Electronic Journal of Probability

Electron. J. Probab. **19** (2014), no. 110, 1–33.
ISSN: 1083-6489 DOI: 10.1214/EJP.v19-3175

On the distances between probability density functions

Vlad Bally* Lucia Caramellino†

Abstract

We give estimates of the distance between the densities of the laws of two functionals F and G on the Wiener space in terms of the Malliavin-Sobolev norm of $F-G$. We actually consider a more general framework which allows one to treat with similar (Malliavin type) methods functionals of a Poisson point measure (solutions of jump type stochastic equations). We use the above estimates in order to obtain a criterion which ensures that convergence in distribution implies convergence in total variation distance; in particular, if the functionals at hand are absolutely continuous, this implies convergence in L^1 of the densities.

Keywords: Integration by parts formulas; Riesz transform; Malliavin calculus; weak convergence; total variation distance.
AMS MSC 2010: 60H07; 60H30.
Submitted to EJP on November 29, 2013, final version accepted on September 10, 2014.

1 Introduction

In this paper we give estimates of the distance between the densities of the laws of two functionals F and G on the Wiener space in terms of the Malliavin-Sobolev norm of $F-G$. Actually, we consider a slightly more general framework defined in [5] or [6] which allows one to treat with similar methods functionals of a Poisson point measure (solutions of jump type stochastic equations). Such estimates may be used in order to study the behavior of a diffusion process in short time as it is done in [3]. But here we focus on a different application: we use the above estimates in order to obtain a criterion which guarantees that convergence in distribution implies convergence in total variation distance; in particular, if the functionals at hand are absolutely continuous, this implies convergence in L^1 of the densities. Moreover, by using some more general distances, we obtain the convergence of the derivatives of the density functions as well. The main estimates are given in Theorem 2.1 in the general framework and in Theorem 2.14 in the case of the Wiener space. The convergence result is given in Theorem 2.11 and, for the Wiener space, in Theorem 2.20.

*Université Paris-Est, LAMA (UMR CNRS, UPEMLV, UPEC), INRIA, F-77454 Marne-la-Vallée, France.
E-mail: bally@univ-mlv.fr
†Dipartimento di Matematica, Università di Roma - Tor Vergata, Italy.
E-mail: caramell@mat.uniroma2.it

The reader interested in the Wiener space case may go directly to Section 2.4. For functionals on the Wiener space we get one more result which is in between the Bouleau-Hirsch absolute continuity criterion and the classical criterion of Malliavin for existence and regularity of the density of the law of a d dimensional functional F: we prove that if $F \in \mathbb{D}^{2,p}$ with $p > d$ and $\mathbb{P}(\det \sigma_F > 0) > 0$ (σ_F denoting the Malliavin covariance matrix of F) then, conditionally to $\{\sigma_F > 0\}$ the law of F is absolutely continuous and the density is lower semi-continuous. This regularity property implies that the law of F is locally lower bounded by the Lebesgue measure and this property turns out to be interesting - see the joint paper [4].

In the last years number of results concerning the weak convergence of functionals on the Wiener space using Malliavin calculus and Stein's method have been obtained by Nourdin, Peccati, Nualart and Poly, see [16], [17] and [19]. In particular in [16] and [19] the authors consider functionals living in a finite (and fixed) direct sum of chaoses and prove that, under a very weak non degeneracy condition, the convergence in distribution of a sequence of such functionals implies the convergence in total variation. Our initial motivation was to obtain similar results for general functionals: we consider a sequence of d dimensional functionals $F_n, n \in \mathbb{N}$, which is bounded in $\mathbb{D}^{3,p}$ for every $p \geq 1$. Under a very weak non degeneracy condition (see (2.39)) we prove that the convergence in distribution of such a sequence implies the convergence in the total variation distance. Moreover we prove that if a sequence $F_n, n \in \mathbb{N}$, is bounded in every $\mathbb{D}^{3,p}$, $p \geq 1$, $\lim_n F_n = F$ in L^2 and $\det \sigma_F > 0$ a.s., then $\lim_n F_n = F$ in total variation. Recently, Malicet and Poly [15] have proved an alternative version of this result: if $\lim_n F_n = F$ in $\mathbb{D}^{1,2}$ and $\det \sigma_F > 0$ a.s. then the convergence takes place in the total variation distance.

The paper is organized as follows. In Section 2.1, following [5], we introduce an abstract framework which permits to obtain integration by parts formulas. In Section 2.2 we give the main estimate (the distance between two density functions) in this framework and in Section 2.3 we obtain the convergence results. In Section 2.4 we come back to the Wiener space framework, so here the objects and the notations are the standard ones from Malliavin calculus (we refer to Nualart [18] for the general theory). Section 3 is devoted to the proof of the main estimate, that is of Theorem 2.1. Finally, in Section 4 we illustrate our convergence criterion with an example of jump type equation coming from [5].

2 Main results

2.1 Abstract integration by parts framework

In this section we briefly recall the construction of integration by parts formulas for functionals of a finite dimensional noise which mimic the infinite dimensional Malliavin calculus as done in [5] and [6]. We are going to introduce operators that represent the finite dimensional variant of the derivative and the divergence operators from the classical Malliavin calculus - and as an outstanding consequence all the constants which appear in the estimates do not depend on the dimension of the noise. So, given some constants $c_i \in \mathbb{R}, i = 1, ..., m$ we denote by $\mathcal{C}(c_1, ..., c_m)$ the family of universal constants which depend on $c_i, i = 1, ..., m$ only. So $C \in \mathcal{C}(c_1, ..., c_m)$ means that C depends on $c_i, i = 1, ..., m$ but on nothing else in the statement. This is crucial in the following theorems.

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we consider a random variable $V = (V_1, ..., V_J)$ which represents the basic noise. Here $J \in \mathbb{N}$ is a deterministic integer. For each $i = 1, ..., J$ we consider two constants $-\infty \leq a_i < b_i \leq \infty$ that are allowed to reach ∞. We denote

$$O_i = \{v = (v_1, ..., v_J) : a_i < v_i < b_i\}, \quad i = 1, \dots, J. \tag{2.1}$$

The basic hypothesis is that the law of V is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^J$ and the density p_J is smooth with respect to v_i on the set O_i. The natural example which comes on in the standard Malliavin calculus is the Gaussian law on $\mathbb{R}^J$, in which $a_i = -\infty$ and $b_i = +\infty$. But we may also (as an example) take V_i independent random variables of exponential law and here, $a_i = 0$ and $b_i = \infty$.

In order to obtain integration by parts formulas for functionals of V, one performs classical integration by parts with respect to $p_J(v)dv$. But in order to nullify the border terms in a_i and b_i, it suffices to take into account suitable "weights"

$$\pi_i \,:\, \mathbb{R}^J \to [0,1], \quad i = 1, ..., J.$$

We give the precise statement of the hypothesis. But let us first set up the notations we are going to use. We set $C^k(\mathbb{R}^d)$ the space of the functions which are continuously differentiable up to order k and $C^\infty(\mathbb{R}^d)$ for functions which are infinitely differentiable. We use the subscripts p, resp. b, to denote functions having polynomial growth, resp. bounded, together with their derivatives, and this gives $C^k_p(\mathbb{R}^d)$, $C^\infty_p(\mathbb{R}^d)$, $C^k_b(\mathbb{R}^d)$ and $C^\infty_b(\mathbb{R}^d)$. For $k \in \mathbb{N}$ and for a multi index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., d\}^k$ we denote $|\alpha| = k$ and $\partial_\alpha f = \partial_{x^{\alpha_1}} ... \partial_{x^{\alpha_k}} f$. The case $k = 0$ is allowed and gives $\partial_\alpha f = f$. We also set $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$.

So, throughout this paper, we assume the following assumption does hold.

Assumption. *The law of the vector $V = (V_1, ..., V_J)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^J$ and we denote with p_J the density; we assume that p_J has polynomial growth. We also assume that*

(H0) *for all $i \in \{1, \ldots, J\}$, $0 \le \pi_i \le 1$, $\pi_i \in C^\infty_b$ and there exist $-\infty \le a_i < b_i \le +\infty$ such that, with O_i defined in (2.1), $\{\pi_i > 0\} \subset O_i$;*

(H1) *the set $\{p_J > 0\}$ is open in $\mathbb{R}^J$ and on $\{p_J > 0\}$ we have $\ln p_J \in C^\infty$.*

We define now the functional spaces and the differential operators.

□ **Simple functionals.** A random variable F is called a simple functional if there exists $f \in C^\infty_p(\mathbb{R}^J)$ such that $F = f(V)$. We denote through $\mathcal{S}$ the set of simple functionals.

□ **Simple processes.** A simple process is a random variable $U = (U_1, \ldots, U_J)$ in $\mathbb{R}^J$ such that $U_i \in \mathcal{S}$ for each $i \in \{1, \ldots, J\}$. We denote by $\mathcal{P}$ the space of the simple processes. On $\mathcal{P}$ we define the scalar product

$$\langle \cdot, \cdot \rangle : \mathcal{P} \times \mathcal{P} \to \mathcal{S}, \quad (U, V) \mapsto \langle U, V \rangle_J = \sum_{i=1}^J U_i V_i.$$

□ **The derivative operator.** We define $D : \mathcal{S} \to \mathcal{P}$ by

$$DF := (D_i F)_{i=1,\ldots,J} \in \mathcal{P} \quad \text{where } D_i F := \pi_i(V) \partial_i f(V). \tag{2.2}$$

□ **The divergence operator.** Let $U = (U_1, \ldots, U_J) \in \mathcal{P}$, so that $U_i \in \mathcal{S}$ and $U_i = u_i(V)$, for some $u_i \in C^\infty_p(\mathbb{R}^J)$, $i = 1, \ldots, J$. We define $\delta \,:\, \mathcal{P} \to \mathcal{S}$ by

$$\delta(U) = \sum_{i=1}^J \delta_i(U), \quad \text{with } \delta_i(U) := -(\partial_{v_i}(\pi_i u_i) + \pi_i u_i 1_{O_i} \partial_{v_i} \ln p_J)(V),\ i = 1, \ldots, J. \tag{2.3}$$

Clearly, both D and δ depend on π so a correct notation should be D^π and δ^π. Since here the weights π_i are fixed, we do not mention them in the notation.

□ **The Malliavin covariance matrix.** For $F \in \mathcal{S}^d$, the Malliavin covariance matrix of F is defined by

$$\sigma_F^{k,k'} = \langle DF^k, DF^{k'} \rangle_J = \sum_{j=1}^J D_j F^k D_j F^{k'}, \quad k, k' = 1, \dots, d.$$

We also denote

$$\gamma_F(\omega) = \sigma_F^{-1}(\omega), \quad \omega \in \{\det \sigma_F > 0\}.$$

□ **The Ornstein Uhlenbeck operator.** We define $L : \mathcal{S} \to \mathcal{S}$ by

$$L(F) = \delta(DF). \tag{2.4}$$

□ **Higher order derivatives and norms.** Let $\alpha = (\alpha_1, \dots, \alpha_k)$ be a multi-index, with $\alpha_i \in \{1, \dots, J\}$, for $i = 1, \dots, k$ and $|\alpha| = k$. For $F \in \mathcal{S}$, we define recursively

$$D_{(\alpha_1,\dots,\alpha_k)} F = D_{\alpha_k}(D_{(\alpha_1,\dots,\alpha_{k-1})} F) \quad \text{and} \quad D^{(k)} F = \big(D_{(\alpha_1,\dots,\alpha_k)} F\big)_{\alpha_i \in \{1,\dots,J\}}. \tag{2.5}$$

We set $D^{(0)} F = F$ and we notice that $D^{(1)} F = DF$. Remark that $D^{(k)} F \in \mathbb{R}^{J \otimes k}$ and consequently we define the norm of $D^{(k)} F$ as

$$|D^{(k)} F| = \Big(\sum_{\alpha_1,\dots,\alpha_k=1}^J |D_{(\alpha_1,\dots,\alpha_k)} F|^2 \Big)^{1/2}. \tag{2.6}$$

Moreover, we introduce the following norms for simple functionals: for $F \in \mathcal{S}$ we set

$$|F|_{1,l} = \sum_{k=1}^l |D^{(k)} F| = \sum_{k=0}^l |D^{(k)} F|, \qquad |F|_l = |F| + |F|_{1,l} \tag{2.7}$$

and for $F = (F^1, \dots, F^d) \in \mathcal{S}^d$, $|F|_{1,l} = \sum_{r=1}^d |F^r|_{1,l}$ and $|F|_l = \sum_{r=1}^d |F^r|_l$. Finally, for $U = (U_1, \dots, U_J) \in \mathcal{P}$, we set $D^{(k)} U = (D^{(k)} U_1, \dots, D^{(k)} U_J)$ and we define the norm of $D^{(k)} U$ as

$$|D^{(k)} U| = \Big(\sum_{i=1}^J |D^{(k)} U_i|^2 \Big)^{1/2}.$$

We allow the case $k = 0$, giving $|U| = \langle U, U \rangle_J^{1/2}$. Similarly to (2.7), we set $|U|_l = \sum_{k=0}^l |D^{(k)} U|$.

□ **Localization functions.** As it will be clear in the sequel we need to introduce some localization random variables as follows. Consider a random variable $\Theta \in \mathcal{S}$ taking values on $[0, 1]$ and set

$$d\mathbb{P}_\Theta = \Theta d\mathbb{P}.$$

$\mathbb{P}_\Theta$ is a non negative measure (but generally not a probability measure) and we set $\mathbb{E}_\Theta$ the expectation (integral) w.r.t. $\mathbb{P}_\Theta$. For $F \in \mathcal{S}$, we define

$$\|F\|_{p,\Theta} = \mathbb{E}_\Theta(|F|^p)^{1/p}$$

and

$$\|F\|_{1,l,p,\Theta}^p = \mathbb{E}_\Theta(|F|_{1,l}^p) \quad \text{and} \quad \|F\|_{l,p,\Theta}^p = \|F\|_{p,\Theta}^p + \|F\|_{1,l,p,\Theta}^p. \tag{2.8}$$

that is $\|\cdot\|_{p,\Theta}$ and $\|\cdot\|_{l,p,\Theta}$ are the standard L^p Sobolev norms in Malliavin calculus with $\mathbb{P}$ replaced by the localized measure $\mathbb{P}_\Theta$. Notice also that $\|F\|_{1,l,p,\Theta}^p$ does not take into account the L^p norm of F itself but only of the derivatives of F. This is the motivation of considering this norm.

Since $|F|_0 = |F|$, one has $\|F\|_{0,p,\mathbf{\Theta}} = \|F\|_{p,\mathbf{\Theta}}$. In the case $\mathbf{\Theta} = 1$ we come back to the standard notation: $\|F\|_p = \mathbb{E}(|F|^p)$ and

$$\|F\|_{1,l,p}^p = \mathbb{E}(|F|_{1,l}^p) \quad \text{and} \quad \|F\|_{l,p}^p = \mathbb{E}(|F|^p + |F|_{1,l}^p). \tag{2.9}$$

Notice also that since $\mathbf{\Theta} \leq 1$ we have

$$\|F\|_{1,l,p,\mathbf{\Theta}} \leq \|F\|_{1,l,p} \quad \text{and} \quad \|F\|_{l,p,\mathbf{\Theta}} \leq \|F\|_{l,p}. \tag{2.10}$$

For $p \in \mathbb{N}$ we set

$$\mathrm{m}_{q,p}(\mathbf{\Theta}) := 1 \vee \|\ln \mathbf{\Theta}\|_{1,q,p,\mathbf{\Theta}}. \tag{2.11}$$

Since $\mathbf{\Theta} > 0$ almost surely with respect to $\mathbb{P}_{\mathbf{\Theta}}$ the above quantity makes sense.

We will work with localization random variables of the following specific form. For $a > 0$, set $\psi_a, \phi_a : \mathbb{R} \to \mathbb{R}_+$ as follows:

$$\begin{aligned} \psi_a(x) &= 1_{|x|\leq a} + \exp\left(1 - \tfrac{a^2}{a^2-(|x|-a)^2}\right) 1_{a<|x|<2a}, \\ \phi_a(x) &= 1_{|x|\geq a} + \exp\left(1 - \tfrac{a^2}{(2|x|-a)^2}\right) 1_{a/2<|x|<a}. \end{aligned} \tag{2.12}$$

The function ψ_a is suited to localize around zero and ϕ_a is suited to localize far from zero. Then $\psi_a, \phi_a \in C_b^\infty(\mathbb{R})$, $0 \leq \psi_a \leq 1$, $0 \leq \phi_a \leq 1$ and we have the following property: for every $p, k \in \mathbb{N}$ there exists a universal constant $C_{k,p}$ such that for every $x \in \mathbb{R}_+$

$$\psi_a(x) \left|(\ln \psi_a)^{(k)}(x)\right|^p \leq \frac{C_{k,p}}{a^{pk}} \quad \text{and} \quad \phi_a(x) \left|(\ln \phi_a)^{(k)}(x)\right|^p \leq \frac{C_{k,p}}{a^{pk}}. \tag{2.13}$$

We consider now $\Theta_i \in \mathcal{S}$ and $a_i > 0, i = 1, ..., l + l'$ and define

$$\mathbf{\Theta} = \prod_{i=1}^{l} \psi_{a_i}(\Theta_i) \times \prod_{i=l+1}^{l+l'} \phi_{a_i}(\Theta_i). \tag{2.14}$$

As an easy consequence of (2.13) we obtain

$$\mathrm{m}_{q,p}(\mathbf{\Theta}) \leq 1 \vee \|\ln \mathbf{\Theta}\|_{1,q,p,\mathbf{\Theta}} \leq 1 \vee C_{p,q} \sum_{i=1}^{l+l'} \frac{1}{a_i^q} \|\Theta_i\|_{1,q,p,\mathbf{\Theta}}. \tag{2.15}$$

In particular, if $\|\Theta_i\|_{1,q,p} < \infty, i = 1, ..., l + l'$ then

$$\mathrm{m}_{q,p}(\mathbf{\Theta}) \leq 1 + C_{p,q} \sum_{i=1}^{l+l'} \frac{1}{a_i^q} \|\Theta_i\|_{1,q,p} < \infty. \tag{2.16}$$

Moreover, given some $q \in \mathbb{N}, p \geq 1$ we denote

$$\mathrm{U}_{q,p,\mathbf{\Theta}}(F) := \max\{1, \mathbb{E}_{\mathbf{\Theta}}((\det \sigma_F)^{-p})(\|F\|_{1,q+2,p,\mathbf{\Theta}} + \|LF\|_{q,p,\mathbf{\Theta}})\}. \tag{2.17}$$

In the case $\mathbf{\Theta} = 1$ we have $\mathrm{m}_{q,p}(\mathbf{\Theta}) = 1$ and

$$\mathrm{U}_{q,p}(F) := \max\{1, \mathbb{E}((\det \sigma_F)^{-p})(\|F\|_{1,q+2,p} + \|LF\|_{q,p})\}. \tag{2.18}$$

Notice that $\mathrm{U}_{q,p,\mathbf{\Theta}}(F)$ and $\mathrm{U}_{q,p}(F)$ do not involve the L^p norm of F but only of its derivatives and of LF.

We are now able to state the main result in our paper.

Theorem 2.1. *Let $q \in \mathbb{N}_*$. We consider the localization random variable $\boldsymbol{\Theta}$ defined in (2.14) and we assume that for every $p \in \mathbb{N}$ one has $\|\Theta_i\|_{q+2,p} < \infty, i = 1, ..., l + l'$. In particular $\mathrm{m}_{q+2,p}(\boldsymbol{\Theta}) < \infty$. Let $\mathrm{U}_{q,p,\boldsymbol{\Theta}}(F)$ be as in (2.17).*

***A**. Let $F \in \mathcal{S}^d$ be such that $\mathrm{U}_{q,p,\boldsymbol{\Theta}}(F) < \infty$ for every $p \in \mathbb{N}$. Then under $\mathbb{P}_{\boldsymbol{\Theta}}$ the law of F is absolutely continuous with respect to the Lebesgue measure. We denote by $p_{F,\boldsymbol{\Theta}}$ its density and we have $p_{F,\boldsymbol{\Theta}} \in C^{q-1}(\mathbb{R}^d)$. Moreover there exist $C, a, b, p \in \mathcal{C}(q, d)$ such that for every $y \in \mathbb{R}^d$ and every multi index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., d\}^k, k \in \{0, ..., q\}$ one has*

$$|\partial_\alpha p_{F,\boldsymbol{\Theta}}(y)| \le C\mathrm{U}^a_{q,p,\boldsymbol{\Theta}}(F) \times \mathrm{m}^a_{q+2,p}(\boldsymbol{\Theta}) \times \big(\mathbb{P}_{\boldsymbol{\Theta}}(|F - y| < 2)\big)^b. \tag{2.19}$$

***B**. Let $F, G \in \mathcal{S}^d$ be such that $\mathrm{U}_{q+1,p,\boldsymbol{\Theta}}(F), \mathrm{U}_{q+1,p,\boldsymbol{\Theta}}(G) < \infty$ for every $p \in \mathbb{N}$ and let $p_{F,\boldsymbol{\Theta}}$ and $p_{G,\boldsymbol{\Theta}}$ be the densities of the laws of F respectively of G under $\mathbb{P}_{\boldsymbol{\Theta}}$. There exist $C, a, b, p \in \mathcal{C}(q, d)$ such that for every $y \in \mathbb{R}^d$ and every multi index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., d\}^k$, $0 \le k \le q$ one has*

$$\begin{aligned} |\partial_\alpha p_{F,\boldsymbol{\Theta}}(y) - \partial_\alpha p_{G,\boldsymbol{\Theta}}(y)| \;\le\;& C\mathrm{U}^a_{q+1,p,\boldsymbol{\Theta}}(F) \times \mathrm{U}^a_{q+1,p,\boldsymbol{\Theta}}(G) \times \mathrm{m}^a_{q+2,p}(\boldsymbol{\Theta}) \times \\ & \times \big(\mathbb{P}_{\boldsymbol{\Theta}}(|F - y| < 2|) + \mathbb{P}_{\boldsymbol{\Theta}}(|G - y| < 2)\big)^b \times \\ & \times \left(\|F - G\|_{q+2,p,\boldsymbol{\Theta}} + \|LF - LG\|_{q,p,\boldsymbol{\Theta}}\right). \end{aligned} \tag{2.20}$$

Remark 2.2. The above result can be written in the case $\boldsymbol{\Theta} = 1$. Here, $\mathrm{m}_{q+2,p}(\boldsymbol{\Theta}) = 1$ and the quantities $\|F - G\|_{q+2,p,\boldsymbol{\Theta}}$ and $\|LF - LG\|_{q,p,\boldsymbol{\Theta}}$ are replaced by $\|F - G\|_{q+2,p}$ and $\|LF - LG\|_{q,p}$ respectively.

Remark 2.3. Since $\mathbb{P}_{\boldsymbol{\Theta}}(A) \le \mathbb{P}(A)$, in (2.19) and (2.20) one can replace $\mathbb{P}_{\boldsymbol{\Theta}}$ with $\mathbb{P}$.

Remark 2.4. Estimates (2.19) and (2.20) may be rewritten in terms of the queues of the law of F and G by noticing that if $|y| > 4$ then $\{|F - y| < 2\} \subset \{|F| > |y|/2\}$ and $\{|G - y| < 2\} \subset \{|G| > |y|/2\}$. But we can do something else. In fact, by using the Markov inequality, for every $\ell \ge 1$ and for $|y| > 4$ we get $\mathbb{P}_{\boldsymbol{\Theta}}(|F - y| < 2) \le \mathbb{P}_{\boldsymbol{\Theta}}(|F| > |y|/2) \le C\,\mathbb{E}_{\boldsymbol{\Theta}}(|F|^\ell)/(1 + |y|)^\ell$, C denoting a universal constant. And by taking into account also the case $|y| \le 4$, for a suitable C we have

$$\mathbb{P}_{\boldsymbol{\Theta}}(|F - y| < 2) \le C\,\frac{1 \vee \mathbb{E}_{\boldsymbol{\Theta}}(|F|^\ell)}{(1 + |y|)^\ell}$$

and similarly for G. Then, the second factors in formulas (2.19) and (2.20) may be written in terms of the above inequality as follows: for every $\ell \ge 1$ and $y \in \mathbb{R}^d$,

$$|\partial_\alpha p_{F,\boldsymbol{\Theta}}(y)| \le C\mathrm{U}^a_{q,p,\boldsymbol{\Theta}}(F) \times \mathrm{m}^a_{q+2,p}(\boldsymbol{\Theta}) \times \frac{(1 + \|F\|^\ell_{\ell,\boldsymbol{\Theta}})^b}{(1 + |y|)^{\ell b}} \tag{2.21}$$

and

$$\begin{aligned} |\partial_\alpha p_{F,\boldsymbol{\Theta}}(y) - \partial_\alpha p_{G,\boldsymbol{\Theta}}(y)| \;\le\;& C\mathrm{U}^a_{q+1,p,\boldsymbol{\Theta}}(F) \times \mathrm{U}^a_{q+1,p,\boldsymbol{\Theta}}(G) \times \mathrm{m}^a_{q+2,p}(\boldsymbol{\Theta}) \times \\ & \times \frac{(1 + \|F\|^\ell_{\ell,\boldsymbol{\Theta}} + \|G\|^\ell_{\ell,\boldsymbol{\Theta}})^b}{(1 + |y|)^{\ell b}} \times \\ & \times \left(\|F - G\|_{q+2,p,\boldsymbol{\Theta}} + \|LF - LG\|_{q,p,\boldsymbol{\Theta}}\right). \end{aligned} \tag{2.22}$$

The proof of Theorem 2.1 is the main effort in our paper and it is postponed for Section 3 (see Proposition 3.8 **C.** and Theorem 3.10).

As a consequence of Theorem 2.1 we obtain the following regularization result. Let γ_δ be the density of the centred normal law of covariance $\delta \times I$ on $\mathbb{R}^d$. Here $\delta > 0$ and I is the identity matrix.

Lemma 2.5. *There exist universal constants $C, p, a \in \mathcal{C}(d)$ such that for every $\varepsilon > 0, \delta > 0$ and every $F \in \mathcal{S}^d$ one has*

$$|\mathbb{E}(f(F)) - \mathbb{E}(f * \gamma_\delta(F))| \leq C \left\| f \right\|_\infty \left(\mathbb{P}(\sigma_F < \varepsilon) + \frac{\sqrt{\delta}}{\varepsilon^p}(1 + \left\| F \right\|_{3,p} + \left\| LF \right\|_{1,p})^a \right) \quad (2.23)$$

for every bounded and measurable $f : \mathbb{R}^d \to \mathbb{R}$. Moreover, if $f \in L^1(\mathbb{R}^d)$

$$|\mathbb{E}(f(F)) - \mathbb{E}(f * \gamma_\delta(F))| \leq C(\left\| f \right\|_\infty + \left\| f \right\|_1)\left(\mathbb{P}(\sigma_F < \varepsilon) + \frac{\sqrt{\delta}}{\varepsilon^p}(1 + \left\| F \right\|_{1,3,p} + \left\| LF \right\|_{1,p})^a \right) \quad (2.24)$$

Notice that in the r.h.s. of (2.24) $\left\| F \right\|_{3,p}$ is replaced by $\left\| F \right\|_{1,3,p}$ so $\left\| F \right\|_p$ is not involved. The price to be paid is that we have to replace $\left\| f \right\|_\infty$ with $\left\| f \right\|_\infty + \left\| f \right\|_1$.

Proof. Along this proof C denotes a constant in $\mathcal{C}(d)$ which may change from a line to another. We construct the localization random variable $\mathbf{\Theta}_\varepsilon = \phi_\varepsilon(\det \sigma_F)$ with ϕ_ε given in (2.12). By (2.15) for every $p \geq 1$

$$\mathrm{m}_{q,p}(\mathbf{\Theta}_\varepsilon) \leq \frac{C}{\varepsilon^q} \| \det \sigma_F \|_{q,p,\mathbf{\Theta}_\varepsilon} \leq \frac{C}{\varepsilon^q} \| F \|_{1,q+1,p}^d. \quad (2.25)$$

We fix $\delta \in (0,1)$ and we define $F_\delta = F + \sqrt{\delta}\,\Delta$ where Δ is a standard Gaussian random variable independent of V. We will use the result in Theorem 2.1, here not with respect to $V = (V_1, ..., V_J)$ but with respect to $(V, \Delta) = (V_1, ..., V_J, \Delta)$. The Malliavin covariance matrix of F with respect to (V, Δ) is the same as the one with respect to V (because F does not depend on Δ) so on the set $\{\mathbf{\Theta}_\varepsilon \neq 0\}$ we have $\det \sigma_F \geq \varepsilon$. We denote by σ_{F_δ} the Malliavin covariance matrix of F_δ computed with respect to (V, Δ). We have $\langle \sigma_{F_\delta} \xi, \xi \rangle = \delta \left| \xi \right|^2 + \langle \sigma_F \xi, \xi \rangle$. By Lemma 7-29, pg 92 in [9], for every symmetric non negative defined matrix Q one has

$$\frac{1}{\det Q} \leq C_1 \int_{\mathbb{R}^d} \left| \xi \right|^d e^{-\langle Q\xi, \xi \rangle} d\xi \leq C_2 \frac{1}{\det Q}$$

where C_1 and C_2 are universal constants. Using these two inequalities we obtain $\det \sigma_{F_\delta} \geq \frac{1}{C} \det \sigma_F \geq \frac{1}{C}\varepsilon$ on the set $\mathbf{\Theta}_\varepsilon > 0$. So for $\varepsilon \in (0,1)$ we have

$$\| (\det \sigma_F)^{-1} \|_{p,\mathbf{\Theta}_\varepsilon} + \| (\det \sigma_{F_\delta})^{-1} \|_{p,\mathbf{\Theta}_\varepsilon} \leq C\varepsilon^{-1}.$$

It is also easy to check that

$$\| F_\delta \|_{1,3,p,\mathbf{\Theta}_\varepsilon} + \| LF_\delta \|_{1,p,\mathbf{\Theta}_\varepsilon} \leq C(1 + \| F \|_{1,3,p,\mathbf{\Theta}_\varepsilon} + \| LF \|_{1,p,\mathbf{\Theta}_\varepsilon})$$

so finally we obtain

$$\begin{aligned} \mathrm{U}_{1,p,,\mathbf{\Theta}_\varepsilon}(F) + \mathrm{U}_{1,p,,\mathbf{\Theta}_\varepsilon}(F_\delta) &\leq C(1 + \varepsilon^{-p}(\| F \|_{1,3,p,\mathbf{\Theta}_\varepsilon} + \| LF \|_{1,p,\mathbf{\Theta}_\varepsilon})) \\ &\leq C(1 + \varepsilon^{-p}(\| F \|_{1,3,p} + \| LF \|_{1,p})). \end{aligned}$$

By using (2.25), we apply Theorem 2.1 and we obtain

$$\begin{aligned} |p_{F,\mathbf{\Theta}_\varepsilon}(y) - p_{F_\delta,\mathbf{\Theta}_\varepsilon}(y)| &\leq C(1 + \varepsilon^{-p}(\| F \|_{1,3,p} + \| LF \|_{1,p}))^a \times \\ &\quad \times (\| F - F_\delta \|_{2,p} + \| LF - LF_\delta \|_{0,p}) \,. \end{aligned} \quad (2.26)$$

The r.h.s. of the above inequality does not depend on y, so its integral over $\mathbb{R}^d$ is infinite. In order to obtain a finite integral we use inequality (2.22) discussed in Remark 2.4 with ℓ large enough: we may find $C, p, a, b \in \mathcal{C}(d)$ such that

$$\begin{aligned} |p_{F,\mathbf{\Theta}_\varepsilon}(y) - p_{F_\delta,\mathbf{\Theta}_\varepsilon}(y)| &\leq C(1 + \varepsilon^{-p}(\| F \|_{3,p} + \| LF \|_{1,p}))^a \times \frac{1}{(1 + |y|)^{2d}} \times \\ &\quad \times (\| F - F_\delta \|_{2,p} + \| LF - LF_\delta \|_{0,p}) \,. \end{aligned} \quad (2.27)$$

But now $\|F\|_p$ comes on and this is why we have to replace $\|F\|_{1,3,p}$ by $\|F\|_{3,p}$.

Moreover one can easily check using directly the definitions that

$$\|F - F_\delta\|_{2,p} + \|LF - LF_\delta\|_{0,p} \le C\delta^{1/2}.$$

So finally we obtain

$$|p_{F,\mathbf{\Theta}_\varepsilon}(y) - p_{F_\delta,\mathbf{\Theta}_\varepsilon}(y)| \le \frac{C}{(1+|y|)^{2d}\varepsilon^p}(1 + \|F\|_{3,p} + \|LF\|_{1,p}))^a \times \sqrt{\delta}. \tag{2.28}$$

We are now ready to start the proof of our Lemma. We take $f \in C(\mathbb{R}^d)$ with $\|f\|_\infty < \infty$ and we write

$$\begin{aligned}\mathbb{E}(f(F)) - \mathbb{E}(f * \gamma_\delta(F)) =& \mathbb{E}(f(F)) - \mathbb{E}(f(F_\delta)) \\ =& \big[\mathbb{E}(f(F)(1-\mathbf{\Theta}_\varepsilon)) - \mathbb{E}(f(F_\delta)(1-\mathbf{\Theta}_\varepsilon))\big] + \\ &+ \big[\mathbb{E}(f(F)\mathbf{\Theta}_\varepsilon) - \mathbb{E}(f(F_\delta)\mathbf{\Theta}_\varepsilon)\big] \\ =&: I(\delta,\varepsilon) + J(\delta,\varepsilon).\end{aligned}$$

We have

$$|I(\delta,\varepsilon)| \le 2\|f\|_\infty \mathbb{E}(|1-\mathbf{\Theta}_\varepsilon|) \le 2\|f\|_\infty \mathbb{P}(\det \sigma_F < \varepsilon).$$

We use (2.28) in order to obtain

$$\begin{aligned}|J(\delta,\varepsilon)| &= |\mathbb{E}_{\mathbf{\Theta}_\varepsilon}(f(F)) - \mathbb{E}_{\mathbf{\Theta}_\varepsilon}(f(F_\delta))| \\ &= \left|\int f(x)(p_{F,\mathbf{\Theta}_\varepsilon}(x) - p_{F_\delta,\mathbf{\Theta}_\varepsilon}(x))dx\right| \le \|f\|_\infty \int |p_{F,\mathbf{\Theta}_\varepsilon}(x) - p_{F_\delta,\mathbf{\Theta}_\varepsilon}(x)|\, dx \\ &\le \frac{C}{\varepsilon^p}\|f\|_\infty (1 + \|F\|_{3,p} + \|LF\|_{1,p}))^a \times \sqrt{\delta} \times \int \frac{1}{(1+|y|)^{2d}}dy\end{aligned}$$

and (2.23) follows. We write now

$$\begin{aligned}|J(\delta,\varepsilon)| &= \left|\int f(x)(p_{F,\mathbf{\Theta}_\varepsilon}(x) - p_{F_\delta,\mathbf{\Theta}_\varepsilon}(x))dx\right| \\ &\le \|p_{F,\mathbf{\Theta}_\varepsilon} - p_{F_\delta,\mathbf{\Theta}_\varepsilon}\|_\infty \|f\|_1.\end{aligned}$$

Using (2.26) we obtain (2.24). □

In the one-dimensional case, the requests in Lemma 2.5 can be weakened: only the Malliavin derivatives up to order 2 are required. And moreover, precise estimates can be given. In fact, one has:

Lemma 2.6. *Let $d = 1$. There exists a universal constant $C > 0$ such that for every $\varepsilon > 0, \delta > 0$ and every $F \in \mathcal{S}$ one has*

$$|\mathbb{E}(f(F)) - \mathbb{E}(f * \gamma_\delta(F))| \le C\|f\|_\infty \Big(\mathbb{P}(\sigma_F < \varepsilon) + \frac{\sqrt{\delta}}{\varepsilon^3}(\|F\|^2_{1,2,2} + \|LF\|)\Big) \tag{2.29}$$

for every bounded and measurable function $f : \mathbb{R} \to \mathbb{R}$.

Proof. The statement can be proved in several ways, we propose here a short proof that makes use of integration by parts formulas and weights that are developed in next Section 3.1.

We use notations as in the proof of Lemma 2.5. So, we take f with $\|f\|_\infty < \infty$ and we write

$$\begin{aligned}\mathbb{E}(f(F)) - \mathbb{E}(f * \gamma_\delta(F)) =& \big[\mathbb{E}(f(F)(1-\mathbf{\Theta}_\varepsilon)) - \mathbb{E}(f * \gamma_\delta(F)(1-\mathbf{\Theta}_\varepsilon))\big] + \\ &+ \big[\mathbb{E}(f(F)\mathbf{\Theta}_\varepsilon) - \mathbb{E}(f * \gamma_\delta(F)\mathbf{\Theta}_\varepsilon)\big] \\ =&: I(\delta,\varepsilon) + J(\delta,\varepsilon).\end{aligned}$$

The term $I(\delta,\varepsilon)$ is handled as before, so

$$|I(\delta,\varepsilon)| \le 2\left\|f\right\|_\infty \mathbb{E}(|1-\mathbf{\Theta}_\varepsilon|) \le 2\left\|f\right\|_\infty \mathbb{P}(\det\sigma_F<\varepsilon).$$

As for $J(\delta,\varepsilon)$, we set $\Psi_\delta(x)=\int_0^x(f-f*\gamma_\delta)(y)dy$, with the convention $\int_a^b(\cdot)=-\int_b^a(\cdot)$ when $a>b$. We also note that

$$\Psi_\delta(x)=\int_{\mathbb{R}}\gamma_\delta(z)\Big(\int_0^x(f(y)-f(y-z))dy\Big)dz=\int_{\mathbb{R}}\gamma_\delta(z)\Big(\int_{x-z}^x f(y)dy-\int_{-z}^0 f(y)dy\Big)dz,$$

so that

$$\|\Psi_\delta\|_\infty\le 2\|f\|_\infty\int_{\mathbb{R}}|z|\gamma_\delta(z)dz\le 2\sqrt{\delta}\left\|f\right\|_\infty.$$

Now, by using the (localized) integration by parts formula in Proposition 3.5, we have

$$J(\delta,\varepsilon)=\mathbb{E}_{\mathbf{\Theta}_\varepsilon}(\Psi_\delta'(F))=\mathbb{E}_{\mathbf{\Theta}_\varepsilon}(\Psi_\delta(F)H_{\mathbf{\Theta}_\varepsilon}(F,1)),$$

where

$$H_{\mathbf{\Theta}_\varepsilon}(F,1)=\gamma_F LF-\langle D\gamma_F,DF\rangle_J-\gamma_F\langle D(\ln\mathbf{\Theta}_\varepsilon),DF\rangle_J, \tag{2.30}$$

in which $\gamma_F=\sigma_F^{-1}$. Therefore, we can write

$$|J(\delta,\varepsilon)|\le\|\Psi_\delta\|_\infty\mathbb{E}_{\mathbf{\Theta}_\varepsilon}(|H_{\mathbf{\Theta}_\varepsilon}(F,1)|)\le 2\sqrt{\delta}\left\|f\right\|_\infty\mathbb{E}_{\mathbf{\Theta}_\varepsilon}(|H_{\mathbf{\Theta}_\varepsilon}(F,1)|).$$

The estimate of the last expectation is developed for general values of d and general localizations $\mathbf{\Theta}$ in Section 3.1. But for $d=1$ and $\mathbf{\Theta}=\mathbf{\Theta}_\varepsilon$, very precise estimates can be given. In fact, since $D\gamma_F=-\sigma_F^{-2}D\sigma_F$ and since on the set $\{\mathbf{\Theta}_\varepsilon\neq 0\}$ one has $\sigma_F>2/\varepsilon$, (2.30) gives

$$\begin{aligned}|H_{\mathbf{\Theta}_\varepsilon}(F,1)| &\le \frac{C}{\varepsilon^2}\big(|LF|+|D\sigma_F|\,|DF|+|D(\ln\mathbf{\Theta}_\varepsilon)|\,|DF|\big)\\ &\le \frac{C}{\varepsilon^2}\big(|LF|+|F|_{1,2}^2+|D(\ln\mathbf{\Theta}_\varepsilon)|\,|F|_{1,1}\big).\end{aligned}$$

Now, by using (2.13) we have

$$\mathbf{\Theta}_\varepsilon|D(\ln\mathbf{\Theta}_\varepsilon)|=\phi_\varepsilon(\sigma_F)|(\ln\phi_\varepsilon)'(\sigma_F)D\sigma_F|\le\frac{C}{\varepsilon}|F|_{1,2}.$$

Therefore,

$$\mathbb{E}_{\mathbf{\Theta}_\varepsilon}(|H_{\mathbf{\Theta}_\varepsilon}(F,1)|)\le\frac{C}{\varepsilon^3}\mathbb{E}\big(|LF|+|F|_{1,2,2}^2\big).\le\frac{C}{\varepsilon^3}(\|LF\|+\|F\|_{2,2}^2)$$

and the statement holds. We finally note that in dimension $d>1$, a similar reasoning would bring to take primitives of $f-f*\gamma_\delta$ in all the directions of the space and in the end one has to do d integration by parts in order to remove the derivatives, and this needs Malliavin derivatives for F up to order $d+1$. The use of the Riesz transform can actually overcome this difficulty. □

2.2 Distances and basic estimate

In this section we discuss the convergence in the total variation distance defined by

$$d_{TV}(F,G)=\sup\{|\mathbb{E}(f(F))-\mathbb{E}(f(G))|:\left\|f\right\|_\infty\le 1\}$$

The convergence in this distance is related to the convergence of the densities of the laws: given a sequence of random variables $F_n \sim p_n(x)dx$ and $F \sim p(x)dx$ then $d_{TV}(F_n, F) \to 0$ is equivalent to

$$\lim_n \int |p(x) - p_n(x)|\, dx = 0.$$

We also consider the Fortet-Mourier distance defined by

$$d_{FM}(F,G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \|f\|_\infty + \|\nabla f\|_\infty \le 1\}$$

and the Wasserstein distance

$$d_W(F,G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \|\nabla f\|_\infty \le 1\}.$$

The convergence in d_W is equivalent to the convergence in distribution plus the convergence of the first order moments. Clearly $d_{FM}(F,G) \le d_W(F,G)$ so convergence in distribution plus the convergence of the first order moments implies convergence in d_{FM}. One also has $d_{FM}(F,G) \le d_{TV}(F,G)$. The aim of this section is to prove a kind of converse type inequality.

We will be interested in a larger class of distances that we define now. For $f \in C_b^m(\mathbb{R}^d)$ we denote

$$\|f\|_{m,\infty} = \|f\|_\infty + \sum_{1 \le |\alpha| \le m} \|\partial_\alpha f\|_\infty .$$

Then we define

$$d_m(F,G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \|f\|_{m,\infty} \le 1\}. \tag{2.31}$$

So

$$d_{FM} = d_1 \qquad \text{and} \qquad d_{TV} = d_0.$$

Our basic estimate is the following. For $F \in \mathcal{S}^d$ we denote

$$A_l(F) := \|F\|_{3,l} + \|LF\|_{1,l} \tag{2.32}$$

Theorem 2.7. *Let $k \in \mathbb{N}$. There exist universal constants $C, l, b \in \mathcal{C}(d,k)$ such that for every $F, G \in \mathcal{S}^d$ with $A_p(F), A_p(G) < \infty, \forall p \in \mathbb{N}$, and every $\varepsilon > 0$ one has*

$$\begin{aligned} d_0(F,G) \quad &\le \quad \frac{C}{\varepsilon^b}(1 + A_l(F) + A_l(G))^b d_k^{\frac{1}{k+1}}(F,G) + \\ &\quad + C\mathbb{P}(\det \sigma_F < \varepsilon) + C\mathbb{P}(\det \sigma_G < \varepsilon). \end{aligned} \tag{2.33}$$

Proof. Let $\delta > 0$ and let $f \in C(\mathbb{R}^d)$ with $\|f\|_\infty \le 1$. Since $\|f * \gamma_\delta\|_{k,\infty} \le C\delta^{-k/2}$ we have

$$|\mathbb{E}(f * \gamma_\delta(F)) - \mathbb{E}(f * \gamma_\delta(G))| \le C\delta^{-k/2} d_k(F,G).$$

Then using (2.23)

$$\begin{aligned} |\mathbb{E}(f(F)) - \mathbb{E}(f(G))| \quad &\le \quad C\delta^{-k/2} d_k(F,G) + C\mathbb{P}(\det \sigma_F < \varepsilon) + C\mathbb{P}(\det \sigma_G < \varepsilon) + \\ &\quad + \frac{C\delta^{1/2}}{\varepsilon^p}(1 + A_l(F) + A_l(G))^a. \end{aligned}$$

We optimize over δ: we take

$$\delta^{(k+1)/2} = d_k(F,G)\Big(\frac{1}{\varepsilon^p}(1 + A_l(F) + A_l(G))^a\Big)^{-1}.$$

We insert this in the previous inequality and we obtain (2.33). □

As a consequence, we have

Corollary 2.8. *Let $F, G \in \mathcal{S}^d$ and suppose that there exists $\alpha > 0$ such that for every $\varepsilon > 0$ one has*

$$\mathbb{P}(\det \sigma_F \leq \varepsilon) + \mathbb{P}(\det \sigma_G \leq \varepsilon) \leq \varepsilon^{\alpha}. \tag{2.34}$$

Then under the hypotheses of Theorem 2.7 one has

$$d_0(F, G) \quad \leq \quad C(1 + A_l(F) + A_l(G))^{\frac{\alpha b}{\alpha + b}}\, d_k^{\frac{\alpha}{\alpha+b} \times \frac{1}{k+1}}(F, G).$$

Proof. It suffices to apply Theorem 2.7 and then to optimize w.r.t. $\varepsilon > 0$. □

Remark 2.9. When $d = 1$, in the proof of Theorem 2.7 and Corollary 2.8 we can use the precise estimate in Lemma 2.6. Therefore, in the one-dimensional case we set

$$A(F) = \|F\|_{2,2} + \|LF\|$$

and Theorem 2.7 becomes: for $k \in \mathbb{N}$, there exists a universal constant $C > 0$ such that for every $F, G \in \mathcal{S}$ and every $\varepsilon > 0$ one has

$$d_0(F, G) \leq \frac{C}{\varepsilon^3}(1 + A(F) + A(G))^2 d_k^{\frac{1}{k+1}}(F, G) + C\mathbb{P}(\sigma_F < \varepsilon) + C\mathbb{P}(\sigma_G < \varepsilon). \tag{2.35}$$

Similarly, Corollary 2.8 can be rephrased as follows: under (2.34), for $k \in \mathbb{N}$, there exists a universal constant $C > 0$ such that for every $F, G \in \mathcal{S}$ one has

$$d_0(F, G) \leq C(1 + A(F) + A(G))^{\frac{2\alpha}{\alpha+3}}\, d_k^{\frac{\alpha}{\alpha+3} \times \frac{1}{k+1}}(F, G). \tag{2.36}$$

2.3 Convergence results

In the previous sections we considered a functional $F \in \mathcal{S}^d$ with $\mathcal{S}$ associated to a certain random variable $V = (V_1, ..., V_J)$. So $F = f(V)$. But the estimates that we have obtained are estimates of the law and so it is not necessary that the random variables at hand are functionals of the same V. We may have $F = f(V)$ and $\overline{F} = \overline{f}(\overline{V})$ with $\overline{V} = (\overline{V}_1, ..., \overline{V}_{\overline{J}})$. Having this in mind, for a fixed random variable $V = (V_1, ..., V_J)$ we denote by $\mathcal{S}(V) = \{F = f(V) : f \in C_b^\infty(\mathbb{R}^J)\}$ the space of the simple functionals associated to V. We denote by $\sigma_F(V)$ the Malliavin covariance matrix and

$$A_p(V, F) := \|F\|_{3,p} + \|LF\|_{1,p}\,. \tag{2.37}$$

Here the norms $\|F\|_{q,l}$ and the operator LF are defined as in (2.7) and (2.4) with respect to V.

In the following we will work with a sequence $(F_n)_{n\in\mathbb{N}}$ of d dimensional functionals $F_n = (F_{n,1}, ..., F_{n,d})$. For each n, $F_{n,i} \in \mathcal{S}(V_{(n)})$, $i = 1, ..., d$ for some random variables $V_{(n)} = (V_{(n),1}, ..., V_{(n),J_n})$. We will use the following two assumptions. First, we consider a regularity assumption:

$$\overline{F}_p := \sup_n A_p(V_{(n)}, F_n) < \infty, \qquad \forall p \geq 1. \tag{2.38}$$

The second one is a (very weak) non degeneracy hypothesis:

$$\lim_{\varepsilon \to 0} \eta(\varepsilon) = 0 \qquad \text{with} \qquad \eta(\varepsilon) := \limsup_n \mathbb{P}(\det \sigma_{F_n}(V_{(n)}) \leq \varepsilon) \tag{2.39}$$

One has

Lemma 2.10. *Let $\overline{F}_p$ be as in (2.38). If $\overline{F}_1 < \infty$ then (2.39) is equivalent to*

$$\lim_{\varepsilon \to 0} \overline{\eta}(\varepsilon) = 0 \qquad \textit{with} \qquad \overline{\eta}(\varepsilon) := \limsup_n \mathbb{P}(\lambda(F_n) \leq \varepsilon) \tag{2.40}$$

where $\lambda(F_n)$ is the smaller eigenvalue of $\sigma_{F_n}(V_{(n)})$.

Proof. The statement is trivial for $d=1$, so we consider the case $d>1$. Since $\det\sigma_{F_n}(V_{(n)})\geq\lambda^d(F_n)$ we have $\mathbb{P}(\det\sigma_{F_n}(V_{(n)})\leq\varepsilon)\leq\mathbb{P}(\lambda(F_n)\leq\varepsilon^{1/d})$ so that $\eta(\varepsilon)\leq\overline{\eta}(\varepsilon^{1/d})$. If $\gamma(F_n)$ is the largest eigenvalue of $\sigma_{F_n}(V_{(n)})$ then $\det\sigma_{F_n}(V_{(n)})\leq\lambda(F_n)\gamma^{d-1}(F_n)$ so that

$$\begin{aligned}\mathbb{P}(\lambda(F_n)\leq\varepsilon)&\leq\mathbb{P}(\det\sigma_{F_n}(V_{(n)})\leq\varepsilon\gamma^{d-1}(F_n))\\&\leq\mathbb{P}(\gamma^{d-1}(F_n)\geq\varepsilon^{-1/2})+\mathbb{P}(\det\sigma_{F_n}(V_{(n)})\leq\varepsilon^{1/2})\end{aligned}$$

But $\gamma(F_n)\leq|DF_n|^2$, so

$$\mathbb{P}(\gamma^{d-1}(F_n)\geq\varepsilon^{-1/2})\leq\varepsilon^{\frac{1}{4(d-1)}}\,\mathbb{E}(|DF_n|)\leq\varepsilon^{\frac{1}{4(d-1)}}\,\overline{F}_1.$$

We conclude that $\overline{\eta}(\varepsilon)\leq\varepsilon^{\frac{1}{4(d-1)}}\overline{F}_1+\eta(\varepsilon^{1/2})$. □

Theorem 2.11. *We consider a sequence of functionals $F_n=(F_{n,1},...,F_{n,d})\in\mathcal{S}^d(V_{(n)})$ and we assume that (2.38) and (2.39) hold. Suppose also that $\lim_n F_n=F$ in distribution and $\lim_n\mathbb{E}(F_n)=\mathbb{E}(F)$. Then*

$$\lim_n d_{TV}(F,F_n)=0. \tag{2.41}$$

In particular if the laws of F and F_n are absolutely continuous with density p_F and p_{F_n} respectively, then

$$\lim_n\int|p_F(x)-p_{F_n}(x)|\,dx=0.$$

Proof. Using (2.33) with $k=1$ we may find some $C,l,b\in\mathcal{C}(d,k)$ such that for every $n,m\in\mathbb{N}$

$$d_0(F_n,F_m)\leq\frac{C}{\varepsilon^b}(1+\overline{F}_l)^b\,d_1^{1/2}(F_n,F_m)+C\mathbb{P}(\det\sigma_{F_n}(V_{(n)})<\varepsilon)+C\mathbb{P}(\det\sigma_{F_m}(V_{(m)})<\varepsilon).$$

Since $\lim_n F_n=F$ in law, one has that $\limsup_{n,m\to\infty}d_1(F_n,F_m)=0$, so that

$$\limsup_{n,m\to\infty}d_0(F_n,F_m)\leq C\eta(\varepsilon).$$

This is true for every $\varepsilon>0$. So using (2.39) we obtain $\limsup_{n,m\to\infty}d_0(F_n,F_m)=0$. □

As a consequence, we can state a similar result under a stronger non degeneracy hypothesis:

Corollary 2.12. *Let $F_n=(F_{n,1},...,F_{n,d})\in\mathcal{S}^d(V_{(n)})$ be a sequence of functionals such that (2.38) holds and such that there exists $\alpha>0$ with*

$$\sup_n\mathbb{P}(\det\sigma_{F_n}(V_{(n)})\leq\varepsilon)\leq\varepsilon^\alpha, \tag{2.42}$$

for every $\epsilon>0$. Suppose also that $\lim_n F_n=F$ in distribution and $\lim_n\mathbb{E}(F_n)=\mathbb{E}(F)$. Then

$$\lim_n d_{TV}(F,F_n)=0. \tag{2.43}$$

Proof. Since (2.42) implies (2.39), the statement follows by applying Theorem 2.11. We also note that, by applying Corollary 2.8 with $k=1$, we can also state that

$$d_0(F_n,F_m)\leq(1+\overline{F}_l)^{\frac{\alpha b}{\alpha+b}}\,d_1^{\frac{\alpha}{\alpha+b}\times 1/2}(F_n,F_m). \tag{2.44}$$

□

Note that (2.44) gives an estimate of the total variation distance in terms of the Fortet-Mourier distance starting from the non-degeneracy condition (2.43). A similar non-degeneracy request has been already discussed in [17], where the convergence in total variation is studied for sequences in a finite sum of Wiener chaoses that converge in distribution. We also note that (2.44) can be generalized to functionals that are not necessarily simple ones by passing to the limit (for this, it is crucial that constants are "universal", that is independent of the functionals).

Remark 2.13. In the one-dimensional case, in the proof of Theorem 2.11 we can use (2.35) instead of (2.33). This gives

$$\begin{aligned} d_0(F_n,F_m) \leq \frac{C}{\varepsilon^3}(1+A(F_n)+A(F_m))^2\, d_1^{1/2}(F_n,F_m) + C\mathbb{P}(\det\sigma_{F_n}(V_{(n)})<\varepsilon) \\ + C\mathbb{P}(\det\sigma_{F_m}(V_{(m)})<\varepsilon), \end{aligned}$$

where $A(F)=\|F\|_{2,2}+\|LF\|$. Therefore, for $d=1$ Theorem 2.11 holds with (2.38) replaced by the following weaker assumption:

$$\overline{F} = \sup_n \big(\|F_n\|_{2,2}+\|LF_n\|\big) < \infty. \tag{2.45}$$

Similarly, in dimension 1 Corollary 2.12 continues to hold if we ask (2.45) instead of (2.38).

2.4 Functionals on the Wiener space

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space where a Brownian motion $W=(W^1,...,W^N)$ is defined. We briefly recall the main notations in Malliavin calculus, for which we refer to Nualart [18]. We denote by $\mathbb{D}^{m,p}$ the space of the random variables which are m times differentiable in Malliavin sense in L^p and for a multi-index $\alpha=(\alpha_1,\dots,\alpha_k)\in\{1,\dots,N\}^k$, $k\leq m$, we denote by $D^\alpha F$ the Malliavin derivative of F corresponding to the multi-index α. Moreover we define

$$|D^{(k)}F|^2 = \sum_{|\alpha|=k}\int_{[0,1)^k}\big|D^\alpha_{s_1,\dots,s_k}F\big|^2\,ds_1,\dots ds_k \quad\text{and}\quad |F|^2_m = |F|^2+\sum_{k=1}^m|D^{(k)}F|^2. \tag{2.46}$$

So, $\mathbb{D}^{m,p}$ is the closure of the space of the simple functionals with respect to the Malliavin Sobolev norm

$$\|F\|^p_{m,p} = \mathbb{E}\big(|F|^p_m\big) \tag{2.47}$$

We set $\mathbb{D}^{m,\infty}=\cap_{p\geq1}\mathbb{D}^{m,p}$ and $\mathbb{D}^\infty=\cap_{m\geq1}\mathbb{D}^{n,\infty}$. Moreover, for $F\in(\mathbb{D}^{1,2})^d$, we let σ_F denote the Malliavin covariance matrix associated to F :

$$\sigma^{i,j}_F = \langle DF^i,DF^j\rangle = \sum_{k=1}^N\int_0^1 D^k_sF^iD^k_sF^j ds,\quad i,j=1,\dots,d.$$

If σ_F is invertible, we denote through γ_F the inverse matrix. Finally, as usual, the notation L will be used for the Ornstein-Uhlenbeck operator and we recall that the Meyer inequality asserts that $\|LF\|_{m,p}\leq C_{m,p}\|F\|_{m+2,p}$, for $F\in(\mathbb{D}^{m+2,\infty})^d$.

Our aim is to rephrase the results from the previous sections in the framework of the Wiener space considered here. We introduce first the localization random variable $\mathbf{\Theta}$. We consider some random variables Θ_i and some numbers $a_i>0, i=1,...,l+l'$ and we define

$$\mathbf{\Theta} = \prod_{i=1}^l\psi_{a_i}(\Theta_i)\times\prod_{i=l+1}^{l+l'}\phi_{a_i}(\Theta_i). \tag{2.48}$$

with ψ_{a_i}, ϕ_{a_i} defined in (2.12). Following what developed in Section 2.1, we define

$$d\mathbb{P}_{\mathbf{\Theta}} = \mathbf{\Theta} d\mathbb{P}$$

and

$$\|F\|_{p,\mathbf{\Theta}}^p = \mathbb{E}_{\mathbf{\Theta}}(|F|^p) \quad \text{and} \quad \|F\|_{l,p,\mathbf{\Theta}}^p = \mathbb{E}_{\mathbf{\Theta}}(|F|_l^p). \tag{2.49}$$

In the case $\mathbf{\Theta} = 1$ we have $\|F\|_{p,\mathbf{\Theta}} = \|F\|_p$ and $\|F\|_{l,p,\mathbf{\Theta}} = \|F\|_{l,p}$. Moreover, given some $q \in \mathbb{N}, p \geq 1$, we denote

$$\mathrm{m}_{q,p}(\mathbf{\Theta}) := 1 \vee \|\ln \mathbf{\Theta}\|_{q,p,\mathbf{\Theta}}$$

and

$$\mathrm{U}_{q,p,\mathbf{\Theta}}(F) := \max\{1, \mathbb{E}_{\mathbf{\Theta}}((\det \sigma_F)^{-p})(\|F\|_{q+2,p,\mathbf{\Theta}} + \|LF\|_{q,p,\mathbf{\Theta}})\}. \tag{2.50}$$

In the case $\mathbf{\Theta} = 1$ we have $\mathrm{m}_{q,p}(\mathbf{\Theta}) = 1$ and

$$\begin{aligned} \mathrm{U}_{q,p}(F) &:= \max\{1, \mathbb{E}((\det \sigma_F)^{-p})(\|F\|_{q+2,p} + \|LF\|_{q,p})\} \\ &\leq C \max\{1, \mathbb{E}((\det \sigma_F)^{-p}) \|F\|_{q+2,p}\} \end{aligned} \tag{2.51}$$

the last inequality being a consequence of Meyer's inequality.

We rephrase now Theorem 2.1:

Theorem 2.14. *Let $q \in \mathbb{N}_*$. We consider the localization random variable $\mathbf{\Theta}$ defined in (2.48) and we assume that for every $p \in \mathbb{N}$ one has $\|\Theta_i\|_{q+2,p} < \infty, i = 1, ..., l + l'$. In particular $\mathrm{m}_{q+2,p}(\mathbf{\Theta}) < \infty$.*

***A**. Let $F \in (\mathbb{D}^{q+2,\infty})^d$ be such that $\mathrm{U}_{q,p,\mathbf{\Theta}}(F) < \infty$. Then under $\mathbb{P}_{\mathbf{\Theta}}$ the law of F is absolutely continuous with respect to the Lebesgue measure. We denote by $p_{F,\mathbf{\Theta}}$ its density and we have $p_{F,\mathbf{\Theta}} \in C^{q-1}(\mathbb{R}^d)$. Moreover there exist $C, a, b, p \in \mathcal{C}(q, d)$ such that for every $y \in \mathbb{R}^d$ and every multi index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., d\}^k, k \in \{0, ..., q\}$ one has*

$$|\partial_\alpha p_{F,\mathbf{\Theta}}(y)| \leq C \mathrm{U}_{q,p,,\mathbf{\Theta}}^a(F) \times \mathrm{m}_{q+2,p}^a(\mathbf{\Theta}) \times \big(\mathbb{P}(|F-y| < 2)\big)^b. \tag{2.52}$$

***B**. Let $F, G \in (\mathbb{D}^{q+2,\infty})^d$ be such that $\mathrm{U}_{q+1,p,\mathbf{\Theta}}(F), \mathrm{U}_{q+1,p,\mathbf{\Theta}}(G) < \infty$ for every $p \in \mathbb{N}$ and let $p_{F,\mathbf{\Theta}}$ and $p_{G,\mathbf{\Theta}}$ be the densities of the laws of F respectively of G under $\mathbb{P}_{\mathbf{\Theta}}$. Then there exist $C, a, b, p \in \mathcal{C}(q, d)$ such that for every $y \in \mathbb{R}^d$ and every multi index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, 2, ...\}^k$, $0 \leq k \leq q$ one has*

$$\begin{aligned} |\partial_\alpha p_{F,\mathbf{\Theta}}(y) - \partial_\alpha p_{G,\mathbf{\Theta}}(y)| \leq \quad & C \mathrm{U}_{q+1,p,\mathbf{\Theta}}^a(F) \times \mathrm{U}_{q+1,p,\mathbf{\Theta}}^a(G) \times \mathrm{m}_{q+2,p}^a(\mathbf{\Theta}) \times \\ & \times \big(\mathbb{P}(|F-y| < 2) + \mathbb{P}(|G-y| < 2)\big)^b \times \\ & \times (\|F-G\|_{q+2,p,\mathbf{\Theta}} + \|LF - LG\|_{q,p,\mathbf{\Theta}}) \,. \end{aligned} \tag{2.53}$$

Remark 2.15. The arguments used in Remark 2.4 can be applied here: the second factor in the estimates (2.52) and (2.53) can be replaced, as $|y| > 4$, with the queue of the law of F and G. Also, by using the Markov inequality, such factors can be over estimated by means of any power of $(1 + |y|)^{-1}$, for every $y \in \mathbb{R}^d$.

Proof of Theorem 2.14. One may prove Theorem 2.14 just by repeating exactly the same reasoning as in the proof of Theorem 2.1: all the arguments are based on the properties of the norms from the finite dimensional calculus and these properties are preserved in the infinite dimensional case. However we give here a different proof: we obtain Theorem 2.14 from Theorem 2.1 by using a convergence argument.

We fix $n \geq 1$. For $k = 1, \ldots, 2^n$ we denote

$$\Delta_n^k = (\Delta_n^{k,1}, ..., \Delta_n^{k,N}), \quad \text{where} \quad \Delta_n^{k,i} = W^i_{\frac{k}{2^n}} - W^i_{\frac{k-1}{2^n}}, \; i = 1, \ldots, N.$$

So, by taking Δ_n^k, $k = 1, \ldots, 2^n$, as the underlying noise $V_1, \ldots, V_J$ and by taking the weights $\pi_{k,i} = 2^{-n/2}$, $k = 1, \ldots, 2^n$ and $i = 1, \ldots, N$, it is easy to see that the finite dimensional Malliavin calculus in Section 2.1 and the standard Malliavin calculus coincide for simple functionals (see e.g. [1] for details). So, we set $\mathcal{S}_n = \{F = \phi(\Delta_n^1, \ldots, \Delta_n^{2^n}) : \phi \in C_p^\infty(\mathbb{R}^{N2^n})\}$ and we take $F_n, G_n, \Theta_{n,i} \in \mathcal{S}_n, n \in \mathbb{N}, i = 1, ..., l + l'$ which approximate $F, G, \Theta_i \in \mathbb{D}^{q+2,\infty}, i = 1, ..., l + l'$.We use Theorem 2.1 for them and then we pass to the limit in order to obtain the conclusion in Theorem 2.14. The fact that the constants which appear in Theorem 2.1 belong to $\mathcal{C}(q, d)$, so do not depend on $n \in \mathbb{N}$, plays here a crucial role. □

We give now a regularity property which is an easy consequence of the above theorem.

Theorem 2.16. ***A***. *Let* $F \in \mathbb{D}^{2,p}, p > d$ *such that* $\mathbb{P}(\det \sigma_F > 0) > 0$. *Then, conditionally to* $\{\det \sigma_F > 0\}$, *the law of* F *is absolutely continuous with respect to the Lebesgue measure and the density is lower semi-continuous.*

B. *In particular the law of* F *is locally lower bounded by the Lebesgue measure* λ *in the following sense: there exist* $\delta > 0$ *and an open set* $D \subset \mathbb{R}^d$ *such that for every Borel set* A

$$\mathbb{P}(F \in A) \geq \delta \lambda(A \cap D).$$

Remark 2.17. . The celebrated theorem of Bouleau and Hirsch [10] says that if $F \in \mathbb{D}^{1,2}$ then, conditionally to $\{\det \sigma_F > 0\}$, the law of F is absolutely continuous. So it requires much less regularity than us. But the new fact is that the conditional density is lower semi-continuous and in particular is locally lower bounded by the Lebesgue measure. This last property turns out to be especially interesting - see the joint paper [4].

Proof of Theorem 2.16. For $\varepsilon > 0$ we consider the localization function ψ_ε defined in (2.12) and we denote $\Theta_\varepsilon = \psi_\varepsilon(\det \sigma_F)$. By Theorem 2.14 we know that under $\mathbb{P}_{\Theta_\varepsilon}$ the law of F is absolutely continuous and has a continuous density p_{Θ_ε}. Let A be a Borel set with $\lambda(A) = 0$ where λ is the Lebesgue measure. Since $\Theta_\varepsilon \uparrow \Theta := 1_{\{\det \sigma_F > 0\}}$ we have

$$\begin{aligned} \mathbb{P}_\Theta(F \in A) = \mathbb{P}_{\{\det \sigma_F > 0\}}(F \in A) &= \frac{1}{\mathbb{P}(\det \sigma_F > 0)} \mathbb{E}(1_{\{F \in A\}} \times \Theta) \\ &= \frac{1}{\mathbb{P}(\det \sigma_F > 0)} \lim_{\varepsilon \to 0} \mathbb{E}(1_{\{F \in A\}} \times \Theta_\varepsilon) = 0. \end{aligned}$$

So we may find p_Θ such that

$$\mathbb{E}(f(F)\Theta) = \int f(x) p_\Theta(x) dx.$$

For $f \geq 0$ we have

$$\int f(x) p_{\Theta_\varepsilon}(x) dx = \mathbb{E}(f(F)\Theta_\varepsilon) \leq \mathbb{E}(f(F)\Theta) = \int f(x) p_\Theta(x) dx,$$

so that $p_\Theta \geq p_{\Theta_\varepsilon}$ a.e. This implies that $p_\Theta \geq \sup_{\varepsilon > 0} p_{\Theta_\varepsilon}$. We claim that

$$p_\Theta = \sup_{\varepsilon > 0} p_{\Theta_\varepsilon}$$

which gives that p_Θ is lower semi-continuous. In fact, set $A = \{x : p_\Theta(x) > \sup_{\varepsilon > 0} p_{\Theta_\varepsilon}(x)\}$. If $\lambda(A) > 0$ then we may find $\delta > 0$ such that $\lambda(A_\delta) > 0$ with $A_\delta = \{x : p_\Theta(x) > \delta + \sup_{\varepsilon > 0} p_{\Theta_\varepsilon}(x)\}$. Then

$$\begin{aligned} \int_{A_\delta} p_\Theta(x) dx = \mathbb{P}_{\{\det \sigma_F > 0\}}(F \in A_\delta) &= \frac{1}{\mathbb{P}(\det \sigma_F > 0)} \lim_{\varepsilon \to 0} \mathbb{E}(1_{\{F \in A_\delta\}} \times \Theta_\varepsilon) \\ &= \lim_{\varepsilon \to 0} \int_{A_\delta} p_{\Theta_\varepsilon}(x) dx \leq \int_{A_\delta} (p_\Theta(x) - \delta) dx \end{aligned}$$

and this gives $\lambda(A_\delta)=0$.

The assertion **B** is immediate: since $p_\Theta=\sup_{\varepsilon>0}p_{\Theta_\varepsilon}$ is not identically null we may find $\varepsilon>0$ and $x_0\in\mathbb{R}^d$ such that $p_{\Theta_\varepsilon}(x_0)>0$. And since p_{Θ_ε} is a continuous function we may find $r,\delta>0$ such that $p_{\Theta_\varepsilon}(x)\geq\delta$ for $x\in B_r(x_0)$. It follows that

$$\begin{aligned}\mathbb{P}(F\in A)&\geq\mathbb{P}(\{F\in A\}\cap\{F\in B_r(x_0)\}\cap\{\sigma_F>0\})\\&=\mathbb{P}(\sigma_F>0)\int_{A\cap B_r(x_0)}p_\Theta(x)dx\\&\geq\mathbb{P}(\sigma_F>0)\int_{A\cap B_r(x_0)}p_{\Theta_\varepsilon}(x)dx\geq\delta\mathbb{P}(\sigma_F>0)\lambda(A\cap B_r(x_0)).\end{aligned}$$

□

We rephrase now other consequences of Theorem 2.14. We begin with the regularization Lemma 2.5. We recall that γ_δ is the centred Gaussian density with variance $\delta>0$.

Lemma 2.18. *There exist universal constants $C,p,a\in\mathcal{C}(d)$ such that for every $\varepsilon>0,\delta>0$ and every $F\in(\mathbb{D}^{3,\infty})^d$ one has*

$$|\mathbb{E}(f(F))-\mathbb{E}(f*\gamma_\delta(F))|\leq C\left\|f\right\|_\infty\left(\mathbb{P}(\sigma_F<\varepsilon)+\frac{\sqrt{\delta}}{\varepsilon^p}(1+\|F\|_{3,p})^a\right) \tag{2.54}$$

for every $f\in C_b(\mathbb{R}^d)$.

Proof. The proof is identical with the one of Lemma 2.5 so we skip it (an approximation procedure may also been used). We mention that due to Meyer's inequalities $\left\|LF\right\|_{1,p}$ does no more appear here. □

We consider now the distances d_m defined in (2.31) and we rewrite Theorem 2.7:

Theorem 2.19. *Let $k\in\mathbb{N}$. There exist universal constants $C,p,b\in\mathcal{C}(d,k)$ such that for every $F,G\in(\mathbb{D}^{3,\infty})^d$ and every $\varepsilon>0$ one has*

$$d_0(F,G)\leq\frac{C}{\varepsilon^b}(1+\|F\|_{3,p}+\|G\|_{3,p}\|)^b d_k^{\frac{1}{k+1}}(F,G)+C\mathbb{P}(\det\sigma_F<\varepsilon)+C\mathbb{P}(\det\sigma_G<\varepsilon). \tag{2.55}$$

Proof. The proof is identical with the one of Theorem 2.7 so we skip it. □

We give now the convergence results.

Theorem 2.20. *We consider a sequence of functionals $F_n=(F_{n,1},...,F_{n,d})\in(\mathbb{D}^{3,\infty})^d,n\in\mathbb{N}$ and we assume that*

$$\begin{aligned}&i)\qquad \sup\nolimits_n\|F_n\|_{3,p}<\infty,\qquad\forall p\geq1,\\&ii)\qquad \limsup_{\varepsilon\to0}\limsup_{n\to\infty}\mathbb{P}(\det\sigma_{F_n}<\varepsilon)=0.\end{aligned} \tag{2.56}$$

Suppose also that $\lim_n F_n=F$ in distribution and $\lim_n\mathbb{E}(F_n)=\mathbb{E}(F)$. Then

$$\lim_n d_{TV}(F,F_n)=0.$$

In particular if the laws of F and F_n are absolutely continuous with densities p_F and p_{F_n} then

$$\lim_n\int|p_F(x)-p_{F_n}(x)|\,dx=0.$$

Proof. The proof is identical with the one of Theorem 2.11 so we skip it. □

In the framework of Wiener functionals we are able to obtain one more result:

Corollary 2.21. *Let $F_n \in (\mathbb{D}^{3,\infty})^d, n \in \mathbb{N}$ such that $\sup_n \|F_n\|_{3,p} < \infty$ for every $p \in \mathbb{N}$. Consider also $F \in (\mathbb{D}^{2,2})^d$ such that $\det \sigma_F > 0$ almost surely. If $\lim_n F_n = F$ in L^2 then $\lim_n F_n = F$ in d_{TV}.*

Proof. We will prove that $\lim_n \langle DF_n^i, DF_n^j \rangle = \langle DF^i, DF^j \rangle$ in probability for every $i, j = 1, ..., d$. This implies $\lim_n \det \sigma_{F_n} = \det \sigma_F$ in probability so that $\limsup_n \mathbb{P}(\det \sigma_{F_n} < \varepsilon) \leq P(\det \sigma_F < 2\varepsilon)$. And since $\lim_{\varepsilon \to 0} P(\det \sigma_F > \varepsilon) = 1$, we obtain (2.56) $ii)$ and this enables us to conclude by applying Theorem 2.20.

We denote by $f_k, k \in \mathbb{N}$, respectively by $f_{k,n}, k \in \mathbb{N}$, the kernels of the chaos expansion of F, respectively of F_n. So we have

$$F = \sum_{k=0}^{\infty} I_k(f_k) \quad \text{and} \quad F_n = \sum_{k=0}^{\infty} I_k(f_{k,n})$$

where I_k denotes the multiple integral of order k. For $N \in \mathbb{N}$ we write $F = S_N + R_N$ with $S_N = \sum_{k=0}^{N} I_k(f_k)$ and $R_N = \sum_{k=N+1}^{N} I_k(f_k)$. With similar notations, we set $F_n = S_{N,n} + R_{N,n}$. We write

$$\left| \langle DF_n^i, DF_n^j \rangle - \langle DF^i, DF^j \rangle \right| \leq (|DF| + |DF_n|) \times |D(F - F_n)| .$$

Since the sequence $|DF_n|, n \in \mathbb{N}$ is bounded in L^1 our conclusion follows as soon as we check that $\lim_n |D(F - F_n)| = 0$ in probability. We fix $\varepsilon > 0$ and we write

$$\begin{aligned} \mathbb{P}(|D(F - F_n)| \geq \varepsilon) &\leq \mathbb{P}\Big(|D(S_N - S_{N,n})| \geq \frac{1}{2}\varepsilon \Big) + \mathbb{P}\Big(|D(R_N - R_{N,n})| \geq \frac{1}{2}\varepsilon \Big) \\ &=: I_{N,\varepsilon,n} + J_{N,\varepsilon,n}. \end{aligned}$$

Using Chebyshev's inequality

$$J_{N,\varepsilon,n} \leq \frac{4}{\varepsilon^2} \mathbb{E} |D(R_N - R_{N,n})|^2 \leq \frac{8}{\varepsilon^2} (\mathbb{E} |DR_N|^2 + \mathbb{E} |DR_{N,n}|^2).$$

Since $\mathbb{E} |DI_k(f_k)|^2 = k \times k! \|f_k\|^2_{L^2[0,T]^k}$ we obtain

$$\begin{aligned} \mathbb{E} |DR_N|^2 &= \sum_{k=N+1}^{\infty} k \times k! \|f_k\|^2_{L^2[0,1]^k} \leq \frac{1}{N+2} \sum_{k=N+1}^{\infty} (k+1) \times k \times k! \|f_k\|^2_{L^2[0,1]^k} \\ &\leq \frac{1}{N+2} \mathbb{E}(|D^2 F|^2) \leq \frac{1}{N+2} \|F\|^2_{2,2} \end{aligned}$$

and a similar inequality holds for $\mathbb{E} |DR_{N,n}|^2$. We conclude that

$$J_{N,\varepsilon,n} \leq \frac{8}{(N+2)\varepsilon^2} (\|F\|_{2,2} + \sup_n \|F_n\|_{2,2}).$$

Moreover, since $\lim_n \|F - F_n\|_2 = 0$ we have $\lim_n \|f_k - f_{k,n}\|^2_{L^2[0,T]^k} = 0$ for every $k \in \mathbb{N}$ and this implies $\lim_n \mathbb{E} |D(S_N - S_{N,n})|^2 = 0$. Finally $\limsup_n I_{N,\varepsilon,n} = 0$ for each fixed N and ε. So we obtain

$$\limsup_n \mathbb{P}(|D(F - F_n)| \geq \varepsilon) \leq \frac{8}{N^2 \varepsilon^2} (\|F\|_{2,2} + \sup_n \|F_n\|_{2,2}).$$

Since this is true for each N the above limit is null. □

Remark 2.22. As an immediate consequence of Corollary 2.21 one may obtain the following result. Let X_t be a diffusion process with coefficients in C_b^∞ and suppose that the weak Hörmander condition holds in $x = X_0$. Consider also the Euler scheme X_t^n of step $\frac{1}{n}$. Then for every $q \in \mathbb{N}$ one has $d_{-q}(\mu_{X_t}, \mu_{X_t^n}) \to 0$. This type of result has already been obtained in [7], [8] and in [12]: there, under more restrictive assumptions (uniform Hörmander condition) one obtains the above result and moreover, one gives a development in Taylor series of the error.

3 Proof of Theorem 2.1

This section is devoted to the proof of Theorem 2.1. We are in the framework defined in Section 2.1 and we use all the notations introduced there. In the following subsection we recall and develop some basic results concerning integration by parts formulas from [5].

3.1 Integration by parts formulae

By using standard integration by parts formulas, one gets the duality between δ and D and the standard computation rules (see [5], Proposition 1 and Lemma 1): under our assumption, for every $F \in \mathcal{S}^d$, $U \in \mathcal{P}$ and $\phi : \mathbb{R}^d \to \mathbb{R}$ smooth,

$$\mathbb{E}(\langle DF, U\rangle_J) = \mathbb{E}(F\delta(U)), \tag{3.1}$$

$$D\phi(F) = \sum_{r=1}^{d} \partial_r \phi(F) DF^r, \tag{3.2}$$

$$\delta(FU) = F\delta(U) - \langle DF, U\rangle_J\,, \tag{3.3}$$

$$L\phi(F) = \sum_{r=1}^{d} \partial_r \phi(F) LF^r - \sum_{r,r'=1}^{d} \partial_{r,r'}\phi(F)\langle DF^r, DF^{r'}\rangle_J, \tag{3.4}$$

with the convention that $d = 1$ in (3.1) and (3.3). Once the above equalities are done, the integration by parts formulas can be stated (see [5], Theorem 1 and 2):

Theorem 3.1. *Let $F \in \mathcal{S}^d$ be such that*

$$\mathbb{E}(|\det \gamma_F|^p) < \infty \quad \forall p \geq 1, \tag{3.5}$$

γ_F denoting the inverse of the Malliavin covariance matrix σ_F. Then, for every $G \in \mathcal{S}$ and for every smooth function $\phi \in C_b^\infty(\mathbb{R}^d)$ one has

$$\mathbb{E}\left(\partial_r\phi(F)G\right) = \mathbb{E}\left(\phi(F)H_r(F,G)\right), \quad r = 1, ..., d, \tag{3.6}$$

with

$$H_r(F,G) = \sum_{r'=1}^{d} \delta(G\gamma_F^{r',r} DF^{r'}) = \sum_{r'=1}^{d} \big(G\delta(\gamma_F^{r',r} DF^{r'}) - \gamma_F^{r',r}\langle DF^{r'}, DG\rangle_J\big). \tag{3.7}$$

Moreover, for every $q \in \mathbb{N}^*$ and multi-index $\beta = (\beta_1, \ldots, \beta_q) \in \{1, \ldots, d\}^q$ then

$$\mathbb{E}\left(\partial_\beta\phi(F)G\right) = \mathbb{E}\left(\phi(F)H^q_\beta(F,G)\right) \tag{3.8}$$

where the weights $H^q_\beta(F,G)$ are defined recursively by (3.7) if $q = 1$ and for $q > 1$,

$$H^q_\beta(F,G) = H_{\beta_1}\big(F, H^{q-1}_{(\beta_2,\ldots,\beta_q)}(F,G)\big). \tag{3.9}$$

3.2 Estimates of the weights

In this section we give estimates of the weights $H^q_\alpha(F,G)$ appearing in the integration by parts formulae of Theorem 3.1 using the norms introduced in (2.8). We first deal with useful estimates for the inverse of the Malliavin covariance matrix. For $F\in\mathcal{S}^d$, we set

$$m_F=\max\Big(1,\frac{1}{\det\sigma_F}\Big). \tag{3.10}$$

Proposition 3.2. ***A***. *If $F\in\mathcal{S}^d$ then $\forall l\in\mathbb{N}$ one has*

$$|\gamma_F|_l\le C_{l,d}{m_F}^{l+1}(1+|F|_{1,l+1}^{2d(l+1)}). \tag{3.11}$$

B. *If $F,\overline{F}\in\mathcal{S}^d$ then $\forall l\in\mathbb{N}$ one has*

$$|\gamma_F-\gamma_{\overline{F}}|_l\le C_{l,d}m_F^{l+1}\,m_{\overline{F}}^{l+1}(1+|F|_{1,l+1}+|\overline{F}|_{1,l+1})^{2d(l+3)}\left|F-\overline{F}\right|_{1,l+1}. \tag{3.12}$$

Proof. **A** is proved in [5], Proposition 2. As for **B**, we use the following estimates proved in [5] (see Lemma 2 and the proof of Proposition 2):

$$|\langle DF,DG\rangle_J|_l\le 2^l\sum_{l_1+l_2\le l}|F|_{1,l_1+1}|G|_{1,l_2+1}, \tag{3.13}$$

$$|F\times G|_l\le 2^l\sum_{l_1+l_2\le l}|F|_{l_1}|G|_{l_2}, \tag{3.14}$$

$$|(\det\sigma_F)^{-1}|_l\le C_{l_1}m_F^{l+1}(1+|F|_{1,l+1}^{2l_1d}), \tag{3.15}$$

$$|\det\sigma_F|_l\le C_l(1+|F|_{1,l+1}^{2l_1d}) \tag{3.16}$$

So, by (3.13), we have

$$\left|\sigma_F^{r,r'}-\sigma_{\overline{F}}^{r,r'}\right|_l\le C_{l,d}\left|F-\overline{F}\right|_{1,l+1}(|F|_{1,l+1}+\left|\overline{F}\right|_{1,l+1})$$

and then, by (3.14) and (3.16)

$$\begin{aligned}|\det\sigma_F-\det\sigma_{\overline{F}}|_l&\le C_{l,d}\left|F-\overline{F}\right|_{1,l+1}(|F|_{1,l+1}+\left|\overline{F}\right|_{1,l+1})^{2d-1}\\ \left|\widehat{\sigma}_F^{r,r'}-\widehat{\sigma}_{\overline{F}}^{r,r'}\right|_l&\le C_{l,d}\left|F-\overline{F}\right|_{1,l+1}(|F|_{1,l+1}+\left|\overline{F}\right|_{1,l+1})^{2d-3},\end{aligned} \tag{3.17}$$

in which $\widehat{\sigma}_F$ denotes the algebraic complement of σ_F. Then, by using also (3.15)

$$\begin{aligned}\left|(\det\sigma_F)^{-1}-(\det\sigma_{\overline{F}})^{-1}\right|_l\le&C_{l,d}\left|(\det\sigma_F)^{-1}\right|_l\left|(\det\sigma_{\overline{F}})^{-1}\right|_l\times\\ &\times|\det\sigma_F-\det\sigma_{\overline{F}}|_l\\ \le&C_{l,d}m_F^{l+1}m_{\overline{F}}^{l+1}|F-\overline{F}|_{1,l+1}\big(1+|F|_{1,l}+|\overline{F}|_{1,l}\big)^{2(l+1)d}.\end{aligned}$$

Since $\gamma^{r,r'}(F)=(\det\sigma_F)^{-1}\widehat{\sigma}^{r,r'}(F)$, (3.12) follows by using the above estimates. □

We define now

$$L_r^\gamma(F)=\sum_{r'=1}^d\delta(\gamma_F^{r',r}DF^{r'})=\sum_{r'=1}^d\big(\gamma_F^{r',r}LF^{r'}-\langle D\gamma_F^{r',r},DF^{r'}\rangle\big). \tag{3.18}$$

Using (3.12) one can easily check that for every $l\in\mathbb{N}$

$$|L_r^\gamma(F)|_l\le C_{l,d}m_F^{l+2}\big(1+|F|_{1,l+1}^{2d(l+2)}+|L(F)|_l^2\big). \tag{3.19}$$

And by using both (3.12) and (3.18), one immediately gets

$$|L_r^\gamma(F)-L_r^\gamma(\overline{F})|_l \leq C_{l,d}Q_l(F,\overline{F})(|F-\overline{F}|_{1,l+2}+|L(F-\overline{F}))|_l), \tag{3.20}$$

where

$$Q_l(F,\overline{F})=m_F^{l+2}m_{\overline{F}}^{l+2}\big(1+|F|_{1,l+2}^{2d(l+4)}+|L(F)|_l^2+\big|\overline{F}\big|_{1,l+2}^{2d(l+4)}+\big|L(\overline{F})\big|_l^2\big). \tag{3.21}$$

For $F\in\mathcal{S}^d$, we define the linear operator $T_r(F,\cdot):\mathcal{S}\to\mathcal{S}, r=1,...,d$ by

$$T_r(F,G)=\langle DG,(\gamma_F DF)^r\rangle\,,$$

where $(\gamma_F DF)^r=\sum_{r'=1}^d\gamma_F^{r',r}DF^{r'}$. Moreover, for a multi-index $\beta=(\beta_1,..,\beta_q)$ we denote $|\beta|=q$ and we define by induction

$$T_\beta(F,G)=T_{\beta_q}(F,T_{(\beta_1,\ldots,\beta_{q-1})}(F,G)).$$

For $l\in\mathbb{N}$ and $F,\overline{F}\in\mathcal{S}^d$, we denote

$$\Theta_l(F)=m_F^l(1+|F|_{1,l+1}^{2d(l+1)})\qquad\text{and} \tag{3.22}$$

$$\Theta_l(F,\overline{F})=m_F^l m_{\overline{F}}^l\big(1+|F|_{1,l}^{2d(l+2)}+|\overline{F}|_{1,l}^{2d(l+2)}\big). \tag{3.23}$$

We notice that $\Theta_l(F)\leq\Theta_l(F,\overline{F})$, $\Theta_l(F)\leq\Theta_{l+1}(F)$ and $\Theta_l(F,\overline{F})\leq\Theta_{l+1}(F,\overline{F})$.

Proposition 3.3. *Let $F,\overline{F}\in\mathcal{S}^d$ and $G,\overline{G}\in\mathcal{S}$. Then for every $l\in\mathbb{N}$ and for every multi index β with $|\beta|=q\geq1$ one has*

$$|T_\beta(F,G)|_l\leq C\Theta_{l+q}^q(F)\,|G|_{1,l+q} \tag{3.24}$$

and

$$\begin{aligned}\big|T_\beta(F,G)-T_\beta(\overline{F},\overline{G})\big|_l \;\leq\; & C\Theta_{l+q}^{\frac{q(q+1)}{2}}(F,\overline{F})\big(1+|G|_{1,l+q}+|\overline{G}|_{1,l+q}\big)^q\times\\ &\times\big(\big|F-\overline{F}\big|_{1,l+q}+\big|G-\overline{G}\big|_{1,l+q}\big)\end{aligned} \tag{3.25}$$

where $C\in\mathcal{C}(l,d,q)$.

Proof. (3.24) follows from [5] (see (26) in the proof of Theorem 3 therein). We prove (3.25) by recurrence. Hereafter, C denotes a constant in $\mathcal{C}(l,d,q)$, possibly varying from a line to another. For $|\beta|=q=1$ we have

$$\begin{aligned}\big|T_r(F,G)-T_r(\overline{F},\overline{G})\big|_l \;\leq\; & \big|G-\overline{G}\big|_{l+1}(|F|_{1,l+1}+\big|\overline{F}\big|_{1,l+1})(|\gamma_F|_l+|\gamma_{\overline{F}}|_l)+\\ &+\big|F-\overline{F}\big|_{1,l+1}(|G|_{1,l+1}+\big|\overline{G}\big|_{1,l+1})(|\gamma_F|_l+|\gamma_{\overline{F}}|_l)+\\ &+|\gamma_F-\gamma_{\overline{F}}|_l\,(|F|_{1,l+1}+\big|\overline{F}\big|_{1,l+1})(|G|_{1,l+1}+\big|\overline{G}\big|_{1,l+1}).\end{aligned}$$

Using (3.12), we obtain (3.25). The case $|\beta|=q\geq2$, easily follows by induction. □

We can now establish estimates for the weights H^q. For $l\geq1$ and $F,\overline{F}\in\mathcal{S}^d$, we set (with the convention $|\cdot|_0=|\cdot|$)

$$A_l(F)=m_F^{l+1}\big(1+|F|_{1,l+1}^{2d(l+2)}+|LF|_{l-1}^2\big) \tag{3.26}$$

$$A_l(F,\overline{F})=m_F^{l+1}m_{\overline{F}}^{l+1}\big(1+|F|_{1,l+1}^{2d(l+3)}+|\overline{F}|_{1,l+1}^{2d(l+3)}+|LF|_{l-1}^2+|L\overline{F}|_{l-1}^2\big). \tag{3.27}$$

Theorem 3.4. ***A***. *For $l \in \mathbb{N}$ and $q \in \mathbb{N}^*$ there exists $C \in \mathcal{C}(l,d,q)$ such that for every $F \in \mathcal{S}^d$, $G \in \mathcal{S}$ and for every multi-index $\beta = (\beta_1,..,\beta_q)$*

$$\Big|H^q_\beta(F,G)\Big|_l \leq CA_{l+q}(F)^q|G|_{l+q}, \tag{3.28}$$

$A_{l+q}(F)$ being defined in (3.26).

B. *There exists $C \in \mathcal{C}(l,d,q)$ such that for every $F,\overline{F} \in \mathcal{S}^d$, $G,\overline{G} \in \mathcal{S}$ and every multi-index $\beta = (\beta_1,..,\beta_q)$*

$$\begin{aligned}\Big|H^q_\beta(F,G) - H^q_\beta(\overline{F},\overline{G})\Big|_l \leq \quad & CA_{l+q}(F,\overline{F})^{\frac{q(q+1)}{2}} \times \big(1+|G|_{l+q}+|\overline{G}|_{l+q}\big)^q \times \\ & \times\big(\big|F-\overline{F}\big|_{l+q+1} + \big|L(F-\overline{F})\big|_{l+q-1} + \big|G-\overline{G}\big|_{l+q}\big).\end{aligned} \tag{3.29}$$

Proof. **A.** Suppose $q=1$ and $\beta = r$. Then,

$$|H_r(F,G)|_l \leq C|G|_l|L^\gamma_r(F)|_l + C|T_r(F,G)|_l.$$

By using (3.19) and (3.24), we can write

$$\begin{aligned}|H_r(F,G)|_l \leq & Cm_F^{l+2}\big(1+|F|^{2d(l+2)}_{1,l+1} + |LF|^2_l\big)|G|_l + m_F^{l+1}\big(1+|F|^{2d(l+3)}_{1,l+2}\big)|G|_{l+1} \\ \leq & A_{l+1}(F)|G|_{l+1}.\end{aligned}$$

So, the statement holds for $q=1$. And for $q>1$ it follows by iteration and by using the fact that $A_{l+1}(F) \leq A_{l+q}(F)$.

B. Suppose $q=1$ and $\beta = r$. Then,

$$\begin{aligned}|H_r(F,G) - H_r(\overline{F},\overline{G})|_l \leq & C\big(|L^\gamma_r(F)|_l|G-\overline{G}|_l + |\overline{G}|_l|L^\gamma_r(F) - L^\gamma_r(\overline{F})|_l + \\ & + |T_r(F,G) - T_r(\overline{F},\overline{G})|_l\big).\end{aligned}$$

Now, estimate (3.29) follows by using (3.19) and (3.20), (3.25). In the iteration for $q>1$, it suffices to observe that $A_l(F) \leq A_l(F,\overline{F})$, $A_l(F) \leq A_{l+1}(F)$ and $A_l(F,\overline{F}) \leq A_{l+1}(F,\overline{F})$. □

3.3 Localized representation formulas for the density

In this section we discuss localized integration by parts formulas with the localization random variable $\mathbf{\Theta}$ defined in (2.14). We will use the norms $\|F\|_{p,\mathbf{\Theta}}$ and $\|F\|_{1,l,p,\mathbf{\Theta}}, \|F\|_{l,p,\mathbf{\Theta}}$ defined in (2.8). We also recall that $\mathrm{m}_{q,p}(\mathbf{\Theta})$ is defined in (2.11) and that an estimate of this quantity is given in (2.15).

We give now the integration by parts formula with respect to $\mathbb{P}_{\mathbf{\Theta}}$ (that is, locally) and we study the regularity of the law starting from the results in [2].

Once for all, in addition to $\mathrm{m}_{q,p}(\mathbf{\Theta})$ we define the following quantities: for $p\geq 1$, $q \in \mathbb{N}$, $F \in \mathcal{S}^d$,

$$\begin{aligned}\mathrm{S}_{F,\mathbf{\Theta}}(p) &= \max\{1, \big\|(\det\sigma_F)^{-1}\big\|_{p,\mathbf{\Theta}}\}, \\ \mathrm{Q}_{F,\mathbf{\Theta}}(q,p) &= 1+\|F\|_{1,q,p,\mathbf{\Theta}} + \|LF\|_{q-2,p,\mathbf{\Theta}} \\ \mathrm{Q}_{F,\overline{F},\mathbf{\Theta}}(q,p) &= 1+\|F\|_{1,q,p,\mathbf{\Theta}} + \|LF\|_{q-2,p,\mathbf{\Theta}} + \big\|\overline{F}\big\|_{1,q,p,\mathbf{\Theta}} + \big\|L\overline{F}\big\|_{q-2,p,\mathbf{\Theta}}\end{aligned} \tag{3.30}$$

with the convention $\mathrm{S}_{F,\mathbf{\Theta}}(p) = +\infty$ if the r.h.s. is not finite.

Proposition 3.5. *Let $\kappa \in \mathbb{N}^*$ and assume that $\mathrm{m}_{\kappa,p}(\mathbf{\Theta}) < \infty$ for all $p \geq 1$. Let $F \in \mathcal{S}^d$ be such that $\mathrm{S}_{F,\mathbf{\Theta}}(p) < \infty$ for every $p \in \mathbb{N}$. Let γ_F be the inverse of σ_F on the set $\{\mathbf{\Theta} \neq 0\}$. Then the following localized integration by parts formula holds: for every $f \in C^\infty_b(\mathbb{R}^d)$, $G \in \mathcal{S}$ and for every multi index α of length equal to $q \leq \kappa$ one has*

$$\mathbb{E}_{\mathbf{\Theta}}(\partial_\alpha f(F)\,G) = \mathbb{E}_{\mathbf{\Theta}}(f(F) H^q_{\alpha,\mathbf{\Theta}}(F,G))$$

where as $r = 1, \dots, d$

$$\begin{aligned} H_{r,\mathbf{\Theta}}(F,G) &= \sum_{r'=1}^{d} G\gamma_F^{r',r} LF^{r'} - \langle D(G\gamma_F^{r',r}), DF^{r'}\rangle_J - G\gamma_F^{r',r}\langle D(\ln\mathbf{\Theta}), DF^{r'}\rangle_J \\ &= H_r(F,G) - G\sum_{r'=1}^{d}\gamma_F^{r',r}\langle D(\ln\mathbf{\Theta}), DF^{r'}\rangle_J \end{aligned} \tag{3.31}$$

and for a general multi index β *with* $|\beta| = q$

$$H^q_{\beta,\mathbf{\Theta}}(F,G) = H_{\beta_q,\mathbf{\Theta}}\big(F, H^{q-1}_{(\beta_1,\dots,\beta_{q-1}),\mathbf{\Theta}}(F,G)\big).$$

Proof. For $|\beta| = 1$, the integration by parts formula immediately follows from the equality $\mathbb{E}_{\mathbf{\Theta}}(\partial_i f(F)G) = \mathbb{E}(\partial_i f(F)G\mathbf{\Theta}) = \mathbb{E}(f(F)H_i(F,G\mathbf{\Theta}))$, so that $H_{i,\mathbf{\Theta}}(F,G) = \frac{1}{\mathbf{\Theta}}H_i(F,G\mathbf{\Theta})$, and this gives the formula for $H_{i,\mathbf{\Theta}}(F,G)$. For higher order integration by parts it suffices to iterate this procedure. □

We give now estimates for the weights in the integration by parts formula.

Proposition 3.6. *Let* $\kappa \in \mathbb{N}^*$ *and* $l \in \mathbb{N}$ *be such that* $\mathrm{m}_{l+\kappa+1,p}(\mathbf{\Theta}) < \infty$ *for all* $p \geq 1$. *Let* $F, \overline{F} \in \mathcal{S}^d$, *with* $\mathrm{S}_{F,\mathbf{\Theta}}(p), \mathrm{S}_{\overline{F},\mathbf{\Theta}}(p) < \infty$ *for every* p, *and* $G, \overline{G} \in \mathcal{S}$. *For* $q \leq \kappa$, *let* $H^q_{\beta,\mathbf{\Theta}}(\cdot,\cdot)$ *be the weight of the integration by parts formula as in Proposition 3.5. Then for every* $p \geq 1$ *one may find two universal constants* $C, p' \in \mathcal{C}(\kappa, d)$ *such that for every multi index* β *with* $|\beta| = q \leq \kappa$

$$\|H^q_{\beta,\mathbf{\Theta}}(F,G)\|_{l,p,\mathbf{\Theta}} \leq C\, B_{l+q,p',\mathbf{\Theta}}(F)^q\, \|G\|_{l+q,p',\mathbf{\Theta}}, \tag{3.32}$$

and

$$\begin{aligned} &\|H^q_{\beta,\mathbf{\Theta}}(F,G) - H^q_{\beta,\mathbf{\Theta}}(\overline{F},\overline{G})\|_{l,p,\mathbf{\Theta}} \leq \\ &\quad \leq C\, B_{l+q,p',\mathbf{\Theta}}(F,\overline{F})^{q(q+1)/2}\big(1 + \|G\|_{l+q,p'} + \|\overline{G}\|_{l+q,p'}\big)\times \\ &\quad\quad \times(\|F-\overline{F}\|_{l+q+1,p,\mathbf{\Theta}} + \|LF - L\overline{F}\|_{l+q-1,p,\mathbf{\Theta}} + \|G-\overline{G}\|_{l+q,p,\mathbf{\Theta}}) \end{aligned} \tag{3.33}$$

where

$$B_{l,p,\mathbf{\Theta}}(F) = \mathrm{S}_{F,\mathbf{\Theta}}(p)^{l+1}\mathrm{Q}_{F,\mathbf{\Theta}}(l+1,p)^{2d(l+2)}\mathrm{m}_{l,p}(\mathbf{\Theta}) \tag{3.34}$$

$$B_{l,p,\mathbf{\Theta}}(F,\overline{F}) = \mathrm{S}_{F,\mathbf{\Theta}}(p)^{l+1}\mathrm{S}_{\overline{F},\mathbf{\Theta}}(p)^{l+1}\mathrm{Q}_{F,\overline{F},\mathbf{\Theta}}(l+1,p)^{2d(l+2)}\mathrm{m}_{l,p}(\mathbf{\Theta}) \tag{3.35}$$

$\mathrm{S}_{\cdot,\mathbf{\Theta}}(p)$, $\mathrm{Q}_{\cdot,\mathbf{\Theta}}(l,p)$ *and* $\mathrm{Q}_{\cdot,\cdot,\mathbf{\Theta}}(l,p)$ *being defined in* (3.30).

Proof. By using the same arguments as in Theorem 3.4, one gets that there exists $C \in \mathcal{C}(q,d)$ such that for every multi index β of length q then

$$|H^q_{\beta,\mathbf{\Theta}}(F,G)|_l \leq C\, A_{l+q}(F)^q\big(1 + |D\ln\mathbf{\Theta}|_{l+q-1}\big)^q|G|_{l+q}$$

and

$$\begin{aligned} |H^q_{\beta,\mathbf{\Theta}}(F,G) - H^q_{\beta,\mathbf{\Theta}}(\overline{F},\overline{G})|_l \leq & C\, A_{l+q}(F,\overline{F})^{\frac{q(q+1)}{2}}\big(1 + |D\ln\mathbf{\Theta}|_{l+q-1}\big)^{\frac{q(q+1)}{2}}\times \\ & \times\big(1 + |G|_{l+q} + |\overline{G}|_{l+q}\big)^q\times \\ & \times\big(|F-\overline{F}|_{l+q+1} + |L(F-\overline{F})|_{l+q-1} + |G-\overline{G}|_{l+q}\big), \end{aligned}$$

where $A_l(F)$ and $A_l(F,\overline{F})$ are defined in (3.26) and (3.27) respectively (as usual, $|\cdot|_0 \equiv |\cdot|$). By using Hölder inequality one gets (3.32) and (3.33). □

In next Lemma we study properties of $H^q_{\beta,\mathbf{\Theta}}(F,G)$ in the case G is a special function of F. We denote with $B_r(0)$ the ball with radius r centered at 0.

Lemma 3.7. *Let $\phi \in C_b^\infty(\mathbb{R}^d)$ be such that $\mathbf{1}_{B_1(0)} \leq \phi \leq \mathbf{1}_{B_2(0)}$ and set $\phi_x(y) = \phi(x-y)$. For $l \in \mathbb{N}$, $p \geq 1$ and $F, \overline{F} \in \mathcal{S}^d$, one has*

$$\|\phi_x(F)\|_{l,p,\mathbf{\Theta}} \leq C(1+\|F\|_{1,l,2p,\mathbf{\Theta}})^l \mathbb{P}_{\mathbf{\Theta}}(F \in B_2(x)), \tag{3.36}$$

$$\|\phi_x(F)-\phi_x(\overline{F})\|_{l,p,\mathbf{\Theta}} \leq C\|F-\overline{F}\|_{l,2p,\mathbf{\Theta}}(1+\|F\|_{1,l,2p,\mathbf{\Theta}}+\|\overline{F}\|_{1,l,2p,\mathbf{\Theta}})^l, \tag{3.37}$$

in which $C \in \mathcal{C}(l,p,d)$. Moreover for every $F \in \mathcal{S}^d$ and $V \in \mathcal{S}$ one may find universal constants $C, a, p' \in \mathcal{C}(q,l,p,d)$ such that for every multi index β with $|\beta| = q$

$$\begin{aligned}\|H^q_{\beta,\mathbf{\Theta}}(F,V\phi_x(F))\|_{l,p,\mathbf{\Theta}} \leq \quad & \mathrm{S}_{F,\mathbf{\Theta}}(p')^a \mathrm{Q}_{F,\mathbf{\Theta}}(l+q+1,p')^a \mathrm{m}_{l+q,p}(\mathbf{\Theta})^a \times \\ & \times \mathbb{P}_{\mathbf{\Theta}}(|F-x|<2)^{\frac{1}{p'}} \times \|V\|_{l,p',\mathbf{\Theta}},\end{aligned} \tag{3.38}$$

where $\mathrm{S}_{F,\mathbf{\Theta}}(p)$, $\mathrm{Q}_{F,\mathbf{\Theta}}(l,p)$ and $\mathrm{m}_{l,p}(\mathbf{\Theta})$ are defined in (3.30).

Proof. We prove (3.37), (3.36) following with similar arguments. First, for a multi-index α with $|\alpha| = k$, one has

$$D_\alpha \phi_x(F) = \sum_{l=1}^k \phi_x^{(l)}(F) \sum_{\beta_1,\ldots,\beta_l \in \mathcal{B}_\alpha} D_{\beta_1}F \cdots D_{\beta_l}F$$

where $C > 0$ depends on l, d only and "$\beta_1,\ldots,\beta_l \in \mathcal{B}_\alpha$" means that $\beta_1,\ldots,\beta_l$ are non empty multi indexes of α running through the list of all of the (non empty) "blocks" of α. So, straightforward computations give

$$\begin{aligned}&D_\alpha\phi_x(F) - D_\alpha\phi_x(\overline{F}) \\ &\quad = \sum_{l=1}^k \big(\phi_x^{(l)}(F) - \phi_x^{(l)}(\overline{F})\big) \sum_{\beta_1,\ldots,\beta_l \in \mathcal{B}_\alpha} D_{\beta_1}F \cdots D_{\beta_l}F + \\ &\qquad + \sum_{l=1}^k \phi_x^{(l)}(\overline{F}) \sum_{\beta_1,\ldots,\beta_l \in \mathcal{B}_\alpha} \sum_{j=1}^l \Big(\prod_{k=1}^{j-1} D_{\beta_k}F\Big)\big(D_{\beta_j}F - D_{\beta_j}\overline{F}\big)\Big(\prod_{k=j+1}^{l} D_{\beta_k}\overline{F}\Big)\end{aligned}$$

with the convention that $\prod_{k=1}^0 (\cdot)_k = 1 = \prod_{k=l+1}^l (\cdot)_k$. Since $\phi_x^{(l)}$ is Lipschitz continuous, with a Lipschitz constant independent of x, it follows that

$$\begin{aligned}|\phi_x(F)-\phi_x(\overline{F})|_l &\leq C|F-\overline{F}|(1+|F|_{1,l})^l + C|F-\overline{F}|_l(1+|F|_{1,l}+|\overline{F}|_{1,l})^{l-1} \\ &\leq C|F-\overline{F}|_l(1+|F|_{1,l}+|\overline{F}|_{1,l})^l\end{aligned}$$

and by using the Hölder inequality one gets (3.37).

As for (3.38), we first note that since $\phi_x(y) \equiv 0$ for $|y-x| > 2$ then (3.31) gives

$$D_\alpha H^q_{\beta,\mathbf{\Theta}}(F,V\phi_x(F)) = D_\alpha H^q_{\beta,\mathbf{\Theta}}(F,V\phi_x(F))\mathbf{1}_{\{|F-x|<2\}} \tag{3.39}$$

for every multi index α. So, for $l \in \mathbb{N}$ we can write

$$|H^q_{\beta,\mathbf{\Theta}}(F,V\phi_x(F))|_l = |H^q_{\beta,\mathbf{\Theta}}(F,V\phi_x(F))|_l \mathbf{1}_{\{|F-x|<2\}}.$$

Therefore, (3.38) is a consequence of the use of the Hölder inequality and of the estimate (3.32). $\square$

We recall that the Poisson kernel Q_d is the solution to the equation $\Delta Q_d = \delta_0$ in $\mathbb{R}^d$ (δ_0 denoting the Dirac mass in $\{0\}$) and has the following explicit form:

$$Q_1(x) = \max\{x, 0\}, \quad Q_2(x) = a_2^{-1} \ln |x| \quad \text{and} \quad Q_d(x) = -a_d^{-1} |x|^{2-d}, d > 2 \tag{3.40}$$

where a_d is the area of the unit sphere in $\mathbb{R}^d$. By using the result in [2], we have the following

Proposition 3.8. *Let $\phi \in C_b^\infty(\mathbb{R}^d)$ be such that $1_{B_1(0)} \leq \phi \leq 1_{B_2(0)}$ and set $\phi_x(y) = \phi(x-y)$. Let $\kappa \in \mathbb{N}^*$ and assume that $\mathrm{m}_{\kappa,p}(\mathbf{\Theta}) < \infty$ for every $p \geq 1$. Let $F \in \mathcal{S}^d$ be such that $\mathrm{S}_{F,\mathbf{\Theta}}(p) < \infty$ for every $p \geq 1$.*

***A**. Let Q_d be the Poisson kernel in $\mathbb{R}^d$ given in (3.40). Then for every $p > d$ there exists a universal constant $C \in \mathcal{C}(d,p)$ such that*

$$\|\nabla Q_d(F-x)\|_{\frac{p}{p-1},U} \leq C \|H_{\mathbf{\Theta}}(F,1)\|_{p,\mathbf{\Theta}}^{k_{p,d}} \tag{3.41}$$

where $k_{p,d} = (d-1)/(1-d/p)$ and $H_{\mathbf{\Theta}}(F,1)$ denotes the vector in $\mathbb{R}^d$ whose ith entry is given be $H_{i,\mathbf{\Theta}}(F,1)$.

***B**. Under $\mathbb{P}_{\mathbf{\Theta}}$, the law of F is absolutely continuous and has a density $p_{F,\mathbf{\Theta}} \in C^{\kappa-1}(\mathbb{R}^d)$ whose derivatives up to order $\kappa - 1$ may be represented as*

$$\partial_\alpha p_{F,\mathbf{\Theta}}(x) = \sum_{i=1}^d \mathbb{E}_{\mathbf{\Theta}}\big(\partial_i Q_d(F-x) H^{q+1}_{(i,\alpha),\mathbf{\Theta}}(F, \phi_x(F))\big) \tag{3.42}$$

for every multi index α with $|\alpha| = q \leq \kappa - 1$.

***C**. Let V a random variable taking values in $(0,1)$ and such that $\mathrm{m}_{\kappa,p}(V) < \infty$ for every $p \geq 1$. Then for $|\alpha| = q \leq \kappa - 1$ one has*

$$|\partial_\alpha p_{F,\mathbf{\Theta} V}(x)| \leq C \mathrm{S}_{F,\mathbf{\Theta}}(p')^a \mathrm{Q}_{F,\mathbf{\Theta}}(q+2, p^{'})^a \mathrm{m}_{q+1,p'}(\Theta)^a \, \|V\|_{q+1,p',\mathbf{\Theta}} \times \mathbb{P}_{\mathbf{\Theta}}(|F-x| < 2)^b, \tag{3.43}$$

in which $C, a, b, p' \in \mathcal{C}(\kappa, d)$.

Proof. **A.** This point is actually Theorem 5 in [2] (recall that $\|1\|_{W^{1,p}_{\mu_F}} \leq \|H(F,1)\|_p$, see Remark 17 in [2]) with $\mathbb{P}$ replaced by $\mathbb{P}_{\mathbf{\Theta}}$.

B. Set $\mu_{F,\mathbf{\Theta}}$ the law of F under $\mathbb{P}_{\mathbf{\Theta}}$ and let α denote a multi index with $|\alpha| = q$. By using the arguments similar to the ones developed in Proposition 10 in [2] one easily gets (notations from that paper)

$$\partial_\alpha p_{F,\mathbf{\Theta}}(x) = (-1)^{|\alpha|+1} \sum_{i=1}^d \int_{\mathbb{R}^d} \partial_i Q_d(y-x) \partial^{\mu_{F,\mathbf{\Theta}}}_{(i,\alpha)} \phi_x(y) \mu_{F,\mathbf{\Theta}}(dy).$$

And by recalling that $(-1)^{|\alpha|+1} \partial^{\mu_{F,\mathbf{\Theta}}}_{(i,\alpha)} \phi_x(F) = \mathbb{E}_{\mathbf{\Theta}}(H^{q+1}_{(i,\alpha),\mathbf{\Theta}}(F, \phi_x(F)) \mid F)$ (see Section 3 of [2]), (3.42) follows.

C. We first note that $\mathrm{m}_{\mathbf{\Theta} V}(\kappa,p) \leq C(\mathrm{m}_{\mathbf{\Theta}}(\kappa,p) + \mathrm{m}_V(\kappa,p))$. So, we can apply (3.42) with localization $\mathbf{\Theta} V$ and we get

$$\partial_\alpha p_{F,\mathbf{\Theta} V}(x) = \sum_{i=1}^d \mathbb{E}_{\mathbf{\Theta}}\big(\partial_i Q_d(F-x) V H^{q+1}_{(i,\alpha),\mathbf{\Theta} V}(F, \phi_x(F))\big).$$

Now, from (3.31) one has $V H_{i,\mathbf{\Theta} V}(F, \phi_x(F)) = H_{i,\mathbf{\Theta}}(F, V\phi_x(F))$ and by iteration it follows that $V H^q_{\beta,\mathbf{\Theta} V}(F, \phi_x(F)) = H^q_{\beta,\mathbf{\Theta}}(F, V\phi_x(F))$. Therefore,

$$\partial_\alpha p_{F,\mathbf{\Theta} V}(x) = \sum_{i=1}^d \mathbb{E}_{\mathbf{\Theta}}\big(\partial_i Q_d(F-x) H^{q+1}_{(i,\alpha),\mathbf{\Theta}}(F, V\phi_x(F))\big)$$

and, by using the Hölder inequality, for $p > d$ we have

$$\begin{aligned}|\partial_\alpha p_{F,\mathbf{\Theta}V}(x)| &\leq \sum_{i=1}^d \|\nabla Q_d(F-x)\|_{\frac{p}{p-1},\mathbf{\Theta}}\|H^{q+1}_{(i,\alpha),\mathbf{\Theta}}(F,V\phi_x(F))\|_{p,\mathbf{\Theta}}\\ &\leq \sum_{i=1}^d \|H_{\mathbf{\Theta}}(F,1)\|^{k_{p,d}}_{p,\mathbf{\Theta}}\|H^{q+1}_{(i,\alpha),\mathbf{\Theta}}(F,V\phi_x(F))\|_{p,\mathbf{\Theta}}\end{aligned}$$

in which we have used (3.41). Now, by using (3.32) to estimate the first term and by applying (3.38) to the second one, (3.43) follows. □

3.4 The distance between density functions and their derivatives

We compare now the probability density functions (and their derivatives) of two random variables under $\mathbb{P}_{\mathbf{\Theta}}$. In a different setting, this problem has already been considered in [13].

Proposition 3.9. *Let $q \in \mathbb{N}$ and assume that $\mathrm{m}_{q+2,p}(\mathbf{\Theta}) < \infty$ for every $p \geq 1$. Let $F, G \in \mathcal{S}^d$ be such that*

$$\mathrm{S}_{F,G,\mathbf{\Theta}}(p) := 1 + \sup_{0\leq\lambda\leq 1} \|(\det \sigma_{G+\lambda(F-G)})^{-1}\|_{p,\mathbf{\Theta}} < \infty, \quad \forall p \in \mathbb{N}. \tag{3.44}$$

Then under $\mathbb{P}_{\mathbf{\Theta}}$ the laws of F and G are absolutely continuous with respect to the Lebesgue measure with density $p_{F,\mathbf{\Theta}}$ and $p_{G,\mathbf{\Theta}}$ respectively and for every multi index α with $|\alpha| = q$ there exist constants $C, a, b, p' \in \mathcal{C}(q,d)$ such that

$$\begin{aligned}|\partial_\alpha p_{F,\mathbf{\Theta}}(y) - \partial_\alpha p_{G,\mathbf{\Theta}}(y)| \quad &\leq \quad C\mathrm{S}_{F,G,\mathbf{\Theta}}(p')^a \mathrm{Q}_{F,G,\mathbf{\Theta}}(q+3,p')^a \mathrm{m}^a_{q+2,p}(\mathbf{\Theta})\times\\ &\times(\|F-G\|_{q+2,p',\mathbf{\Theta}} + \|LF-LG\|_{q,p',\mathbf{\Theta}})\times\\ &\times(\mathbb{P}_{\mathbf{\Theta}}(|F-y|<2) + \mathbb{P}_{\mathbf{\Theta}}(|G-y|<2))^b\end{aligned} \tag{3.45}$$

with $\mathrm{m}_{k,p}(\mathbf{\Theta})$ and $\mathrm{Q}_{F,G,\mathbf{\Theta}}(k,p)$ given in (2.11) and (3.30) respectively.

Proof. Throughout this proof, $C, p', a, b \in \mathcal{C}(q,d)$ will denote constants that can vary from line to line. By applying Lemma 3.8, under $\mathbb{P}_{\mathbf{\Theta}}$ the laws of F and G are both absolutely continuous with respect to the Lebesgue measure and for every multi index α with $|\alpha| = q$ one has

$$\begin{aligned}\partial_\alpha p_{F,\mathbf{\Theta}}(y) - \partial_\alpha p_{G,\mathbf{\Theta}}(y) &= \sum_{j=1}^d \mathbb{E}_{\mathbf{\Theta}}\big(\big(\partial_j Q_d(F-y) - \partial_j Q_d(G-y)\big)H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))\big)+\\ &\quad+\sum_{j=1}^d \mathbb{E}_{\mathbf{\Theta}}\big(\partial_j Q_d(F-y)\big(H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(F,\phi_y(F)) - H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))\big)\big)\\ &=: \sum_{j=1}^d I_j + \sum_{j=1}^d J_j.\end{aligned}$$

By using (3.41), for $p > d$ we obtain

$$\begin{aligned}|J_j| &\leq C\,\|\nabla Q_d(F-y)\|_{\frac{p}{p-1},U}\|H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(F,\phi_y(F)) - H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))\|_{p,\mathbf{\Theta}}\\ &\leq C\,\|H_{\mathbf{\Theta}}(F,1)\|^{k_{d,p}}_{p,\mathbf{\Theta}}\|H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(F,\phi_y(F)) - H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))\|_{p,\mathbf{\Theta}}.\end{aligned}$$

Now, from (3.39) (with $\alpha = \emptyset$) it follows that the above term is null on $\{|F-y| \geq 2\} \cap \{|G-y| \geq 2\}$. So

$$\begin{aligned}&|H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(F,\phi_y(F)) - H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))|\\ &\quad\leq |H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(F,\phi_y(F)) - H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))|\mathbf{1}_{\{|F-y|<2\}}+\\ &\qquad + |H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(F,\phi_y(F)) - H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))|\mathbf{1}_{\{|G-y|<2\}}\end{aligned}$$

so that the Hölder inequality gives

$$\begin{aligned}|J_j| \leq & C\,\|H_{\mathbf{\Theta}}(F,1)\|^{k_{d,p}}_{p,U}\|H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(F,\phi_y(F)) - H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))\|_{2p,U}\times\\ &\times \big(\mathbb{P}_{\mathbf{\Theta}}(|F-x|<2) + \mathbb{P}_{\mathbf{\Theta}}(|G-x|<2)\big)^{\frac{1}{2p}}.\end{aligned}$$

So, by applying (3.32) and (3.33), there exists $p' > p > d$ such that

$$\begin{aligned}|J_j| \leq & CB_{q+1,p',\mathbf{\Theta}}(F,G)^{k_{d,p}+\frac{(q+1)(q+2)}{2}}\times\\ &\times \big(\|F-G\|_{q+2,p',\mathbf{\Theta}} + \|L(F-G)\|_{q,p',\mathbf{\Theta}} + \|\phi_y(F)-\phi_y(G)\|_{q+1,p',\mathbf{\Theta}}\big)\times\\ &\times \big(\mathbb{P}_{\mathbf{\Theta}}(|F-x|<2) + \mathbb{P}_{\mathbf{\Theta}}(|G-x|<2)\big)^{\frac{1}{2p'}},\end{aligned}$$

$B_{q+1,p',\mathbf{\Theta}}(F,G)$ being defined in (3.34). By using (3.37) and the quantities $\mathrm{S}_{F,G,\mathbf{\Theta}}(p)$ and $\mathrm{Q}_{F,G,\mathbf{\Theta}}(k,p)$, for a suitable $a > 1$ and $p' > d$ we can write

$$\begin{aligned}|J_j| \leq & C\mathrm{S}_{F,G,\mathbf{\Theta}}(p')^a\mathrm{Q}_{F,G,\mathbf{\Theta}}(q+2,p')^a\mathrm{m}_{q+1,p'}(\mathbf{\Theta})^a\times\\ &\times \big(\|F-G\|_{q+2,p',\mathbf{\Theta}} + \|L(F-G)\|_{q,p',\mathbf{\Theta}}\big)\times\\ &\times \big(\mathbb{P}_{\mathbf{\Theta}}(|F-x|<2) + \mathbb{P}_{\mathbf{\Theta}}(|G-x|<2)\big)^{\frac{1}{2p'}}.\end{aligned}$$

We study now I_j. For $\lambda \in [0,1]$ we denote $F_\lambda = G + \lambda(F-G)$ and we use Taylor's expansion to obtain

$$I_j = \sum_{k=1}^d R_{k,j} \quad \text{with} \quad R_{k,j} = \int_0^1 \mathbb{E}_{\mathbf{\Theta}}\big(\partial_k\partial_j Q_d(F_\lambda - y)H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))(F-G)_k\big)d\lambda.$$

Let $V_{k,j} = H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))(F-G)_k$. Since for $\lambda \in [0,1]$ then $\mathbb{E}_{\mathbf{\Theta}}((\det\sigma_{F_\lambda})^{-p})) < \infty$ for every p, we can use the integration by parts formula with respect to F_λ, so

$$R_{k,j} = \int_0^1 \mathbb{E}_{\mathbf{\Theta}}\big(\partial_j Q_d(F_\lambda - y)H_{k,\mathbf{\Theta}}(F_\lambda, V_{k,j})\big)d\lambda.$$

Therefore, by taking $p > d$ and by using again (3.41), (3.32) and (3.33), we get

$$\begin{aligned}|R_{k,j}| \leq & \int_0^1 \|\partial_j Q_d(F_\lambda - y)\|_{\frac{p}{p-1},\mathbf{\Theta}}\|H_{k,\mathbf{\Theta}}(F_\lambda,V_{k,j})\|_{p,\mathbf{\Theta}}d\lambda\\ \leq & C\int_0^1 \|H_{\mathbf{\Theta}}(F_\lambda,1)\|^{k_{d,p}}_{p,U}\|H_{k,\mathbf{\Theta}}(F_\lambda,V_{k,j})\|_{p,\mathbf{\Theta}}d\lambda\\ \leq & C\int_0^1 B_{1,p',\mathbf{\Theta}}(F_\lambda)^{k_{d,p}+1}\|V_{k,j}\|_{1,p'\mathbf{\Theta}}\,d\lambda\end{aligned}$$

in which we have used (3.32). Now, from (3.44) and (3.34) it follows that

$$B_{1,p,\mathbf{\Theta}}(F_\lambda) \leq C\mathrm{S}_{F,G,\mathbf{\Theta}}(p)^4\mathrm{Q}_{F,G,\mathbf{\Theta}}(2,p)^{8d}\times \mathrm{m}_{1,p}(\mathbf{\Theta}).$$

Moreover,

$$\|V_{k,j}\|_{1,p,\mathbf{\Theta}} = \|H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))(F-G)_k\|_{1,p,\mathbf{\Theta}} \leq \|H^{q+1}_{(j,\alpha),\mathbf{\Theta}}(G,\phi_y(G))\|_{1,2p,\mathbf{\Theta}}\|F-G\|_{1,2p,\mathbf{\Theta}}$$

and from (3.38) and (3.32) we get

$$\|V_{k,j}\|_{1,p,\mathbf{\Theta}} \leq C\, B_{q+2,p',\mathbf{\Theta}}(G)^{q+1}\,\|\phi_y(G)\|_{q+2,p',U}\mathbb{P}_{\mathbf{\Theta}}(|G-y|<2)^{\frac{1}{p'}}\,\|F-G\|_{1,p',\mathbf{\Theta}}$$

By using (3.34),

$$B_{q+2,p',\mathbf{\Theta}}(G) \leq \mathrm{S}_{F,G,\mathbf{\Theta}}(p)^{2q+4}\mathrm{Q}_{G,\mathbf{\Theta}}(q+3,p')^{2d(q+3)} \times \mathrm{m}_{q+2,p'}(\mathbf{\Theta}),$$

$\mathrm{Q}_{G,\mathbf{\Theta}}(l,p)$ being given in (3.30). We use also (3.36) and, by inserting everything, we can resume by writing

$$|I_j| \leq C\mathrm{S}_{F,G,\mathbf{\Theta}}(p')^a\mathrm{Q}_{F,G,\mathbf{\Theta}}(q+3,p')^a\mathrm{m}_{q+2,p'}(\mathbf{\Theta})^a\|F-G\|_{1,p',\mathbf{\Theta}} \times \mathbb{P}_{\mathbf{\Theta}}(|G-x|<2)^b$$

and the statement follows. □

Using the localizing function in (2.12) and by applying Proposition 3.9 we get the following result.

Theorem 3.10. *Let $q \in \mathbb{N}$. Assume that $\mathrm{m}_{q+2,p}(\mathbf{\Theta}) < \infty$ for every $p \geq 1$. Let $F, G \in \mathcal{S}^d$ be such that $\mathrm{S}_{F,\mathbf{\Theta}}(p), \mathrm{S}_{G,\mathbf{\Theta}}(p) < \infty$ for every $p \in \mathbb{N}$. Then under $\mathbb{P}_{\mathbf{\Theta}}$, the laws of F and G are absolutely continuous with respect to the Lebesgue measure, with densities $p_{F,\mathbf{\Theta}}$ and $p_{G,\mathbf{\Theta}}$ respectively. Moreover, there exist constants $C, a, b, p' \in \mathcal{C}(q,d)$ such that for every multi index α of length q one has*

$$\begin{aligned}|\partial_\alpha p_{F,\mathbf{\Theta}}(y) - \partial_\alpha p_{G,\mathbf{\Theta}}(y)| \quad \leq \quad & C\mathrm{S}_{F,\mathbf{\Theta}}(p')^a\mathrm{S}_{G,\mathbf{\Theta}}(p')^a\mathrm{Q}_{F,G,\mathbf{\Theta}}(q+3,p')^a\mathrm{m}^a_{q+2,p}(\mathbf{\Theta})\times\\ &\times(\|F-G\|_{q+2,p',\mathbf{\Theta}} + \|LF-LG\|_{q,p',\mathbf{\Theta}})\times\\ &\times(\mathbb{P}_{\mathbf{\Theta}}(|F-y|<2) + \mathbb{P}_{\mathbf{\Theta}}(|G-y|<2))^b\end{aligned} \tag{3.46}$$

with $\mathrm{m}_{k,p}(\mathbf{\Theta})$ and $\mathrm{Q}_{F,G,\mathbf{\Theta}}(k,p)$ given in (2.11) and (3.30) respectively.

Proof. The proof consists in proving that (3.45) holds.

Set $R = F - G$. We use the deterministic estimate (3.17) on the distance between the determinants of two Malliavin covariance matrices: for every $\lambda \in [0,1]$ we can write

$$\begin{aligned}|\det\sigma_{G+\lambda R} - \det\sigma_G| &\leq C_d\,|DR|\,(|DG|+|DF|)^{2d-1}\\ &\leq \big(\alpha_d|DR|^2(|DG|^2+|DF|^2)^{\frac{2d-1}{2}}\big)^{1/2},\end{aligned}$$

so that

$$\det\sigma_{G+\lambda R} \geq \det\sigma_G - \alpha_d\big(|DR|^2(|DG|^2+|DF|^2)^{\frac{2d-1}{2}}\big)^{1/2}. \tag{3.47}$$

For ψ_a as in (2.12), we define

$$V = \psi_{1/8}(H) \quad \text{with} \quad H = |DR|^2\,\frac{(|DG|^2+|DF|^2)^{\frac{2d-1}{2}}}{(\det\sigma_G)^2},$$

so that

$$V \neq 0 \quad \Rightarrow \quad \det\sigma_{G+\lambda R} \geq \frac{1}{2}\det\sigma_G. \tag{3.48}$$

Before continuing, let us give the following estimate for the Sobolev norm of H. First, coming back to the notation $|\cdot|_l$ as in (2.7), by using (3.14) one easily get

$$|H|_l \leq C|(\det\sigma_G)^{-1}|_l^2\big(1+\big||DF|^2\big|_l+\big||DG|^2\big|_l\big)^d\big||DR|^2\big|_l.$$

By using the estimate concerning the determinant from (3.15) and the straightforward estimate $\big||DF|^2\big|_l \leq C|F|^{l+1}_{1,l+1}$, we have

$$|H|_l \leq C|(\det\sigma_G)^{-1}|^{2(l+1)}\big(1+|F|_{1,l+1}+|G|_{1,l+1}\big)^{8dl}\,|R|^{l+1}_{1,l+1}.$$

As a consequence, by using the Hölder inequality we obtain

$$\|H\|_{l,p,\mathbf{\Theta}} \leq C\mathrm{S}_{G,\mathbf{\Theta}}(\bar{p})^{\bar{a}}\mathrm{Q}_{F,G,\mathbf{\Theta}}(l+1,\bar{p})^{\bar{a}}\|F-G\|_{l+1,\bar{p},\mathbf{\Theta}}^{l+1} \tag{3.49}$$

where $C, \bar{p}, \bar{a}$ depends on l, d, p.

Now, because of (3.48), we have $\mathrm{S}_{F,G,\mathbf{\Theta}V}(p) \leq C\mathrm{S}_{G,\mathbf{\Theta}}(p)$, C denoting a suitable positive constant (which will vary in the following lines). We also have $\mathrm{m}_{k,p}(\mathbf{\Theta}V) \leq C(\mathrm{m}_{k,p}(\mathbf{\Theta}) + \mathrm{m}_{k,p}(V))$. By (2.15) and (3.49) we have

$$\mathrm{m}_{q+2,p}(V) \leq C\,\mathrm{S}_{G,\mathbf{\Theta}}(\bar{p})^{\bar{a}}\mathrm{Q}_{F,G,\mathbf{\Theta}}(q+3,\bar{p})^{\bar{a}}$$

for some $\bar{p}, \bar{a}$, so that $m_{q+2,p}(\mathbf{\Theta}V) \leq C\,\mathrm{S}_{G,\mathbf{\Theta}}(\bar{p})^{\bar{a}}\mathrm{Q}_{F,G,\mathbf{\Theta}}(q+3,\bar{p})^{\bar{a}}\mathrm{m}_{q+2,\bar{p}}(\mathbf{\Theta})^{\bar{a}}$. So, we can apply (3.45) with localization $\mathbf{\Theta}V$ and we get

$$\begin{aligned}|\partial_\alpha p_{F,\mathbf{\Theta}\mathbf{V}}(y) - \partial_\alpha p_{G,\mathbf{\Theta}\mathbf{V}}(y)| \quad \leq \quad & C\mathrm{S}_{G,\mathbf{\Theta}}(p')^a\mathrm{Q}_{F,G,\mathbf{\Theta}}(q+3,p')^a\mathrm{m}_{q+2,p}^a(\mathbf{\Theta})\times \\ & \times(\|F-G\|_{q+2,p',\mathbf{\Theta}} + \|LF-LG\|_{q,p',\mathbf{\Theta}})\times \\ & \times(\mathbb{P}_{\mathbf{\Theta}}(|F-y|\leq 2) + \mathbb{P}_{\mathbf{\Theta}}(|G-y|\leq 2))^b\end{aligned}$$

with $p' > d$ and $C, a, b > 0$ depending on q, d. We write now

$$\begin{aligned}|\partial_\alpha p_{F,\mathbf{\Theta}}(y) - \partial_\alpha p_{G,\mathbf{\Theta}}(y)| \leq & |\partial_\alpha p_{F,\mathbf{\Theta}V}(y) - \partial_\alpha p_{G,\mathbf{\Theta}V}(y)| + \\ & + \left|\partial_\alpha p_{F,\mathbf{\Theta}(1-V)}(y)\right| + |\partial_\alpha p_{G,\mathbf{\Theta}V}(y)|\,,\end{aligned}$$

and we have already seen that the first addendum on the r.h.s. behaves as desired. So, it suffices to show that also the remaining two terms have the right behavior. To this purpose, we use (3.43). We have

$$|\partial_\alpha p_{F,\mathbf{\Theta}(1-V)}(x)| \leq C\mathrm{S}_{F,\mathbf{\Theta}}(p)^a\mathrm{Q}_{F,\mathbf{\Theta}}(q+2,p)^a\mathrm{m}_{q+1,p}(\mathbf{\Theta})^a\,\|1-V\|_{q+1,p,\mathbf{\Theta}}\times\mathbb{P}_{\mathbf{\Theta}}(|F-x|<2)^b.$$

Now, we can write

$$\|1-V\|_{q+1,p,\mathbf{\Theta}}^p = \mathbb{E}_{\mathbf{\Theta}}(|1-V|^p) + \|DV\|_{q,p,\mathbf{\Theta}}.$$

But $1-V \neq 0$ implies that $H \geq 1/8$. Moreover, from (2.13), $\|DV\|_{q,p,\mathbf{\Theta}} \leq \|V\|_{q+1,p,\mathbf{\Theta}} \leq C\|H\|_{q+1,p(q+1),\mathbf{\Theta}}^{q+1}$. So, we have

$$\begin{aligned}\|1-V\|_{q+1,p,\mathbf{\Theta}} & \leq C\big(\mathbb{P}_{\mathbf{\Theta}}(H>1/8)^{1/p} + \|DV\|_{q,p,\mathbf{\Theta}}\big) \\ & \leq C\big(\|H\|_{p,\mathbf{\Theta}} + \|H\|_{q+1,p(q+1),\mathbf{\Theta}}^{q+1}\big) \leq C\|H\|_{q+1,p(q+1),\mathbf{\Theta}}^{q+1}\end{aligned}$$

and by using (3.49) one gets

$$\|1-V\|_{q+1,p,\mathbf{\Theta}} \leq C\big(\mathrm{S}_{G,\mathbf{\Theta}}(\bar{p})^{2(q+2)}\mathrm{Q}_{F,G,\mathbf{\Theta}}(q+2,\bar{p})^{8d(q+1)}\|F-G\|_{q+2,\bar{p},\mathbf{\Theta}}^{q+2}\big)^{q+1}.$$

But $\|F-G\|_{q+2,\bar{p},\mathbf{\Theta}} \leq \mathrm{Q}_{F,G,\mathbf{\Theta}}(q+2,\bar{p})$, and we get

$$\begin{aligned}\left|\partial_\alpha p_{F,\mathbf{\Theta}(1-V)}(y)\right| \leq & C\big(\mathrm{S}_{F,\mathbf{\Theta}}(p') \vee \mathrm{S}_{G,\mathbf{\Theta}}(p')\big)^a\mathrm{Q}_{F,G,\mathbf{\Theta}}(p')^a\mathrm{m}_{q+2,p'}(\mathbf{\Theta})^a\|F-G\|_{q+2,p',\mathbf{\Theta}}\times \\ & \times\,\mathbb{P}_{\mathbf{\Theta}}(|F-x|<2)^b\end{aligned}$$

for $p' > d$ and suitable constants $C > 0$ and $a > 1$ depending on q, d. And similarly we get

$$\begin{aligned}\left|\partial_\alpha p_{G,\mathbf{\Theta}(1-V)}(y)\right| \leq & C\,\mathrm{S}_{G,\mathbf{\Theta}}(p')^a\mathrm{Q}_{F,G,\mathbf{\Theta}}(q+2,p')^a\mathrm{m}_{q+2,p'}(\mathbf{\Theta})^a\|F-G\|_{q+2,p',\mathbf{\Theta}}\times \\ & \times\,\mathbb{P}_{\mathbf{\Theta}}(|G-x|<2)^b,\end{aligned}$$

with the same constraints for p', C, a. The statement now follows. □

4 Stochastic equations with jumps

In this section we consider a jump type stochastic differential equation which has already been considered in [5]. It is closely related to piecewise deterministic Markov processes (in fact it is a particular case of this type of processes). We consider a Poisson point process p with state space $(E, \mathcal{B}(E))$, where $E = \mathbb{R}^d \times \mathbb{R}_+$. We refer to [14] for the notations. We denote by N the counting measure associated to p, we have $N([0,t) \times A) = \#\{0 \le s < t; p_s \in A\}$ for $t \ge 0$ and $A \in \mathcal{B}(E)$. We assume that the associated intensity measure is given by $\widehat{N}(dt, dz, du) = dt \times \mu(dz) \times 1_{[0,\infty)}(u)du$ where $(z,u) \in E = \mathbb{R}^d \times \mathbb{R}_+$ and $\mu(dz) = h(z)dz$.

We are interested in the solution to the d dimensional stochastic equation

$$X_t = x + \int_0^t \int_E c(z, X_{s-}) 1_{\{u < \gamma(z, X_{s-})\}} N(ds, dz, du) + \int_0^t g(X_s) ds. \tag{4.1}$$

We remark that the infinitesimal generator of the Markov process X_t is given by

$$L\psi(x) = g(x)\nabla\psi(x) + \int_{\mathbb{R}^d} (\psi(x + c(z,x)) - \psi(x)) K(x, dz)$$

where $K(x, dz) = \gamma(z,x)h(z)dz$ depends on the variable $x \in \mathbb{R}^d$. See [11] for the proof of existence and uniqueness of the solution to (4.1).

We describe now our approximation procedure. We consider a non-negative and smooth function $\varphi : \mathbb{R}^d \to \mathbb{R}_+$ such that $\varphi(z) = 0$ for $|z| > 1$ and $\int_{\mathbb{R}^d} \varphi(z)dz = 1$. And for $M \in \mathbb{N}$ we denote $\Phi_M(z) = \varphi * 1_{B_M}$ with $B_M = \{z \in \mathbb{R}^d : |z| < M\}$. Then $\Phi_M \in C_b^\infty$ and we have $1_{B_{M-1}} \le \Phi_M \le 1_{B_{M+1}}$. We denote by X_t^M the solution of the equation

$$X_t^M = x + \int_0^t \int_E c(z, X_{s-}^M) 1_{\{u < \gamma(z, X_{s-}^M)\}} \Phi_M(z) N(ds, dz, du) + \int_0^t g(X_s^M) ds. \tag{4.2}$$

In the following we will assume that $|\gamma(z,x)| \le \overline{C}$ for some constant $\overline{C}$. Let $N_M(ds, dz, du) := 1_{B_{M+1}}(z) \times 1_{[0, 2\overline{C}]}(u) N(ds, dz, du)$. Since $\{u < \gamma(z, X_{s-}^M)\} \subset \{u < 2\overline{C}\}$ and $\Phi_M(z) = 0$ for $|z| > M+1$, we may replace N by N_M in the above equation and consequently X_t^M is solution to the equation

$$X_t^M = x + \int_0^t \int_E c_M(z, X_{s-}^M) 1_{\{u < \gamma(z, X_{s-}^M)\}} N_M(ds, dz, du) + \int_0^t g(X_s^M) ds, \quad \text{with}$$
$$c_M(z,x) = \Phi_M(z) c(z,x).$$

Since the intensity measure $\widehat{N}_M$ is finite we may represent the random measure N_M by a compound Poisson process. Let $\lambda_M = 2\overline{C} \times \mu(B_{M+1}) = t^{-1}\mathbb{E}(N_M(t,E))$ and let J_t^M a Poisson process of parameter λ_M. We denote by $T_k^M, k \in \mathbb{N}$ the jump times of J_t^M. We also consider two sequences of independent random variables $(Z_k^M)_{k \in \mathbb{N}}$ in $\mathbb{R}^d$ and $(U_k)_{k \in \mathbb{N}}$ in $\mathbb{R}_+$ which are independent of J^M and such that

$$Z_k \sim \frac{1}{\mu(B_{M+1})} 1_{B_{M+1}}(z) d\mu(z) \quad \text{and} \quad U_k \sim \frac{1}{2\overline{C}} 1_{[0, 2\overline{C}]}(u) du.$$

To simplify the notation, we omit the dependence on M for the variables (T_k^M) and (Z_k^M). Then equation (4.2) may be written as

$$X_t^M = x + \sum_{k=1}^{J_t^M} c_M(Z_k, X_{T_k-}^M) 1_{(U_k, \infty)}(\gamma(Z_k, X_{T_k-}^M)) + \int_0^t g(X_s^M) ds. \tag{4.3}$$

In [5] it is proved that $X_t^M \to X_t$ in L^1. We study here the convergence in finite variation. Let us give our hypothesis.

Hypothesis 4.1. *We assume that γ, g, h and c are infinitely differentiable functions in both variables z and x. Moreover we assume that g and its derivatives are bounded and that $\ln h$ has bounded derivatives*

Hypothesis 4.2. *We assume that there exist two functions $\overline{\gamma}, \underline{\gamma} : \mathbb{R}^d \to \mathbb{R}_+$ and a constant $\overline{C}$ such that*

$$\overline{C} \geq \overline{\gamma}(z) \geq \gamma(z,x) \geq \underline{\gamma}(z) \geq 0, \quad \forall x \in \mathbb{R}^d$$

Hypothesis 4.3. ***i)*** *We assume that there exists a non negative and bounded function $\overline{c} : \mathbb{R}^d \to \mathbb{R}_+$ such that $\int_{\mathbb{R}^d} \overline{c}(z) d\mu(z) < \infty$ and*

$$\left\| \nabla_x c \times (I + \nabla_x c)^{-1}(z,x) \right\| + |c(z,x)| + \left| \partial_z^\beta \partial_x^\alpha c(z,x) \right| \leq \overline{c}(z) \quad \forall z, x \in \mathbb{R}^d.$$

ii) *There exists a non negative function $\underline{c} : \mathbb{R}^d \to \mathbb{R}_+$ such that for every $z \in \mathbb{R}^d$*

$$\sum_{r=1}^{d} \langle \partial_{z_r} c(z,x), \xi \rangle^2 \geq \underline{c}^2(z) \left| \xi \right|^2, \quad \forall \xi \in \mathbb{R}^d$$

and we assume that there exists $\theta > 0$ such that

$$\underline{\lim}_{a \to +\infty} \frac{1}{\ln a} \int_{\{\underline{c}^2 \geq 1/a\}} \underline{\gamma}(z) d\mu(z) = \theta. \tag{4.4}$$

Hypothesis 4.4. *We assume that*

$$\begin{aligned} &i) && \sup_{x,z} \sup_{1 \leq |\beta| \leq l} |\partial_{\beta,z} \ln \gamma(z,x)| < \infty, \\ &ii) && \sup_{z^* \in \mathbb{R}^d} \int_{B(z^*,1)} \overline{\gamma}(z) d\mu(z) < +\infty, \\ &iii) && \int_{\mathbb{R}^d} \overline{\gamma}_{\ln}^{x,l}(z) \overline{\gamma}(z) d\mu(z) < \infty \end{aligned}$$

with $\overline{\gamma}_{\ln}^{x,l}(z) = \sup_x \sup_{1 \leq |\beta| \leq l} |\partial_{\beta,x} \ln \gamma(z,x)|$.

We are now able to give our convergence result.

Theorem 4.5. *Suppose that Hypothesis 4.1-4.4 hold. Then for every $t > 0$ one has*

$$\lim_{M \to \infty} d_{TV}(X_t, X_t^M) = 0.$$

Proof. The proof is an easy consequence of the results from [5], we use the estimates obtained there.

Step 1. In [5] Lemma 4 one proves that $X_t^M \to X_t$ in L^1 and then $\lim_{M \to \infty} d_1(X_t, X_t^M) = 0$.

Step 2. Following [5], we consider an alternative representation of the law of X_t^M. The random variable X_t^M solution to (4.3) is a function of $(Z_1 \ldots, Z_{J_t^M})$ but it is not a simple functional, as defined in Section 2.1, because the coefficient $c_M(z,x) 1_{(u,\infty)}(\gamma(z,x))$ is not differentiable with respect to z. In order to avoid this difficulty we use the following alternative representation. Let $z_M^* \in \mathbb{R}^d$ such that $|z_M^*| = M + 3$. We define

$$\begin{aligned} q_M(z,x) &:= \varphi(z - z_M^*) \theta_{M,\gamma}(x) + \frac{1}{2\overline{C} \mu(B_{M+1})} 1_{B_{M+1}}(z) \gamma(z,x) h(z) \\ \theta_{M,\gamma}(x) &:= \frac{1}{\mu(B_{M+1})} \int_{\{|z| \leq M+1\}} \Big(1 - \frac{1}{2\overline{C}} \gamma(z,x)\Big) \mu(dz). \end{aligned}$$

We recall that φ is a non-negative and smooth function with $\int \varphi = 1$ and which is null outside the unit ball. Moreover since, $0 \leq \gamma(z,x) \leq \overline{C}$ and then $1 \geq \theta_{M,\gamma}(x) \geq 1/2$. By construction the function q_M satisfies $\int q_M(x,z)dz = 1$. Hence we can easily check (see [5] for a complete proof) that

$$\mathbb{E}(f(X^M_{T_k}) \mid X^M_{T_k-} = x) = \int_{\mathbb{R}^d} f(x + c_M(z,x))q_M(z,x)dz. \tag{4.5}$$

From the relation (4.5) we construct a process $(\overline{X}^M_t)$ equal in law to (X^M_t) in the following way. We denote by $\Psi_t(x)$ the solution of $\Psi_t(x) = x + \int_0^t g(\Psi_s(x))ds$. We assume that the times $T_k, k \in \mathbb{N}$ are fixed and we consider a sequence $(z_k)_{k\in\mathbb{N}}$ with $z_k \in \mathbb{R}^d$. Then we define $x_t, t \geq 0$ by $x_0 = x$ and, if x_{T_k} is given, then

$$\begin{aligned} x_t &= \Psi_{t-T_k}(x_{T_k}) \quad T_k \leq t < T_{k+1}, \\ x_{T_{k+1}} &= x_{T^-_{k+1}} + c_M(z_{k+1}, x_{T^-_{k+1}}). \end{aligned}$$

We remark that for $T_k \leq t < T_{k+1}, x_t$ is a function of $z_1, ..., z_k$. Notice also that x_t solves the equation

$$x_t = x + \sum_{k=1}^{J^M_t} c_M(z_k, x_{T^-_k}) + \int_0^t g(x_s)ds.$$

We consider now a sequence of random variables $(\overline{Z}_k), k \in \mathbb{N}^*$ and we denote $\mathcal{G}_k = \sigma(T_p, p \in \mathbb{N}) \vee \sigma(\overline{Z}_p, p \leq k)$ and $\overline{X}^M_t = x_t(\overline{Z}_1, ..., \overline{Z}_{J^M_t})$. We assume that the law of $\overline{Z}_{k+1}$ conditionally on $\mathcal{G}_k$ is given by

$$P(\overline{Z}_{k+1} \in dz \mid \mathcal{G}_k) = q_M(x_{T^-_{k+1}}(\overline{Z}_1, ..., \overline{Z}_k), z)dz = q_M(\overline{X}^M_{T^-_{k+1}}, z)dz.$$

Clearly $\overline{X}^M_t$ satisfies the equation

$$\overline{X}^M_t = x + \sum_{k=1}^{J^M_t} c_M(\overline{Z}_k, \overline{X}^M_{T_k-}) + \int_0^t g(\overline{X}^M_s)ds. \tag{4.6}$$

Notice that $\overline{X}^M_t$ is a piecewise deterministic Markov process, but not a completely general one because the intensity of the law of the jump times $T_k, k \in \mathbb{N}$ does not depend on the position of the particle $\overline{X}^M_t$. We think that the more general case may also be considered using similar arguments but we leave this out here.

Step 3. We will use the integration by parts formulae from Section 2.1 with the random variable $V = (V_1, ..., V_J)$ replaced by $(\overline{Z}_1, ..., \overline{Z}_{J^M_t})$ with fixed M and $t > 0$. We use the weight $\pi_{k,r} = \Phi_M(\overline{Z}_k), k \in \mathbb{N}, r = 1, ..., d$ and the Malliavin derivative is

$$D_{k,r} = \pi_{k,r}\partial_{\overline{Z}^r_k}.$$

In fact we will work conditionally to the time grid $T_k, k \in \mathbb{N}$ but all the constants coming on are independent of the time grid (as well as on M and $t > 0$) so we do not mention this in the notation.

We will use several estimates obtained in [5]. First, by Lemma 7 and Lemma 13 in [5], for every $p \geq 1, l \in \mathbb{N}$ we have

$$\mathbb{E}(|\overline{X}^M_t|^p_l) + \mathbb{E}(|L\overline{X}^M_t|^p_l) \leq C_{l,p} \tag{4.7}$$

with $C_{l,p} \in \mathcal{C}(d,l,p)$. So hypothesis (2.38) holds for every r.

We discuss now the non degeneracy property. We consider the tangent flow Y_t^M solution to

$$Y_t^M = I + \sum_{k=1}^{J_t^M} \nabla_x c_M(\overline{Z}_k, \overline{X}_{T_k-}^M) Y_{T_k-}^M + \int_0^t \nabla_x g(\overline{X}_s^M) Y_s^M ds. \tag{4.8}$$

Since $\left\|\nabla_x c \times (I+\nabla_x c)^{-1}(z,x)\right\| \leq \overline{c}(z)$ it follows that Y_t^M is invertible; we denote by $\widehat{Y}_t^M$ its inverse. Then it is proved in [5] that $D_{k,r}\overline{X}_t^{M,r'} = \pi_k(Y_t^M \nabla_z c_M(\overline{Z}_k, \overline{X}_{T_k^-}^M))_{r',r}$ and moreover, if λ_t denotes the lower eigenvalue of the Malliavin covariance matrix of $\overline{X}_t^M$ we have

$$\rho_t^M \geq \|\widehat{Y}_t^M\|^{-2} \sum_{k=1}^{J_t^M} 1_{B_{M-1}}(\overline{Z}_k) \underline{c}^2(\overline{Z}_k).$$

Then

$$\begin{aligned}
\mathbb{P}(\sigma_{\overline{X}_t^M} \leq \varepsilon) &\leq \mathbb{P}\big(\|\widehat{Y}_t^M\|^{-2} \sum_{k=1}^{J_t^M} 1_{B_{M-1}}(\overline{Z}_k) \underline{c}^2(\overline{Z}_k) \leq \varepsilon^{1/d}\big) \\
&\leq \mathbb{P}\Big(\sum_{k=1}^{J_t^M} 1_{B_{M-1}}(\overline{Z}_k) \underline{c}^2(\overline{Z}_k) \leq \varepsilon^{1/2d}\Big) + \mathbb{P}\big(\|\widehat{Y}_t^M\|^{-2} \leq \varepsilon^{1/2d}\big) \\
&\leq \mathbb{P}\Big(\sum_{k=1}^{J_t^M} \Phi_M(\overline{Z}_k) \underline{c}^2(\overline{Z}_k) \leq \varepsilon^{1/2d}\Big) + \mathbb{P}\big(\|Y_t^M\|^2 \geq \varepsilon^{1/2d}\big).
\end{aligned}$$

In [5] one proves that $\sup_M \mathbb{E}\|Y_t^M\|^{2p} < \infty$ for every $p \geq 1$ so that

$$\limsup_{\varepsilon\to 0} \limsup_{M\to 0} P(\left\|Y_t^M\right\|^2 \geq \varepsilon^{1/2d}) = 0.$$

One also proves in Lemma 5 from [5] that $\sum_{k=1}^{J_t^M} \Phi_M(\overline{Z}_k)\underline{c}^2(\overline{Z}_k)$ has the same law as

$$\sum_{k=1}^{J_t^M} \Phi_M(Z_k) \underline{c}^2(Z_k) 1_{[0,U_k]}\big(|\gamma(Z_k, X_{T_k-}^M)|\big),$$

so

$$\mathbb{P}\Big(\sum_{k=1}^{J_t^M} \Phi_M(\overline{Z}_k) \underline{c}^2(\overline{Z}_k) \leq \varepsilon^{1/2d}\Big) = \mathbb{P}\Big(\sum_{k=1}^{J_t^M} \Phi_M(Z_k) \underline{c}^2(Z_k) 1_{[0,U_k]}(\left|\gamma(Z_k, X_{T_k-}^M)\right|) \leq \varepsilon^{1/2d}\Big).$$

Let us denote

$$N_M(t) = \sum_{k=1}^{J_t^M} \Phi_M(Z_k) \underline{c}^2(Z_k) 1_{[0,U_k]}(\left|\gamma(Z_k, X_{T_k-}^M)\right|) \quad \text{and} \quad U_M(t) = t \int_{B_{M-1}^c} \underline{c}^2(z) \underline{\gamma}(z) d\mu(z).$$

In the final part of the proof of Lemma 16 in [5] one shows that if $p/t < \theta$ (with θ from (4.4)) then $\mathbb{E}((N_M(t)+U_M(t))^{-p}) \leq C_p$. Since $\lim_{M\to\infty} U_M(t) = 0$ one has for every fixed $\varepsilon > 0$

$$\limsup_{M\to\infty} \mathbb{P}(N_M(t) < \varepsilon) = \limsup_{M\to\infty} \mathbb{P}(N_M(t)+U_M(t) < \varepsilon) \leq \varepsilon^p \mathbb{E}((N_M(t)+U_M(t))^{-p}) \leq C_p \varepsilon^p.$$

So if $\theta > 0$ we take $p < \theta t$ and we obtain

$$\limsup_{\varepsilon\to 0} \limsup_{M\to\infty} \mathbb{P}(N_M(t) < \varepsilon) = 0$$

so that hypothesis (2.39) is also verified. Now the conclusion follows from Theorem 2.11. $\square$

References

[1] Bally, V.: *An elementary introduction to Malliavin calculus.* Research report INRIA n. 00071868, `http://hal.inria.fr/inria-00071868`, (2003).

[2] Bally, V. and Caramellino, L.: Riesz transform and integration by parts formulas for random variables. *Stochastic Process. Appl.* **121**, (2011), 1332–1355. MR-2794979

[3] Bally, V. and Caramellino, L.: Positivity and lower bounds for the density of Wiener functionals. *Potential Analysis*, **39**, (2013), 141–168. MR-3078335

[4] Bally, V., and Caramellino, L.: Asymptotic development for the CLT in total variation distance, arXiv:1407.0896

[5] Bally, V. and Clément, E. : Integration by parts formula and applications to equations with jumps. *Probab. Theory Related Fields*, **151**, (2011), 613–657. MR-2851695

[6] Bally, V. and Clément, E.: Integration by parts formula with respect to jump times for stochastic differential equations. *Stochastic analysis 2010*, (2011), 7–29, Springer, Heidelberg. MR-2789077

[7] Bally, V. and Talay, D.: The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Related Fields* **104**, (1996), 43-60. MR-1367666

[8] Bally, V. and Talay, D.: The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density. *Monte Carlo Methods Appl.* **2**, (1996), 93-128. MR-1401964

[9] Bichtler, K., Gravereaux, J.-B. and Jacod J.: *Malliavin calculus for processes with jumps*. Gordon and Breach Science Publishers, (1987). MR-1008471

[10] Bouleau, N. and Hirsch F.: Dirichlet forms and analysis on the Wiener space. In: *de Gruyter Studies in Math.*, vol. **14**, Walter de Gruyter, (1991). MR-1133391

[11] Fournier N.: Jumping SDE's: absolute continuity using monotonicity. *Stochastic Process. Appl.* **98**, (2002), 317-330. MR-1887538

[12] Guyon, J.: Euler scheme and tempered distributions. *Stochastic Process. Appl.* **116**, (2006), 877-904. MR-2254663

[13] Hu, Y., Lu, F. and Nualart, D.: Convergence of densities of some functionals of Gaussian processes. *J. Funct. Anal.* **266**, (2014), 814-875. MR-3132731

[14] Ikeda, N. and Watanabe, S.: *Stochastic Differential Equations and Diffusion processes*. North-Holland Mathematical Library 24, (1989). MR-1011252

[15] Malicet, D. and Poly, G.: Properties of convergence in Dirichlet structures. *J. Funct. Anal.* **264**, (2013), 2077–2096. MR-3029147

[16] Nourdin, I. and Peccati, G.: *Normal Approximations Using Malliavin Calculus: from Stein's Method to Universality*. Cambridge Tracts in Mathematics, 192, (2012). MR-2962301

[17] Nourdin, I. and Poly, G.: Convergence in total variation on Wiener chaos. *Stochastic Process. Appl.* **123**, (2013), 651–674. MR-3003367

[18] Nualart, D.: *The Malliavin calculus and related topics. Second Edition*. Springer-Verlag, (2006). MR-2200233

[19] Nourdin, I., Nualart, D. and Poly, G.: Absolute continuity and convergence of densities for random vectors on Wiener chaos. *Electron. J. Probab.* **18**, no. 22, (2013), 1-19. MR-3035750