_k^- + \sqrt{1 - U_k} a(T_k, X_{T_k}) N_k^+)_k$  (which corresponds to  $\tilde{R} = 0$ ).

It is well known that in parametric models, the Hajek's convolution theorem usually requires a regularity assumption on the estimator (see [12], [22]). Here, our theorem does not require any assumption on the estimator, apart its convergence with rate  $\sqrt{n}$ . This comes from the fact that the target  $J$  of the estimator is random, yielding to some additional regularity properties, compared with the usual parametric setting (see a related situation in Jeganathan [16]).

**Remark 1** *We can observe that*

$$\left( \sqrt{U_k} a(T_k, X_{T_k-}) N_k^- + \sqrt{1 - U_k} a(T_k, X_{T_k}) N_k^+ \right)_k \stackrel{\text{law}}{=} (I^{\text{opt}})^{-1/2} N,$$

where  $I^{\text{opt}}$  is the diagonal random matrix of size  $K \times K$ , defined on the extended probability space  $\tilde{\Omega}$ , with diagonal entries:

$$I_k^{\text{opt}} = [U_k a(T_k, X_{T_k-})^2 + (1 - U_k) a(T_k, X_{T_k})^2]^{-1}, \quad \text{for } k = 1, \dots, K. \quad (5)$$

The vector  $N$  is a standard Gaussian vector on  $\mathbb{R}^K$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$  and consequently  $N$  is independent of  $I^{\text{opt}}$ .

**Remark 2** *The Theorem 1 states, in particular, that any estimator of the jumps with rate  $\sqrt{n}$  must have an asymptotic conditional variance greater than  $(I^{\text{opt}})^{-1}$ .*

*Let us stress that if the rate of convergence is faster than  $\sqrt{n}$ , then (3) is still true with  $\tilde{Z} = 0$  and consequently the Theorem 1 proves that a convergence faster than  $\sqrt{n}$  is impossible.*

Now if instead estimating  $J$ , we estimate a function of the vector of jumps, we can prove in a similar way the following result. For the sake of shortness, we will omit the proof of the following proposition.

**Proposition 1** *Assume H0–H3. Let  $F$  be a  $C^1$  function from  $\mathbb{R}^K$  to  $\mathbb{R}$  and let  $\tilde{F}^n$  be any sequence of estimators of  $F(J)$  such that*

$$\sqrt{n}(\tilde{F}^n - F(J)) \xrightarrow[\text{law}]{n \rightarrow \infty} \tilde{Z}_F \quad (6)$$

for some variable  $\tilde{Z}_F$ . Then, the law of  $\tilde{Z}_F$  is necessarily a convolution:

$$\tilde{Z}_F \stackrel{\text{law}}{=} \sum_{k=1}^K \frac{\partial F}{\partial x_k}(J) \left( \sqrt{U_k} a(T_k, X_{T_k-}) N_k^- + \sqrt{1 - U_k} a(T_k, X_{T_k}) N_k^+ \right) + \tilde{R}_F, \quad (7)$$

where the real random variable  $\tilde{R}_F$  is independent of  $(N_k^-, N_k^+)_k$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$ .

**Remark 3** *From the results of Jacod (Theorems 2.11 and 2.12 in [13]), we deduce that the lower bound of Proposition 1 is optimal, and that the estimators of [13] are efficient.*

### 2.2.2 Random number of jumps

If the number of jumps is random, we need to modify some assumptions accordingly.

**H0.** We note  $K = \text{card}\{k \mid T_k \in [0, 1]\}$ . Conditionally on  $K$  the law of the vector of jump times  $T = (T_1, \dots, T_K)$  admits a density.

**H3.** Conditionally on  $K$ , the law of  $(\Lambda_1, \dots, \Lambda_K)$  is absolutely continuous with respect to the Lebesgue measure. We assume also  $\forall(x, \theta) \in \mathbb{R} \times \mathbb{R} \quad \dot{c}(x, \theta) \neq 0$ .

We can extend Theorem 1.

**Corollary 1** *Assume  $\widetilde{H0}$ ,  $H1$ ,  $H2$  and  $\widetilde{H3}$ . Let  $\widetilde{J}^n$  be any sequence of estimators with values in  $\mathbb{R}^N$  such that*

$$\sqrt{n}(\widetilde{J}^n - J) \xrightarrow[\text{law}]{n \rightarrow \infty} \widetilde{Z}$$

for some variable  $\widetilde{Z}$ . Then, the law of  $\widetilde{Z}$  admits the decomposition:

$$\widetilde{Z} \stackrel{\text{law}}{=} \left( [\sqrt{U_k} a(T_k, X_{T_k-}) N_k^- + \sqrt{1 - U_k} a(T_k, X_{T_k}) N_k^+] 1_{\{1 \leq k \leq K\}} \right)_k + \widetilde{R},$$

where conditionally on  $(K, (T_k)_{1 \leq k \leq K}, (\Lambda_k)_{1 \leq k \leq K}, (W_t)_{t \in [0, 1]}, (U_k)_{1 \leq k \leq K})$  the random vector of the  $K$  first components of  $\widetilde{R}$  is independent of  $(N_k^-, N_k^+)_{1 \leq k \leq K}$ .

In Section 3, we consider a parametric model related to the process (2), and enounce the associated LAMN property. This is the key step before proving Theorem 1 and Corollary 1. Remark that, directly considering the values of the jumps size as the parameter, is not the right choice. The reason is that the jump sizes are not independent of the Brownian motion  $(W_t)_t$ . Instead, we prefer to consider the values of the marks  $\Lambda$  as the statistical parameter.

## 3 LAMN property in an associated parametric model

We focus on the parametric model where the values of the marks  $\Lambda$  are considered as the unknown (deterministic) parameters, and  $K$  is deterministic. This is the crucial step before proving our convolution theorem.

More precisely, our aim is to obtain the LAMN property for the parametric model

$$X_t^\lambda = x_0 + \int_0^t b(s, X_s^\lambda) ds + \int_0^t a(s, X_s^\lambda) dW_s + \sum_{k=1}^K c(X_{T_k-}^\lambda, \lambda_k) 1_{t \geq T_k} \quad (8)$$

where the parameter  $\lambda = (\lambda_1, \dots, \lambda_K) \in \mathbb{R}^K$ . We note  $T = (T_1, \dots, T_K)$  the vector of jump times such that  $0 < T_1 < \dots < T_K < 1$ . Let us remark that, under the assumption H0, the



solutions of (8) might be defined on the probability space  $\Omega^1 \times \mathbb{R}^K$  endowed with the product of the Wiener measure and the law of the jumps times. But, to avoid new notation, we can assume that, for all  $\lambda \in \mathbb{R}^K$ , the process  $(X_t^\lambda)_{t \in [0,1]}$  is defined on the space  $(\Omega, \mathcal{F}, \mathbb{P})$  of Section 2.

In this model, we assume that we observe both the regular discretization  $(X_{i/n}^\lambda)_{1 \leq i \leq n}$  of the process solution of (8) on the time interval  $[0, 1]$  and the jump times vector  $T$ . The observation of  $T$  leads to a more tractable computation of the likelihood. This is not restrictive to add some observations to the statistical experiment, since our aim is to derive an asymptotic lower bound. Under H0 and H1 the law of the observations  $(T, (X_{i/n}^\lambda)_{1 \leq i \leq n})$  admits a density  $\mathbf{p}^{n,\lambda}$ . We note  $\mathbf{p}^{n,\lambda,T}$  the density of  $(X_{i/n}^\lambda)_{1 \leq i \leq n}$  conditionally on  $T$ . For  $h = (h_1, \dots, h_K) \in \mathbb{R}^K$  we introduce the log-likelihood ratio:

$$Z_n(\lambda, \lambda + h/\sqrt{n}, T, x_1, \dots, x_n) = \log \frac{\mathbf{p}^{n,\lambda+h/\sqrt{n}}}{\mathbf{p}^{n,\lambda}}(T, x_1, \dots, x_n) \quad (9)$$

**Theorem 2** *Assume H0, H1 and H2. Then, the statistical experiment  $(\mathbf{p}^{n,\lambda})_{\lambda \in \mathbb{R}^K}$  satisfies a LAMN property. For  $\lambda \in \mathbb{R}^K$ ,  $h \in \mathbb{R}^K$  we have:*

$$Z_n(\lambda, \lambda + h/\sqrt{n}, T, X_{1/n}^\lambda, \dots, X_1^\lambda) = \sum_{k=1}^K h_k I_n(\lambda)_k^{1/2} N_n(\lambda)_k - \frac{1}{2} \sum_{k=1}^K h_k^2 I_n(\lambda)_k + o_{\mathbf{p}^{n,\lambda}}(1) \quad (10)$$

where  $I_n(\lambda)$  is a diagonal random matrix and  $N_n(\lambda)$  are random vectors in  $\mathbb{R}^K$  such that

$$(I_n(\lambda), N_n(\lambda)) \xrightarrow[\text{law}]{n \rightarrow \infty} (I(\lambda), N),$$

with:

$$I(\lambda)_k = \frac{\dot{c}(X_{T_k-}^\lambda, \lambda_k)^2}{a^2(T_k, X_{T_k-}^\lambda)[1 + c'(X_{T_k-}^\lambda, \lambda_k)]^2 U_k + a^2(T_k, X_{T_k-}^\lambda + c(X_{T_k-}^\lambda, \lambda_k))(1 - U_k)}, \quad (11)$$

where  $U = (U_1, \dots, U_K)$  is a vector of independent uniform laws on  $[0, 1]$  such that  $U$ ,  $T$  and  $(W_t)_{t \in [0,1]}$  are independent, and conditionally on  $(U, T, (W_t)_{t \in [0,1]})$ ,  $N$  is a standard Gaussian vector in  $\mathbb{R}^K$ .

Actually, we can complete the statement of the theorem by giving explicit expressions for  $I_n(\lambda)$  and  $N_n(\lambda)$ :

$$I_n(\lambda)_k = \frac{\dot{c}(X_{i_k/n}^\lambda, \lambda_k)^2}{n D^{n,\lambda_k,k}(X_{i_k/n}^\lambda)}, \quad N_n(\lambda)_k = \frac{\sqrt{n}(X_{i_k+1/n}^\lambda - X_{i_k/n}^\lambda - c(X_{i_k/n}^\lambda, \lambda_k))}{\sqrt{n D^{n,\lambda_k,k}(X_{i_k/n}^\lambda)}},$$

$$D^{n,\lambda_k,k}(X_{i_k/n}^\lambda) = a^2\left(\frac{i_k}{n}, X_{i_k/n}^\lambda\right)(1 + c'(X_{i_k/n}^\lambda, \lambda_k))^2(T_k - \frac{i_k}{n}) + a^2\left(\frac{i_k}{n}, X_{i_k/n}^\lambda + c(X_{i_k/n}^\lambda, \lambda_k)\right)\left(\frac{i_k+1}{n} - T_k\right),$$

where  $i_k$  is the integer part of  $nT_k$ .

**Remark 4** We remark that from a direct application of Hajek's Theorem (see Van der Vaart [22], Corollary 9.9 p.132), any regular estimator of  $\lambda$  has an asymptotic conditional variance greater than  $I(\lambda)^{-1}$ . Here, an estimator of  $\lambda$  is a measurable function of  $(T, (X_{i/n}^\lambda)_{1 \leq i \leq n})$ , and so we deduce that, a fortiori, any measurable function of  $(X_{i/n}^\lambda)_{1 \leq i \leq n}$  satisfies the same asymptotic lower bound.

## 4 Efficient estimator of the jumps

We use the notation of Section 2.1 and since we just propose an estimator of the jumps, we can weaken the assumptions of the previous sections.

**A1 (smoothness assumption).** The functions  $a : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $b : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$  and  $c : \mathbb{R}^2 \rightarrow \mathbb{R}$  are continuous.

**A2 (identifiability of the jumps).** We have almost surely:  $c(X_{T_k-}, \Lambda_k) \neq 0$ ,  $\forall k \in \{1, \dots, K\}$ .

This last condition ensures that the jump times of  $X$  are exactly the times  $T_k$ .

Recall that  $J = (J_k)_{k \geq 1}$  is the sequence of jumps of  $X$  (on  $[0, 1]$ ): we set  $J_k = \Delta X_{T_k} = X_{T_k} - X_{T_k-}$  for  $k \leq K$  and we define  $J_k = 0$  for  $k > K$ .

We construct an estimator of  $J$  following the threshold estimation method proposed by Mancini [19], [20].

Let  $(u_n)_n$  be a sequence of positive numbers tending to 0. We set  $\hat{i}_1^n = \inf\{0 \leq i \leq n-1 : |X_{\frac{i+1}{n}} - X_{\frac{i}{n}}| \geq u_n\}$  with the convention  $\inf \emptyset = +\infty$ . We recursively define for  $k \geq 2$ ,

$$\hat{i}_k^n = \inf\{\hat{i}_{k-1}^n < i \leq n-1 : |X_{\frac{i+1}{n}} - X_{\frac{i}{n}}| \geq u_n\}. \quad (12)$$

We set  $\hat{K}_n = \sup\{k \geq 1 : \hat{i}_k^n < \infty\}$  the number of increments of the jump diffusion exceeding the threshold  $u_n$ . We then define for  $k \geq 1$ ,

$$\hat{J}_k^n = \begin{cases} X_{\frac{\hat{i}_{k+1}^n}{n}} - X_{\frac{\hat{i}_k^n}{n}}, & \text{if } k \leq \hat{K}_n \\ 0, & \text{if } k > \hat{K}_n \end{cases} \quad (13)$$

The sequence  $(\hat{J}_k^n)_n$  is an estimator of the vector of jumps  $J$ , and  $(\hat{K}_n)_n$  estimates the number of jumps.

**Proposition 2** Let us assume  $\widetilde{H}_0$ , A1, A2 and  $u_n \sim n^{-\varpi}$  with  $\varpi \in (0, 1/2)$ . Then, we have

almost surely,

$$\begin{aligned} \widehat{K}_n &= K \text{ for } n \text{ large enough} \\ \text{if } k \leq K, \quad \widehat{J}_k^n &\xrightarrow{n \rightarrow \infty} J_k = \Delta X_{T_k}, \\ \text{if } k > K, \quad \widehat{J}_k^n &= 0 \text{ for } n \text{ large enough.} \end{aligned}$$

The consistency result concerning the estimator  $\widehat{K}_n$  is a special case of Mancini ([20], Theorem 1) and the jump sizes  $(J_k)_k$  were consistently estimated in Mancini [19] with exactly the same estimator but when the observation time goes to infinity.

We now describe the asymptotic law of the error between  $\widehat{J}^n$  and  $J$ . Note that Theorem 3 in [20] gives the asymptotic distribution of the estimator of the sum of the jumps assuming that the diffusion coefficient  $a$  is independent of the Brownian process  $W$  and the jump part, this is the reason why the uniform laws do not appear in the asymptotic law. The situation is completely different here, since the diffusion coefficient  $a$  depends on the process  $X$ , and is more related to Jacod's results (see [13]).

**Theorem 3** *Let us assume  $\widetilde{H}_0$ , A1, A2 and  $u_n \sim n^{-\varpi}$  with  $\varpi \in (0, 1/2)$ . Then  $\sqrt{n}(\widehat{J}^n - J)$  converges in law to  $Z = (Z_k)_{k \geq 1}$  where the limit can be described on the extended space  $\widetilde{\Omega}$  by:*

$$\begin{aligned} Z_k &= \sqrt{U_k} a(T_k, X_{T_k-}) N_k^- + \sqrt{1 - U_k} a(T_k, X_{T_k}) N_k^+, \text{ for } k \leq K, \\ Z_k &= 0, \text{ for } k > K. \end{aligned}$$

Moreover the convergence is stable with respect to the sigma-field  $\mathcal{F}$ . Let us precise that, here, the convergence in law of the infinite dimensional vector  $\sqrt{n}(\widehat{J}^n - J)$  means the convergence of any finite dimensional marginals.

**Remark 5** *The Theorem 3 shows that the error for the estimation of the jump  $\Delta X_{T_k}$  is asymptotically conditionally Gaussian and that the estimator  $\widehat{J}^n$  is efficient. In particular, the conditional variance on  $(T, \Lambda, K, (W_t)_{t \in [0,1]}, (U_k)_k)$  of the error is equal to the lower bound  $(I^{\text{opt}})^{-1} = U_k a(T_k, X_{T_k-})^2 + (1 - U_k) a(T_k, X_{T_k})^2$ , and consequently this lower bound is optimal.*

## 5 Proof section

We divide the proofs into three sections.

We first prove the LAMN property of the parametric model in Section 5.1. Then, the convolution result is established in Section 5.2. Finally, the Section 5.3 is devoted to the proof of the convergence and normality of the estimator  $\widehat{J}^n$ .

We first state a Lemma which will be useful in the next sections.

**Lemma 1** *Let  $K_0 \in \mathbb{N} \setminus \{0\}$  and consider  $T = (T_1, \dots, T_{K_0})$  a random variable on  $[0, 1]^{K_0}$  with density  $f_T$ . For  $k = 1, \dots, K_0$ , we note  $i_k = [nT_k]$  the integer part of  $nT_k$ . Let  $(W_t)_{t \in [0,1]}$  be a standard Brownian motion independent of  $T$ .*

*Then, we have the convergence in law of the variables*

$$(T, (n(T_k - \frac{i_k}{n}))_k, (\sqrt{n}(W_{T_k} - W_{\frac{i_k}{n}}))_k, (\sqrt{n}(W_{\frac{i_k+1}{n}} - W_{T_k}))_k, (W_t)_{t \in [0,1]})$$

to

$$(T, (U_k)_k, (\sqrt{U_k}N_k^-)_k, (\sqrt{1 - U_k}N_k^+)_k, (W_t)_{t \in [0,1]})$$

where  $U = (U_1, \dots, U_{K_0})$  is a vector of independent uniform laws on  $[0, 1]$ ,  $N^- = (N_1^-, \dots, N_{K_0}^-)$  and  $N^+ = (N_1^+, \dots, N_{K_0}^+)$  are independent standard Gaussian vectors such that  $T, U, N^-, N^+$  and  $(W_t)_t$  are independent.

**Proof** The convergence of the vector  $(T, (\sqrt{n}(W_{T_k} - W_{\frac{i_k}{n}}))_k, (\sqrt{n}(W_{\frac{i_k+1}{n}} - W_{T_k}))_k, (W_t)_{t \in [0,1]})$  is a direct consequence of Lemma 6.2 in [14] ( see also Lemma 5.8 in [13] ) and following this proof (which is simpler in our case), there is no difficulty to add the variables  $(n(T_k - \frac{i_k}{n}))_k$  in the vector.

◇

## 5.1 LAMN property: proof of Theorem 2

We use the framework of Section 3 and we introduce some more notation. For  $k = 1, \dots, K$ , we note  $i_k = [nT_k]$  the integer part of  $nT_k$  and for  $t \in [i_k/n, (i_k + 1)/n]$ , we note  $(X_t^{\theta,k})$  the process solution of the following jump-diffusion equation with only one jump at time  $T_k$ :

$$X_t^{\theta,k} = X_0 + \int_0^t b(s, X_s^{\theta,k})ds + \int_0^t a(s, X_s^{\theta,k})dW_s + c(X_{T_k-}^{\theta,k}, \theta)1_{t \geq T_k}. \quad (14)$$

Under H1 and H2 and conditionally on  $T$ , this process admits a strictly positive conditional density, which is  $\mathcal{C}^1$  with respect to  $\theta$ . We will note  $p^{\theta,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y)$  the density of  $X_{\frac{i_k+1}{n}}^{\theta,k}$  conditionally on  $T$  and  $X_{\frac{i_k}{n}}^{\theta,k} = x$  and  $\dot{p}^{\theta,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y)$  its derivative with respect to  $\theta$ .

We observe that the log-likelihood ratio  $Z_n$  only involves the transition densities of  $X^\lambda$  on a time interval where a jump occurs. This transition is  $p^{\theta,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y)$  if there is exactly one jump in the corresponding interval. Then, one can easily see that the following decomposition holds for  $Z_n$ :

$$\begin{aligned} Z_n(\lambda, \lambda + h/\sqrt{n}, T, x_1, \dots, x_n) 1_{\mathcal{T}_n}(T) &= \sum_{k=1}^K \ln \frac{p^{\lambda_k+h_k/\sqrt{n}, T}}{p^{\lambda_k, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x_{i_k}, x_{i_k+1} \right) 1_{\mathcal{T}_n}(T) \\ &= \sum_{k=1}^K \int_{\lambda_k}^{\lambda_k+h_k/\sqrt{n}} \frac{\dot{p}^{\theta, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x_{i_k}, x_{i_k+1} \right) d\theta 1_{\mathcal{T}_n}(T) \end{aligned} \quad (15)$$

where  $1_{\mathcal{T}_n}(T)$  is the indicator function that there is at most one jump in each time interval  $[i/n, (i+1)/n)$  for  $i = 0, \dots, n-1$ .

We have now to study the asymptotic behaviour of (15). This is divided into several lemmas. The lemmas 2–4 give an expansion for the score function, with a uniform control in  $\theta$ . We deduce then an explicit expansion for  $\int_{\lambda_k}^{\lambda_k+h_k/\sqrt{n}} \frac{\dot{p}^{\theta, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x_{i_k}, x_{i_k+1} \right) d\theta$  in Lemma 5, and conclude by passing through the limit in Lemma 6.

We begin with a representation of  $\frac{\dot{p}^{\theta, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, y \right)$  as a conditional expectation, using Malliavin calculus. We refer to Nualart [21] for a detailed presentation of Malliavin calculus. The Malliavin calculus techniques to derive LAMN properties have been introduced by Gobet [10] in the case of multi-dimensional diffusion processes and then used by Gloter and Gobet [9] for integrated diffusions.

In all what follows, we will denote by  $C_p$  a constant (independent on  $n, k$  and  $\theta$ ) which value may change from line to line.

**Lemma 2** *Assuming H1 and H2, we have  $\forall (x, y) \in \mathbb{R}^2$ :*

$$\frac{\dot{p}^{\theta, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, y \right) = E^{x, T, k} \left( \delta(P^{n, \theta, k}) | X_{\frac{i_k+1}{n}}^{\theta, k} = y \right),$$

where  $E^{x, T, k}$  is the conditional expectation on  $T$  and  $X_{\frac{i_k}{n}}^{\theta, k} = x$ ,  $\delta$  is the Malliavin divergence operator and  $P^{n, \theta, k}$  is the process given on  $[\frac{i_k}{n}, \frac{i_k+1}{n}]$  by

$$P_s^{n, \theta, k} = \frac{(Y_{T_k}^{\theta, k} Y_s^{\theta, k})^{-1} (1 + c'(X_{T_k-}^{\theta, k}, \theta) 1_{s \leq T_k}) a(s, X_s^{\theta, k}) \dot{c}(X_{T_k-}^{\theta, k}, \theta)}{\int_{\frac{i_k}{n}}^{(i_k+1)/n} (Y_u^{\theta, k})^{-2} a^2(u, X_u^{\theta, k}) (1 + c'(X_{T_k-}^{\theta, k}, \theta) 1_{u \leq T_k})^2 du}$$

where  $(Y_t^{\theta, k})_t$  is the process solution of

$$Y_t^{\theta, k} = 1 + \int_0^t b'(s, X_s^{\theta, k}) Y_s^{\theta, k} ds + \int_0^t a'(s, X_s^{\theta, k}) Y_s^{\theta, k} dW_s. \quad (16)$$

We remark that under H1, the process  $(Y_t^{\theta,k})_t$  and its inverse satisfy  $\forall p \geq 1$ ,

$$(E(\sup_{0 \leq t \leq 1} |Y_t^{\theta,k}|^p))^{1/p} \leq C_p, \quad (E(\sup_{0 \leq t \leq 1} |Y_t^{\theta,k}|^{-p}))^{1/p} \leq C_p. \quad (17)$$

**Proof** The proof of Lemma 2 is based on Malliavin calculus on the time interval  $[i_k/n, (i_k+1)/n]$ , conditionally on  $T$  and  $(W_t)_{t \leq i_k/n}$ . We first observe that under H1 and H2, the process  $(X_t^{\theta,k})$  solution of (14) admits a derivative with respect to  $\theta$  that we will denote by  $(\dot{X}_t^{\theta,k})$  (see for example Kunita [18] since this problem is similar to the derivative with respect to the initial condition). Moreover  $(X_t^{\theta,k})$  and  $(\dot{X}_t^{\theta,k})$  belong respectively to the Malliavin spaces  $\mathbb{D}^{2,p}$  and  $\mathbb{D}^{1,p}$ ,  $\forall p \geq 1$ . Now, let  $\varphi$  be a smooth function with compact support, we have :

$$\frac{\partial}{\partial \theta} E^{x,T,k} \varphi(X_{\frac{i_k+1}{n}}^{\theta,k}) = E^{x,T,k} \varphi'(X_{\frac{i_k+1}{n}}^{\theta,k}) \dot{X}_{\frac{i_k+1}{n}}^{\theta,k}.$$

Using the integration by part formula (see Nualart [21] Proposition 2.1.4 p.100), we can write

$$E^{x,T,k} \varphi'(X_{\frac{i_k+1}{n}}^{\theta,k}) \dot{X}_{\frac{i_k+1}{n}}^{\theta,k} = E^{x,T,k} \varphi(X_{\frac{i_k+1}{n}}^{\theta,k}) H(X_{\frac{i_k+1}{n}}^{\theta,k}, \dot{X}_{\frac{i_k+1}{n}}^{\theta,k})$$

where the weight  $H$  can be expressed in terms of the Malliavin derivative of  $X_{\frac{i_k+1}{n}}^{\theta,k}$ , the inverse of its Malliavin variance-covariance matrix and the divergence operator as follows:

$$H(X_{\frac{i_k+1}{n}}^{\theta,k}, \dot{X}_{\frac{i_k+1}{n}}^{\theta,k}) = \delta \left( \dot{X}_{\frac{i_k+1}{n}}^{\theta,k} \gamma^{\theta,k} D X_{\frac{i_k+1}{n}}^{\theta,k} \right)$$

where

$$\gamma^{\theta,k} = \left( \int_{i_k/n}^{(i_k+1)/n} \left( D_u X_{\frac{i_k+1}{n}}^{\theta,k} \right)^2 du \right)^{-1}. \quad (18)$$

On the other hand, from Lebesgue derivative theorem, we have:

$$\frac{\partial}{\partial \theta} E^{x,T,k} \varphi(X_{\frac{i_k+1}{n}}^{\theta,k}) = \int \varphi(y) \dot{p}^{\theta,T} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, y \right) dy,$$

this leads to the following representation

$$\dot{p}^{\theta,T} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, y \right) = E^{x,T,k} (\delta(\dot{X}_{\frac{i_k+1}{n}}^{\theta,k} \gamma^{\theta,k} D X_{\frac{i_k+1}{n}}^{\theta,k}) | X_{\frac{i_k+1}{n}}^{\theta,k} = y) p^{\theta,T} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, y \right).$$

It remains to give a more tractable expression of  $\dot{X}_{\frac{i_k+1}{n}}^{\theta,k} \gamma^{\theta,k} D X_{\frac{i_k+1}{n}}^{\theta,k}$ . We first observe that:

$$\dot{X}_{\frac{i_k+1}{n}}^{\theta,k} = \dot{c}(X_{T_k^-}^{\theta,k}, \theta) + \int_{T_k}^{(i_k+1)/n} b'(u, X_u^{\theta,k}) \dot{X}_u^{\theta,k} du + \int_{T_k}^{(i_k+1)/n} a'(u, X_u^{\theta,k}) \dot{X}_u^{\theta,k} dW_u,$$

and consequently

$$\dot{X}_{\frac{i_k+1}{n}}^{\theta,k} = Y_{\frac{i_k+1}{n}}^{\theta,k} (Y_{T_k}^{\theta,k})^{-1} \dot{c}(X_{T_k-}^{\theta,k}, \theta), \quad (19)$$

where  $(Y_t^{\theta,k})$  is solution of (16). Turning to the Malliavin derivative of  $X_{\frac{i_k+1}{n}}^{\theta,k}$ , we first observe that  $DX_{\frac{i_k+1}{n}}^{\theta,k} \in L^2([i_k/n, (i_k+1)/n])$  and so we just have to explicit  $D_s X_{\frac{i_k+1}{n}}^{\theta,k}$  for  $s \neq T_k$ . Assuming first that  $T_k < s \leq (i_k+1)/n$ , we have for  $u \in [s, (i_k+1)/n]$ :

$$D_s X_u^{\theta,k} = a(s, X_s^{\theta,k}) + \int_s^u b'(v, X_v^{\theta,k}) D_s X_v^{\theta,k} dv + \int_s^u a'(v, X_v^{\theta,k}) D_s X_v^{\theta,k} dW_v,$$

and then  $D_s X_u^{\theta,k} = Y_u^{\theta,k} (Y_s^{\theta,k})^{-1} a(s, X_s^{\theta,k})$ .

Now, if  $i_k/n \leq s < T_k$ , we have for  $u \geq s$

$$\begin{aligned} D_s X_u^{\theta,k} &= a(s, X_s^{\theta,k}) + c'(X_{T_k-}^{\theta,k}, \theta) D_s X_{T_k-}^{\theta,k} 1_{u \geq T_k} \\ &\quad + \int_s^u b'(v, X_v^{\theta,k}) D_s X_v^{\theta,k} dv + \int_s^u a'(v, X_v^{\theta,k}) D_s X_v^{\theta,k} dW_v, \end{aligned}$$

and we deduce that  $D_s X_u^{\theta,k} = Y_u^{\theta,k} (1 + c'(X_{T_k-}^{\theta,k}, \theta) 1_{u \geq T_k}) (Y_s^{\theta,k})^{-1} a(s, X_s^{\theta,k})$ .

It follows that :

$$D_s X_{\frac{i_k+1}{n}}^{\theta,k} = Y_{\frac{i_k+1}{n}}^{\theta,k} (1 + c'(X_{T_k-}^{\theta,k}, \theta) 1_{s \leq T_k}) (Y_s^{\theta,k})^{-1} a(s, X_s^{\theta,k}). \quad (20)$$

From (19) and (20), we obtain

$$\dot{X}_{\frac{i_k+1}{n}}^{\theta,k} \gamma^{\theta,k} D_s X_{\frac{i_k+1}{n}}^{\theta,k} = \frac{(Y_{T_k}^{\theta,k} Y_s^{\theta,k})^{-1} (1 + c'(X_{T_k-}^{\theta,k}, \theta) 1_{s \leq T_k}) a(s, X_s^{\theta,k}) \dot{c}(X_{T_k-}^{\theta,k}, \theta)}{\int_{i_k/n}^{(i_k+1)/n} (Y_u^{\theta,k})^{-2} a^2(u, X_u^{\theta,k}) (1 + c'(X_{T_k-}^{\theta,k}, \theta) 1_{u \leq T_k})^2 du} = P_s^{n,\theta,k}, \quad (21)$$

and the Lemma 2 is proved.  $\diamond$

In the next Lemma, we explicit the conditional expectation appearing in the decomposition of  $\frac{\dot{p}^{\theta,T}}{p^{\theta,T}}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y)$ .

**Lemma 3** *Assuming H1 and H2, we have*

$$E^{x,T,k} \left( \delta(P^{n,\theta,k}) | X_{\frac{i_k+1}{n}}^{\theta,k} = y \right) = \frac{(y - x - c(x, \theta)) \dot{c}(x, \theta)}{D^{n,\theta,k}(x)} + E^{x,T,k}(Q^{n,\theta,k} | X_{\frac{i_k+1}{n}}^{\theta,k} = y), \quad (22)$$

with

$$D^{n,\theta,k}(x) = a^2(\frac{i_k}{n}, x) (1 + c'(x, \theta))^2 (T_k - \frac{i_k}{n}) + a^2(\frac{i_k}{n}, x + c(x, \theta)) (\frac{i_k+1}{n} - T_k) \quad (23)$$

and where  $Q^{n,\theta,k}$  satisfies

$$\forall p \geq 1 \quad (E^{x,T,k} |Q^{n,\theta,k}|^p)^{1/p} \leq C_p,$$

for a constant  $C_p$  independent of  $x, n$  and  $\theta$ .

The first term in the right hand side of (22) is the main term and we will prove later that the contribution of the conditional expectation of  $Q^{n,\theta,k}$  is negligible.

**Proof** We first give an approximation of the process  $P^{n,\theta,k}$  which depends on the position of  $s$  with respect to the jump time  $T_k$ . We have:

$$P_s^{n,\theta,k} = \frac{\left( (1 + c'(X_{\frac{i_k}{n}}^{\theta,k}, \theta))a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k})1_{[\frac{i_k}{n}, T_k]}(s) + a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k}) + c(X_{\frac{i_k}{n}}^{\theta,k}, \theta)1_{(T_k, \frac{i_{k+1}}{n}]}(s) \right) \dot{c}(X_{\frac{i_k}{n}}^{\theta,k}, \theta)}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})} + U_s^{n,\theta,k}, \quad (24)$$

where  $D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})$  is defined by (23) and  $U_s^{n,\theta,k}$  is a remainder term. We deduce then that

$$\delta(P^{n,\theta,k}) = \frac{\left( (1 + c'(X_{\frac{i_k}{n}}^{\theta,k}, \theta))a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k})(W_{T_k} - W_{\frac{i_k}{n}}) + a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k}) + c(X_{\frac{i_k}{n}}^{\theta,k}, \theta)(W_{\frac{i_{k+1}}{n}} - W_{T_k}) \right) \dot{c}(X_{\frac{i_k}{n}}^{\theta,k}, \theta)}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})} + \delta(U^{n,\theta,k}). \quad (25)$$

Now, we can approximate  $X_{\frac{i_{k+1}}{n}}^{\theta,k}$  in the following way:

$$X_{\frac{i_{k+1}}{n}}^{\theta,k} = X_{\frac{i_k}{n}}^{\theta,k} + c(X_{T_k-}^{\theta,k}, \theta) + a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k})(W_{T_k} - W_{\frac{i_k}{n}}) + a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k}) + c(X_{\frac{i_k}{n}}^{\theta,k}, \theta)(W_{\frac{i_{k+1}}{n}} - W_{T_k}) + R_1^{n,\theta,k},$$

but observing that

$$c(X_{T_k-}^{\theta,k}, \theta) = c(X_{\frac{i_k}{n}}^{\theta,k}, \theta) + c'(X_{\frac{i_k}{n}}^{\theta,k}, \theta)a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k})(W_{T_k} - W_{\frac{i_k}{n}}) + R_2^{n,\theta,k},$$

we finally obtain

$$X_{\frac{i_{k+1}}{n}}^{\theta,k} = X_{\frac{i_k}{n}}^{\theta,k} + c(X_{\frac{i_k}{n}}^{\theta,k}, \theta) + (1 + c'(X_{\frac{i_k}{n}}^{\theta,k}, \theta))a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k})(W_{T_k} - W_{\frac{i_k}{n}}) + a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k}) + c(X_{\frac{i_k}{n}}^{\theta,k}, \theta)(W_{\frac{i_{k+1}}{n}} - W_{T_k}) + R^{n,\theta,k}, \quad (26)$$

with  $R^{n,\theta,k} = R_1^{n,\theta,k} + R_2^{n,\theta,k}$ .

Putting together (25) and (26), this yields

$$\delta(P^{n,\theta,k}) = \frac{(X_{\frac{i_{k+1}}{n}}^{\theta,k} - X_{\frac{i_k}{n}}^{\theta,k} - c(X_{\frac{i_k}{n}}^{\theta,k}, \theta))\dot{c}(X_{\frac{i_k}{n}}^{\theta,k}, \theta)}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})} - R^{n,\theta,k} \frac{\dot{c}(X_{\frac{i_k}{n}}^{\theta,k}, \theta)}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})} + \delta(U^{n,\theta,k}). \quad (27)$$

Letting  $Q^{n,\theta,k}$  be the random variable defined by

$$Q^{n,\theta,k} = \delta(U^{n,\theta,k}) - R^{n,\theta,k} \frac{\dot{c}(X_{\frac{i_k}{n}}^{\theta,k}, \theta)}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})}, \quad (28)$$



where  $U^{n,\theta,k}$  and  $R^{n,\theta,k}$  are respectively defined by (24) and (26), we deduce easily the first part of Lemma 3. It remains to bound  $E^{x,T,k}|Q^{n,\theta,k}|^p, \forall p \geq 1$ .

We remark that from H1 and H2

$$0 \leq \frac{|\dot{c}(X_{\frac{i_k}{n}}^{\theta,k}, \theta)|}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})} \leq nC, \quad (29)$$

for a constant  $C$  independent on  $n, k$  and  $\theta$ . Moreover, we have,

$$(E \sup_{\frac{i_k}{n} \leq s \leq T_k - \frac{i_k}{n}} |X_s^{\theta,k} - X_{\frac{i_k}{n}}^{\theta,k}|^p)^{1/p} \leq \frac{C_p}{\sqrt{n}} \quad \text{and} \quad (E \sup_{T_k \leq s \leq \frac{i_{k+1}}{n}} |X_s^{\theta,k} - X_{T_k}^{\theta,k}|^p)^{1/p} \leq \frac{C_p}{\sqrt{n}} \quad (30)$$

So, one can easily deduce that, assuming H1,

$$(E^{x,T,k}|R^{n,\theta,k}|^p)^{1/p} \leq C_p/n,$$

and combining this with (29), we derive

$$E^{x,T,k} \left( |R^{n,\theta,k}| \frac{|\dot{c}(X_{\frac{i_k}{n}}^{\theta,k}, \theta)|}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})} \right)^p \leq C_p$$

Turning to  $\delta(U^{n,\theta,k})$ , we first recall that, from the continuity property of the divergence operator (see Nualart [21] proposition 1.5.8 p.80), we have

$$(E^{x,T,k}|\delta(U^{n,\theta,k})|^p)^{1/p} \leq C_p(\|U^{n,\theta,k}\|_p + \|DU^{n,\theta,k}\|_p), \quad (31)$$

where

$$\|U^{n,\theta,k}\|_p^p = E^{x,T,k} \left( \int_{i_k/n}^{(i_k+1)/n} |U_s^{n,\theta,k}|^2 ds \right)^{p/2}, \quad (32)$$

$$\|DU^{n,\theta,k}\|_p^p = E^{x,T,k} \left( \int_{i_k/n}^{(i_k+1)/n} \int_{i_k/n}^{(i_k+1)/n} |D_v U_s^{n,\theta,k}|^2 ds dv \right)^{p/2}. \quad (33)$$

To bound  $U^{n,\theta,k}$ , we first observe that from (32)

$$\|U^{n,\theta,k}\|_p^p \leq \left(\frac{1}{n}\right)^{p/2} E^{x,T,k} \sup_{\frac{i_k}{n} \leq s \leq \frac{i_{k+1}}{n}} |U_s^{n,\theta,k}|^p,$$

so we just have to prove

$$(E^{x,T,k} \sup_{\frac{i_k}{n} \leq s \leq \frac{i_{k+1}}{n}} |U_s^{n,\theta,k}|^p)^{1/p} \leq C_p \sqrt{n}. \quad (34)$$

The error term  $U^{n,\theta,k}$  is defined by (24) as the difference between  $P_s^{\theta,n,k}$ , given in (21), and an explicit ratio :

$$U_s^{n,\theta,k} = \frac{(1 + c'(X_{T_k-}^{\theta,k}, \theta)1_{s \leq T_k})a(s, X_s^{\theta,k})\dot{c}(X_{T_k-}^{\theta,k}, \theta)}{Y_{T_k}^{\theta,k}Y_s^{\theta,k} \int_{i_k/n}^{(i_k+1)/n} (Y_u^{\theta,k})^{-2} a^2(u, X_u^{\theta,k})(1 + c'(X_{T_k-}^{\theta,k}, \theta)1_{u \leq T_k})^2 du} \\ \frac{\left( (1 + c'(X_{\frac{i_k}{n}}^{\theta,k}, \theta))a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k})1_{[\frac{i_k}{n}, T_k]}(s) + a(\frac{i_k}{n}, X_{\frac{i_k}{n}}^{\theta,k}) + c(X_{\frac{i_k}{n}}^{\theta,k}, \theta)1_{(T_k, \frac{i_k+1}{n}]}(s) \right) \dot{c}(X_{\frac{i_k}{n}}^{\theta,k}, \theta)}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})}.$$

Since  $\dot{c}$  and  $c'$  are bounded, we see easily from (30) that the difference between the numerators is of order  $1/\sqrt{n}$ . Now, we remark that

$$(E \sup_{\frac{i_k}{n} \leq s, u \leq \frac{i_k+1}{n}} |Y_{T_k}^{\theta,k} Y_s^{\theta,k} (Y_u^{\theta,k})^{-2} - 1|^p)^{1/p} \leq \frac{C_p}{\sqrt{n}}, \quad (35)$$

and that, using the non-degeneracy assumption H2

$$\int_{i_k/n}^{(i_k+1)/n} (Y_u^{\theta,k})^{-2} a^2(u, X_u^{\theta,k})(1 + c'(X_{T_k-}^{\theta,k}, \theta)1_{u \leq T_k})^2 du \geq \frac{\underline{a}^2 \min(1, \underline{a}^2)}{n \sup_{\frac{i_k}{n} \leq u \leq \frac{i_k+1}{n}} (Y_u^{\theta,k})^2}. \quad (36)$$

So, combining (17), (29), (35) and (36), we obtain

$$(E \sup_{\frac{i_k}{n} \leq s \leq \frac{i_k+1}{n}} \left| \frac{1}{Y_{T_k}^{\theta,k} Y_s^{\theta,k} \int_{i_k/n}^{(i_k+1)/n} (Y_u^{\theta,k})^{-2} a^2(u, X_u^{\theta,k})(1 + c'(X_{T_k-}^{\theta,k}, \theta)1_{u \leq T_k})^2 du} - \frac{1}{D^{n,\theta,k}(X_{\frac{i_k}{n}}^{\theta,k})} \right|^p)^{1/p} \\ \leq C_p \sqrt{n}$$

This proves (34) and consequently

$$\|U^{n,\theta,k}\|_p \leq C_p. \quad (37)$$

It remains to bound the Malliavin derivative of  $U^{n,\theta,k}$ . From (24) and (21), we have for  $v \in [i_k/n, (i_k+1)/n]$

$$D_v U_s^{n,\theta,k} = D_v P_s^{n,\theta,k} = D_v (\dot{X}_{\frac{i_k+1}{n}}^{\theta,k} \gamma^{\theta,k} D_s X_{\frac{i_k+1}{n}}^{\theta,k}).$$

Under H1, the Malliavin derivatives of  $\dot{X}_{\frac{i_k+1}{n}}^{\theta,k}$  and  $D_s X_{\frac{i_k+1}{n}}^{\theta,k}$  are bounded in  $L^p$ . Turning to the inverse of the Malliavin variance-covariance matrix  $\gamma^{\theta,k}$ , given by (18), we have

$$\gamma^{\theta,k} = \frac{1}{\int_{i_k/n}^{(i_k+1)/n} (Y_{\frac{i_k+1}{n}}^{\theta,k})^2 (Y_u^{\theta,k})^{-2} a^2(u, X_u^{\theta,k})(1 + c'(X_{T_k-}^{\theta,k}, \theta)1_{u \leq T_k})^2 du}$$

and from (17) and (36), it is easy to see that

$$(E^{x,T,k} |\gamma^{\theta,k}|^p)^{1/p} \leq n C_p \quad \text{and} \quad (E^{x,T,k} \sup_{\frac{i_k}{n} \leq v \leq \frac{i_k+1}{n}} |D_v \gamma^{\theta,k}|^p)^{1/p} \leq n C_p. \quad (38)$$

Putting this together, we obtain

$$(E^{x,T,k} \sup_{\frac{i_k}{n} \leq s, v \leq \frac{i_{k+1}}{n}} |D_v U_s^{n,\theta,k}|^p)^{1/p} \leq nC_p,$$

and then

$$\|DU^{n,\theta,k}\|_p \leq C_p. \quad (39)$$

From (31), (37) and (39), we deduce

$$(E^{x,T,k} |\delta(U^{n,\theta,k})|^p)^{1/p} \leq C_p,$$

and the Lemma 3 is proved.  $\diamond$

The bound on  $Q^{n,\theta,k}$  given in Lemma 3 is not sufficient, since to obtain the LAMN property, we have to compute the conditional expectation with  $x = X_{\frac{i_k}{n}}^\lambda$  and  $y = X_{\frac{i_{k+1}}{n}}^\lambda$ . So we complete the Lemma 3 with the following bound.

**Lemma 4** *With the assumptions and notations of Lemma 3, we have for  $\theta$  such that  $|\theta - \lambda_k| \leq C/\sqrt{n}$*

$$E^{x,T,k} |E^{x,T,k}(Q^{n,\theta,k} |X_{\frac{i_k+1}{n}}^{\theta,k} = X_{\frac{i_k+1}{n}}^\lambda)| \leq C',$$

where the constant  $C'$  is independent of  $x, n$  and  $\theta$ .

**Proof** We first remark that

$$E^{x,T,k} |E^{x,T,k}(Q^{n,\theta,k} |X_{\frac{i_k+1}{n}}^{\theta,k} = X_{\frac{i_k+1}{n}}^\lambda)| \leq E^{x,T,k} |Q^{n,\theta,k} | \frac{p^{\lambda_k, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, X_{\frac{i_k+1}{n}}^{\theta,k} \right). \quad (40)$$

From Hölder's inequality and Lemma 3, we obtain for  $p > 1, q > 1$  such that  $1/p + 1/q = 1$

$$E^{x,T,k} |E^{x,T,k}(Q^{n,\theta,k} |X_{\frac{i_k+1}{n}}^{\theta,k} = X_{\frac{i_k+1}{n}}^\lambda)| \leq C_p \left( E^{x,T,k} \left( \frac{p^{\lambda_k, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, X_{\frac{i_k+1}{n}}^{\theta,k} \right) \right)^q \right)^{1/q}, \quad (41)$$

and the result of Lemma 4 reduces to prove that there exists  $q_0 > 1$  such that

$$E^{x,T,k} \left( \frac{p^{\lambda_k, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, X_{\frac{i_k+1}{n}}^{\theta,k} \right) \right)^{q_0} \leq C \quad (42)$$

where  $C$  is independent of  $n, x$  and  $\theta$ . We can write:

$$E^{x,T,k} \left( \frac{p^{\lambda_k, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, X_{\frac{i_k+1}{n}}^{\theta,k} \right) \right)^{q_0} = \int p^{\lambda_k, T} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, y \right)^{q_0} p^{\theta, T} \left( \frac{i_k}{n}, \frac{i_k+1}{n}, x, y \right)^{1-q_0} dy, \quad (43)$$

and we can express the transition  $p^{\theta,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y)$  by decomposing it in terms on the transitions of a diffusion without jump on the time intervals  $(\frac{i_k}{n}, T_k)$  and  $(T_k, \frac{i_k+1}{n})$

$$p^{\theta,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y) = \int p^{\theta,T}(\frac{i_k}{n}, T_k, x, z) p^{\theta,T}(T_k, \frac{i_k+1}{n}, z + c(z, \theta), y) dz. \quad (44)$$

Now, assuming H1 and H2, we have the following classical estimates of the transition probabilities of a diffusion process (see Azencott [5] p.478), for some constants  $C_1, C_2$ :

$$C_1 G(x, \underline{a}^2(T_k - \frac{i_k}{n}), z) \leq p^{\theta,T}(\frac{i_k}{n}, T_k, x, z) \leq C_2 G(x, \bar{a}^2(T_k - \frac{i_k}{n}), z),$$

$$C_1 G(z + c(z, \theta), \underline{a}^2(\frac{i_k+1}{n} - T_k), y) \leq p^{\theta,T}(T_k, \frac{i_k+1}{n}, z + c(z, \theta), y) \leq C_2 G(z + c(z, \theta), \bar{a}^2(\frac{i_k+1}{n} - T_k), y)$$

where  $G(m, \sigma^2, y)$  denotes the density of the Gaussian law with mean  $m$  and variance  $\sigma^2$ . To simplify the notation, we note  $\sigma_{k,n}^- = T_k - \frac{i_k}{n}$  and  $\sigma_{k,n}^+ = \frac{i_k+1}{n} - T_k$ . Plugging this in (44), we obtain

$$p^{\theta,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y) \geq C_1 \int G(x, \underline{a}^2 \sigma_{k,n}^-, z) G(z + c(z, \theta), \underline{a}^2 \sigma_{k,n}^+, y) dz := I_1. \quad (45)$$

We get analogously,

$$p^{\lambda_k,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y) \leq C_2 \int G(x, \bar{a}^2 \sigma_{k,n}^-, z) G(z + c(z, \lambda_k), \bar{a}^2 \sigma_{k,n}^+, y) dz := I_2. \quad (46)$$

Observe that, in order to bound (43), we have to compute an upper bound for  $p^{\lambda_k,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y)$  and a lower bound for  $p^{\theta,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y)$ , since  $1 - q_0 < 0$ .

Our aim now is to give more tractable bounds for the transition density  $p^{\theta,T}(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y)$ . For this, we make the following change of variables in the integrals  $I_1$  and  $I_2$  defined in (46) and (45). We put  $u = \varphi(z) = z + c(z, \theta) - x - c(x, \theta)$ . We observe that  $\varphi(x) = 0$ . Moreover, from H1 and H2,  $\varphi$  is invertible and its derivative satisfies, for some constant  $c_0$  :

$$0 < \underline{a} \leq |\varphi'(z)| \leq c_0,$$

and consequently

$$\frac{1}{c_0} |z| \leq |\varphi^{-1}(z) - \varphi^{-1}(0)| \leq \frac{1}{\underline{a}} |z|.$$

So we obtain, for some constant  $C_1$

$$\begin{aligned} I_1 &\geq C_1 \int G(0, \underline{a}^2 \sigma_{k,n}^-, \varphi^{-1}(u) - \varphi^{-1}(0)) G(u + x + c(x, \theta), \underline{a}^2 \sigma_{k,n}^+, y) du, \\ &\geq C_1 \int G(0, \underline{a}^4 \sigma_{k,n}^-, u) G(x + c(x, \theta), \underline{a}^2 \sigma_{k,n}^+, y - u) du, \\ &= C_1 G(x + c(x, \theta), \underline{a}^4 \sigma_{k,n}^- + \underline{a}^2 \sigma_{k,n}^+, y). \end{aligned} \quad (47)$$

Proceeding similarly,

$$I_2 \leq C_2 G(x + c(x, \lambda_k), c_0^2 \bar{a}^2 \sigma_{k,n}^- + \bar{a}^2 \sigma_{k,n}^+, y). \quad (48)$$

Turning back to (43), it follows that

$$E^{x,T,k} \left( \frac{p^{\lambda_k, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k + 1}{n}, x, X_{\frac{i_k+1}{n}}^{\theta, k} \right) \right)^{q_0} \leq C \int G^{q_0}(x + c(x, \lambda_k), \sigma_{k,n}^1, y) G^{1-q_0}(x + c(x, \theta), \sigma_{k,n}^2, y) dy, \quad (49)$$

where  $\sigma_{k,n}^1 = c_0^2 \bar{a}^2 \sigma_{k,n}^- + \bar{a}^2 \sigma_{k,n}^+$  and  $\sigma_{k,n}^2 = \underline{a}^4 \sigma_{k,n}^- + \underline{a}^2 \sigma_{k,n}^+$ . Since  $\sigma_{k,n}^- + \sigma_{k,n}^+ = 1/n$ , we check that  $\sigma_{k,n}^1$  and  $\sigma_{k,n}^2$  are both lower and upper bound by some constants over  $n$ . Moreover, we have

$$\sigma_{k,n}^1 - \sigma_{k,n}^2 = (c_0^2 \bar{a}^2 - \underline{a}^4) \sigma_{k,n}^- + (\bar{a}^2 - \underline{a}^2) \sigma_{k,n}^+,$$

with  $c_0^2 > \bar{a}^2$  and  $\bar{a}^2 > \underline{a}^2$ , so  $0 < \sigma_{k,n}^2 / \sigma_{k,n}^1 < 1$ .

Turning back to the right hand side term of (49), we have to bound

$$\int \frac{e^{-q_0 \frac{(y-x-c(x, \lambda_k))^2}{2\sigma_{k,n}^1}}}{(2\pi\sigma_{k,n}^1)^{q_0/2}} \frac{e^{-(1-q_0) \frac{(y-x-c(x, \theta))^2}{2\sigma_{k,n}^2}}}{(2\pi\sigma_{k,n}^2)^{(1-q_0)/2}} dy$$

with  $1 < q_0$ . First we observe that this integral is finite if  $q_0 / \sigma_{k,n}^1 + (1 - q_0) / \sigma_{k,n}^2 > 0$ , that is  $1 < q_0 < \sigma_{k,n}^1 / (\sigma_{k,n}^1 - \sigma_{k,n}^2)$ . This choice of  $q_0$  is possible since  $0 < \sigma_{k,n}^2 / \sigma_{k,n}^1 < 1$ . After some calculus, we get

$$\int \frac{e^{-q_0 \frac{(y-x-c(x, \lambda_k))^2}{2\sigma_{k,n}^1}}}{(2\pi\sigma_{k,n}^1)^{q_0/2}} \frac{e^{-(1-q_0) \frac{(y-x-c(x, \theta))^2}{2\sigma_{k,n}^2}}}{(2\pi\sigma_{k,n}^2)^{(1-q_0)/2}} dy = \frac{\sqrt{\frac{2\pi}{q_0/\sigma_{k,n}^1 + (1-q_0)/\sigma_{k,n}^2}}}{(2\pi\sigma_{k,n}^1)^{q_0/2} (2\pi\sigma_{k,n}^2)^{(1-q_0)/2}} e^{+\frac{c_n}{2} (c(x, \theta) - c(x, \lambda_k))^2},$$

with

$$c_n = \left( \frac{q_0(q_0 - 1)}{\sigma_{k,n}^1 \sigma_{k,n}^2} \right) / \left( \frac{q_0}{\sigma_{k,n}^1} + \frac{(1 - q_0)}{\sigma_{k,n}^2} \right) > 0$$

Recalling that  $\sigma_{k,n}^1$  and  $\sigma_{k,n}^2$  are of order  $1/n$ , we observe that  $c_n$  is bounded by some constant times  $n$  and assuming that  $|\theta - \lambda_k| \leq C/\sqrt{n}$ , we finally obtain

$$E^{x,T,k} \left( \frac{p^{\lambda_k, T}}{p^{\theta, T}} \left( \frac{i_k}{n}, \frac{i_k + 1}{n}, x, X_{\frac{i_k+1}{n}}^{\theta, k} \right) \right)^{q_0} \leq C',$$

for a constant  $C'$  independent on  $x$ ,  $n$  and  $\theta$  and the Lemma 4 is proved.  $\diamond$

**Lemma 5** *Assuming H1 and H2, we have:*

$$\int_{\lambda_k}^{\lambda_k+h_k/\sqrt{n}} \frac{\dot{p}^{\theta,T}}{p^{\theta,T}}\left(\frac{i_k}{n}, \frac{i_k+1}{n}, X_{\frac{i_k}{n}}^\lambda, X_{\frac{i_k+1}{n}}^\lambda\right) d\theta = h_k \frac{\sqrt{n}(X_{\frac{i_k+1}{n}}^\lambda - X_{\frac{i_k}{n}}^\lambda - c(X_{\frac{i_k}{n}}^\lambda, \lambda_k)) \dot{c}(X_{\frac{i_k}{n}}^\lambda, \lambda_k)}{nD^{n,\lambda_k,k}(X_{\frac{i_k}{n}}^\lambda)} - \frac{h_k^2}{2} \frac{\dot{c}(X_{\frac{i_k}{n}}^\lambda, \lambda_k)^2}{nD^{n,\lambda_k,k}(X_{\frac{i_k}{n}}^\lambda)} + o_{\mathbf{P}^{n,\lambda}}(1).$$

**Proof** We deduce easily from Lemmas 2 and 3 that

$$\int_{\lambda_k}^{\lambda_k+h_k/\sqrt{n}} \frac{\dot{p}^{\theta,T}}{p^{\theta,T}}\left(\frac{i_k}{n}, \frac{i_k+1}{n}, x, y\right) d\theta = \int_{\lambda_k}^{\lambda_k+h_k/\sqrt{n}} \frac{(y-x-c(x,\theta))\dot{c}(x,\theta)}{D^{n,\theta,k}(x)} d\theta + \int_{\lambda_k}^{\lambda_k+h_k/\sqrt{n}} E^{x,T,k}(Q^{n,\theta,k}|X_{\frac{i_k+1}{n}}^{\theta,k} = y) d\theta$$

with  $(x, y) = (X_{\frac{i_k+1}{n}}^\lambda, X_{\frac{i_k}{n}}^\lambda)$ . From Lemma 4, the second term on the right hand side of the preceding equation tends to zero in probability. Now, from a Taylor expansion of  $c$ , we have the approximation for  $\theta \in [\lambda_k, \lambda_k + h_k/\sqrt{n}]$

$$\frac{(y-x-c(x,\theta))\dot{c}(x,\theta)}{D^{n,\theta,k}(x)} = \frac{(y-x-c(x,\lambda_k)) - (\theta-\lambda_k)\dot{c}(x,\lambda_k)}{D^{n,\lambda_k,k}(x)} \dot{c}(x,\lambda_k) + \varepsilon^{n,\theta,\lambda_k}(x, y). \quad (50)$$

From H1, and using (29), we have  $\forall \theta \in [\lambda_k, \lambda_k + h_k/\sqrt{n}]$

$$\left| \frac{\dot{c}(x,\theta)}{D^{n,\theta,k}(x)} - \frac{\dot{c}(x,\lambda_k)}{D^{n,\lambda_k,k}(x)} \right| \leq C\sqrt{n}, \quad (51)$$

where  $C$  does not depend on  $x$ . So we deduce that  $\forall \theta \in [\lambda_k, \lambda_k + h_k/\sqrt{n}]$

$$|\varepsilon^{n,\theta,\lambda_k}(x, y)| \leq C(1 + \sqrt{n}|y-x-c(x,\lambda_k)|),$$

for a constant  $C$  independent on  $x$  and  $y$ . Consequently, it follows that  $\int_{\lambda_k}^{\lambda_k+h_k/\sqrt{n}} \varepsilon^{n,\theta,\lambda_k}(X_{\frac{i_k}{n}}^\lambda, X_{\frac{i_k+1}{n}}^\lambda) d\theta$  goes to zero in probability as  $n$  goes to infinity, and the thesis follows.  $\diamond$

**Lemma 6** *Let us assume H0–H2. Let  $I_n(\lambda)$  be the diagonal matrix of size  $K \times K$ , and  $N_n(\lambda)$  be the random vector of size  $K$ , defined by the entries,*

$$I_n(\lambda)_k = \frac{\dot{c}(X_{\frac{i_k}{n}}^\lambda, \lambda_k)^2}{nD^{n,\lambda_k,k}(X_{\frac{i_k}{n}}^\lambda)}, \quad N_n(\lambda)_k = \frac{\sqrt{n}(X_{\frac{i_k+1}{n}}^\lambda - X_{\frac{i_k}{n}}^\lambda - c(X_{\frac{i_k}{n}}^\lambda, \lambda_k))}{\sqrt{nD^{n,\lambda_k,k}(X_{\frac{i_k}{n}}^\lambda)}}. \quad (52)$$

Then, we have,

$$(I_n(\lambda), N_n(\lambda)) \xrightarrow[\text{law}]{n \rightarrow \infty} (I(\lambda), N),$$

with  $I(\lambda)$  the diagonal matrix,

$$I(\lambda)_k = \frac{\dot{c}(X_{T_k-}^\lambda, \lambda_k)^2}{a^2(T_k, X_{T_k-}^\lambda)[1 + c'(X_{T_k-}^\lambda, \lambda_k)]^2 U_k + a^2(T_k, X_{T_k-}^\lambda + c(X_{T_k-}^\lambda, \lambda_k))(1 - U_k)},$$

and  $U = (U_1, \dots, U_K)$  is a vector of independent uniform laws on  $[0, 1]$  such that  $U$ ,  $T$  and  $(W_t)_{t \in [0,1]}$  are independent, and conditionally on  $(U, T, (W_t)_{t \in [0,1]})$ ,  $N$  is a standard Gaussian vector in  $\mathbb{R}^K$ .

**Proof** We just have to prove the convergence in law of the couple  $(nD^{n,\lambda_k,k}(X_{\frac{i_k}{n}}^\lambda), \sqrt{n}(X_{\frac{i_k+1}{n}}^\lambda - X_{\frac{i_k}{n}}^\lambda - c(X_{\frac{i_k}{n}}^\lambda, \lambda_k)))$ . We have from (23)

$$D^{n,\lambda_k,k}(X_{\frac{i_k}{n}}^\lambda) = a^2\left(\frac{i_k}{n}, X_{\frac{i_k}{n}}^\lambda\right)(1 + c'(X_{\frac{i_k}{n}}^\lambda, \lambda_k))^2(T_k - \frac{i_k}{n}) + a^2\left(\frac{i_k}{n}, X_{\frac{i_k}{n}}^\lambda + c(X_{\frac{i_k}{n}}^\lambda, \lambda_k)\right)\left(\frac{i_k+1}{n} - T_k\right)$$

and from (26)

$$\begin{aligned} X_{\frac{i_k+1}{n}}^\lambda &= X_{\frac{i_k}{n}}^\lambda + c(X_{\frac{i_k}{n}}^\lambda, \lambda_k) + (1 + c'(X_{\frac{i_k}{n}}^\lambda, \lambda_k))a\left(\frac{i_k}{n}, X_{\frac{i_k}{n}}^\lambda\right)(W_{T_k} - W_{\frac{i_k}{n}}) \\ &\quad + a\left(\frac{i_k}{n}, X_{\frac{i_k}{n}}^\lambda + c(X_{\frac{i_k}{n}}^\lambda, \lambda_k)\right)(W_{\frac{i_k+1}{n}} - W_{T_k}) + R^{n,\lambda,k}, \end{aligned}$$

where  $R^{n,\lambda,k}$  is bounded in  $L^p$  by  $C/n$  (see the proof of Lemma 3). So as a straightforward consequence of Lemma 1, we obtain that  $(nD^{n,\lambda_k,k}(X_{\frac{i_k}{n}}^\lambda), \sqrt{n}(X_{\frac{i_k+1}{n}}^\lambda - X_{\frac{i_k}{n}}^\lambda - c(X_{\frac{i_k}{n}}^\lambda, \lambda_k)))$  converges in law to

$$(D^{\lambda_k,k}(X_{T_k-}^\lambda), (1 + c'(X_{T_k-}^\lambda, \lambda_k))a(T_k, X_{T_k-}^\lambda)\sqrt{U_k}N_k^- + a(T_k, X_{T_k-}^\lambda + c(X_{T_k-}^\lambda, \lambda_k))\sqrt{1 - U_k}N_k^+),$$

with

$$D^{\lambda_k,k}(X_{T_k-}^\lambda) = a^2(T_k, X_{T_k-}^\lambda)[1 + c'(X_{T_k-}^\lambda, \lambda_k)]^2 U_k + a^2(T_k, X_{T_k-}^\lambda + c(X_{T_k-}^\lambda, \lambda_k))(1 - U_k).$$

This gives the result of Lemma 6.  $\diamond$

As noticed earlier, the proof of Theorem 2 follows from the decomposition (15) with  $\mathbb{P}(T \in \mathcal{T}_n) \xrightarrow{n \rightarrow \infty} 1$ , and Lemmas 5 and 6.

## 5.2 Proof of the convolution theorem

In this section, we prove the Theorem 1 and some related results.

We recall the framework described in Section 2.

$(\Omega, \mathcal{F}, \mathbb{P})$  is the canonical product space, on which are defined the independent variables  $(W_t)_{t \in [0,1]}$ ,  $T = (T_1, \dots, T_K)$ ,  $\Lambda = (\Lambda_1, \dots, \Lambda_K)$ . The probability  $\mathbb{P}$  is the simple product of the corresponding probabilities. From this simple disintegration of the measure  $\mathbb{P}$  as a product, we can introduce  $\mathbb{P}^\lambda$  the probability  $\mathbb{P}$  conditional on  $\Lambda = \lambda \in \mathbb{R}^K$ . The process  $X$  is solution of (2), and we may assume that for any  $\lambda \in \mathbb{R}^K$  the law of  $X$  under  $\mathbb{P}^\lambda$  is equal to the law of  $X^\lambda$  solution of (8). We recall that  $\tilde{\Omega}$  is the extension of  $\Omega$  which contains the uniform variables  $U_1, \dots, U_K$ , and the Gaussian variables,  $N_1^-, \dots, N_K^-, N_1^+, \dots, N_K^+$ .

With these notations, the LAMN expansion of Theorem 2 writes,

$$Z_n(\lambda, \lambda + h/\sqrt{n}, T, X_{1/n}, \dots, X_1) = \sum_{k=1}^K h_k I_n(\lambda)_k^{1/2} N_n(\lambda)_k - \frac{1}{2} \sum_{k=1}^K h_k^2 I_n(\lambda)_k + o_{\mathbb{P}^\lambda}(1) \quad (53)$$

with

$$I_n(\lambda)_k = \frac{\dot{c}(X_{\frac{i_k}{n}}, \lambda_k)^2}{n D^{n, \lambda_k, k}(X_{\frac{i_k}{n}})}, \quad N_n(\lambda)_k = \frac{\sqrt{n}(X_{\frac{i_k+1}{n}} - X_{\frac{i_k}{n}} - c(X_{\frac{i_k}{n}}, \lambda_k))}{\sqrt{n D^{n, \lambda_k, k}(X_{\frac{i_k}{n}})}}, \quad (54)$$

$$D^{n, \lambda_k, k}(X_{\frac{i_k}{n}}) = a^2\left(\frac{i_k}{n}, X_{\frac{i_k}{n}}\right) \left(1 + c'(X_{\frac{i_k}{n}}, \lambda_k)\right)^2 \left(T_k - \frac{i_k}{n}\right) + a^2\left(\frac{i_k}{n}, X_{\frac{i_k}{n}} + c(X_{\frac{i_k}{n}}, \lambda_k)\right) \left(\frac{i_k+1}{n} - T_k\right)$$

The Theorem 2 states the convergence in law of  $(I_n(\lambda), N_n(\lambda))$  to  $(I(\lambda), N)$  under  $\mathbb{P}^\lambda$ . Actually, from the proof of Lemma 6, we get the following convergence result under  $\mathbb{P}$ .

**Proposition 3** *Assuming H0–H2, we have the convergence*

$$\begin{aligned} & ((nT_k - i_k)_{k=1, \dots, K}, (\sqrt{n}(W_{T_k} - W_{i_k/n}))_{k=1, \dots, K}, (\sqrt{n}(W_{(i_k+1)/n} - W_{T_k}))_{k=1, \dots, K}, I_n(\Lambda), N_n(\Lambda)) \\ & \xrightarrow[\text{law}]{n \rightarrow \infty} ((U_k)_{k=1, \dots, K}, (\sqrt{U_k} N_k^-)_{k=1, \dots, K}, (\sqrt{1 - U_k} N_k^+)_{k=1, \dots, K}, I(\Lambda), N(\Lambda)) \end{aligned} \quad (55)$$

where  $N(\Lambda)$  is distributed as a standard Gaussian variable in  $\mathbb{R}^K$ . Moreover this convergence is stable with respect to  $\mathcal{F}$ , and the last two limit variables can be represented on the extended space  $\tilde{\Omega}$  as,

$$I(\Lambda)_k = \frac{\dot{c}(X_{T_k-}, \Lambda_k)^2}{a^2(T_k, X_{T_k-})(1 + c'(X_{T_k-}, \Lambda_k))^2 U_k + a^2(T_k, X_{T_k})(1 - U_k)}, \quad (56)$$

$$N(\Lambda)_k = \frac{a(T_k, X_{T_k-})(1 + c'(X_{T_k-}, \Lambda_k))\sqrt{U_k} N_k^- + a(T_k, X_{T_k})\sqrt{1 - U_k} N_k^+}{[a^2(T_k, X_{T_k-})(1 + c'(X_{T_k-}, \Lambda_k))^2 U_k + a^2(T_k, X_{T_k})(1 - U_k)]^{1/2}}. \quad (57)$$



Remark that the matrix  $I(\Lambda)$  is not equal to the matrix  $I^{\text{opt}}$  appearing in the statement of the convolution Theorem 1. Comparing the expression (5) of  $I^{\text{opt}}$  with the expression (11) of  $I(\lambda)$ , we see that in the parametric case, the information is proportional to  $(\dot{c}(X_{T_k-}, \lambda_k))^2$ . This is quite natural. If instead of estimating the "mark"  $\lambda_k$  we estimate the jump, equal to  $c(X_{T_k-}, \lambda_k)$  in the parametric model, we can expect that the effect of  $(\dot{c}(X_{T_k-}, \lambda_k))^2$  vanishes (by a simple first order expansion of the error of estimation). This gives some insight on why  $\dot{c}(X_{T_k-}, \Lambda_k)^2$  disappears in the expression of  $I^{\text{opt}}$ .

On the other hand, it is not immediate why the expression of the parametric information involves the quantity  $c'(X_{T_k-}, \lambda_k)$ , which is not present in the expression of  $I^{\text{opt}}$ . We will see that it is due to the fact that the value of the jump  $c(X_{T_k-}, \lambda_k)$  depends on the unobserved quantity  $X_{T_k-}$  and thus is not a simple functional of the parameter  $\lambda_k$ .

If  $c$  does not depend on  $X$ , the situation is simpler and the proof of Theorem 1 is much easier. For this reason, in the next Section we prove the convolution theorem in this easier setting. The general proof is given in Section 5.2.3 and some intermediate results are stated in Section 5.2.2.

### 5.2.1 Proof of Theorem 1 when $c(x, \theta) = c(\theta)$

We start with a simple lemma.

**Lemma 7** *Assume H0–H2 then for all  $\lambda, h \in \mathbb{R}^K$ ,*

$$\begin{aligned} I_n(\lambda + \frac{h}{\sqrt{n}}) - I_n(\lambda) &\xrightarrow{n \rightarrow \infty} 0, \text{ in } \mathbb{P}^\lambda \text{ probability,} \\ N_n(\lambda + \frac{h}{\sqrt{n}}) - N_n(\lambda) + I_n(\lambda)^{1/2}h &\xrightarrow{n \rightarrow \infty} 0, \text{ in } \mathbb{P}^\lambda \text{ probability.} \end{aligned}$$

**Proof** This follows easily from the expressions (54). ◇

Assume that  $\tilde{J}^n$  is a sequence of estimators (based on  $(X_{i/n})_{i=0, \dots, n}$ ) such that  $\sqrt{n}(\tilde{J}^n - J) \xrightarrow{n \rightarrow \infty} \tilde{Z}$  in law under  $\mathbb{P}$ .

Then, the Theorem 1 is an immediate consequence of the following result.

**Theorem 4** *Assume H0–H3 and that  $c(x, \theta) = c(\theta)$ . Denote  $\dot{C}(\Lambda)$  the diagonal matrix of size  $K \times K$  such that  $\dot{C}(\Lambda)_k = \dot{c}(\Lambda_k)$ .*

*Then, we have the decomposition for all  $n$ ,*

$$\sqrt{n}(\tilde{J}^n - J) = \dot{C}(\Lambda)I_n(\Lambda)^{-1/2}N_n(\Lambda) + R_n, \tag{58}$$

for  $(R_n)_n$  a sequence of random variables with values in  $\mathbb{R}^K$ .

Along a subsequence  $(n)$  we have the convergence in law,

$$(\dot{C}(\Lambda)I_n(\Lambda)^{-1/2}N_n(\Lambda), R_n) \xrightarrow{(n) \rightarrow \infty} (\dot{C}(\Lambda)I(\Lambda)^{-1/2}N(\Lambda), R) = ((I^{\text{opt}})^{-1/2}N, R) \quad (59)$$

where  $N = N(\Lambda)$  is Gaussian, and  $R$  is independent of  $N$  conditionally on  $I^{\text{opt}}$ .

In particular, we have  $\tilde{Z} = \lim_{(n)} \sqrt{n}(\tilde{J}^n - J) = (I^{\text{opt}})^{-1/2}N + R$ .

**Proof** We set  $R_n = \sqrt{n}(\tilde{J}^n - J) - \dot{C}(\Lambda)I_n(\Lambda)^{-1/2}N_n(\Lambda)$  and define,

$$R_n(\lambda) = \sqrt{n}(\tilde{J}^n - c(\lambda_k)_k) - \dot{C}(\lambda)I_n(\lambda)^{-1/2}N_n(\lambda), \quad (60)$$

so that  $R_n = R_n(\Lambda)$ . Since  $\tilde{J}^n$  is a measurable function of the  $(X_{i/n})_i$ ,  $J = (c(\Lambda_k))_k$  and  $\dot{C}(\Lambda)$  are measurable functions of the marks, and from the expression (54), we deduce that  $R_n = f_n((X_{i/n})_i, T, \Lambda)$  for some borelian function  $f_n$ .

Using Lemma 7 and the expression (60), we easily get:

$$R_n(\lambda + \frac{h}{\sqrt{n}}) - R_n(\lambda) \xrightarrow{n \rightarrow \infty} 0, \text{ in } \mathbb{P}^\lambda \text{ probability for any } \lambda, h \in \mathbb{R}^K.$$

Remark now that by Proposition 3 and the convergence of  $\sqrt{n}(\tilde{J}^n - J)$ , we get that  $(R_n)_n$  is a tight sequence of variables.

Hence, we can apply Proposition 4 below. We deduce that,

$$(I_n(\Lambda), N_n(\Lambda), R_n) \xrightarrow[\text{law}]{n \rightarrow \infty} (I(\Lambda), N(\Lambda), R),$$

where the limit can be represented on an extension  $\tilde{\Omega} \times \mathbb{R}^K$  of the space  $\tilde{\Omega}$ , and the convergence is stable with respect to  $(T, \Lambda, (W_t)_{t \in [0,1]})$ . On this extension, the variable  $R$  is independent of  $N(\Lambda)$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$ . This implies (59), and thus the theorem.  $\diamond$

**Proposition 4** *Assume H0–H3. Let  $R_n = f_n((X_{i/n})_i, T, \Lambda) \in \mathbb{R}^K$  where  $(f_n)_n$  is a sequence of borelian functions. Set  $R_n(\lambda) = f_n((X_{i/n})_i, T, \lambda)$ , and assume:*

- $R_n(\lambda + \frac{h}{\sqrt{n}}) - R_n(\lambda) \xrightarrow{n \rightarrow \infty} 0$ , in  $\mathbb{P}^\lambda$  probability for any  $\lambda, h \in \mathbb{R}^K$ ,
- the sequence  $(R_n)_n$  is tight.

Then, one has the convergence in law, along a subsequence,

$$\begin{aligned} & ((nT_k - i_k)_{k=1, \dots, K}, (\sqrt{n}(W_{T_k} - W_{i_k/n}))_{k=1, \dots, K}, (\sqrt{n}(W_{(i_k+1)/n} - W_{T_k}))_{k=1, \dots, K}, I_n(\Lambda), N_n(\Lambda), R_n) \\ & \xrightarrow[\text{law}]{(n) \rightarrow \infty} ((U_k)_{k=1, \dots, K}, (\sqrt{U_k} N_k^-)_{k=1, \dots, K}, (\sqrt{1 - U_k} N_k^+)_{k=1, \dots, K}, I(\Lambda), N(\Lambda), R). \end{aligned} \quad (61)$$

The limit can be represented on an extension  $\tilde{\Omega} \times \mathbb{R}^K$  of the space  $\tilde{\Omega}$ . On this space, the variable  $R$  is independent of  $N(\Lambda)$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$ . Moreover the convergence (61) is stable with respect to  $(T, \Lambda, (W_t)_{t \in [0,1]})$ .

**Proof** Consider the joint law of the random variables,

$$\begin{aligned} & (T, \Lambda, (W_t)_{t \in [0,1]}, (nT_k - i_k)_{k=1, \dots, K}, (\frac{W_{T_k} - W_{i_k/n}}{\sqrt{T_k - i_k/n}})_{k=1, \dots, K}, \\ & (\frac{W_{(i_k+1)/n} - W_{T_k}}{\sqrt{(i_k+1)/n - T_k}})_{k=1, \dots, K}, I_n(\Lambda), N_n(\Lambda), R_n) \end{aligned} \quad (62)$$

defined on the corresponding canonical product space, endowed with the usual product topology. From the assumption, all the components of this vector are tight, and thus the joint law is tight. Along some subsequence, it converges in law to some limit, and thus (61) holds true. The stability of the convergence with respect to  $T, \Lambda, (W_t)_{t \in [0,1]}$  is immediate. Remark that from Proposition 3, the law of the limit  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_{k=1, \dots, K}, (N_k^-)_{k=1, \dots, K}, (N_k^+)_{k=1, \dots, K}, I(\Lambda), N(\Lambda), R)$  is known, apart for the last component  $R$ . It can be clearly represented on an extension  $\tilde{\Omega} \times \mathbb{R}^K$  of  $\tilde{\Omega}$

To determine some information on the law of  $R$ , we use techniques inspired from the proof of convolution theorems in [17].

Consider the following set of random variables defined on the space  $\Omega$ ,

$$\left. \begin{aligned} & G = g(X_{s_1}, \dots, X_{s_r}), \text{ with } r \geq 1 \text{ and } (s_1, \dots, s_r) \in [0, 1]^r, \\ & G_n = g(X_{\frac{[ns_1]}{n}}, \dots, X_{\frac{[ns_r]}{n}}), \\ & \kappa = k(T_1, \dots, T_K), \\ & L_n = l(nT_1 - i_1, \dots, nT_K - i_K), \\ & M = m(\Lambda_1, \dots, \Lambda_K), \end{aligned} \right\} \quad (63)$$

where  $g, k, l, m$  are bounded continuous functions.

For  $(\mu_1, \mu_2) \in \mathbb{R}^{2K}$  we set

$$\varphi_n(\mu_1, \mu_2) = \mathbb{E} \left[ e^{i\mu_1 \cdot R_n} e^{i\mu_2 \cdot N_n(\Lambda)} G_n \kappa L_n M \right].$$

Clearly  $G_n \rightarrow G$  in probability, and from the convergence, along a subsequence, of (62), it is simple to show

$$\varphi_n(\mu_1, \mu_2) \xrightarrow{(n) \rightarrow \infty} \mathbb{E} \left[ e^{i\mu_1 \cdot R} e^{i\mu_2 \cdot N(\Lambda)} G \kappa l(U_1, \dots, U_K) M \right]. \quad (64)$$

By conditioning on the variable  $\Lambda$ , whose law admits a density, we have

$$\varphi_n(\mu_1, \mu_2) = \int_{\mathbb{R}^K} \mathbb{E}^\lambda \left[ e^{i\mu_1 \cdot R_n(\lambda)} e^{i\mu_2 \cdot N_n(\lambda)} G_n \kappa L_n m(\lambda) \right] f_\Lambda(\lambda) d\lambda,$$

For  $h \in \mathbb{R}^K$ , we make a simple change of variable in the integral,

$$\varphi_n(\mu_1, \mu_2) = \int_{\mathbb{R}^K} \mathbb{E}^{\lambda+h/\sqrt{n}} \left[ e^{i\mu_1 \cdot R_n(\lambda+h/\sqrt{n})} e^{i\mu_2 \cdot N_n(\lambda+h/\sqrt{n})} G_n \kappa L_n \right] m(\lambda+h/\sqrt{n}) f_\Lambda(\lambda+h/\sqrt{n}) d\lambda.$$

Now the translation is a continuous operator in  $\mathbf{L}^1(\mathbb{R})$  and by assumption  $\lambda \mapsto m(\lambda) f_\Lambda(\lambda)$  is integrable. Thus, we easily deduce,

$$\varphi_n(\mu_1, \mu_2) = \int_{\mathbb{R}^K} \mathbb{E}^{\lambda+h/\sqrt{n}} \left[ e^{i\mu_1 \cdot R_n(\lambda+h/\sqrt{n})} e^{i\mu_2 \cdot N_n(\lambda+h/\sqrt{n})} G_n \kappa L_n \right] m(\lambda) f_\Lambda(\lambda) d\lambda + o(1)$$

From the assumptions, we know the expansion  $R_n(\lambda+h/\sqrt{n}) = R_n(\lambda) + o_{\mathbb{P}^\lambda}(1)$ , and from Lemma 7, we have the expansion  $N_n(\lambda + \frac{h}{\sqrt{n}}) = N_n(\lambda) - I_n(\lambda)^{1/2} h + o_{\mathbb{P}^\lambda}(1)$ . In these expansions, all the random variables are only depending on  $((X_{i/n})_i, T)$ . But, from the LAMN property, we know that the measures  $\mathbb{P}^\lambda$  and  $\mathbb{P}^{\lambda+h/\sqrt{n}}$ , restricted to  $((X_{i/n})_i, T)$ , are contiguous. Hence, in these expansions, one can replace  $o_{\mathbb{P}^\lambda}(1)$  with  $o_{\mathbb{P}^{\lambda+h/\sqrt{n}}}(1)$ . Then, using dominated convergence Theorem, one can get,

$$\varphi_n(\mu_1, \mu_2) = \int_{\mathbb{R}^K} \mathbb{E}^{\lambda+h/\sqrt{n}} \left[ e^{i\mu_1 \cdot R_n(\lambda)} e^{i\mu_2 \cdot (N_n(\lambda) - I_n(\lambda)^{1/2} h)} G_n \kappa L_n \right] m(\lambda) f_\Lambda(\lambda) d\lambda + o(1).$$

Remark that the random variables appearing in the inner expectation only depend on the observations  $((X_{i/n})_i, T)$ , and thus the likelihood ratio  $\frac{\mathbf{P}^{n, \lambda+h/\sqrt{n}}}{\mathbf{P}^{n, \lambda}}(T, (X_{i/n})_i) = \exp(Z_n(\lambda, \lambda + h/\sqrt{n}, T, (X_{i/n})_i))$  might be used to change the measure,

$$\begin{aligned} \varphi_n(\mu_1, \mu_2) &= \int_{\mathbb{R}^K} \mathbb{E}^\lambda \left[ e^{i\mu_1 \cdot R_n(\lambda)} e^{i\mu_2 \cdot (N_n(\lambda) - I_n(\lambda)^{1/2} h)} e^{Z_n(\lambda, \lambda+h/\sqrt{n}, T, (X_{i/n})_i)} G_n \kappa L_n \right] \\ &\quad m(\lambda) f_\Lambda(\lambda) d\lambda + o(1). \end{aligned} \quad (65)$$

We deduce,

$$\varphi_n(\mu_1, \mu_2) = \mathbb{E} \left[ e^{i\mu_1 \cdot R_n} e^{i\mu_2 \cdot (N_n(\Lambda) - I_n(\Lambda)^{1/2}h)} e^{Z_n(\Lambda, \Lambda + h/\sqrt{n}, T, (X_{i/n})_i)} G_n \kappa L_n M \right] + o(1).$$

But from the LAMN expansion (53), one can easily get

$$Z_n(\Lambda, \Lambda + h/\sqrt{n}, T, (X_{i/n})_i) = h^* I_n(\Lambda)^{1/2} N_n(\Lambda) - \frac{1}{2} h^* I_n(\Lambda) h + o_{\mathbb{P}}(1),$$

where  $h^*$  is the transpose of the vector  $h$ . Hence, using the convergence in law of (62), and uniform integrability of the sequence  $Z_n(\Lambda, \Lambda + h/\sqrt{n}, T, (X_{i/n})_i)$ , it can be seen that

$$\varphi_n(\mu_1, \mu_2) \xrightarrow{(n) \rightarrow \infty} E \left[ e^{i\mu_1 \cdot R} e^{i\mu_2 \cdot (N(\Lambda) - I(\Lambda)^{1/2}h)} e^{h^* I(\Lambda)^{1/2} N(\Lambda) - \frac{1}{2} h^* I(\Lambda) h} G \kappa l(U_1, \dots, U_K) M \right]. \quad (66)$$

Comparing the expressions (64) and (66), it comes  $\forall \mu_1, \mu_2, h$ ,

$$E \left[ e^{i\mu_1 \cdot R} e^{i\mu_2 \cdot N(\Lambda)} G \kappa l(U_1, \dots, U_K) M \right] = E \left[ e^{i\mu_1 \cdot R} e^{i\mu_2 \cdot (N(\Lambda) - I(\Lambda)^{1/2}h)} e^{h^* I(\Lambda)^{1/2} N(\Lambda) - \frac{1}{2} h^* I(\Lambda) h} G \kappa l(U_1, \dots, U_K) M \right].$$

We deduce that  $\forall \mu_1, \mu_2, h$ , the two following conditional expectations are almost surely equal,

$$E \left[ e^{i\mu_1 \cdot R} e^{i\mu_2 \cdot N(\Lambda)} \mid X, T, (U_k)_k, \Lambda \right] = E \left[ e^{i\mu_1 \cdot R} e^{i\mu_2 \cdot (N(\Lambda) - I(\Lambda)^{1/2}h)} e^{h^* I(\Lambda)^{1/2} N(\Lambda) - \frac{1}{2} h^* I(\Lambda) h} \mid X, T, (U_k)_k, \Lambda \right].$$

But from continuity and analyticity arguments, it can be seen that this equality holds, almost surely, for any  $\mu_1 \in \mathbb{R}^K, \mu_2 \in \mathbb{R}^K, h \in \mathbb{C}^K$ .

Hence, we can set  $h = -iI(\Lambda)^{-1/2} \mu_2$  in the above relation, and find

$$E \left[ e^{i\mu_1 \cdot R} e^{i\mu_2 \cdot N(\Lambda)} \mid X, T, (U_k)_k, \Lambda \right] = E \left[ e^{i\mu_1 \cdot R} \mid X, T, (U_k)_k, \Lambda \right] e^{-\mu_2^* \mu_2 / 2}.$$

This precisely states that, conditionally on  $(X, T, (U_k)_k, \Lambda)$ , the random variables  $R$  and  $N(\Lambda)$  are independent. The Proposition is proved after remarking that the Brownian motion  $(W_t)_t$  can be recovered as a measurable functional of  $X, T, \Lambda$ .  $\diamond$

### 5.2.2 Intermediate results

The assumption  $c(x, \theta) = c(\theta)$  is crucial for the proof of Theorem 4. Indeed if  $c$  depends on the jump-diffusion, then  $J_k = c(X_{T_k^-}, \lambda_k)$ , and instead of (60), we have  $R_n(\lambda) = \sqrt{n}(\tilde{J}^n -$

$c(X_{T_k-}, \lambda_k)_k - \dot{C}(X, \lambda)I_n(\lambda)^{-1/2}N_n(\lambda)$ , where  $\dot{C}(X, \lambda) = \text{diag}(\dot{c}(X_{T_k-}, \lambda_k)_k)$ . This quantity depends on  $X_{T_k-}$  which is unobserved. However, the assumption that  $R_n(\lambda)$  is only function of  $((X_{i/n})_i, T)$  is essential in the Proposition 4 (at the step just before equation (65)).

But if instead of  $R_n(\lambda)$  we consider  $R_n^{\text{obs}}(\lambda) = \sqrt{n}(\tilde{J}^n - c(X_{i_k/n}, \lambda_k)_k) - \dot{C}_n^{\text{obs}}(\lambda)I_n(\lambda)^{-1/2}N_n(\lambda)$ , where  $\dot{C}_n^{\text{obs}}(\lambda) = \text{diag}(\dot{c}(X_{i_k/n}, \lambda_k)_k)$ , then, the Proposition 4 can be applied, and we can prove the following modification of Theorem 4.

**Theorem 5** *Let  $\tilde{J}^n$  be any sequence of estimators such that  $\sqrt{n}(\tilde{J}^n - (c(X_{i_k/n}, \Lambda_k))_k) \xrightarrow[\text{law}]{(n) \rightarrow \infty} \bar{Z}$  for some variable  $\bar{Z}$ . Then, the law of  $\bar{Z}$  is necessarily a convolution,  $\bar{Z} \stackrel{\text{law}}{=} \dot{C}(X, \Lambda)I(\Lambda)^{-1/2}N(\Lambda) + R$ , where  $N(\Lambda)$  is a standard Gaussian vector independent of  $\dot{C}(X, \Lambda)^{-2}I(\Lambda)$ , and  $R$  is some random variable independent of  $N(\Lambda)$  conditionally on  $\dot{C}(X, \Lambda)^{-2}I(\Lambda)$ . A simple expression for the entries of the diagonal matrix  $\dot{C}(X, \Lambda)^{-2}I(\Lambda)$  is*

$$I_k = [U_k a(T_k, X_{T_k-})^2 (1 + c'(T_k, X_{T_k-}))^2 + (1 - U_k) a(T_k, X_{T_k})^2]^{-1}, \quad \text{for } k = 1, \dots, K. \quad (67)$$

Actually, to prove the convolution theorem when the coefficient  $c(x, \theta)$  depends on  $x$ , we need a strengthened version of the Proposition 4. Indeed, we will show that the variable  $R$ , in the statement of Proposition 4, is independent of  $N$  conditionally on any variable that can be obtained as a limit of the observations. This yields some additional knowledge on the dependence between the variable  $R$  and the other variables.

**Proposition 5** *Let us make the same assumptions as in Proposition 4. Assume furthermore that there exist a continuous function  $\Psi$  with values in  $\mathbb{R}^K$  and  $(A_n)_n$  a sequence of random variables depending on the observations  $(T, (X_{i/n})_i)$ , such that*

$$A_n - \Psi((nT_k - i_k)_{k=1, \dots, K}, (\sqrt{n}(W_{T_k} - W_{i_k/n}))_{k=1, \dots, K}, (\sqrt{n}(W_{(i_k+1)/n} - W_{T_k}))_{k=1, \dots, K}) \xrightarrow{n \rightarrow \infty} 0, \text{ in } \mathbb{P} \text{ probability.}$$

*Then, in the description of the limit (61), the variable  $R$  is independent of  $N(\Lambda)$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$  and  $\Psi((U_k)_k, (\sqrt{U_k}N_k^-)_k, ((\sqrt{1 - U_k}N_k^+)_k)$ .*

**Proof** The proof is a slight modification of the proof of Proposition 4. We simply add to the list of random variables (63), the new one  $S_n = s(A_n)$ , with  $s$  being any continuous bounded function. Accordingly, we set  $\varphi_n(\mu_1, \mu_2) = \mathbb{E} [e^{i\mu_1 \cdot R_n} e^{i\mu_2 \cdot N_n(\Lambda)} S_n G_n K L_n M]$ . Then, the proof follows the same lines as the proof of Proposition 4.  $\diamond$

### 5.2.3 Proof of Theorem 1. The general case

We prove the Theorem 1 in the general situation where  $c(x, \theta)$  depends on  $x$ .

As seen in the previous section, a difficulty comes from the fact that the target of the estimator  $J = (\Delta X_{T_k})_k = (c(X_{T_k-}, \Lambda_k))_k$  depends on the unobserved value  $X_{T_k-}$ . We introduce  $\bar{J}^n = (c(X_{\frac{i_k}{n}}, \Lambda_k))_k$ , and with simple computations, one can write the following expansion, for any sequence of estimators  $\tilde{J}^n$ ,

$$\sqrt{n}(\tilde{J}_k^n - J_k) = \sqrt{n}(\tilde{J}_k^n - \bar{J}_k^n) - c'(X_{\frac{i_k}{n}}, \Lambda_k)\sqrt{n}(X_{T_k-} - X_{\frac{i_k}{n}}) + o_{\mathbb{P}}(1).$$

If  $\sqrt{n}(\tilde{J}_k^n - \bar{J}_k^n)$  is tight we can use Theorem 5 and deduce,  $\lim_{(n)} \sqrt{n}(\tilde{J}_k^n - J_k) = \tilde{Z}_k = \dot{C}(X, \Lambda)I(\Lambda)_k^{-1/2}N(\Lambda)_k - c'(X_{T_k-}, \Lambda_k)a(T_k, X_{T_k-})\sqrt{U_k}N_k^- + R_k$ . After a few algebra, involving the expressions (56)–(57), it could be seen that this reduces to the algebraic relation (4), with  $N$  being some standard normal variable. However by this method, we can not deduce the conditional independence of  $R$  with  $N$ . Indeed, only the conditional independence of  $R$  with  $N(\Lambda)$  is known, and we have no information about the joint law of  $R$  and  $N^-$ .

To solve this problem, we consider two new statistical experiments where we add the observation of the jump-diffusion just before (or just after) the jump. We first state the LAMN properties for these new experiments. We omit the proof, which is similar to the proof of Theorem 2.

**Proposition 6 (LAMN property adding the observations before the jumps.)** *Assume*

*H0, H1 and H2. Denote  $(\mathbf{p}^{n,\lambda,\text{aug}^-})$  the density on  $\mathbb{R}^{n+2K}$  of the augmented vector of observations  $\mathcal{O}^{\text{aug}^-} = ((X_{i/n})_i, (T_k)_k, (X_{T_k-})_k)$  under  $\mathbb{P}^\lambda$ . For  $\lambda \in \mathbb{R}^K$ ,  $h \in \mathbb{R}^K$ , define the log-likelihood ratio  $Z_n^{\text{aug}^-}(\lambda, \lambda + h/\sqrt{n}, \mathcal{O}^{\text{aug}^-}) = \ln \frac{\mathbf{p}^{n,\lambda+h/\sqrt{n},\text{aug}^-}(\mathcal{O}^{\text{aug}^-})}{\mathbf{p}^{n,\lambda,\text{aug}^-}(\mathcal{O}^{\text{aug}^-})}$ .*

*We have the expansion:*

$$Z_n(\lambda, \lambda + h/\sqrt{n}, \mathcal{O}^{\text{aug}^-}) = \sum_{k=1}^K h_k I_n^{\text{aug}^-}(\lambda)_k^{1/2} N_n^{\text{aug}^-}(\lambda)_k - \frac{1}{2} \sum_{k=1}^K h_k^2 I_n^{\text{aug}^-}(\lambda)_k + o_{\mathbb{P}^\lambda}(1) \quad (68)$$

where

$$I_n^{\text{aug}^-}(\lambda)_k = \frac{\dot{c}(X_{T_k-}, \lambda_k)^2}{nD^{n,\lambda_k,k,\text{aug}^-}(X_{T_k-})}, \quad N_n^{\text{aug}^-}(\lambda)_k = \frac{\sqrt{n}(X_{\frac{i_k+1}{n}} - X_{T_k-} - c(X_{T_k-}, \lambda_k))}{\sqrt{nD^{n,\lambda_k,k,\text{aug}^-}(X_{T_k-})}}, \quad (69)$$

$$D^{n,\lambda_k,k,\text{aug}^-}(X_{T_k-}) = a^2(T_k, X_{T_k-} + c(X_{T_k-}, \lambda_k))\left(\frac{i_k+1}{n} - T_k\right).$$

Moreover

$$(I_n^{\text{aug}^-}(\lambda), N_n^{\text{aug}^-}(\lambda)) \xrightarrow[\text{law}]{n \rightarrow \infty} (I^{\text{aug}^-}(\lambda), N^{\text{aug}^-})$$

where  $I^{\text{aug}^-}(\lambda)$  is the diagonal information matrix whose entries are

$$I^{\text{aug}^-}(\lambda)_k = \frac{\dot{c}(X_{T_k-}, \lambda_k)^2}{a^2(T_k, X_{T_k})(1 - U_k)}.$$

and  $N^{\text{aug}^-}$  is a standard Gaussian vector in  $\mathbb{R}^K$ .

**Proposition 7 (LAMN property adding the observations after the jumps.)** Assume H0,

H1 and H2. Denote  $(\mathbf{p}^{n,\lambda,\text{aug}^+})$  the density on  $\mathbb{R}^{n+2K}$  of the augmented vector of observations

$\mathcal{O}^{\text{aug}^+} = ((X_{i/n})_i, (T_k)_k, (X_{T_k})_k)$  under  $\mathbb{P}^\lambda$ . For  $\lambda \in \mathbb{R}^K$ ,  $h \in \mathbb{R}^K$ , define the log-likelihood ratio

$$Z_n^{\text{aug}^+}(\lambda, \lambda + h/\sqrt{n}, \mathcal{O}^{\text{aug}^+}) = \ln \frac{\mathbf{p}^{n,\lambda+h/\sqrt{n},\text{aug}^+}(\mathcal{O}^{\text{aug}^+})}{\mathbf{p}^{n,\lambda,\text{aug}^+}(\mathcal{O}^{\text{aug}^+})}.$$

We have the expansion:

$$Z_n(\lambda, \lambda + h/\sqrt{n}, \mathcal{O}^{\text{aug}^+}) = \sum_{k=1}^K h_k I_n^{\text{aug}^+}(\lambda)_k^{1/2} N_n^{\text{aug}^+}(\lambda)_k - \frac{1}{2} \sum_{k=1}^K h_k^2 I_n^{\text{aug}^+}(\lambda)_k + o_{\mathbb{P}^\lambda}(1) \quad (70)$$

where

$$I_n^{\text{aug}^+}(\lambda)_k = \frac{\dot{c}(X_{\frac{i_k}{n}}, \lambda_k)^2}{n D^{n,\lambda_k,k,\text{aug}^+}(X_{\frac{i_k}{n}})}, \quad N_n^{\text{aug}^+}(\lambda)_k = \frac{\sqrt{n}(X_{T_k} - X_{\frac{i_k}{n}} - c(X_{\frac{i_k}{n}}, \lambda_k))}{\sqrt{n D^{n,\lambda_k,k,\text{aug}^+}(X_{\frac{i_k}{n}})}}, \quad (71)$$

$$D^{n,\lambda_k,k,\text{aug}^+}(X_{\frac{i_k}{n}}) = a^2\left(\frac{i_k}{n}, X_{\frac{i_k}{n}}\right)(1 + c'(X_{\frac{i_k}{n}}, \lambda_k))^2(T_k - \frac{i_k}{n}).$$

Moreover

$$(I_n^{\text{aug}^+}(\lambda), N_n^{\text{aug}^+}(\lambda)) \xrightarrow[\text{law}]{n \rightarrow \infty} (I^{\text{aug}^+}(\lambda), N^{\text{aug}^+}),$$

where  $I^{\text{aug}^+}(\lambda)$  is the diagonal information matrix whose entries are

$$I^{\text{aug}^+}(\lambda)_k = \frac{\dot{c}(X_{T_k-}, \lambda_k)^2}{a^2(T_k, X_{T_k-})(1 + c'(X_{T_k-}, \lambda_k))^2 U_k}$$

and  $N^{\text{aug}^+}$  is a standard Gaussian vector in  $\mathbb{R}^K$ .

We now deduce convolution results from these LAMN properties.

**Proposition 8** Let  $\tilde{J}^n$  be a sequence of estimator based on the observations of  $(X_{i/n})_i$  and denote  $\bar{J}^n = (c(X_{\frac{i_k}{n}}, \Lambda))_k$ . Suppose that the sequence  $\sqrt{n}(\tilde{J}^n - \bar{J}^n)$  is tight and define  $R_n^{\text{aug}^-}$  and  $R_n^{\text{aug}^+}$  by the following expansions

$$\sqrt{n}(\tilde{J}^n - \bar{J}^n) = \dot{C}_n^{\text{obs}}(\Lambda) I_n^{\text{aug}^-}(\Lambda)^{-1/2} N_n^{\text{aug}^-}(\Lambda) + R_n^{\text{aug}^-}, \quad (72)$$



$$\sqrt{n}(\tilde{J}^n - \bar{J}^n) = \dot{C}_n^{\text{obs}}(\Lambda) I_n^{\text{aug}^+}(\Lambda)^{-1/2} N_n^{\text{aug}^+}(\Lambda) + R_n^{\text{aug}^+}, \quad (73)$$

where  $I_n^{\text{aug}^-}(\Lambda)$  (resp.  $I_n^{\text{aug}^+}(\Lambda)$ ) is the diagonal matrix with entries  $(I_n^{\text{aug}^-}(\Lambda))_k$  (resp.  $(I_n^{\text{aug}^+}(\Lambda))_k$ ) and  $\dot{C}_n^{\text{obs}}(\Lambda)$  is diagonal with entries  $\dot{c}(X_{\frac{i_k}{n}}, \Lambda_k)$ .

Then, we have the convergence in law

$$\begin{aligned} & [\sqrt{n}(X_{\frac{i_k+1}{n}} - X_{T_k-} - c(X_{T_k-}, \Lambda_k))_k, \sqrt{n}(X_{T_k} - X_{\frac{i_k}{n}} - c(X_{\frac{i_k}{n}}, \Lambda_k))_k, R_n^{\text{aug}^-}, R_n^{\text{aug}^+}] \xrightarrow{(n) \rightarrow \infty} \\ & [(a(T_k, X_{T_k})\sqrt{1 - U_k}N_k^+)_k, (a(T_k, X_{T_k-})(1 + c'(X_{T_k-}, \Lambda_k))\sqrt{U_k}N_k^-)_k, R^{\text{aug}^-}, R^{\text{aug}^+}]. \end{aligned} \quad (74)$$

This convergence holds jointly with (55) and the limit variables can be represented on an extension of  $\tilde{\Omega}$ . On this space, one has

$$R_k^{\text{aug}^+} = R_k^{\text{aug}^-} - a(T_k, X_{T_k-})(1 + c'(X_{T_k-}, \Lambda_k))\sqrt{U_k}N_k^- + a(T_k, X_{T_k})\sqrt{1 - U_k}N_k^+, \quad \forall k \in \{1, \dots, K\}. \quad (75)$$

Moreover, conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k, (N_k^-)_k)$ , the variable  $R^{\text{aug}^-}$  is independent of  $(N_k^+)_k$ . In a symmetric way, conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k, (N_k^+)_k)$ , the variable  $R^{\text{aug}^+}$  is independent of  $(N_k^-)_k$ .

**Proof** From the definition of the variables  $R_n^{\text{aug}^-}$  and  $R_n^{\text{aug}^+}$  given by equations (72) and (73), we deduce immediately the relations

$$\sqrt{n}(\tilde{J}^n - \bar{J}^n) = [\sqrt{n}(X_{\frac{i_k+1}{n}} - X_{T_k-} - c(X_{T_k-}, \Lambda_k))_k + R_n^{\text{aug}^-} + o_{\mathbb{P}}(1), \quad (76)$$

$$\sqrt{n}(\tilde{J}^n - \bar{J}^n) = [\sqrt{n}(X_{T_k} - X_{\frac{i_k}{n}} - c(X_{\frac{i_k}{n}}, \Lambda_k))_k + R_n^{\text{aug}^+}. \quad (77)$$

By a tightness argument the joint convergence, along a subsequence, of (55) and (74) is clear. The relation (75) is a consequence of the equality between the quantities (76) and (77).

Now, we can deduce, from the LAMN property (Proposition 6), a result analogous to Proposition 4. Hence  $R^{\text{aug}^-}$  is independent of the limit of  $N_n^{\text{aug}^-}(\Lambda)$ , conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$ . Moreover, remark that in the experiment  $\mathcal{O}^{\text{aug}^-}$ , the sequence of variables

$$A_n = \frac{\sqrt{n}(X_{T_k-} - X_{\frac{i_k}{n}})}{a(T_k, X_{\frac{i_k}{n}})}$$

is observed. But  $A_n - \sqrt{n}(W_{T_k-} - W_{\frac{i_k}{n}})$  converges to zero in  $\mathbb{P}$ -probability. Showing a result analogous to Proposition 5, we deduce that  $R^{\text{aug}^-}$  is independent of the limit of  $N_n^{\text{aug}^-}(\Lambda)$ , conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k, (\sqrt{U_k}N_k^-)_k)$ . This shows immediately that  $R^{\text{aug}^-}$  is independent of  $(N_k^+)$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k, (N_k^-)_k)$ , since the sigma-fields generated by the two vectors are the same.

The conditional independence between  $R^{\text{aug}^+}$  and  $(N_k^-)_k$  is obtained in a symmetric way: one uses the LAMN property of Proposition 7, and the fact that the sequence

$$A'_n = \frac{\sqrt{n}(X_{\frac{i_k+1}{n}} - X_{T_k})}{a(T_k, X_{T_k})}$$

is observed in the experiment based on  $\mathcal{O}^{\text{aug}^+}$ .  $\diamond$

Finally, we are able to prove Theorem 1.

### Proof of Theorem 1.

First, we write

$$\begin{aligned} \sqrt{n}(\tilde{J}_k^n - J_k) &= \sqrt{n}(\tilde{J}_k^n - \bar{J}_k^n) - \sqrt{n}(J_k - \bar{J}_k^n) \\ &= \sqrt{n}(\tilde{J}_k^n - \bar{J}_k^n) - c'(X_{\frac{i_k}{n}}, \Lambda_k) \sqrt{n}(X_{T_k-} - X_{\frac{i_k}{n}}) + o_{\mathbb{P}}(1). \end{aligned} \quad (78)$$

But the sequence  $\sqrt{n}(\tilde{J}^n - \bar{J}^n)$  is tight, and we can apply Proposition 8. Using (76), (74), and (78) we deduce

$$\sqrt{n}(\tilde{J}_k^n - J_k) \xrightarrow[\text{law}]{n \rightarrow \infty} -a(T_k, X_{T_k-})c'(X_{T_k-}, \Lambda_k) \sqrt{U_k}N_k^- + a(T_k, X_{T_k}) \sqrt{1 - U_k}N_k^+ + R_k^{\text{aug}^-}.$$

We write the last equation as

$$\sqrt{n}(\tilde{J}_k^n - J_k) \xrightarrow[\text{law}]{n \rightarrow \infty} a(T_k, X_{T_k-}) \sqrt{U_k}N_k^- + a(T_k, X_{T_k}) \sqrt{1 - U_k}N_k^+ + \tilde{R}_k,$$

where  $\tilde{R}_k = R_k^{\text{aug}^-} - (a(T_k, X_{T_k-})(1 + c'(X_{T_k-}, \Lambda_k)) \sqrt{U_k}N_k^-)$ . Using Proposition 8, we deduce that  $\tilde{R}$  is independent of  $N^+$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k, (N_k^-)_k)$ .

From (75), we have  $\tilde{R}_k = R_k^{\text{aug}^+} - (a(T_k, X_{T_k}) \sqrt{1 - U_k}N_k^+)_k$  and we deduce that  $\tilde{R}$  is independent of  $N^-$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k, (N_k^+)_k)$ .

Remarking that  $N^-$  and  $N^+$  are independent conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$ , we deduce that  $\tilde{R}$  is independent of  $(N^-, N^+)$  conditionally on  $(T, \Lambda, (W_t)_{t \in [0,1]}, (U_k)_k)$ .

The theorem is proved.  $\diamond$

### 5.2.4 Proof of Corollary 1

We introduce the conditional probability  $\widehat{\mathbb{P}}^{K_0} = \frac{1_{\{K=K_0\}}}{\mathbb{P}(K=K_0)}\mathbb{P}$  for any  $K_0 \in \mathbb{N}$  such that  $\mathbb{P}(K = K_0) > 0$ . For any  $K_0 \geq 0$ , the sequence  $\sqrt{n}(\tilde{J}^n - J)$  is tight (for the product topology on  $\mathbb{R}^{\mathbb{N}}$ ) under  $\widehat{\mathbb{P}}^{K_0}$ . So, on a subsequence, one has the convergence in law  $\sqrt{n}(\tilde{J}^n - J) \xrightarrow[\text{law}]{\widehat{\mathbb{P}}^{K_0}} \tilde{Z}^{K_0}$ , moreover the subsequence may be chosen independent of  $K_0$  from a diagonal extraction argument.

Fix  $K_0 \geq 1$ , under the probability  $\widehat{\mathbb{P}}^{K_0}$ , the assumptions H0–H3 are satisfied and we can apply Theorem 1 to the  $K_0$  first components of the vector  $\tilde{Z}^{K_0}$ . The Corollary follows from the decomposition of the law of  $\tilde{Z} \stackrel{\text{law}}{=} \sum_{K_0 \geq 0} 1_{\{K=K_0\}} \tilde{Z}^{K_0}$ .  $\diamond$

## 5.3 Study of the estimator $\widehat{J}^n$ : proof of Proposition 2 and Theorem 3

### 5.3.1 Proof of Proposition 2

For  $k \in \{1, \dots, K\}$ , let us note  $i_k$  the integer such that  $i_k/n \leq T_k < (i_k + 1)/n$ . We set  $\mathcal{I} = \{i_1, \dots, i_K\}$  and consider a variable which counts the number of false discovery of a jump by the estimator,

$$E_n = \sum_{i=0}^{n-1} 1_{|X_{(i+1)/n} - X_{i/n}| \geq u_n} 1_{i \notin \mathcal{I}} \quad (79)$$

For  $M > 0$ , we define  $\Omega_M$  as the event  $\Omega_M = \{\sup_{s \in [0,1]} [|b(s, X_s)| + |a(s, X_s)|] \leq M\}$ .

We have

$$\begin{aligned} \mathbb{P}(\{E_n \geq 1\} \cap \Omega_M) &\leq \mathbb{E}[E_n 1_{\Omega_M}] \\ &= \sum_{i=0}^{n-1} \mathbb{E}[1_{|X_{(i+1)/n} - X_{i/n}| \geq u_n} 1_{i \notin \mathcal{I}} 1_{\Omega_M}] \\ &\leq \sum_{i=0}^{n-1} \mathbb{P} \left[ \left\{ \left| \int_{i/n}^{(i+1)/n} a(s, X_s) dW_s + \int_{i/n}^{(i+1)/n} b(s, X_s) ds \right| \geq u_n \right\} \cap \Omega_M \right] \\ &\leq \sum_{i=0}^{n-1} \mathbb{P} \left[ \left\{ \left| \int_{i/n}^{(i+1)/n} a(s, X_s) dW_s \right| \geq u_n - \frac{M}{n} \right\} \cap \Omega_M \right]. \end{aligned} \quad (80)$$

With  $a_M = (a \wedge M) \vee (-M)$  one has, using Markov and Burkholder-Davis-Gundy inequalities:

$$\begin{aligned} \mathbb{P} \left[ \left\{ \left| \int_{i/n}^{(i+1)/n} a(s, X_s) dW_s \right| \geq u_n - \frac{M}{n} \right\} \cap \Omega_M \right] &\leq \mathbb{P} \left[ \left\{ \left| \int_{i/n}^{(i+1)/n} a_M(s, X_s) dW_s \right| \geq u_n - \frac{M}{n} \right\} \right] \\ &\leq C_p \left(u_n - \frac{M}{n}\right)^{-p} n^{-p/2} \quad \forall p > 0, \\ &= C_p n^{p(\varpi-1/2)} \quad \forall p > 0. \end{aligned} \quad (81)$$

Since  $\varpi < 1/2$ , we get, from (80) and (81) by choosing  $p$  large enough,  $\sum_{n \geq 1} \mathbb{P}(\{E_n \geq 1\} \cap \Omega_M) < \infty$ , and by Borel Cantelli's lemma we deduce that  $\mathbb{P}(\bigcap_{n \geq 1} \bigcup_{p \geq n} (\{E_p \geq 1\} \cap \Omega_M)) = 0$ . It immediately implies  $\mathbb{P}(\left(\bigcap_{n \geq 1} \bigcup_{p \geq n} \{E_p \geq 1\}\right) \cap \Omega_M) = 0$  and since  $\bigcup_{M \geq 1} \Omega_M = \Omega$ , we easily deduce that almost surely, there exists  $n$ , such that  $\forall p \geq n, E_p = 0$ . Recalling the definitions (12) and (79), we conclude that almost surely, if  $n$  is large enough,  $\{\hat{i}_1^n, \dots, \hat{i}_{\hat{K}_n}^n\} \subset \mathcal{I}$  and, as a consequence,  $\hat{K}_n \leq K$ .

Now, remark that we have almost surely the convergence, for all  $k \leq K$ ,

$$X_{(i_k+1)/n} - X_{i_k/n} \xrightarrow{n \rightarrow \infty} X_{T_k} - X_{T_k-} = c(X_{T_k-}, \Lambda_k). \quad (82)$$

From the assumption A2, we have  $c(X_{T_k-}, \Lambda_k) \neq 0$  and using that  $u_n \rightarrow 0$ , we deduce that for  $n$  large enough,  $\mathcal{I} \subset \{\hat{i}_1^n, \dots, \hat{i}_{\hat{K}_n}^n\}$ .

As a consequence, we have shown that,

$$\text{almost surely, for } n \text{ large enough, } \hat{K}_n = K \text{ and } \hat{i}_k^n = i_k, \forall k \leq K. \quad (83)$$

Eventually, the proposition follows from (13), (82) and (83).

### 5.3.2 Proof of Theorem 3

We use the notation introduced in the proof of Proposition 2: for  $k \in \{1, \dots, K\}$ , we have  $i_k/n \leq T_k < (i_k + 1)/n$ . Let us define for  $1 \leq k \leq K$ ,  $G_k^n = X_{(i_k+1)/n} - X_{i_k/n} - \Delta X_{T_k}$  and  $G_k^n = 0$  for  $k > K$ . Using (13) and (83), we see that, almost surely, for  $n$  large enough, we have  $\hat{J}^n - J = G^n$ . Hence it is sufficient to study the limit in law of  $\sqrt{n}G^n$ .

Consider any  $K_0 \in \mathbb{N}$  such that  $\mathbb{P}(K = K_0) > 0$  and define  $\hat{\mathbb{P}}^{K_0} = \frac{1_{\{K=K_0\}}}{\mathbb{P}(K=K_0)} \mathbb{P}$ , the conditional probability. Actually, we will prove the convergence of  $\sqrt{n}G^n$  conditionally on the event  $\{K = K_0\}$ , to the law of  $Z$  conditional on  $\{K = K_0\}$ , which is sufficient to prove the theorem.

For  $k > K_0$  we have  $G_k^n = 0$ , hence we focus only on the components  $G_k^n$  with  $k \leq K_0$ .

Define  $\hat{\Omega}^n = \{X \text{ has at most one jump on each interval of size } 1/n\}$ . We have  $\lim_{n \rightarrow \infty} \hat{\mathbb{P}}^{K_0}(\hat{\Omega}^n) = 1$ . On  $\hat{\Omega}^n$ , the following decomposition holds true  $\hat{\mathbb{P}}^{K_0}$  almost surely, for any  $k \leq K_0$ ,

$$\sqrt{n}G_k^n = a(i_k/n, X_{i_k/n})\alpha_{k,n}^- + a(T_k, X_{T_k})\alpha_{k,n}^+ + e_{n,k}$$

where

$$\begin{aligned} \alpha_{k,n}^- &= \sqrt{n}(W_{T_k} - W_{i_k/n}), \quad \alpha_{k,n}^+ = \sqrt{n}(W_{(i_k+1)/n} - W_{T_k}), \\ e_{n,k} &= \sqrt{n} \int_{i_k/n}^{T_k} (a(s, X_s) - a(i_k/n, X_{i_k/n})) dW_s + \sqrt{n} \int_{T_k}^{(i_k+1)/n} (a(s, X_s) - a(T_k, X_{T_k})) dW_s \\ &\quad + \sqrt{n} \int_{i_k/n}^{(i_k+1)/n} b(s, X_s) ds. \end{aligned} \tag{84}$$

First, we show that  $e_{n,k}$  converges to zero in  $\widehat{\mathbb{P}}^{K_0}$  probability as  $n \rightarrow \infty$ . Using A1, the ordinary integral converges almost surely to zero. It remains to see that the two stochastic integrals converge to zero.

Using that the jumps times are  $\mathcal{F}_0$ -measurable, we can write the stochastic integral  $\sqrt{n} \int_{i_k/n}^{T_k} (a(s, X_s) - a(i_k/n, X_{i_k/n})) dW_s$  as a local martingale increment  $\int_0^1 \sqrt{n} 1_{[i_k/n, T_k]}(s) (a(s, X_s) - a(i_k/n, X_{i_k/n})) dW_s$ . The bracket of this local martingale is  $\int_{i_k/n}^{T_k} n (a(s, X_s) - a(i_k/n, X_{i_k/n}))^2 ds$ , which converges to zero almost surely, using the right continuity of the process  $X$ . We deduce that  $\sqrt{n} \int_{i_k/n}^{T_k} (a(s, X_s) - a(i_k/n, X_{i_k/n})) dW_s$  converge to zero in probability. We proceed in the same way to prove that  $\sqrt{n} \int_{T_k}^{(i_k+1)/n} (a(s, X_s) - a(T_k, X_{T_k})) dW_s \xrightarrow{n \rightarrow \infty} 0$  in probability. This yields to the relation,

$$\sqrt{n} G_k^n = a(i_k/n, X_{i_k/n}) \alpha_{k,n}^- + a(T_k, X_{T_k}) \alpha_{k,n}^+ + o_{\widehat{\mathbb{P}}^{K_0}}(1), \quad \text{for } k \leq K_0$$

Using  $\widetilde{H}0$ , and the independence between  $(W_t)_{t \in [0,1]}$  and  $T$  under  $\widehat{\mathbb{P}}^{K_0}$ , we can apply Lemma 1. We get the convergence in law, under  $\widehat{\mathbb{P}}^{K_0}$ ,

$$\begin{aligned} &((T_k)_{k=1, \dots, K_0}, (\alpha_{k,n}^-)_{k=1, \dots, K_0}, (\alpha_{k,n}^+)_{k=1, \dots, K_0}, (W_t)_{t \in [0,1]}) \\ &\xrightarrow{n \rightarrow \infty} ((T_k)_{k=1, \dots, K_0}, (\sqrt{U_k} N_k^-)_{k=1, \dots, K_0}, (\sqrt{1 - U_k} N_k^+)_{k=1, \dots, K_0}, (W_t)_{t \in [0,1]}). \end{aligned}$$

Since the marks  $(\Lambda_k)_k$ , the Brownian motion, and the jump times are independent, we have that, under  $\widehat{\mathbb{P}}^{K_0}$ ,  $(\alpha_{k,n}^-, \alpha_{k,n}^+)_{k \leq K_0}$  converges in law to  $(\sqrt{U_k} N_k^-, \sqrt{1 - U_k} N_k^+)_{k \leq K_0}$  stably with respect to the sigma-field generated by  $(W_t)_{t \in [0,1]}$ ,  $(T_k)_k$  and  $(\Lambda_k)_k$ . The limit can be represented on the extended space  $\widetilde{\Omega}$  endowed with the probability  $\widetilde{\mathbb{P}}$  conditional on  $K = K_0$ .

But the process  $X$  is measurable with respect to  $\mathcal{F}_1$ , and we deduce the stable convergence,

$$\sqrt{n} G_k^n = a(i_k/n, X_{i_k/n}) \alpha_{k,n}^- + a(T_k, X_{T_k}) \alpha_{k,n}^+ \xrightarrow{n \rightarrow \infty} a(T_k, X_{T_k^-}) \sqrt{U_k} N_k^- + a(T_k, X_{T_k}) \sqrt{1 - U_k} N_k^+,$$

for  $k = 1, \dots, K_0$ , under  $\widehat{\mathbb{P}}^{K_0}$ .

By simple computations, this implies the convergence of  $(\sqrt{n} G_k^n)_k$  under  $\mathbb{P}$ , and the theorem is proved.  $\diamond$

**Acknowledgements.** We would like to thank the referees for their careful reading and suggestions which improved the presentation of the paper.

## References

- [1] Yacine Aït-Sahalia. Telling from discrete data whether the underlying continuous-time model is a diffusion. *J. Finance*, 57:2075–2112, 2002.
- [2] Yacine Aït-Sahalia, Jianqing Fan, and Heng Peng. Nonparametric transition-based tests for jump diffusions. *J. Amer. Statist. Assoc.*, 104(487):1102–1116, 2009.
- [3] Yacine Aït-Sahalia and Jean Jacod. Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.*, 37(5A):2202–2244, 2009.
- [4] Yacine Aït-Sahalia and Jean Jacod. Testing for jumps in a discretely observed process. *Ann. Statist.*, 37(1):184–222, 2009.
- [5] Robert Azencott. Densité des diffusions en temps petit: développements asymptotiques. I. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 402–498. Springer, Berlin, 1984.
- [6] Ole E. Barndorff-Nielsen and Neil Shephard. Econometrics of testing for jumps in financial economics using bipower variation. *J. Financial Econometrics*, 4:1–30, 2006.
- [7] Ole E. Barndorff-Nielsen, Neil Shephard, and Matthias Winkel. Limit theorems for multi-power variation in the presence of jumps. *Stochastic Process. Appl.*, 116(5):796–806, 2006.
- [8] Rama Cont and Mancini Cecilia. Nonparametric tests for pathwise properties of semi-martingales. *Bernoulli*, 17:781–813, 2011.
- [9] Arnaud Gloter and Emmanuel Gobet. LAMN property for hidden processes: the case of integrated diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(1):104–128, 2008.
- [10] Emmanuel Gobet. Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach. *Bernoulli*, 7(6):899–912, 2001.
- [11] Xin Huang and George Tauchen. The relative contribution of jumps of total price variance. *J. Financial Econometrics*, 4:456–499, 2006.

- [12] I. A. Ibragimov and R. Z. Has'minskiĭ. *Statistical estimation*, volume 16 of *Applications of Mathematics*. Springer-Verlag, New York, 1981. Asymptotic theory, Translated from the Russian by Samuel Kotz.
- [13] Jean Jacod. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Process. Appl.*, 118(4):517–559, 2008.
- [14] Jean Jacod and Philip Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26(1):267–307, 1998.
- [15] Jean Jacod and Viktor Todorov. Do price and volatility jump together? *Ann. Appl. Probab.*, 20(4):1425–1469, 2010.
- [16] P. Jeganathan. On a decomposition of the limit distribution of a sequence of estimators. *Sankhyā Ser. A*, 43(1):26–36, 1981.
- [17] P. Jeganathan. On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A*, 44(2):173–212, 1982.
- [18] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997. Reprint of the 1990 original.
- [19] Cecilia Mancini. Estimation of the characteristics of the jumps of a general Poisson-diffusion model. *Scand. Actuar. J.*, (1):42–52, 2004.
- [20] Cecilia Mancini. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.*, 36(2):270–296, 2009.
- [21] David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.
- [22] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.
- [23] Jeannette H. C. Woerner. Power and multipower variation: inference for high frequency data. In *Stochastic finance*, pages 343–364. Springer, New York, 2006.