HAL
open science

Convergence and regularity of probability laws by using an interpolation method

Vlad Bally, Lucia Caramellino

To cite this version:

Vlad Bally, Lucia Caramellino. Convergence and regularity of probability laws by using an interpolation method. Annals of Probability, 2017, 45 (2). hal-01109276v2

HAL Id: hal-01109276

https://hal.science/hal-01109276v2

Submitted on 8 Jan 2018

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

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

Convergence and regularity of probability laws by using an interpolation method

Vlad Bally*
Lucia Caramellino†

Abstract

Fournier and Printems [Bernoulli, 2010] have recently established a methodology which allows to prove the absolute continuity of the law of the solution of some stochastic equations with Hölder continuous coefficients. This is of course out of reach by using already classical probabilistic methods based on Malliavin calculus. By employing some Besov space techniques, Debussche and Romito [Probab. Theory Related Fields, 2014] have substantially improved the result of Fournier and Printems. In our paper we show that this kind of problem naturally fits in the framework of interpolation spaces: we prove an interpolation inequality (see Proposition 2.5) which allows to state (and even to slightly improve) the above absolute continuity result. Moreover it turns out that the above interpolation inequality has applications in a completely different framework: we use it in order to estimate the error in total variance distance in some convergence theorems.

AMS: 46B70, 60H07.

Keywords: Regularity of probability laws, Orlicz spaces, Hermite polynomials, interpolation spaces, Malliavin calculus, integration by parts formulas.

1 Introduction

In this paper we prove an interpolation type inequality which leads to three main applications. First we give a criteria for the regularity of the law μ of a random variable. This was the first aim of the integration by parts formulas constructed in the Malliavin calculus (in the Gaussian framework, and of many other variants of this calculus, in a more general case). But our starting point was the paper of N. Fournier and J. Printems [18] who noticed that some regularity of the law may be obtained even if no integration by parts formula holds for μ itself: they just use a sequence $\mu_n \to \mu$ and assume that an integration by parts formula of type $\int f' d\mu_n = \int f h_n d\mu_n$ holds for each μ_n. If $\sup_n \int |h_n| \, d\mu_n < \infty$ we are close to Malliavin calculus. But the interesting point is that one may obtain some regularity for μ even if $\sup_n \int |h_n| \, d\mu_n = \infty$ - so we are out of the domain of application of Malliavin calculus. The key point is that one establishes an equilibrium between the speed of convergence of $\mu_n \to \mu$ and the blow up $\int |h_n| \, d\mu_n \uparrow \infty$. The approach of Fournier and Printems is based on Fourier transforms and more

*Université Paris-Est, LAMA (UMR CNRS, UPEMLV, UPEC), MathRisk INRIA, F-77454 Marne-la-Vallée, France. Email: bally@univ-mlv.fr

†Dipartimento di Matematica, Università di Roma "Tor Vergata", and INDAM-GNAMPA, Via della Ricerca Scientifica 1, I-00133 Roma, Italy. Email: caramell@mat.uniroma2.it.

recently Debussche and Romito [11] obtained a much more powerful version of this type of criteria based on Besov space techniques. This methodology has been used in several recent papers (see [5], [6], [7], [12], [10] and [17]) in order to obtain the absolute continuity of the law of the solution of some stochastic equations with weak regularity assumptions on the coefficients: as a typical example, one proves that, under uniform ellipticity conditions, diffusion processes with Hölder continuous coefficients have absolute continuous law at any time $t > 0$. In the present paper we use a different approach, based on an interpolation argument and on Orlicz spaces, which allows one to go further and to treat, for example, diffusion processes with log-Hölder coefficients.

The second application concerns the regularity of the density with respect to a parameter. We illustrate this direction by giving sufficient conditions in order that $(x, y) \mapsto p_t(x, y)$ is smooth with respect to (x, y) where $p_t(x, y)$ is the density of the law of $X_t(x)$ which is a piecewise deterministic Markov process starting from x.

The third application concerns estimates of the speed of convergence $\mu_n \to \mu$ in total variation distance, and under some stronger assumptions, the speed of convergence of the derivatives of the densities of μ_n to the corresponding derivative of the density of μ. Such results appear in a natural way as soon as the suited interpolation framework is settled.

Let us give our main results. We work with the following weighted Sobolev norms on $C^\infty(\mathbb{R}^d; \mathbb{R})$:

$$\|f\|_{k,m,p} = \sum_{0\le|\alpha|\le k} \Big(\int (1 + |x|)^m \, |\partial_\alpha f(x)|^p \, dx \Big)^{1/p}, \qquad p > 1,$$

where α is a multi index, $|\alpha|$ denotes its length and ∂_α is the corresponding derivative. In the case $m = 0$ we have the standard Sobolev norm that we denote by $\|f\|_{k,p}$. We will also consider the weaker norm

$$\|f\|_{k,m,1+} = \sum_{0\le|\alpha|\le k} \int (1 + |x|)^m \, |\partial_\alpha f(x)| \, (1 + \ln^+ |x| + \ln^+ |f(x)|) dx,$$

with $\ln^+(x) = \max\{0, \ln |x|\}$. Moreover, for two measures μ and ν we consider the distances

$$d_k(\mu, \nu) = \sup \Big\{ \Big| \int \phi d\mu - \int \phi d\nu \Big| : \sum_{0\le|\alpha|\le k} \|\partial_\alpha \phi\|_\infty \le 1 \Big\}.$$

For $k = 0$ this is the total variation distance and for $k = 1$ this is the Fortet Mourier distance.

Our key estimate is the following. Let $m, q, k \in \mathbb{N}$ and $p > 1$ be given and let p_* be the conjugate of p. We consider a function $\phi \in C^{q+2m}(\mathbb{R}^d)$ and a sequence of functions $\phi_n \in C^{q+2m}(\mathbb{R}^d), n \in \mathbb{N}$ and we denote $\mu(dx) = \phi(x)dx$ and $\mu_n(dx) = \phi_n(x)dx$. We prove that there exists a universal constant C such that

$$\|\phi\|_{q,p} \le C \Big(\sum_{n=0}^{\infty} 2^{n(q+k+d/p_*)} d_k(\mu, \mu_n) + \sum_{n=0}^{\infty} \frac{1}{2^{2mn}} \|\phi_n\|_{q+2m,2m,p} \Big) \tag{1.1}$$

and

$$\|\phi\|_{q,1+} \le C \Big(\sum_{n=0}^{\infty} n 2^{n(q+k)} d_k(\mu, \mu_n) + \sum_{n=0}^{\infty} \frac{1}{2^{2mn}} \|\phi_n\|_{q+2m,2m,1+} \Big). \tag{1.2}$$

This is Proposition 2.5 and the proof is based on a development in Hermite series and on a powerful estimate for mixtures of Hermite kernels inspired from [27]. This inequality fits in the general theory

of interpolation spaces (we thank to D. Elworthy for a useful remark in this sense). Many interpolation results between Sobolev spaces of positive and negative indexes are known but they are not relevant from a probabilistic point of view: convergence in distribution is characterized by the Fortet Mourier distance and this amounts to convergence in the dual of $W^{1,\infty}$. So we are not concerned with Sobolev spaces associated to L^p norms but to L^∞ norms. This is a limit case which is more delicate and we have not found in the literature classical interpolation results which may be used in our framework.

Once we have (1.1) and (1.2) we obtain the following regularity criteria. Let μ be a finite non negative measure. Suppose that there exists a sequence of functions $\phi_n \in C^{q+2m}(\mathbb{R}^d), n \in \mathbb{N}$ such that

$$d_k(\mu, \mu_n) \times \|\phi_n\|^{\alpha}_{q+2m,2m,p} \leq C, \qquad \alpha > \frac{q+k+d/p_*}{2m}. \tag{1.3}$$

with $\mu_n(dx) = \phi_n(x)dx$. Then $\mu(dx) = \phi(x)dx$ and $\phi \in W^{q,p}$ (the standard Sobolev space).

In terms of $\|\phi\|_{q,m,1+}$ the statement is the following: suppose that there exists $m \in \mathbb{N}$ such that

$$d_1(\mu, \mu_n) \times \|\phi_n\|^{1/2m}_{2m,2m,1+} \leq \frac{C}{(\ln n)^{2+1/2m}}. \tag{1.4}$$

Then μ is absolutely continuous with respect to the Lebesgue measure.

The statement of the corresponding results are Theorem 2.10 **A** and Theorem 2.9 respectively. These are two significant particular cases of a more general result stated in terms of Orlicz norms in Theorem 2.6. The proof is, roughly speaking, as follows: let γ_ε be the Gaussian density of variance $\varepsilon > 0$ and let $\mu^\varepsilon = \mu * \gamma_\varepsilon$ and $\mu^\varepsilon_n = \mu_n * \gamma_\varepsilon$. Then $\mu^\varepsilon(dx) = \phi^\varepsilon(x)dx$ and $\mu^\varepsilon_n(x) = \phi^\varepsilon_n(x)dx$. Using (1.1) for ϕ^ε and $\phi^\varepsilon_n, n \in \mathbb{N}$ one proves that $\sup_\varepsilon \|\phi^\varepsilon\|_{q,p} < \infty$. And then one employs an argument of relative compactness in $W^{q,p}$ in order to produce the density ϕ of μ.

We give now the convergence result (see Theorem 2.11). Suppose that (1.3) holds for some $\alpha > \frac{q+k+d/p_*}{m}$. Then $\mu(dx) = \phi(x)dx$ and, for every $n \in \mathbb{N}$,

$$\|\phi - \phi_n\|_{W^{q,p}} \leq C d_k^\theta(\mu, \mu_n) \quad \text{with} \quad \theta = \frac{1}{\alpha} \wedge (1 - \frac{q+k+d/p_*}{\alpha m}). \tag{1.5}$$

Roughly speaking this inequality is obtained by using (1.1) with μ replaced by $\mu - \mu_n$.

In the statements of (1.3) we do not use $d_k(\mu, \mu_n)$ and $\|\phi_n\|_{q+2m,2m,p}$ directly, but some function λ having some nice properties such that $\lambda(1/n) \geq \|\phi_n\|_{1+q+2m,2m,p}$. But this is a technical point which we leave out in this introduction.

The paper is organized as follows. In Section 2 we introduce the Orlicz spaces, we give the general result and the criteria concerning the absolute continuity and the regularity of the density. We also give in Section 2.5 the convergence criteria mentioned above. In Section 2.6 we translate the results in terms of integration by parts formulae. In Section 3.1 (respectively Section 3.2) we prove absolute continuity for the law of the solution to a SDE (respectively to a SPDE) with log-Hölder continuous coefficients. Moreover, in Section 3.3 we discuss an example concerning piecewise deterministic Markov processes: we assume that the coefficients are smooth and we prove existence of the density of the law of the solution together with regularity with respect to the initial condition. We also consider an approximation scheme and we use (1.5) in order to estimate the error. Finally, we add some appendices containing technical results: Appendix A is devoted to the proof of the main estimate (1.1) based on a development in Hermite series; in Appendix B we discuss the relation with interpolation spaces; in Appendix C we give some auxiliary estimates concerning super kernels.

2 Criterion for the regularity of a probability law

2.1 Notations

We work on $\mathbb{R}^d$ and we denote by $\mathcal{M}$ the set of the finite signed measures on $\mathbb{R}^d$ with the Borel σ algebra. Moreover $\mathcal{M}_a \subset \mathcal{M}$ is the set of the measures which are absolutely continuous with respect to the Lebesgue measure. For $\mu \in \mathcal{M}_a$ we denote by p_μ the density of μ with respect to the Lebesgue measure. And for a measure $\mu \in \mathcal{M}$ we denote by L^p_μ the space of the measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ such that $\int |f|^p \, d|\mu| < \infty$. For $f \in L^1_\mu$ we denote $f\mu$ the measure $(f\mu)(A) = \int_A f d\mu$. For a bounded function $\phi : \mathbb{R}^d \to \mathbb{R}$ we denote $\mu * \phi$ the measure defined by $\int f d\mu * \phi = \int f * \phi d\mu = \int \int \phi(x-y) f(y) dy d\mu(x)$. Then $\mu * \phi \in \mathcal{M}_a$ and $p_{\mu*\phi}(x) = \int \phi(x-y) d\mu(y)$.

We denote by $\alpha = (\alpha_1, ..., \alpha_d) \in \mathbb{N}^d$ a multi index and we put $|\alpha| = \sum_{i=1}^d \alpha_i$. Here $\mathbb{N} = \{0, 1, 2, ...\}$ is the set of non negative integers and we put $\mathbb{N}_* = \mathbb{N} \setminus \{0\}$. For a multi index α with $|\alpha| = k$ we denote ∂_α the corresponding derivative that is $\partial^{\alpha_1}_{x_1} ... \partial^{\alpha_d}_{x_d}$ with the convention that $\partial^{\alpha_i}_{x_i} f = f$ if $\alpha_i = 0$. In particular if α is the null multi index then $\partial_\alpha f = f$.

We denote by $\|f\|_p = (\int |f(x)|^p \, dx)^{1/p}, p \geq 1$ and $\|f\|_\infty = \sup_{x \in \mathbb{R}^d} |f(x)|$. Then $L^p = \{f : \|f\|_p < \infty\}$ are the standard L^p spaces with respect to the Lebesgue measure.

2.2 Orlicz spaces

In the following we will work in Orlicz spaces, so we briefly recall the notations and the results we will use, for which we refer to [20].

A function $\mathbf{e} : \mathbb{R} \to \mathbb{R}_+$ is said to be a Young function if it is symmetric, strictly convex, non negative and $\mathbf{e}(0) = 0$. In the following we will consider Young functions having the two supplementary properties:

$$\begin{aligned} &i) \qquad \text{there exists } \lambda > 0 \text{ such that } \mathbf{e}(2s) \leq \lambda \mathbf{e}(s), \\ &ii) \qquad s \mapsto \frac{\mathbf{e}(s)}{s} \text{ is non decreasing.} \end{aligned} \tag{2.1}$$

The property $i)$ is known as the Δ_2 condition or doubling condition (see [20]). Through the whole paper we work with Young functions which satisfy (2.1). We set $\mathcal{E}$ the space of these functions:

$$\mathcal{E} = \{\mathbf{e} \, : \, \mathbf{e} \text{ is a Young function satisfying (2.1)}\}. \tag{2.2}$$

For $\mathbf{e} \in \mathcal{E}$ and $f : \mathbb{R}^d \to \mathbb{R}$, we define the norm

$$\|f\|_{\mathbf{e}} = \inf \Big\{ c > 0 : \int \mathbf{e}\Big(\frac{1}{c} f(x)\Big) dx \leq 1 \Big\}. \tag{2.3}$$

This is the so called Luxembourg norm which is equivalent to the Orlicz norm (see [20] p 227 Th 7.5.4). It is convenient for us to work with this norm (instead of the Orlicz norm). The space $L^{\mathbf{e}} = \{f : \|f\|_{\mathbf{e}} < \infty\}$ is the Orlicz space.

Remark 2.1. *Let $u_l(x) = (1 + |x|)^{-l}$. As a consequence of (2.1) ii), for every $\mathbf{e} \in \mathcal{E}$ and $l > d$ one has $u_l \in L^{\mathbf{e}}$ and moreover,*

$$\|u_l\|_{\mathbf{e}} \leq (\mathbf{e}(1) \|u_l\|_1) \vee 1 < \infty. \tag{2.4}$$

Indeed (2.1) ii) implies that for $t \leq 1$ one has $\mathbf{e}(t) \leq \mathbf{e}(1)t$. For $c \geq (\mathbf{e}(1) \|u_l\|_1) \vee 1$ one has $\frac{1}{c} u_l(x) \leq u_l(x) \leq 1$ so that

$$\int \mathbf{e}\Big(\frac{1}{c} u_l(x)\Big) dx \leq \frac{\mathbf{e}(1)}{c} \int u_l(x) dx = \frac{\mathbf{e}(1)}{c} \|u_l\|_1 \leq 1.$$

For $a > 0$, we define $\mathbf{e}^{-1}(a) = \sup\{c : \mathbf{e}(c) \leq a\}$ and:

$$\phi_{\mathbf{e}}(r) = \frac{1}{\mathbf{e}^{-1}\left(\frac{1}{r}\right)} \quad \text{and} \quad \beta_{\mathbf{e}}(R) = \frac{R}{\mathbf{e}^{-1}(R)} = R\phi_{\mathbf{e}}\Big(\frac{1}{R}\Big), \quad r, R > 0. \tag{2.5}$$

Remark 2.2. *The function $\phi_{\mathbf{e}}$ is the "fundamental function" of $L^{\mathbf{e}}$ equipped with the Luxembourg norm (see [9] Lemma 8.17 pg 276). In particular $\frac{1}{r}\phi_{\mathbf{e}}(r)$ is decreasing (see [9] Corollary 5.2 pg 67). It follows that $\beta_{\mathbf{e}}$ is increasing. For the sake of completeness we give here the argument. By (2.1), ii), if $a > 1$ then $\mathbf{e}(ax) \geq a\mathbf{e}(x)$ so that $ax \geq \mathbf{e}^{-1}(a\mathbf{e}(x))$. Taking $y = \mathbf{e}(x)$ we obtain $a\mathbf{e}^{-1}(y) \geq \mathbf{e}^{-1}(ay)$ which gives*

$$\beta_{\mathbf{e}}(ay) = \frac{ay}{\mathbf{e}^{-1}(ay)} \geq \frac{ay}{a\mathbf{e}^{-1}(y)} = \beta_{\mathbf{e}}(y).$$

One defines the conjugate of $\mathbf{e}$ by

$$\mathbf{e}_*(s) = \sup\{st - \mathbf{e}(t) : t \in \mathbb{R}\}.$$

$\mathbf{e}_*$ is a Young function as well, so the corresponding Luxembourg norm $\|f\|_{\mathbf{e}_*}$ is given by (2.3) with $\mathbf{e}$ replaced by $\mathbf{e}_*$. And one has the following Hölder inequality:

$$\left|\int fg(x)dx\right| \leq 2\|f\|_{\mathbf{e}}\|g\|_{\mathbf{e}_*}. \tag{2.6}$$

(see Theorem 7.2.1 page 215 in [20]; we stress that the factor 2 does not appear in that reference but in the right hand side of the inequality in the statement of Theorem 7.2.1 in [20] one has the Orlicz norm of g and by using the equivalence between the Orlicz and the Luxembourg norm we can replace the Orlicz norm by $2\|g\|_{\mathbf{e}_*}$).

We will now define Sobolev norms and Sobolev spaces associated to an Young function $\mathbf{e}$. Let us denote by L^1_{loc} the space of measurable functions which are integrable on compact sets and by $W^{k,1}_{loc}$ the space of measurable functions which are k times weakly differentiable and have locally integrable derivatives. More precisely this means that $f \in W^{k,1}_{loc}$ if for every multi index α with $|\alpha| \leq k$ one may find a function $f_\alpha \in L^1_{loc}$ (determined dx almost surely) such that $\int g(x)f_\alpha(x)dx = (-1)^{|\alpha|}\int \partial_\alpha g(x)f(x)dx$ for every $g \in C^\infty_c(\mathbb{R}^d, \mathbb{R})$. In this case we denote $\partial_\alpha f = f_\alpha$. Notice that

$$L^{\mathbf{e}} \subset L^1_{loc}. \tag{2.7}$$

In order to prove this we take $R > 0$ and we notice that for $|x| \leq R$ one has $(1+R)^{d+1}u_{d+1}(x) \geq 1$. Then using (2.6) and (2.4), for every $f \in L^{\mathbf{e}}$

$$\begin{aligned}\int_{B_R} |f(x)|\,dx &\leq (1+R)^{d+1}\int_{R^d} u_{d+1}(x)\,|f(x)|\,dx \\ &\leq (1+R)^{d+1}\|u_{d+1}\|_{e_*}\|f\|_e < \infty.\end{aligned}$$

For $f \in W^{k,1}_{loc}$, we introduce the norms

$$\|f\|_{k,\mathbf{e}} = \sum_{0\leq|\alpha|\leq k}\|\partial_\alpha f\|_{\mathbf{e}} \quad \text{and} \quad \|f\|_{k,\infty} = \sum_{0\leq|\alpha|\leq k}\|\partial_\alpha f\|_\infty \tag{2.8}$$

and we denote

$$W^{k,\mathbf{e}} = \{f \in W^{k,1}_{loc} : \|f\|_{k,\mathbf{e}} < \infty\} \text{ and } W^{k,\infty} = \{f \in W^{k,1}_{loc} : \|f\|_{k,\infty} < \infty\}.$$

For a multi index γ we denote $x^\gamma = \prod_{i=1}^d x_i^{\gamma_i}$ and for two multi indexes α, γ we denote $f_{\alpha,\gamma}$ the function

$$f_{\alpha,\gamma}(x) = x^\gamma \partial_\alpha f(x).$$

Then we consider the norms

$$\|f\|_{k,l,\mathbf{e}} = \sum_{0\le|\alpha|\le k} \sum_{0\le|\gamma|\le l} \|f_{\alpha,\gamma}\|_{\mathbf{e}} \quad \text{and} \quad W^{k,l,\mathbf{e}} = \{f : \|f\|_{k,l,\mathbf{e}} < \infty\}. \tag{2.9}$$

We stress that in $\|\cdot\|_{k,l,\mathbf{e}}$ the first index k is related to the order of the derivatives which are involved while the second index l is connected to the power of the polynomial multiplying the function and its derivatives up to order k.

Finally we recall that if $\mathbf{e}$ satisfies the Δ_2 condition (that is (2.1) i)) then $L^{\mathbf{e}}$ is reflexive (see [20], Theorem 7.7.1, p 234). In particular, in this case, any bounded subset of $L^{\mathbf{e}}$ is weakly relatively compact.

Let us propose two examples of Young functions, that represent the leading ones in our approach.

Example 1. If we take $\mathbf{e}_p(x) = |x|^p, p > 1$, then $\|f\|_{\mathbf{e}_p}$ is the usual L^p norm and the corresponding Orlicz space is the standard L^p space on $\mathbb{R}^d$. Clearly $\beta_{\mathbf{e}_p}(t) = t^{1/p_*}$ with p_* the conjugate of p.

Example 2. Set $\mathbf{e}_{\log}(t) = (1 + |t|)\ln(1 + |t|)$. Since the norm from $\mathbf{e}_{\log}$ is not explicit we replace it by the following quantities:

$$\begin{aligned} \|f\|_{p,1+} &= \int (1+|x|)^p |f(x)| (1 + \ln^+|x| + \ln^+|f(x)|)dx \\ \|f\|_{k,p,1+} &= \sum_{0\le|\alpha|\le k} \|\partial_\alpha f\|_{p,1+} \end{aligned} \tag{2.10}$$

with $\ln^+(x) = \max\{0, \ln|x|\}$. We stress that $\|f\|_{p,1+}$ is not a norm.

We will need the following:

Lemma 2.3. *For each $k \in \mathbb{N}$ and $p \ge 0$ there exists a constant C depending on k, p only such that*

$$\|f\|_{k,p,\mathbf{e}_{\log}} \le C(1 \vee \|f\|_{k,p,1+}). \tag{2.11}$$

Moreover

$$\limsup_{t\to\infty} \frac{\beta_{\mathbf{e}_{\log}}(t)}{\ln t} \le 2. \tag{2.12}$$

Proof. The inequality (2.11) is an immediate consequence of the following simpler one:

$$\|f\|_{\mathbf{e}_{\log}} \le 2\Big(1 \vee \int |f(x)| (1 + \ln^+|f(x)|)dx\Big). \tag{2.13}$$

Let us prove it. We assume that $f \ge 0$ and we take $c \ge 2$ and we write

$$\int \mathbf{e}_{\log}\Big(\frac{1}{c}f(x)\Big)dx \le \int_{\{f\le c\}} \mathbf{e}_{\log}\Big(\frac{1}{c}f(x)\Big)dx + \int_{\{f>c\}} \mathbf{e}_{\log}\Big(\frac{1}{c}f(x)\Big)dx =: I + J.$$

Using the inequality $\ln(1+y) \le y$ we obtain $I \le 2\int \ln(1+\frac{1}{c}f) \le \frac{2}{c}\int f$. And if $f \ge c \ge 2$ then $\frac{f}{c}+1 \le \frac{2}{c}f \le f$. Then $\mathbf{e}_{\log}(\frac{1}{c}f(x)) \le \frac{2}{c}f\ln f$. It follows that $J \le \frac{2}{c}\int_{\{f>c\}} f\ln^+ f$ and finally $\int \mathbf{e}_{\log}(\frac{1}{c}f)) \le \frac{2}{c}\int_{\{f>c\}}(1+f)\ln^+ f$. We conclude that for $c \ge 2\int f(1+\ln^+ f)$ we have $\int \mathbf{e}_{\log}(\frac{1}{c}f) \le 1$ which by the very definition means that $\|f\|_{\mathbf{e}_{\log}} \le 2\int f(1+\ln^+ f)$.

Let us prove (2.12). We denote $e(t) = 2t\ln(2t)$ and we notice that for large t one has $\mathbf{e}_{\log}(t) \le e(t)$. It follows that

$$\beta_{\mathbf{e}_{\log}}(t) \le \frac{t}{e^{-1}(t)}.$$

Using the change of variable $R = e(t)$ we obtain

$$\lim_{R\to\infty} \frac{R}{e^{-1}(R)\ln R} = \lim_{t\to\infty} \frac{e(t)}{t\ln e(t)} = 2.$$

So for large R we have $\beta_{\mathbf{e}_{\log}}(R) \le R/e^{-1}(R) \le 2\ln R$. □

Remark 2.4. *We recall that the LlogL space of Zygmund is the space of the functions f such that $\int |f(x)|\ln^+|f(x)|\,dx < \infty$ (see [9]). Then $L^{\mathbf{e}_{\log}} = L^1 \cap LlogL$. The inequality (2.13) already gives one inclusion. The converse inclusion is a consequence of the following inequalities. Let $\varepsilon_* > 0$ be such that $t \le 2\ln(1+t)$ for $0 < t \le \varepsilon_*$ and let $C_* = 2 + 1/\ln(1+\varepsilon_*)$. Then*

$$\begin{aligned} &i) \qquad \int |f(x)|\,dx \le C_*\,\|f\|_{\mathbf{e}_{\log}} \quad \text{and} \\ &ii) \qquad \int |f(x)|\ln^+|f(x)|\,dx \le \|f\|_{\mathbf{e}_{\log}}(1+2C_*\ln^+\|f\|_{\mathbf{e}_{\log}}). \end{aligned} \tag{2.14}$$

In order to prove i) we denote $g = \|f\|_{\mathbf{e}_{\log}}^{-1}|f|$ and we write

$$\begin{aligned} \int g &= \int_{\{g\le\varepsilon_*\}} g + \int_{\{g>\varepsilon_*\}} g \\ &\le 2\int_{\{g\le\varepsilon_*\}} \ln(1+g) + \frac{1}{\ln(1+\varepsilon_*)}\int_{\{g>\varepsilon_*\}} g\ln(1+g) \\ &\le C_*\int (1+g)\ln(1+g) = C_*\int \mathbf{e}_{\log}(g) = C_*. \end{aligned}$$

In order to prove ii) we notice that $\int g\ln^+ g \le \int \mathbf{e}_{\log}(g) = 1$ so that

$$\int |f|\ln^+\frac{|f|}{\|f\|_{\mathbf{e}_{\log}}} \le \|f\|_{\mathbf{e}_{\log}}.$$

Then we write

$$\begin{aligned} \int |f|\ln^+|f| &= \int_{\{|f|\ge 1\vee\|f\|_{\mathbf{e}_{\log}}\}} |f|\ln^+|f| + \int_{\{|f|<1\vee\|f\|_{\mathbf{e}_{\log}}\}} |f|\ln^+|f| \\ &=: I + J. \end{aligned}$$

If $|f| \geq 1 \vee \|f\|_{\mathbf{e}_{\log}}$ *then* $\ln^+ |f| = \ln |f| = \ln^+(\frac{|f|}{\|f\|_{\mathbf{e}_{\log}}}) + \ln \|f\|_{\mathbf{e}_{\log}}$. *So, by using the previous inequality,*

$$I \leq \|f\|_{\mathbf{e}_{\log}} + \ln \|f\|_{\mathbf{e}_{\log}} \int |f| \leq \|f\|_{\mathbf{e}_{\log}} (1 + C_* \ln \|f\|_{\mathbf{e}_{\log}})$$

the last inequality being a consequence of i). And

$$J \leq \ln^+ \|f\|_{\mathbf{e}_{\log}} \int |f| \leq C_* \|f\|_{\mathbf{e}_{\log}} \ln^+ \|f\|_{\mathbf{e}_{\log}} .$$

2.3 Main results

We consider the following distances between two measures $\mu, \nu \in \mathcal{M}$: for $k \in \mathbb{N}$, we set

$$d_k(\mu, \nu) = \sup \Big\{ \Big| \int \phi d\mu - \int \phi d\nu \Big| : \phi \in C^\infty(\mathbb{R}^d), \|\phi\|_{k,\infty} \leq 1 \Big\}. \tag{2.15}$$

Notice that d_0 is the total variation distance and d_1 is the bounded variation distance (also called Fortet Mourier distance) which is related to the convergence in law of probability measures. The distances d_k with $k \geq 2$ are less often used. We mention however that people working in approximation theory (for diffusion process for example - see [30] or [24]) use such distances in an implicit way: indeed, they study the speed of convergence of certain schemes but they are able to obtain their estimates for test functions $f \in C^k$ with k sufficiently large - so d_k comes on. We also recall that for $k = 1, 2, 3$, d_k plays an important role in the so-called Stein's method for normal approximation (see e.g. [25]).

We fix now a Young function $\mathbf{e} \in \mathcal{E}$ (see (2.2)), and we recall the function $\beta_{\mathbf{e}}$ (see (2.5) and Remark 2.2 respectively).

Let $q, k \in \mathbb{N}$ and $m \in \mathbb{N}_*$. For $\mu \in \mathcal{M}$ and for a sequence $\mu_n \in \mathcal{M}_a, n \in \mathbb{N}$ we define

$$\pi_{q,k,m,\mathbf{e}}(\mu, (\mu_n)_n) = \sum_{n=0}^{\infty} 2^{n(q+k)} \beta_{\mathbf{e}}(2^{nd}) d_k(\mu, \mu_n) + \sum_{n=0}^{\infty} \frac{1}{2^{2nm}} \|p_{\mu_n}\|_{2m+q,2m,\mathbf{e}} . \tag{2.16}$$

Here and in the sequel we make the convention that $\|p_{\mu_n}\|_{2m+q,2m,\mathbf{e}} = \infty$ if $p_{\mu_n} \notin W^{2m+q,1}_{loc}$. Moreover we define

$$\rho_{q,k,m,\mathbf{e}}(\mu) = \inf \pi_{q,k,m,\mathbf{e}}(\mu, (\mu_n)_n) \tag{2.17}$$

the infimum being over all the sequences of measures $\mu_n, n \in \mathbb{N}$ which are absolutely continuous. It is easy to check that $\rho_{q,k,m,\mathbf{e}}$ is a norm on the space $\mathcal{S}_{q,k,m,\mathbf{e}}$ defined by

$$\mathcal{S}_{q,k,m,\mathbf{e}} = \{\mu \in \mathcal{M} : \rho_{q,k,m,\mathbf{e}}(\mu) < \infty\}. \tag{2.18}$$

The following result gives the key estimate in our paper. We prove it in Appendix A.

Proposition 2.5. *Let* $q, k \in \mathbb{N}, m \in \mathbb{N}_*$ *and* $\mathbf{e} \in \mathcal{E}$. *There exists a universal constant* C *(depending on* q, k, m, d *and* $\mathbf{e}$*) such that for every* $f \in C^{2m+q}(\mathbb{R}^d)$ *one has*

$$\|f\|_{q,\mathbf{e}} \leq C \rho_{q,k,m,\mathbf{e}}(\mu) \tag{2.19}$$

where $\mu(dx) = f(x)dx$.

We state now our main theorem:

Theorem 2.6. *Let $q, k \in \mathbb{N}, m \in \mathbb{N}_*$ and let $\mathbf{e} \in \mathcal{E}$.*

i) Take $q = 0$. Then

$$\mathcal{S}_{0,k,m,\mathbf{e}} \subset L^{\mathbf{e}}$$

in the sense that if $\mu \in \mathcal{S}_{0,k,m,\mathbf{e}}$ then μ is absolutely continuous and the density p_μ belongs to $L^{\mathbf{e}}$. Moreover there exists a universal constant C such that

$$\|p_\mu\|_{L^{\mathbf{e}}} \leq C\rho_{0,k,m,\mathbf{e}}(\mu).$$

ii) Take $q \geq 1$. Then

$$\mathcal{S}_{q,k,m,\mathbf{e}} \subset W^{q,\mathbf{e}} \quad \text{and} \quad \|p_\mu\|_{q,\mathbf{e}} \leq C\rho_{q,k,m,\mathbf{e}}(\mu), \quad \mu \in \mathcal{S}_{q,k,m,\mathbf{e}}.$$

Proof. We consider a function $\phi \in C_b^\infty(\mathbb{R}^d)$ such that $0 \leq \phi \leq 1_{B_1}$ and $\int_{\mathbb{R}} \phi(x)dx = 1$. For $\delta \in (0,1)$, we define $\phi_\delta(x) = \delta^{-d}\phi(\delta^{-1}x)$. For a measure μ we define $\mu * \phi_\delta$ by $\int f d\mu * \phi_\delta = \int f * \phi_\delta d\mu$. Since $\|f * \phi_\delta\|_{k,\infty} \leq \|f\|_{k,\infty}$ it follows that $d_k(\mu * \phi_\delta, \nu * \phi_\delta) \leq d_k(\mu, \nu)$. We will also prove that

$$\|f * \phi_\delta\|_{2m+q,2m,\mathbf{e}} \leq 2^{2m} \|f\|_{2m+q,2m,\mathbf{e}}. \tag{2.20}$$

Suppose for a moment that (2.20) holds. Then

$$\pi_{q,k,m,\mathbf{e}}(\mu * \phi_\delta, (\mu_n * \phi_\delta)_n) \leq 2^{2m}\pi_{q,k,m,\mathbf{e}}(\mu, (\mu_n)_n) \leq 2^{2m}\rho_{q,k,m,\mathbf{e}}(\mu).$$

Let $p_\delta \in C^\infty(\mathbb{R}^d, \mathbb{R})$ be the density of the measure $\mu * \phi_\delta$. The above inequality and (2.19) prove that

$$\sup_{0<\delta\leq 1} \|p_\delta\|_{q,e} \leq C\rho_{q,k,m,\mathbf{e}}(\mu) < \infty.$$

For each multi index α with $|\alpha| \leq q$ the family $\partial_\alpha p_\delta, \delta \in (0,1)$ is bounded in $L^{\mathbf{e}}$ which is a reflexive space, so it is weakly relatively compact. Then we may find $p_\alpha \in L^{\mathbf{e}} \subset L^1_{loc}$ (see (2.7) for the above inclusion) and a sequence $\delta_n \to 0$ such that $\partial_\alpha p_{\delta_n} \to p_\alpha$ weakly, for every multi index α with $0 \leq |\alpha| \leq q$ (in the same time). Since $\int g\partial_\alpha p_{\delta_n} = (-1)^{|\alpha|}\int p\partial_\alpha g$, by passing to the limit we obtain $\int gp_\alpha = (-1)^{|\alpha|}\int p\partial_\alpha g$ so $\partial_\alpha p = p_\alpha \in L^{\mathbf{e}}$ and this means that $p \in W^{q,\mathbf{e}}$. Since $\mu * \phi_{\delta_n} \to \mu$ weakly, one has $\mu(dx) = p(x)dx$. And since $\|\partial_\alpha p\|_{\mathbf{e}} \leq \sup_{n\in\mathbb{N}} \|\partial_\alpha p_{\delta_n}\|_{\mathbf{e}} \leq C\rho_{q,k,m,\mathbf{e}}(\mu)$ it follows that $\|p\|_{q,\mathbf{e}} \leq C\rho_{q,k,m,\mathbf{e}}(\mu)$. So the proof is completed.

Let us check (2.20). For $\lambda > 0$ we denote $g_\lambda(x) = (1 + |x|)^\lambda g(x)$. Notice that for $\delta \leq 1$

$$\begin{aligned}
|(g * \phi_\delta)_\lambda(x)| &\leq (1 + |x|)^\lambda \int |g(x-y)|\, \phi_\delta(y)dy \\
&\leq \int (1 + |x-y| + \delta)^\lambda |g(x-y)|\, \phi_\delta(y)dy \\
&\leq 2^\lambda \int (1 + |x-y|)^\lambda |g(x-y)|\, \phi_\delta(y)dy = 2^\lambda |g_\lambda| * \phi_\delta(x).
\end{aligned}$$

Then, by (A.6) $\|(g * \phi_\delta)_\lambda\|_{\mathbf{e}} \leq 2^\lambda \||g_\lambda| * \phi_\delta\|_{\mathbf{e}} \leq 2^\lambda \|\phi_\delta\|_1 \||g_\lambda|\|_{\mathbf{e}} = 2^\lambda \|g_\lambda\|_{\mathbf{e}}$. Using this inequality (with $\lambda = 2m$) for $g = \partial_\alpha f$ we obtain (2.20). ∎

We consider now a special class of Orlicz norms which verify a supplementary condition: given $\alpha, \gamma \geq 0$ we define

$$\mathcal{E}_{\alpha,\gamma} = \Big\{ \mathbf{e} : \limsup_{R\to\infty} \frac{\beta_{\mathbf{e}}(R)}{R^{\alpha}(\ln R)^{\gamma}} < \infty \Big\}. \tag{2.21}$$

In this case we have:

Theorem 2.7. *Let $q, k \in \mathbb{N}, m \in \mathbb{N}_*$ and let $\mathbf{e} \in \mathcal{E}_{\alpha,\gamma}$. If $2m > d$, $\gamma \geq 0$ and $0 \leq \alpha < \frac{2m+q+k}{d(2m-1)}$ then*

$$W^{q+1,2m,\mathbf{e}} \subset \mathcal{S}_{q,k,m,\mathbf{e}} \subset W^{q,\mathbf{e}}$$

and there exists some constant C such that

$$\frac{1}{C} \left\| p_\mu \right\|_{q,\mathbf{e}} \leq \rho_{q,k,m,\mathbf{e}}(\mu) \leq C \left\| p_\mu \right\|_{q+1,2m,\mathbf{e}}. \tag{2.22}$$

In particular this is true for $\mathbf{e}_{\log}$ and for $\mathbf{e}_p$ with $\frac{p-1}{p} < \frac{2m+q+k}{d(2m-1)}$.

Proof. The first inequality in (2.22) is proved in Theorem 2.6. As for the second, we use Lemma C.3 in Appendix C. Let $f \in W^{q+1,2m,\mathbf{e}}$ and $\mu_f(dx) = f(x)dx$. We have to prove that $\rho_{q,k,m,\mathbf{e}}(\mu_f) < \infty$. We consider a super kernel ϕ (see (C.1)) and we define $f_\delta = f * \phi_\delta$. We take $\delta_n = 2^{-\theta n}$ with θ to be chosen in a moment and we choose n_* such that for $n \geq n_*$ one has $\beta_{\mathbf{e}}(2^{nd}) \leq C 2^{nd\alpha} n^{\gamma}$ because $\mathbf{e} \in \mathcal{E}_{\alpha,\gamma}$. Using (C.3) with $l = 2m$, we obtain $d_k(\mu_f, \mu_{f_{\delta_n}}) \leq C_{k,q} \left\| f \right\|_{q+1,2m,\mathbf{e}} \delta_n^{q+k+1}$ and using (C.4) we obtain $\left\| f_{\delta_n} \right\|_{2m+q,2m,\mathbf{e}} \leq C_{2m+q,2m} \left\| f \right\|_{q+1,2m,\mathbf{e}} \delta_n^{-(2m-1)}$ (the constant C depends on k and q, which are fixed). Then we can write

$$\begin{aligned}
&\pi_{q,k,m,\mathbf{e}}(\mu_f, \mu_{f_{\delta_n}}) \\
&= \sum_{n=0}^{\infty} 2^{n(q+k)} \beta_{\mathbf{e}}(2^{nd}) d_k(\mu_f, \mu_{f_{\delta_n}}) + \sum_{n=0}^{\infty} \frac{1}{2^{2nm}} \left\| f_{\delta_n} \right\|_{2m+q,2m,\mathbf{e}} \\
&\leq C_{q,k,m} \left\| f \right\|_{q+1,2m,\mathbf{e}} \times \\
&\qquad \times \Big(1 + \sum_{n \geq n_*}^{\infty} 2^{n(q+k+d\alpha-\theta(q+k+1))} n^{\gamma} + \sum_{n=0}^{\infty} \frac{1}{2^{n(2m-\theta(2m-1))}} \Big),
\end{aligned}$$

$C_{q,k,m} > 0$ denoting a constant depending on q, k, m. In order to obtain the convergence of the above series we need to choose θ such that

$$\frac{q+k+d\alpha}{q+k+1} < \theta < \frac{2m}{2m-1}$$

and this is possible under our restriction on α. □

We give now a criterion in order to check that $\mu \in \mathcal{S}_{q,k,m,\mathbf{e}}$.

Theorem 2.8. *Let $q, k \in \mathbb{N}, m \in \mathbb{N}_*$ and let $\mathbf{e} \in \mathcal{E}_{\alpha,\gamma}$. We consider a non negative finite measure μ and we suppose that there exists a family of measures $\mu_\delta(dx) = f_\delta(x)dx, \delta > 0$ which verifies the following assumptions. There exist $C, r > 0$ and a function $\lambda_{q,m}(\delta)$, $\delta \in (0,1)$, which is right-continuous and non increasing such that*

$$\|f_\delta\|_{2m+q,2m,\mathbf{e}} \leq \lambda_{q,m}(\delta) \leq C\delta^{-r}.$$

We consider some $\eta > 0$ and $\kappa \geq 0$ and we assume that

$$\lambda^{\eta}_{q,m}(\delta) d_k(\mu, \mu_\delta) \leq \frac{C}{(\ln(1/\delta))^{\kappa}}. \tag{2.23}$$

If (2.23) holds with

$$\eta > \frac{q+k+\alpha d}{2m}, \qquad \kappa = 0 \tag{2.24}$$

then

$$\mu \in \mathcal{S}_{q,k,m,\mathbf{e}} \subset W^{q,\mathbf{e}}.$$

The same conclusion holds if

$$\eta = \frac{q+k+\alpha d}{2m} \qquad \textit{and} \qquad \kappa > 1 + \gamma + \eta. \tag{2.25}$$

Proof. Let $\varepsilon_0 > 0$. We define

$$\delta_n = \inf\{\delta > 0 : \lambda_{q,m}(\delta) \leq \frac{2^{2mn}}{n^{1+\varepsilon_0}}\}.$$

Let $0 < \theta < 2m/r$ where r is the one in the growth condition on $\lambda_{q,m}$. Since $\delta^r \lambda_{q,m}(\delta) \leq C$, we have

$$\lambda_{q,m}(2^{-\theta n}) \leq C 2^{n\theta r} \leq \frac{2^{2mn}}{n^{1+\varepsilon_0}}$$

which means that $\delta_n \leq 2^{-\theta n}$. Since $\mathbf{e} \in \mathcal{E}_{\alpha,\gamma}$ we have

$$\begin{aligned} &\pi_{q,k,m,\mathbf{e}}(\mu, (\mu_{\delta_n})_n) \\ &\leq C \sum_{n=1}^{\infty} 2^{n(q+k+\alpha d)} n^{\gamma} d_k(\mu, \mu_{\delta_n}) + C \sum_{n=1}^{\infty} 2^{-2mn} \left\| f_{\delta_n} \right\|_{2m+q,2m,\mathbf{e}}. \end{aligned}$$

Since $\lambda_{q,m}$ is right continuous, $\lambda_{q,m}(\delta_n) = 2^{2mn} n^{-(1+\varepsilon_0)}$ so

$$\sum_{n=1}^{\infty} \frac{1}{2^{2mn}} \lambda_{q,m}(\delta_n) < \infty.$$

By recalling that $\ln(1/\delta_n) \geq C\theta n$ and by using (2.23), we obtain

$$\begin{aligned} 2^{n(q+k+\alpha d)} n^{\gamma} d_k(\mu, \mu_{\delta_n}) &\leq 2^{n(q+k+\alpha d)} \tfrac{C n^{\gamma}}{\lambda^{\eta}_{q,m}(\delta_n)(\ln(1/\delta_n))^{\kappa}} \\ &\leq C \times 2^{n(q+k+\alpha d - 2m\eta)} n^{\gamma + \eta(1+\varepsilon_0) - \kappa}. \end{aligned} \tag{2.26}$$

If $q+k+\alpha d < 2\eta m$ the series with the general term given in (2.26) is convergent. If $q+k+\alpha d = 2\eta m n$ we need that $\kappa > 1 + \gamma + \eta(1+\varepsilon_0)$ in order to obtain the convergence of the series. If $\kappa > 1 + \gamma + \eta$ then we may choose ε_0 sufficiently small in order to have $\gamma + \eta(1+\varepsilon_0) - \kappa > 1$ and we are done. □

There are two important examples: $\mathbf{e} = \mathbf{e}_p$ that we discuss in a special subsection below and $\mathbf{e} = \mathbf{e}_{\log}$ which we discuss now. We recall that $\mathbf{e}_{\log} \in \mathcal{E}_{\alpha,\gamma}$ with $\alpha = 0$ and $\gamma = 1$ and $\|f_\delta\|_{2m,2m,\mathbf{e}_{\log}} \leq C1 \vee \|f_\delta\|_{2m,2m,1+}$ where $\|f_\delta\|_{2m,2m,1+}$ is defined in (2.10). Then as a particular case of the previous theorem we obtain:

Theorem 2.9. *We consider a non negative finite measure μ and we suppose that there exists a family of measures $\mu_\delta(dx) = f_\delta(x)dx, \delta > 0$ which verifies the following assumptions. There exist $m \in \mathbb{N}_*$, $C, r, \varepsilon > 0$ and a function $\lambda_m(\delta)$, $\delta \in (0,1)$, which is right-continuous and non increasing such that*

$$\begin{aligned} &\|f_\delta\|_{2m,2m,1+} \le \lambda_m(\delta) \le C\delta^{-r} \text{ and} \\ &\lambda_m^{\frac{1}{2m}}(\delta) d_1(\mu, \mu_\delta) \le \frac{C}{(\ln(1/\delta))^{2+\frac{1}{2m}+\varepsilon}}. \end{aligned} \tag{2.27}$$

Then $\mu(dx) = f(x)dx$ with $f \in L^{\mathbf{e}_{\log}}$.

2.4 The L^p criterion

In the case of the L^p norms, that is $\mathbf{e} = \mathbf{e}_p$, our result fits in the general theory of the interpolation spaces and we may give a more precise characterization of the space $\mathcal{S}_{q,k,m,\mathbf{e}_p} =: \mathcal{S}_{q,k,m,p}$. We come back to the standard notation and we denote $\|\cdot\|_p$ instead of $\|\cdot\|_{\mathbf{e}_p}$, $W^{q,p}$ instead of $W^{q,\mathbf{e}_p}$ and so on. In Appendix B we prove that in this case the space $\mathcal{S}_{q,k,m,p}$ is related to the following interpolation space. Let $X = W_*^{k,\infty}$ where $W_*^{k,\infty}$ is the dual of $W^{k,\infty}$ (notice that one may look to $\mu \in \mathcal{M}$ as to an element of $W_*^{k,\infty}$ and then $d_k(\mu,\nu) = \|\mu - \nu\|_{W_*^{k,\infty}}$). We also take $Y = W^{q+2m,2m,p}$ and for $\gamma \in (0,1)$ we denote by $(X,Y)_\gamma$ the real interpolation space of order γ between X and Y (see Appendix B for notations). Then we have

$$\mathcal{S}_{q,k,m,p} = (X,Y)_\gamma \qquad \text{with} \qquad \gamma = \frac{q+k+d/p_*}{2m}.$$

So Theorem 2.7 reads

$$W^{q+1,2m,p} \subset (W_*^{k,\infty}, W^{q+2m,2m,p})_\gamma \subset W^{q,p}.$$

We go now further and we notice that if (2.24) holds then the convergence of the series in (2.26) is very fast. This allows us to obtain some more regularity.

Theorem 2.10. *Let $q, k \in \mathbb{N}, m \in \mathbb{N}_*$, $p > 1$ and set*

$$\eta > \frac{q+k+d/p_*}{2m}. \tag{2.28}$$

We consider a non negative finite measure μ and a family of finite non negative measures $\mu_\delta(dx) = f_\delta(x)dx, \delta > 0$.

A. *We assume that there exist $C, r > 0$ and a right-continuous and non increasing function $\lambda_{q,m}(\delta)$, $\delta \in (0,1)$, such that*

$$\|f_\delta\|_{2m+q,2m,p} \le \lambda_{q,m}(\delta) \le C\delta^{-r}$$

and moreover, with η given in (2.28),

$$\lambda_{q,m}(\delta)^\eta d_k(\mu, \mu_\delta) \le C. \tag{2.29}$$

Then $\mu(dx) = f(x)dx$ with $f \in W^{q,p}$.

B. *We assume that (2.29) holds with $q+1$ instead of q, that is*

$$\lambda_{q+1,m}(\delta)^\eta d_k(\mu, \mu_\delta) \le C.$$

We denote

$$s_\eta(q,k,m,p) = \frac{2m\eta - (q+k+d/p_*)}{2m\eta} \wedge \frac{\eta}{1+\eta}. \tag{2.30}$$

Then for every multi index α with $|\alpha| = q$ and every $s < s_\eta(q,k,m,p)$ we have $\partial_\alpha f \in \mathcal{B}^{s,p}$ where $\mathcal{B}^{s,p}$ is the Besov space of index s.

Proof. **A.** The fact that (2.29) implies $\mu(dx) = f(x)dx$ with $f \in W^{q,p}$ is an immediate consequence of Theorem 2.8.

B. We prove the regularity property: $g := \partial_\alpha f \in \mathcal{B}^{s,p}$ for $|\alpha| = q$ and $s < s_\eta(q,k,m)$. In order to do it we will use Lemma B.1 so we have to check (B.4).

Step 1. We begin with the point $i)$ in (B.4) so we have to estimate $\|g * \partial_i \phi_\varepsilon\|_\infty$. The reasoning is analogous with the one in the proof of Theorem 2.8 but we will use the first inequality in (2.22) with q replaced by $q+1$ and k replaced by $k-1$. So we define $\delta_n = \inf\{\delta > 0 : \lambda_{q+1,m}(\delta) \le n^{-2} 2^{2mn}\}$ and we have $\delta_n \le 2^{-\theta n}$ for $\theta < 2m/r$. We obtain

$$\begin{aligned}
\|g * \partial_i \phi_\varepsilon\|_p &= \|\partial_i \partial_\alpha (f * \phi_\varepsilon)\|_p \le \|f * \phi_\varepsilon\|_{q+1,p} \le \rho_{q+1,k-1,m,p}(\mu * \phi_\varepsilon) \\
&\le \sum_{n=1}^\infty 2^{n(q+k+d/p_*)} d_{k-1}(\mu * \phi_\varepsilon, \mu_{\delta_n} * \phi_\varepsilon) \\
&\quad + \sum_{n=1}^\infty 2^{-2mn} \|f_{\delta_n} * \phi_\varepsilon\|_{2m+q+1,2m,p}.
\end{aligned}$$

By the choice of δ_n,

$$\|f_{\delta_n} * \phi_\varepsilon\|_{2m+q+1,2m,p} \le \|f_{\delta_n}\|_{2m+q+1,2m,p} \le \lambda_{q+1,m}(\delta_n) \le \frac{1}{n^2} 2^{2nm}$$

so the second series is convergent. We estimate now the first sum. Since $\|f * \phi_\varepsilon\|_{k,\infty} \le \varepsilon^{-1} \|f\|_{k-1,\infty}$, one has $d_{k-1}(\mu * \phi_\varepsilon, \mu_{\delta_n} * \phi_\varepsilon) \le \varepsilon^{-1} d_k(\mu, \mu_{\delta_n})$. Then, using (2.29) (with $q = 1$ instead of q) and the choice of δ_n we obtain

$$\begin{aligned}
2^{n(q+k+d/p_*)} d_{k-1}(\mu * \phi_\varepsilon, \mu_{\delta_n} * \phi_\varepsilon) &\le \frac{C}{\varepsilon} 2^{n(q+1+d/p_*)} d_k(\mu, \mu_{\delta_n}) \\
&\le \frac{C}{\varepsilon} 2^{n(q+1+d/p_*)} \lambda_{q+1,m}^{-\eta}(\delta_n) \\
&\le \frac{C n^{2\eta}}{\varepsilon} 2^{n(q+1+d/p_* - 2m\eta)}.
\end{aligned}$$

We fix now $\varepsilon > 0$, we take some $n_\varepsilon \in \mathbb{N}$ (to be chosen in the sequel) and we write

$$\begin{aligned}
&\sum_{n=1}^\infty 2^{n(q+k+d/p_*)} d_{k-1}(\mu * \phi_\varepsilon, \mu_{\delta_n} * \phi_\varepsilon) \\
&\quad \le C \sum_{n=1}^{n_\varepsilon} 2^{n(q+k+d/p_*)} + \frac{C}{\varepsilon} \sum_{n=n_\varepsilon+1}^\infty n^{2\eta} 2^{n(q+k+d/p_* - 2\eta m)}.
\end{aligned}$$

We take $a > 0$ and we upper bound the above series by

$$2^{n_\varepsilon(q+k+d/p_*)} + \frac{C}{\varepsilon} 2^{n_\varepsilon(q+k+d/p_* + a - 2\eta m)}.$$

In order to optimize we take n_ε such that $2^{2mn_\varepsilon} = \frac{1}{\varepsilon}$. With this choice we obtain

$$2^{n_\varepsilon(q+k+d/p_*+a)} \leq C\varepsilon^{-\frac{q+k+d/p_*+a}{2m\eta}}.$$

We conclude that

$$\|g * \partial_i \phi_\varepsilon\|_p \leq C\varepsilon^{-\frac{q+k+d/p_*+a}{2m\eta}}$$

which means (B.4) $i)$ holds for $s < 1 - \frac{q+k+d/p_*}{2m\eta}$.

Step 2. We check now (B.4) $ii)$ so we have to estimate $\left\|g * \phi^i_\varepsilon\right\|_p$ with $\phi^i_\varepsilon(x) = x^i\phi_\varepsilon(x)$. We take $u \in (0,1)$ (to be chosen in a moment) and we define

$$\delta_{n,\varepsilon} = \inf\{\delta > 0 : \lambda_{q+1,m}(\delta) \leq n^{-2}2^{2mn} \times \varepsilon^{-(1-u)}\}.$$

Then we proceed as in the previous step:

$$\begin{aligned}\left\|\partial_i(g * \phi^i_\varepsilon)\right\|_p \leq & \rho_{q+1,k-1,m,p}(\mu * \phi^i_\varepsilon) \\ \leq & \sum_{n=1}^{\infty} 2^{n(q+k+d/p_*)} d_{k-1}(\mu * \phi^i_\varepsilon, \mu_{\delta_{n,\varepsilon}} * \phi^i_\varepsilon) \\ & + \sum_{n=1}^{\infty} 2^{-2mn} \left\|f_{\delta_{n,\varepsilon}} * \phi^i_\varepsilon\right\|_{2m+q+1,2m,p}.\end{aligned}$$

It is easy to check that for every $h \in L^p$ one has $\left\|h * \phi^i_\varepsilon\right\|_p \leq \varepsilon \|h\|_p$ so that, by our choice of $\delta_{n,\varepsilon}$ we obtain

$$\left\|f_{\delta_{n,\varepsilon}} * \phi^i_\varepsilon\right\|_{2m+q+1,2m,p} \leq \varepsilon \left\|f_{\delta_{n,\varepsilon}}\right\|_{2m+q+1,2m,p} \leq \varepsilon \times \frac{2^{2mn}}{n^2} \times \varepsilon^{-(1-u)}.$$

It follows that the second sum is upper bounded by $C\varepsilon^u$.

Since $\left\|\partial_j h * \phi^i_\varepsilon\right\|_\infty \leq C\|h\|_\infty$ it follows that

$$d_{k-1}(\mu * \phi^i_\varepsilon, \mu_{\delta_{n,\varepsilon}} * \phi^i_\varepsilon) \leq C d_k(\mu, \mu_{\delta_{n,\varepsilon}}) \leq \frac{C}{\lambda^\eta_{q+1,m}(\delta_{n,\varepsilon})} = \frac{Cn^2}{2^{2mn\eta}}\varepsilon^{\eta(1-u)}.$$

Since $2m\eta > q + k + d/p_*$ the first sum is convergent also and is upper bounded by $C\varepsilon^{\eta(1-u)}$. We conclude that

$$\left\|\partial_i(g * \phi^i_\varepsilon)\right\|_p \leq C\varepsilon^{\eta(1-u)} + C\varepsilon^u.$$

In order to optimize we take $u = \frac{\eta}{1+\eta}$. □

2.5 Convergence criteria in $W^{q,p}$ and $W^{q,\mathbf{e}_{\log}}$

For a function f, we denote $\mu_f(dx) = f(x)dx$.

Theorem 2.11. *Let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be a non decreasing function and $a \geq 1$ be such that*

$$\lim_{n\to\infty} \eta(n) = +\infty \quad \text{and} \quad \eta(n+1) \leq a\eta(n), \quad \text{for every } n \in \mathbb{N}. \tag{2.31}$$

Let $m, k, q \in \mathbb{N}$ be fixed. Let f_n, $n \in \mathbb{N}$, be a sequence of functions and $\mu \in \mathcal{M}$.

i) Let $p \geq 1$. If there exists $\alpha > \frac{q+k+d/p_*}{m}$ such that

$$\|f_n\|_{q+2m,2m,p} \leq \eta^{1/\alpha}(n) \quad \textit{and} \quad d_k(\mu, \mu_{f_n}) \leq \frac{1}{\eta(n)}, \tag{2.32}$$

then $\mu(dx) = f(x)dx$ for some $f \in W^{q,p}$. Moreover, there exists a constant C depending on a, α such that for every $n \in \mathbb{N}$

$$\|f - f_n\|_{q,p} \leq C\eta^{-\theta}(n) \quad \textit{with} \quad \theta = \frac{1}{\alpha} \wedge (1 - \frac{q+k+d/p_*}{\alpha m}). \tag{2.33}$$

ii) If there exists $\alpha > \frac{q+k}{m}$ such that

$$\|f_n\|_{q+2m,2m,1+} \leq \eta^{1/\alpha}(n) \quad \textit{and} \quad d_k(\mu, \mu_{f_n}) \leq \frac{1}{\eta(n)}, \tag{2.34}$$

then $\mu(dx) = f(x)dx$ for some $f \in W^{q,\mathbf{e}_{\log}}$. Moreover, there exists a constant C depending on a, α such that for every $n \in \mathbb{N}$

$$\|f - f_n\|_{q,\mathbf{e}_{\log}} \leq C(\eta^{-1/\alpha}(n) + (\log_2 \eta(n))\eta^{-(1-\frac{q+k}{\alpha m})}(n)) =: \varepsilon_n(\alpha). \tag{2.35}$$

And if $\varepsilon_n(\alpha) \leq 1$ then

$$\sum_{0 \leq |\alpha| \leq q} \int |(\partial_\alpha f - \partial_\alpha f_n)(x)| \left(1 + \ln^+ |(\partial_\alpha f - \partial_\alpha f_n)(x)|\right) dx \leq 2C_* \varepsilon_n(\alpha). \tag{2.36}$$

Proof. i) **Step 1**. For $r \in \mathbb{N}$, we define

$$n_r = \min\{n : \eta(n) \geq 2^{\alpha r m}\} \quad \text{and} \quad r_n = \min\{r \in \mathbb{N} : n_r \geq n\}.$$

Then we have

$$\frac{1}{a}\eta(n) \leq 2^{\alpha r_n m} \leq C\eta(n). \tag{2.37}$$

Since $\{r \in \mathbb{N} : n_r \geq n\}$ is a discrete set, its minimum r_n belongs to this set, so $n_{r_n} \geq n$. Then $\eta(n) \leq \eta(n_{r_n}) \leq a\eta(n_{r_n} - 1) \leq a2^{\alpha r_n m}$. On the other hand, since $r_n - 1 \notin \{r \in \mathbb{N} : n_r \geq n\}$ one has $n > n_{r_n - 1}$ and then $\eta(n) \geq \eta(n_{r_n-1}) \geq 2^{\alpha(r_n-1)m} = C^{-1}2^{\alpha r_n m}$ with $C = 2^{\alpha m}$. So, (2.37) holds.

Step 2. We fix $n \in \mathbb{N}$ and for $r \in \mathbb{N}$ we define

$$g_r = 0 \text{ if } r < r_n \text{ and } g_r = f_{n_r} - f_n \text{ if } r \geq r_n$$

and $\nu(dx) = \mu(dx) - f_n(x)dx, \nu_r(dx) = g_r(x)dx$. Using (2.19) (recall that $\beta_{\mathbf{e}_p} = t^{1/p_*}$) we get

$$\rho_{q,k,m,p}(\nu) \leq \sum_{r=1}^{\infty} 2^{r(q+k+d/p_*)} d_k(\nu, \nu_r) + \sum_{r=1}^{\infty} 2^{-2mr} \|g_r\|_{q+2m,2m,p} =: S_1 + S_2.$$

We estimate S_1. For $r < r_n$ we have $\nu_r = 0$ so that $d_k(\nu, \nu_r) = d_k(\nu, 0) = d_k(\mu, \mu_{f_n}) \leq \eta^{-1}(n)$. And for $r \geq r_n$ we have

$$d_k(\nu,\nu_r)=d_k(\mu,\mu_{f_{n_r}})\le\frac{1}{\eta(n_r)}\le\frac{1}{2^{rm\alpha}}.$$

So, we obtain

$$S_1\le 2^{r_n(q+k+d/p_*)}\eta^{-1}(n)+\frac{C}{2^{r_n m\alpha(1-\frac{(q+k+d/p_*}{\alpha m})}}$$

and using (2.37),

$$S_1\le C\eta^{-(1-\frac{q+k+d/p_*}{\alpha m})}(n).$$

We estimate now S_2. We have $g_r=0$ for $r<r_n$ and for $r\ge r_n$

$$\|g_r\|_{q+2m,2m,p}\le\|f_{n_r}\|_{q+2m,2m,p}+\|f_n\|_{q+2m,2m,p}\le\eta(n_r)^{1/\alpha}+\eta(n)^{1/\alpha}.$$

But $\eta(n_r)\le a\eta(n_r-1)\le a2^{\alpha rm}$, so that

$$\|g_r\|_{q+2m,2m,p}\le a^{1/\alpha}2^{rm}+\eta(n)^{1/\alpha}.$$

It follows that

$$S_2\le a^{1/\alpha}\sum_{r\ge r_n}2^{-rm}+\eta(n)^{1/\alpha}\sum_{r\ge r_n}2^{-2rm}\le C\big(2^{-r_n m}+\eta(n)^{1/\alpha}2^{-2r_n m}\big)$$

and using (2.37) we get

$$S_2\le C\eta(n)^{-1/\alpha}.$$

Then, we obtain

$$\rho_{q,k,m,p}(\nu)\le C(\eta^{-1/\alpha}(n)+\eta^{-(1-\frac{q+k+d/p_*}{\alpha m})}(n))$$

and Theorem 2.6 allows one to conclude.

ii) We take n_r and r_n as in Step 1 above, giving (2.37), and we take g_r, ν, ν_r as in Step 2 above. Then, by using (2.19) we get

$$\rho_{q,k,m,\mathbf{e}_{\log}}(\nu)\le\sum_{r=1}^{\infty}2^{r(q+k)}\beta_{\mathbf{e}_{\log}}(2^{rd})d_k(\nu,\nu_r)+\sum_{r=1}^{\infty}2^{-2mr}\,\|g_r\|_{q+2m,2m,\mathbf{e}_{\log}}.$$

By (2.11) and (2.12), we can write

$$\begin{aligned}\rho_{q,k,m,\mathbf{e}_{\log}}(\nu)&\le C\sum_{r=1}^{\infty}2^{r(q+k)}rd_k(\nu,\nu_r)+\sum_{r=1}^{\infty}2^{-2mr}1\vee\|g_r\|_{q+2m,2m,1+}\\&=:S_1+S_2.\end{aligned}$$

Concerning S_1, for $r<r_n$ we have $d_k(\nu,\nu_r)=d_k(\nu,0)=d_k(\mu,\mu_{f_n})\le\eta^{-1}(n)$ and for $r\ge r_n$ we have $d_k(\nu,\nu_r)\le\frac{1}{\eta(n_r)}\le\frac{1}{2^{rm\alpha}}$. So, we obtain

$$S_1\le C\Big(r_n2^{r_n(q+k)}\eta^{-1}(n)+\frac{r_n}{2^{r_n m\alpha(1-\frac{q+k}{\alpha m})}}\Big).$$

Using (2.37),

$$S_1 \leq C r_n \eta^{-(1-\frac{q+k}{\alpha m})}(n) \leq C(\log_2 \eta(n))\, \eta^{-(1-\frac{q+k}{\alpha m})}(n).$$

As for S_2, we proceed as in Step 2 above and we obtain $S_2 \leq C\eta(n)^{-1/\alpha}$. Then,

$$\rho_{q,k,m,\mathbf{e}_{\log}}(\nu) \leq C(\eta^{-1/\alpha}(n) + \eta^{-(1-\frac{q+k+d/p_*}{\alpha m})}(n))$$

and the statement again follows from Theorem 2.6. So (2.35) is proved. In order to check (2.36) we use (2.14) (notice that, since $\|f - f_n\|_{q,\mathbf{e}_{\log}} \leq \varepsilon_n(\alpha) \leq 1$, we have $\ln^+ \|f - f_n\|_{q,\mathbf{e}_{\log}} = 0$). □

2.6 Random variables and integration by parts

In this section we work in the framework of random variables. For a random variable F we denote by μ_F the law of F and if μ_F is absolutely continuous we denote by p_F its density. We will use Theorem 2.10 for μ_F so we will look for a family of random variables $F_\delta, \delta > 0$ such that μ_{F_δ} satisfy the hypothesis of this theorem. Sometimes it is easy to construct such a family with explicit densities p_{F_δ} and then one may check (2.29) directly (this is the case in the examples in Section 3.1 and 3.2). But sometimes one does not know p_{F_δ} and then it is useful to use the integration by parts machinery in order to prove (2.29) - this is the case in the example given is Section 3.3 or the application to a kind of generalization of the Hörmander condition to general Wiener functionals developed in [4].

We briefly recall the abstract definition of integration by parts formulae and we give some useful properties (coming essentially from [1]). We consider two random variables $F = (F_1, ..., F_d)$ and G. Given a multi index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., d\}^k$ and for $p \geq 1$ we say that $\mathrm{IP}_{\alpha,p}(F, G)$ holds if we may find a random variable $H_\alpha(F; G) \in L^p$ such that for every $f \in C_b^\infty(\mathbb{R}^d)$ one has

$$\mathbb{E}(\partial_\alpha f(F)G) = \mathbb{E}(f(F)H_\alpha(F; G)). \tag{2.38}$$

The weight $H_\alpha(F; G)$ is not uniquely determined: the one with the lowest variance is $\mathbb{E}(H_\alpha(F; G) \mid \sigma(F))$. This quantity is uniquely determined. So we denote

$$\theta_\alpha(F, G) = \mathbb{E}(H_\alpha(F; G) \mid \sigma(F)). \tag{2.39}$$

For $m \in \mathbb{N}$ and $p \geq 1$ we denote by $\mathcal{R}_{m,p}$ the class of random variables F in $\mathbb{R}^d$ such that $\mathrm{IP}_{\alpha,p}(F, 1)$ holds for every multi index α with $|\alpha| \leq m$. We define

$$T_{m,p}(F) = \|F\|_p + \sum_{|\alpha| \leq m} \|\theta_\alpha(F, 1)\|_p. \tag{2.40}$$

Notice that by Hölder's inequality $\|\mathbb{E}(H_\alpha(F; 1) \mid \sigma(F))\|_p \leq \|H_\alpha(F; 1)\|_p$. It follows that for every choice of the weights $H_\alpha(F; 1)$ one has

$$T_{m,p}(F) \leq \|F\|_p + \sum_{|\alpha| \leq m} \|H_\alpha(F; 1)\|_p. \tag{2.41}$$

Theorem 2.12. *Let $m, l \in \mathbb{N}$ and $p > d$. If $F \in \mathcal{R}_{m+1,p}$ then the law of F is absolutely continuous and the density p_F belongs to $C^m(\mathbb{R}^d)$. Moreover, suppose that $F \in \mathcal{R}_{m+1,2(d+1)}$. There exists a universal constant C (depending on d, l and m only) such that for every multi index α with $|\alpha| \leq m$*

$$|\partial_\alpha p_F(x)| \leq C T_{1,2(d+1)}^{d^2-1}(F) T_{m+1,2(d+1)}(F)(1 + \|F\|_l)(1 + |x|)^{-l}. \tag{2.42}$$

In particular, for every $q \geq 1, k \in \mathbb{N}$ there exists a universal constant C (depending on d, m, k, p and q) such that

$$\|p_F\|_{m,k,q} \leq C T_{1,2(d+1)}^{d^2-1}(F) T_{m+1,2(d+1)}(F)(1 + \|F\|_{d+k+1}). \tag{2.43}$$

Proof. The proof is an immediate consequence of the results in [1]. In order to see this we have to give the relation between the notation used in that paper and the notation used here: we work with the probability measure $\mu_F(dx) = \mathbb{P}(F \in dx)$ and in [1] we use the notation $\partial_\alpha^{\mu_F} g(x) = \mathbb{E}(H_\alpha(F; g(F)) \mid F = x)$.

The fact that $F \in \mathcal{R}_{m+1,p}$ implies that $F \sim p_F(x)dx$ with $p_F \in C^m(\mathbb{R}^d)$ is proved in [1] Proposition 9. We consider now a function $\psi \in C_b^\infty(\mathbb{R}^d)$ such that $1_{B_1} \leq \psi \leq 1_{B_2}$. In [1] Theorem 8 we have given the following representation formula:

$$\partial_\alpha p_F(x) = \sum_{i=1}^d \mathbb{E}(\partial_i Q_d(F - x)\theta_{(\alpha,i)}(F; \psi(F - x))1_{B_2}(F - x))$$

where B_r denotes the ball centered at 0 with radius r, Q_d is the Poisson kernel on $\mathbb{R}^d$ and, if $\alpha = (\alpha_1, ..., \alpha_k)$, then $(\alpha, i) = (\alpha_1, ..., \alpha_k, i)$. Using Hölder's inequality we obtain (with p_* the conjugate of p)

$$|\partial_\alpha p_F(x)| \leq \sum_{i=1}^d \|\partial_i Q_d(F - x)\|_p \left\|\theta_{(\alpha,i)}(F; \psi(F - x))1_{B_2}(F - x)\right\|_{p_*}.$$

We take $p = d + 1$ so that $p_* = (d + 1)/d \leq 2$. In [1] Theorem 5 we proved that

$$\|\partial_i Q_d(F - x)\|_p \leq C T_{1,2(d+1)}^{d^2-1}(F).$$

Moreover we have the following computational rule (Lemma 9 in [1])

$$\theta_i(F, fg(F)) = f(F)\theta_i(F, g(F)) + (g\partial_i f)(F).$$

Since $\psi \in C_b^\infty(\mathbb{R}^d)$ we may use the above formula in order to get

$$\begin{aligned}
&\left\|\theta_{(\alpha,i)}(F; \psi(F - x))1_{B_2}(F - x)\right\|_{p_*} \\
&\quad \leq \left\|\theta_{(\alpha,i)}(F; \psi(F - x))\right\|_{2p_*} \sqrt{\mathbb{P}(|F - x| \leq 2)} \\
&\quad \leq C_\psi T_{|\alpha|+1,2p_*}(F) \sqrt{\mathbb{P}(|F - x| \leq 2)}.
\end{aligned}$$

For $|x| \geq 4$

$$\mathbb{P}(|F - x| \leq 2) \leq \mathbb{P}\Big(|F| \geq \frac{1}{2}|x|\Big) \leq \frac{2^k}{|x|^k}\mathbb{E}(|F|^k)$$

so the proof of (2.42) is completed. □

We are now ready to rewrite Theorem 2.10:

Theorem 2.13. *Let $k, q \in \mathbb{N}$, $m \in \mathbb{N}_*$, $p > 1$ and let*

$$\eta > \frac{q + k + d/p_*}{2m},$$

p_* denoting the conjugate of p. Let F, F_δ,$\delta > 0$, be random variables and let μ_F, μ_{F_δ}, $\delta > 0$, denote the associated laws.

A. *Suppose that $F_\delta \in \mathcal{R}_{2m+q+1,2(d+1)}$, $\delta > 0$ are uniformly bounded in L^{2m+d+1} and that there exist $C > 0$ and $\theta > 0$ such that*

$$T_{2m+q+1,2(d+1)}(F_\delta) \leq C\delta^{-\theta(2m+q+1)}, \tag{2.44}$$

$$d_k(\mu_F, \mu_{F_\delta}) \leq C\delta^{\theta\eta d^2(2m+q+1)}. \tag{2.45}$$

Then $\mu_F(dx) = p_F(x)dx$ with $p_F \in W^{q,p}$.

B. *Suppose that $F_\delta \in \mathcal{R}_{2m+q+2,2(d+1)}$, $\delta > 0$, and (2.44) holds with $q+1$ instead of q.Then for every multi index α with $|\alpha| = q$ and every $s < s_\eta(q,k,m,p)$ we have $\partial_\alpha p_F \in \mathcal{B}^{s,p}$ where $\mathcal{B}^{s,p}$ is the Besov space of index s and $s_\eta(q,k,m,p)$ is given in (2.30).*

Proof. **A.** Let $n, l \in \mathbb{N}$ and $p > 1$ be fixed. By using (2.44) and (2.43) we obtain $\|p_{F_\delta}\|_{2m+q,2m,p} \leq C\delta^{-\theta d^2(2m+q+1)}$. So, as a consequence of (2.45) we obtain $\|p_{F_\delta}\|^\eta_{2m+q,2m,p}\, d_k(\mu_F, \mu_{F_\delta}) \leq C$. And we apply Theorem 2.10 **A**. Similarly, **B** follows by applying Theorem 2.10 **B.** □

3 Examples

3.1 Path dependent SDE's

In this section we look to the SDE

$$dX_t = \sum_{j=1}^n \sigma_j(t,X)dW_t^j + b(t,X)dt \tag{3.1}$$

where $W = (W^1, ..., W^n)$ is a standard Brownian motion and

$$\sigma_j, b : C(\mathbb{R}_+;\mathbb{R}^d) \to C(\mathbb{R}_+;\mathbb{R}^d), \quad j = 1, ..., n.$$

For a function $\varphi \in C(\mathbb{R}_+;\mathbb{R}^d)$, we use the notation $\sigma_j(t,\varphi) = \sigma_j(\varphi)(t)$ and $b(t,\varphi) = b(\varphi)(t)$. If σ_j and b satisfy some Lipschitz continuity property with respect to the sup-norm on $C(\mathbb{R}_+;\mathbb{R}^d)$ then this equation has a unique solution. But we do not want to make such an hypothesis here so we just consider an adapted process $X_t, t \geq 0$ which verifies the above equation.

We set $\Delta_{s,t}(w) := \sup_{s\leq u\leq t} |w_u - w_s|$

Theorem 3.1. *Let b and σ_j, $j = 1, \ldots, n$, be bounded. Suppose that there exists $\varepsilon, C > 0$ such that*

$$|\sigma_j(t,w) - \sigma_j(s,w)| \leq C\Big(\ln\Big(\frac{1}{\Delta_{s,t}(w)}\Big)\Big)^{-(2+\varepsilon)}, \qquad \forall j = 1, ..., n \tag{3.2}$$

and that there exists $\lambda_* > 0$ such that

$$\sigma\sigma^*(t,w) \geq \lambda_* \qquad \forall t \geq 0, w \in C(\mathbb{R}_+;\mathbb{R}^d). \tag{3.3}$$

Then for every $T > 0$ the law of X_T is absolutely continuous with respect to the Lebesgue measure and the density belongs to $L^{\mathbf{e}_{\log}}$.

Remark 3.2. *We note that in the particular case of standard SDE's we have $\sigma_j(t,w) = \sigma_j(w_t)$ and a sufficient condition in order that (3.2) holds is $|\sigma_j(x) - \sigma_j(y)| \leq C(\ln(\frac{1}{|x-y|}))^{-(2+\varepsilon)}$. This is weaker than Hölder continuity.*

Proof. During the proof, we set $\lambda^* > 0$ a constant such that $\lambda^* \geq \sigma\sigma^*(t,w)$.

For $\delta > 0$ we construct

$$X_T^\delta = X_{T-\delta} + \sum_{j=1}^n \sigma_j(T-\delta, X)(W_T^j - W_{T-\delta}^j).$$

We will use Theorem 2.9 so we check the hypotheses there.

Step 1. We write $X_T - X_T^\delta = \sum_{j=1}^n I_\delta^j + J_\delta$ with

$$I_\delta^j = \int_{T-\delta}^T (\sigma_j(t,X) - \sigma_j(T-\delta,X))dW_t^j \quad \text{and} \quad J_\delta = \int_{T-\delta}^T b(t,W)dt.$$

Since b is bounded, we have

$$\mathbb{E}(|J_\delta|) \leq C\delta. \tag{3.4}$$

Let $\bar{a}_\delta = \sqrt{\delta}\,\ln\frac{1}{\delta}$ and $A_\delta = \{\Delta_{T-\delta,T}(X) \leq \bar{a}_\delta\}$. We write $\mathbb{E}(|I_\delta^j|^2) = K_\delta + L_\delta$ with

$$\begin{aligned} K_\delta &= \int_{T-\delta}^T \mathbb{E}(1_{A_\delta^c}\,|\sigma_j(t,X) - \sigma_j(T-\delta,X)|^2)dt \\ L_\delta &= \int_{T-\delta}^T \mathbb{E}(1_{A_\delta}\,|\sigma_j(t,X) - \sigma_j(T-\delta,X)|^2)dt. \end{aligned}$$

By using the Bernstein's inequality we obtain $\mathbb{P}(A_\delta^c) \leq C\exp(-\frac{\bar{a}_\delta^2}{C'\delta})$. And since σ_j is bounded, for any small δ we get

$$K_\delta \leq C\delta\mathbb{P}(A_\delta^c) \leq C\delta\exp(-\frac{\bar{a}_\delta^2}{2C'\delta}) \leq C\delta^{\frac{3}{2}}.$$

Moreover using (3.2) and again for δ small enough,

$$L_\delta \leq \frac{C\delta}{(\ln\frac{1}{a_\delta})^{2(2+\varepsilon)}} \leq \frac{C'\delta}{(\ln\frac{1}{\delta})^{2(2+\varepsilon)}}$$

(notice that $\ln(\frac{1}{\delta})/\ln\frac{1}{\bar{a}_\delta} \to \frac{1}{2} > 0$ for $\delta \to 0$). We conclude that

$$\mathbb{E}(|I_\delta^j|^2) \leq \frac{C\delta}{(\ln\frac{1}{\delta})^{2(2+\varepsilon)}}$$

so that, if μ is the law of X_T and μ_δ is the law of X_T^δ then for every δ small,

$$d_1(\mu,\mu_\delta) \leq \mathbb{E}(|X_T - X_T^\delta|) \leq \frac{C\delta^{1/2}}{(\ln\frac{1}{\delta})^{2+\varepsilon}}. \tag{3.5}$$

Step 2. Given a positive definite matrix a, we denote

$$\gamma_{\delta,a}(y) = \frac{1}{(2\pi\delta)^{d/2}(\det a)^{1/2}}\exp\Big(-\frac{1}{2\delta}\langle a^{-1}y,y\rangle\Big).$$

With μ_δ denoting the law of X_T^δ, we have $\mu_\delta(dy) = p_\delta(y)dy$ where

$$p_\delta(y) = \mathbb{E}(\gamma_{\delta,a_{T-\delta}(X)}(y - X_{T-\delta})) \quad \text{with} \quad a_t(X) = \sigma\sigma^*(t, X).$$

Let α denote a multi index $|\alpha| = q$, $k \in \mathbb{N}$ and $\delta \leq 1$. By using (3.3) we have

$$\begin{aligned} |\partial_\alpha p_\delta(y)| &\leq C\delta^{-q/2}\mathbb{E}\Big(\Big(1 + \frac{|y - X_{T-\delta}|}{\delta^{1/2}}\Big)^q \gamma_{\delta,a_{T-\delta}(X)}(y - X_{T-\delta})\Big) \\ &\leq C\delta^{-q/2}\mathbb{E}\Big(\Big(1 + \frac{|y - X_{T-\delta}|}{\delta^{1/2}}\Big)^q \gamma_{\delta,\lambda^* I}(y - X_{T-\delta})\Big). \end{aligned} \tag{3.6}$$

We use the fact that $0 < x \mapsto (1 + x)^q e^{-x^2}$ is bounded. This gives

$$|\partial_\alpha p_\delta(y)| \leq C\delta^{-(d+q)/2},$$

so that, for small values of δ,

$$\ln^+ |\partial_\alpha p_\delta(y)| \leq C\Big(1 + \ln\frac{1}{\delta}\Big) \leq C\ \ln\frac{1}{\delta}. \tag{3.7}$$

Let $m \in \mathbb{N}$. Using (3.6) and (3.7) we obtain

$$\begin{aligned} &\|\partial_\alpha p_\delta\|_{2m,1+} = \int (1 + |y|)^{2m} |\partial_\alpha p_\delta(y)| (1 + \ln^+ |y| + \ln^+ |\partial_\alpha p_\delta(y)|)dy \\ &\leq C\delta^{-q/2}\ \ln\frac{1}{\delta}\mathbb{E}\Big(\int (1 + |y|)^{2m+1}\Big(1 + \frac{|y - X_{T-\delta}|}{\delta^{1/2}}\Big)^q \gamma_{\delta,\lambda^* I}(y - X_{T-\delta})dy\Big) \\ &= C\delta^{-q/2}\ \ln\frac{1}{\delta}\mathbb{E}\Big(\int (1 + |X_{T-\delta} + \delta^{1/2}z|)^{2m+q+1}\gamma_{1,\lambda^* I}(z)dz\Big) \\ &\leq C\delta^{-q/2}\ln\frac{1}{\delta}. \end{aligned}$$

We conclude that

$$\|p_\delta\|_{2m,2m,1+} = \sum_{0\leq|\alpha|\leq 2m} \|\partial_\alpha p_\delta\|_{2m,1+} \leq C\delta^{-m}\ln\frac{1}{\delta}. \tag{3.8}$$

Step 3. We are now ready to check (2.27): the exists $\delta_0 \leq 1$ such that for $\delta < \delta_0$ one has

$$\begin{aligned} \|p_\delta\|_{2m,2m,1+}^{1/2m}\, d_1(\mu, \mu_\delta) &\leq C\delta^{-1/2}\Big(\ln\frac{1}{\delta}\Big)^{1/2m} \times \frac{\delta^{1/2}}{(\ln\frac{1}{\delta})^{2+\varepsilon}} \\ &= \frac{C}{(\ln\frac{1}{\delta})^{2+\varepsilon-\frac{1}{2m}}} \leq \frac{C}{(\ln\frac{1}{\delta})^{2+\frac{1}{2m}+\varepsilon/2}} \end{aligned}$$

the last inequality holding true as soon as $\frac{1}{m} \leq \varepsilon/2$. So (2.27) holds and the conclusion follows from Theorem 2.9. □

3.2 Stochastic heat equation

In this section we investigate the regularity of the law of the solution to the stochastic heat equation introduced by Walsh in [33]. Formally this equation is

$$\partial_t u(t,x) = \partial_x^2 u(t,x) + \sigma(u(t,x))W(t,x) + b(u(t,x)) \tag{3.9}$$

where W denotes a white noise on $\mathbb{R}_+ \times [0,1]$. We consider Neumann boundary conditions that is $\partial_x u(t,0) = \partial_x u(t,1) = 0$ and the initial condition is $u(0,x) = u_0(x)$. The rigorous formulation to this equation is given by the mild form constructed as follows. Let $G_t(x,y)$ be the fundamental solution to the deterministic heat equation $\partial_t v(t,x) = \partial_x^2 v(t,x)$ with Neumann boundary conditions. Then u satisfies

$$\begin{aligned} u(t,x) &= \int_0^1 G_t(x,y)u_0(y)dy + \int_0^t \int_0^1 G_{t-s}(x,y)\sigma(u(s,y))dW(s,y) \\ &\quad + \int_0^t \int_0^1 G_{t-s}(x,y)b(u(s,y))dyds \end{aligned} \tag{3.10}$$

where $dW(s,y)$ is the Itô integral introduced by Walsh. The function $G_t(x,y)$ is explicitly known (see [33] or [8]) but here we will use just few properties that we list below (see the appendix in [8] for the proof). More precisely, for $0 < \varepsilon < t$ we have

$$\int_{t-\varepsilon}^t \int_0^1 G_{t-s}^2(x,y)dyds \leq C\varepsilon^{1/2} \tag{3.11}$$

Moreover, for $0 < x_1 < ... < x_d < 1$ there exists a constant C depending on $\min_{i=1,d}(x_i - x_{i-1})$ such that

$$C\varepsilon^{1/2} \geq \inf_{|\xi|=1} \int_{t-\varepsilon}^t \int_0^1 \left(\sum_{i=1}^d \xi_i G_{t-s}(x_i,y) \right)^2 dyds \geq C^{-1}\varepsilon^{1/2}. \tag{3.12}$$

This is an easy consequence of the inequalities (A2) and (A3) from [8].

In [28] one gives sufficient conditions in order to obtain the absolute continuity of the law of $u(t,x)$ for $(t,x) \in (0,\infty) \times [0,1]$ and in [8], under appropriate hypotheses, one obtains a C^∞ density for the law of the vector $(u(t,x_1),...,u(t,x_d))$ with $(t,x_i) \in (0,\infty) \times \{\sigma \neq 0\}, i = 1,...,d$. The aim of this section is to obtain the same type of results but under much weaker regularity hypothesis on the coefficients. One may first discuss the absolute continuity of the law and further, under more regularity hypothesis on the coefficients, one may discuss the regularity of the density. Here, in order to avoid technicalities, we restrict ourselves to the absolute continuity property. We assume global ellipticity that is

$$\sigma(x) \geq c_\sigma > 0 \qquad \text{for every } x \in [0,1]. \tag{3.13}$$

A local ellipticity condition may also be used but again, this gives more technical complications that we want to avoid. This is somehow a benchmark for the efficiency of the method developed in the previous sections.

We assume the following regularity hypothesis: σ, b are measurable and bounded functions and there exists $h > 0$ such that

$$|\sigma(x) - \sigma(y)| \leq |\ln|x-y||^{-(2+h)}, \quad \text{for every } x,y \in [0,1]. \tag{3.14}$$

This hypothesis is not sufficient in order to ensure existence and uniqueness for the solution to (3.10) (one needs σ and b to be globally Lipschitz continuous in order to obtain it) - so in the following we will just consider a random field $u(t,x), (t,x) \in (0,\infty) \times [0,1]$ which is adapted to the filtration generated by W (see Walsh [33] for precise definitions) and which solves (3.10).

Proposition 3.3. *Suppose that (3.13) and (3.14) hold. Then for every $0 < x_1 < ... < x_d < 1$ and $T > 0$, the law of the random vector $U = (u(T,x_1), ...u(T,x_d))$ is absolutely continuous with respect to the Lebesgue measure.*

Proof. Given $0 < \varepsilon < T$ we decompose

$$u(T,x) = u_\varepsilon(T,x) + I_\varepsilon(T,x) + J_\varepsilon(T,x) \tag{3.15}$$

with

$$\begin{aligned}
u_\varepsilon(T,x) &= \int_0^1 G_t(x,y)u_0(y)dy \\
&\quad + \int_0^T \int_0^1 G_{T-s}(x,y)\sigma(u(s\wedge(T-\varepsilon),y))dW(s,y) \\
&\quad + \int_0^{T-\varepsilon} \int_0^1 G_{T-s}(x,y)b(u(s,y))dyds, \\
I_\varepsilon(T,x) &= \int_{T-\varepsilon}^T \int_0^1 G_{T-s}(x,y)(\sigma(u(s,y)) - \sigma(u(s\wedge(T-\varepsilon),y)))dW(s,y), \\
J_\varepsilon(T,x) &= \int_{T-\varepsilon}^T \int_0^1 G_{T-s}(x,y)b(u(s,y))dyds.
\end{aligned}$$

Step 1. We prove that

$$\mathbb{E}\left|I_\varepsilon(T,x)\right|^2 + \mathbb{E}\left|J_\varepsilon(T,x)\right|^2 \leq C\left|\ln\varepsilon\right|^{-2(2+h)}\varepsilon^{1/2}. \tag{3.16}$$

Let μ be the law of $U = (u(T,x_1), ..., u(T,x_d))$ and let and μ_ε be the law of $U_\varepsilon = (u_\varepsilon(T,x_1), ..., u_\varepsilon(T,x_d))$. Using the above estimate one easily obtains

$$d_1(\mu,\mu_\varepsilon) \leq C\left|\ln\varepsilon\right|^{-(2+h)}\varepsilon^{1/4}. \tag{3.17}$$

Using the isometry property

$$\mathbb{E}\left|I_\varepsilon(T,x)\right|^2 = \int_{T-\varepsilon}^T \int_0^1 G^2_{T-s}(x,y)\mathbb{E}(\sigma(u(s,y) - \sigma(u(s\wedge(T-\varepsilon),y)))^2)dyds.$$

We consider the set $\Lambda_{\varepsilon,\eta}(s,y) = \{|u(s,y) - u(s\wedge(T-\varepsilon),y)| \leq \eta\}$ and we split the above term as $\mathbb{E}\left|I_\varepsilon(T,x)\right|^2 = A_{\varepsilon,\eta} + B_{\varepsilon,\eta}$ with

$$\begin{aligned}
A_{\varepsilon,\eta} &= \int_{T-\varepsilon}^T \int_0^1 G^2_{T-s}(x,y)\mathbb{E}(\sigma(u(s,y) - \sigma(u(s\wedge(T-\varepsilon),y)))^2 1_{\Lambda_{\varepsilon,\eta}(s,y)})dyds \\
B_{\varepsilon,\eta} &= \int_{T-\varepsilon}^T \int_0^1 G^2_{T-s}(x,y)\mathbb{E}(\sigma(u(s,y) - \sigma(u(s\wedge(T-\varepsilon),y)))^2 1_{\Lambda^c_{\varepsilon,\eta}(s,y)})dyds.
\end{aligned}$$

Using (3.14)

$$A_{\varepsilon,\eta} \leq C|\ln\eta|^{-2(2+h)} \int_{T-\varepsilon}^{T}\int_0^1 G^2_{T-s}(x,y)dyds \leq C\,|\ln\eta|^{-2(2+h)}\,\varepsilon^{1/2}$$

the last inequality being a consequence of (3.11). Moreover, coming back to (3.10), we have

$$\begin{aligned}\mathbb{P}(\Lambda^c_{\varepsilon,\eta}(s,y)) &\leq \frac{1}{\eta^2}\mathbb{E}\,|u(s,y)-u(s\wedge(T-\varepsilon),y)|^2\\ &\leq \frac{C}{\eta^2}\int_{T-\varepsilon}^{s}\int_0^1 G^2_{s-r}(y,z)dzdr \leq \frac{C\varepsilon^{1/2}}{\eta^2}\end{aligned}$$

so that

$$B_{\varepsilon,\eta} \leq \frac{C\varepsilon^{1/2}}{\eta^2}\int_{T-\varepsilon}^{T}\int_0^1 G^2_{T-s}(x,y)dyds \leq \frac{C\varepsilon}{\eta^2}.$$

Taking $\eta=\varepsilon^{1/16}$ we obtain

$$\mathbb{E}\,|I_\varepsilon(T,x)|^2 \leq C(|\ln\varepsilon|^{-2(2+h)}+\varepsilon^{1/4})\varepsilon^{1/2} \leq C\,|\ln\varepsilon|^{-2(2+h)}\,\varepsilon^{1/2}.$$

We estimate now

$$|J_\varepsilon(T,x)| \leq \|b\|_\infty \int_{T-\varepsilon}^{T}\int_0^1 G_{T-s}(x,y)dyds = \|b\|_\infty\,\varepsilon$$

so (3.16) is proved.

Step 2. Conditionally to $\mathcal{F}_{T-\varepsilon}$ the random vector $U_\varepsilon=(u_\varepsilon(T,x_1),...,\,u_\varepsilon(T,x_d))$ is Gaussian of covariance matrix

$$\Sigma^{i,j}(U_\varepsilon) = \int_{T-\varepsilon}^{T}\int_0^1 G_{T-s}(x_i,y)G_{T-s}(x_j,y)\sigma^2(u(s\wedge(T-\varepsilon),y))dyds,$$

for $i,j=1,...,d$. By (3.12)

$$C\sqrt{\varepsilon} \geq \Sigma(U_\varepsilon) \geq \frac{1}{C}\sqrt{\varepsilon}$$

where C is a constant which depends on the upper bounds of σ and on c_σ.

We use now the criterion given in Theorem 2.9 . Let p_{U_ε} be the density of the law of U_ε. Conditionally to $\mathcal{F}_{T-\varepsilon}$ this is a Gaussian density and the same reasoning as in the proof of (3.8) gives

$$\|p_{U_\varepsilon}\|_{2m,2m,1+} \leq C(\varepsilon^{-1/4})^{2m}\ln\frac{1}{\varepsilon}.$$

So (2.27) reads

$$\begin{aligned}\|p_{U_\varepsilon}\|^{1/2m}_{2m,2m,1+}\,d_1(\mu,\mu_\varepsilon) &\leq C\varepsilon^{-1/4}(\ln\frac{1}{\varepsilon})^{1/2m}\times|\ln\varepsilon|^{-(2+h)}\,\varepsilon^{1/4}\\ &= C\frac{1}{(\ln\frac{1}{\varepsilon})^{2+h-1/2m}} \leq C\frac{1}{(\ln\frac{1}{\varepsilon})^{2+1/2m}}\end{aligned}$$

the last inequality being true as soon as $h>\frac{1}{m}$. □

3.3 Piecewise deterministic Markov Processes

In this section we deal with a jump type stochastic differential equation which has already been considered in [5]: it is an example of piecewise deterministic Markov 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 [21] for the notations. We denote by N the counting measure associated to p, that is $N([0,t) \times A) = \#\{0 \leq s < t; p_s \in A\}$ for $t \geq 0$ and $A \in \mathcal{B}(E)$. We assume that the associated intensity measure is given by $\widehat{N}(dt, dz, du) = dt \times dz \times 1_{[0,\infty)}(u)du$ where $(z,u) \in E = \mathbb{R}^d \times \mathbb{R}_+$. We are interested in the solution to the d dimensional stochastic equation

$$X_t = x + \int_0^t \int_E c(z, X_{s-}) 1_{\{u < \gamma(z, X_{s-})\}} N(ds, dz, du) + \int_0^t g(X_s) ds. \tag{3.18}$$

The coefficients c, g, γ are smooth functions (see the hypothesis $(H_i), i = 0, 1, 2$ below). We remark that the infinitesimal generator of the Markov process X_t is given by

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

See [15] for the proof of existence and uniqueness of the solution to (3.18). We will deal with two problems related to this equation.

First we give sufficient conditions in order that $\mathbb{P}(X_t(x) \in dy) = p_t(x,y)dy$ where $X_t(x)$ is the solution to (3.18) which starts from x, so $X_0(x) = x$. And we prove that, if the coefficients of the equation are smooth, then $(x,y) \mapsto p_t(x,y)$ is smooth. Notice that the methodology from [15], [11], [10] and [17] seems difficult to implement in order to prove the regularity with respect to the initial condition x. So this is the main point here.

The second result concerns convergence. In [5] it is constructed an approximation scheme which allows one to compute $\mathbb{E}(f(X_t(x))$ using a Monte Carlo method. And it is proved that the convergence takes place in total variation distance. We use here the method developed in our paper in order to prove that the density functions and their derivatives converge as well and to estimate the error.

In [5] one gives a Malliavin type approach to the equation (3.18) which we recall and which we will heavily use here. We describe first the 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 $\Phi_M = \varphi * 1_{B_M}$ with $B_M = \{z \in \mathbb{R}^d : |z| < M\}$. Then $\Phi_M \in C_b^\infty$ and we have $1_{B_{M-1}} \leq \Phi_M \leq 1_{B_{M+1}}$. We denote by X_t^M the solution of the equation

$$X_t^M = x + \int_0^t \int_E c(z, X_{s-}^M) 1_{\{u < \gamma(z, X_{s-}^M)\}} \Phi_M(z) N(ds, dz, du) + \int_0^t g(X_s^M) ds. \tag{3.19}$$

In the following we will assume that $|\gamma(z,x)| \leq \overline{\gamma}$ for some constant $\overline{\gamma}$. Let $N_M(ds, dz, du) := 1_{B_{M+1}}(z) \times 1_{[0,2\overline{\gamma}]}(u) N(ds, dz, du)$. Since $\{u < \gamma(z, X_{s-}^M)\} \subset \{u < 2\overline{\gamma}\}$ 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

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

Since the intensity measure $\widehat{N}_M$ is finite we may represent the random measure N_M by a compound Poisson process. Let $\lambda_M = 2\overline{\gamma} \times \mu(B_{M+1}) = t^{-1}\mathbb{E}(N_M(t,E))$ (with μ the Lebesgue measure) 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 $(Z_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^d$ and $(U_k)_{k\in\mathbb{N}}$ in $\mathbb{R}_+$ which are independent of J^M and such that

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

To simplify the notation, we omit the dependence on M for the variables (T_k^M). Then equation (3.19) may be written as

$$X_t^M = x + \sum_{k=1}^{J_t^M} c_M(Z_k, X_{T_k^-}^M) 1_{(U_k,\infty)}(\gamma(Z_k, X_{T_k^-}^M)) + \int_0^t g(X_s^M)ds. \tag{3.20}$$

Now X_t^M is an explicit functional of the $Z_k, k \in \mathbb{N}$ but, because of the indicator function, this functional is not differentiable. In order to overcome this difficulty, following [5], we consider an alternative representation of the law of X_t^M. Let $z_M^* \in \mathbb{R}^d$ such that $|z_M^*| = M+3$. We define

$$\begin{aligned} q_M(x,z) &:= \varphi(z - z_M^*)\theta_{M,\gamma}(x) + \frac{1}{2\overline{\gamma}\mu(B_{M+1})} 1_{B_{M+1}}(z)\gamma(z,x), \text{ with} \\ \theta_{M,\gamma}(x) &:= \frac{1}{\mu(B_{M+1})} \int_{\{|z|\le M+1\}} \Big(1 - \frac{1}{2\overline{\gamma}}\gamma(z,x)\Big)dz. \end{aligned} \tag{3.21}$$

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, $0 \le \gamma(z,x) \le \overline{\gamma}$ one has $1 \ge \theta_{M,\gamma}(x) \ge 1/2$. By construction the function q_M satisfies $\int q_M(x,z)dz = 1$. Hence we can easily check (see [5] for the proof) that

$$\mathbb{E}(f(X_{T_k}^M) \mid X_{T_k^-}^M = x) = \int_{\mathbb{R}^d} f(x + c_M(z,x))q_M(x,z)dz. \tag{3.22}$$

From the relation (3.22) we construct a process $(\overline{X}_t^M)$, equal in law to (X_t^M), 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 $x_t, t \ge 0$ by $x_0 = x$ and, if x_{T_k} is given, then

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

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

$$x_t = x + \sum_{k=1}^{J_t^M} c_M(z_k, x_{T_k^-}) + \int_0^t g(x_s)ds. \tag{3.23}$$

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

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

Clearly $\overline{X}_t^M$ satisfies the equation

$$\overline{X}_t^M = x + \sum_{k=1}^{J_t^M} c_M(\overline{Z}_k, \overline{X}_{T_k^-}^M) + \int_0^t g(\overline{X}_s^M)ds. \tag{3.24}$$

And by (3.22) the law of $\overline{X}_t^M$ coincides with the law of X_t^M. So now on we work with $\overline{X}_t^M$ which is a smooth functional of $\overline{Z}_k, k \in \mathbb{N}$. But one more difficulty remains: if $T_1 > t$ then $\overline{X}_t^M$ is deterministic, so this functional is not non-degenerated. In order to contouring this last difficulty we add a small noise. We define

$$F_t^M(x) = \overline{X}_t^M(x) + \sqrt{TU_M} \times \Delta, \quad 0 \le t \le T,$$

where $\overline{X}_t^M(x)$ is the solution to (3.24) which starts from x, Δ is a standard normal random variable which is independent of T_k and $\overline{Z}_k, k \in \mathbb{N}$ and

$$U_M = \underline{\gamma} \int_{B_{M-1}^c} \underline{c}^2(z)dz \tag{3.25}$$

with $\underline{\gamma}$ and $\underline{c}$ from (3.26) and (3.28) below. The approximation scheme for $X_t(x)$ is given by $F_t^M(x)$.

Let us give our hypotheses.

(H_0) We assume that γ, g and c are infinitely differentiable functions in both variables z and x. Moreover we assume that g and its derivatives are bounded.

(H_1) There exist $\overline{\gamma} \ge \underline{\gamma}$, such that

$$\overline{\gamma} \ge \gamma(z,x) \ge \underline{\gamma} \ge 0, \quad \forall x \in \mathbb{R}^d \tag{3.26}$$

and, for every $l \in \mathbb{N}$ there exists $\overline{\gamma}_l$ and $\overline{\gamma}_{\ln,l}$ such that for $|\alpha| + |\beta| \le l$

$$\left|\partial_x^\alpha \partial_z^\beta \gamma(x,z)\right| \le \overline{\gamma}_l, \qquad \left|\partial_x^\alpha \partial_z^\beta \ln \gamma(x,z)\right| \le \overline{\gamma}_{\ln,l}. \tag{3.27}$$

(H_2) Setting, for $0 < a < b$ and $r > 0$,

$$\underline{c}(z) = \frac{a}{1+|z|^r}, \qquad \overline{c}(z) = \frac{b}{1+|z|^r},$$

we assume that, for every $z, x, \xi \in \mathbb{R}^d$,

$$\left\|\nabla_x c \times (I + \nabla_x c)^{-1}(z,x)\right\| + |c(z,x)| + \left|\partial_z^\beta \partial_x^\alpha c(z,x)\right| \le \overline{c}(z) \tag{3.28}$$

$$\sum_{j=1}^d \left\langle \partial_{z_j} c(z,x), \xi \right\rangle^2 \ge \underline{c}^2(z) \left|\xi\right|^2. \tag{3.29}$$

Remark 3.4. *The above hypotheses represent a particular case of the hypotheses from [5], corresponding to Example 1,ii) page 634 in that paper. More general hypotheses may be considered (see [5]) but our aim is just to give an example in order to illustrate our method, so we restrict ourself to this case.*

The basic estimate in our approach is the following:

Theorem 3.5. *Suppose that Hypotheses $(H_i), i = 0, 1, 2$ hold. Consider a function $\psi \in C_b^\infty(\mathbb{R}^d)$ such that $1_{B_1} \leq \psi \leq 1_{B_2}$. Then for every $t, R > 0, q \in \mathbb{N}$ and every multi indexes α, β with $|\alpha| + |\beta| \leq q$, one has*

$$\sup_{|x|\leq R, |y|\leq R} \left| \partial_x^\alpha \mathbb{E}((\partial^\beta \phi)(F_t^M(x))\psi(F_t^M(x) - y) \right| \leq C\|\phi\|_\infty M^{dq}. \tag{3.30}$$

Here C is a constant which depends on t, R, q but not on M. In particular the density $p_t^M(x, y)$ of the law of $F_t^M(x)$ verifies

$$\sup_{|x|\leq R, |y|\leq R} \left| \partial_x^\alpha \partial_y^\beta p_t^M(x, y) \right| \leq CM^{d(q+d)}. \tag{3.31}$$

The above theorem is an extension of estimate (42) in Proposition 4 page 640 in [5] and the proof is similar, except for one point: here we consider derivatives ∂_x^α also (while in [5] ∂_y^β only appears). So we just sketch the proof and focus on this supplementary difficulty.

We use an integration by parts formula based on $\overline{Z}_k, k \in \mathbb{N}_*$ and on $\overline{Z}_0 = \Delta$ which is constructed as follows (we follow [5]). Here $J = J_t^M$ and T_k are fixed, so they appear as constants. A simple functional is a random variable of the form $F = f(\overline{Z}_0, \overline{Z}_1, ..., \overline{Z}_J)$ where f is a smooth function. We use the weights $\pi_k = \Phi_M(\overline{Z}_k), k \in \mathbb{N}_*, \pi_0 = 1$ and the Malliavin derivative is defined as

$$D_{k,j} = \pi_k \partial_{\overline{Z}_k^j}.$$

For a multi index $\alpha = (\alpha_1, ..., \alpha_q)$ with $\alpha_i = (k_i, j_i)$ one defines the iterated derivative

$$D_\alpha = D_{\alpha_q} ... D_{\alpha_1}.$$

Then one defines the Sobolev norms:

$$|F|_q^2 = |F|^2 + \sum_{1\leq|\alpha|\leq q} |D_\alpha F|^2, \qquad \|F\|_{q,p} = (\mathbb{E}(|F|_q^p))^{1/p}.$$

For $F = (F^1, ..., F^d)$ the Malliavin covariance matrix is given by

$$\sigma_F^{i,j} = \left\langle DF^i, DF^j \right\rangle = \sum_{k=0}^J \sum_{l=1}^d D_{k,l}F^i \times D_{k,l}F^j.$$

We introduce now the operator L. Notice that the law of $\overline{Z} = (\overline{Z}_0, \overline{Z}_1, ..., \overline{Z}_J)$ is absolutely continuous and has the density

$$p_{J,x}(z_0, z_1, ..., z_J) = N(z_0) \prod_{k=1}^J q_M(x_{T_k}(x, z_1, ..., z_{k-1}), z_k) \tag{3.32}$$

where N is the density of the standard normal law (so of Δ), q_M is defined in (3.21) and $x_{T_k}(x, z_1, ..., z_{k-1})$ is the solution of (3.23) which starts from x. Then we define

$$LF = \sum_{k=0}^J \sum_{j=1}^d D_{k,j}D_{k,j}F + D_{k,j}F \times D_{k,j} \ln p_{J,x}(\overline{Z}_k).$$

The basic duality relation is the following: for two simple functionals F, G

$$\mathbb{E}(FLG) = \mathbb{E}(GLF) = \mathbb{E}(\langle DF, DG\rangle).$$

Having these objects at hand one proves the following integration by parts formula. Let $F = (F^1, ..., F^d)$ and G be simple functionals and let $\beta = (\beta_1, ..., \beta_q) \in \{1, ..., d\}^q$ be multi index of length q. Then for every $\phi \in C^\infty(\mathbb{R}^d)$

$$\mathbb{E}(\partial_\beta \phi(F)G) = \mathbb{E}(\phi(F)H_\beta(F, G)) \tag{3.33}$$

where $H_\beta(F, G)$ is a random variable which verifies

$$\|H_\beta(F, G\|_p \leq C \left\|(\det \sigma_F)^{-1}\right\|_{4p}^{3q-1} (1 + \|F\|_{q+1,4p}^{(6d+1)q})(1 + \|LF\|_{q-1,4p}^q) \|G\|_{q,4p}. \tag{3.34}$$

This result is proved in Theorem 2 and Theorem 3 in [5]. Before going on we need the following estimates.

Lemma 3.6. *For every multi index $\beta = (\beta_1, ..., \beta_q) \in \{1, ..., d\}^q$ and every $p, R, T \geq 1$*

$$\sup_{|x|\leq R} \mathbb{E}(\sup_{t\leq T} \left|\partial_x^\beta F_t^M(x)\right|_l^p) \leq C \tag{3.35}$$

and

$$\sup_{|x|\leq R} \left\|\partial_x^\beta \ln p_{J,x}(\overline{Z})\right\|_{l,q} \leq CM^d. \tag{3.36}$$

Proof. The proof of (3.35) is analogous to the proof of Lemma 7 and Lemma 9 in [5] so we leave it out. Let us prove (3.36). Notice first that

$$\partial_x^\beta \ln p_{J,x}(z_0, z_1, ..., z_J) = \sum_{k=1}^J \partial_x^\beta \ln q_M(x_{T_k}(x, z_1, ..., z_{k-1}), z_k).$$

On the set $\{q_M > 0\}$ we have

$$\begin{aligned}\partial_x^\beta \ln q_M(x_{T_k}(x, z_1, ..., z_{k-1}), z_k) &= 1_{B_{M+1}}(z_k)\partial_x^\beta \ln \gamma(x_{T_k}(x, z_1, ..., z_{k-1}), z_k) \\ &\quad + 1_{B_{M+1}^c}(z_k)\partial_x^\beta \ln \theta_{M,\gamma}(x_{T_k}(x, z_1, ..., z_{k-1}).\end{aligned}$$

We will use the following easy inequality: for any function $f \in C_b^l$ and every simple functional F in $\mathbb{R}^d$ one has $|f(F)|_l \leq C \|f\|_{l,\infty} |F|_l$ where $\|f\|_{l,\infty} = \sup_x \max_{|\alpha|\leq l} |\partial^\alpha f(x)|$. Notice that for every multi index α one has

$$\partial_x^\beta \theta_{M,\gamma}(x) = -\frac{1}{2\overline{\gamma}\mu(B_{M+1})} \int_{B_{M+1}} \partial_x^\beta \gamma(x, z)dz$$

and moreover $\theta_{M,\gamma}(x) \geq 1/2$. It follows that $\|\ln \theta_{M,\gamma}\|_{l,\infty} \leq C\overline{\gamma}_l/\overline{\gamma}$. One also has $\left\|\partial_x^\beta \ln \gamma\right\|_{l,\infty} \leq \overline{\gamma}_{l+|\beta|}$ so finally $\| \ln q_M(\cdot, z)\|_{l,\infty} \leq C$ with C a constant which depends on $\overline{\gamma}, \overline{\gamma}_l, \overline{\gamma}_{\ln l}$. Then, using the above remark we obtain

$$\left|\partial_x^\beta \ln q_M(x_{T_k}(x, \overline{Z}_1, ..., \overline{Z}_{k-1}), \overline{Z}_k)\right|_l \leq C \left|F_{T_k}^M(x)\right|_l.$$

Consequently

$$\left|\partial_x^\beta \ln p_{J,x}(\overline{Z}_1,...,\overline{Z}_{J_t^M})\right|_l \leq C\sum_{k=1}^{J_t^M}\left|F_{T_k}^M(x)\right|_l \leq J_t^M \times \sup_{s\leq t}\left|F_s^M(x)\right|_l$$

Since $(\mathbb{E}(|J_t^M|^2))^{1/2} = CM^d$ this, together with (3.35), gives

$$\left\|\partial_x^\beta \ln p_{J,x}(\overline{Z}_1,...,\overline{Z}_{J_t^M})\right\|_{l,p} \leq CM^d.$$

□

We are now ready to proceed to the

Proof of Theorem 3.5. In order to avoid notational complications we just look to a particular case (the general case is obviously similar). We assume that we are in the one dimensional case $d = 1$ and $|\alpha| = |\beta| = 1$. Then we look to

$$\partial_x^\alpha \mathbb{E}((\partial^\beta\phi)(F_t^M(x))\psi(F_t^M(x)-y)) = \partial_x\mathbb{E}(\phi'(F_t^M(x))\psi(F_t^M(x)-y)).$$

Let $\nu(du)$ be the standard normal law and $z = (z_1,...,z_J)$. Then, with $\delta = \sqrt{TU_M}$ and $J = J_t^M$, we have

$$\begin{aligned}
& \partial_x\mathbb{E}(\phi'(F_t^M(x))\psi(F_t^M(x)-y)) \\
= \; & \partial_x\mathbb{E}\int\nu(du)\int\phi'(\delta u + x_t(x,z))\psi(\delta u + x_t(x,z)-y))p_{J,x}(z)dz \\
= \; & I_1 + I_2 + I_3
\end{aligned}$$

with

$$\begin{aligned}
I_1 &= \mathbb{E}\int\nu(du)\int\phi''(\delta u + x_t(x,z))\partial_x x_t(x,z)\psi(\delta u + x_t(x,z)-y))p_{J,x}(z)dz \\
I_2 &= \mathbb{E}\int\nu(du)\int\phi'(\delta u + x_t(x,z))\psi'(\delta u + x_t(x,z)-y))\partial_x x_t(x,z)p_{J,x}(z)dz \\
I_3 &= \mathbb{E}\int\nu(du)\int\phi'(\delta u + x_t(x,z))\psi(\delta u + x_t(x,z)-y))\partial_x p_{J,x}(z)dz.
\end{aligned}$$

We stress that $x_t(x,z)$ is defined as the solution of the equation (3.23) and so it depends on $T_k, k \leq J_t^M$. This is why $\mathbb{E}$ appears in the previous expressions. Let us treat I_1. Using the integration by parts formula (3.33)

$$\begin{aligned}
I_1 &= \mathbb{E}(\phi''(F_t^M(x))\partial_x F_t^M(x)\psi(F_t^M(x)-y)) \\
&= \mathbb{E}((\phi(F_t^M(x))H_2(F_t^M(x),\partial_x F_t^M(x)\psi(F_t^M(x)-y)).
\end{aligned}$$

We use now some results from [5]: according to Lemma 13 from we have

$$\left\|LF_t^M(x)\right\|_{l,p} \leq CM; \tag{3.37}$$

according to Lemma 9 we have

$$\left\|F_t^M(x)\right\|_{l,p} \leq C; \tag{3.38}$$

Lemma 16 gives

$$\mathbb{E}((\det \sigma_{F_t^M(x)})^{-p}) \leq C \tag{3.39}$$

(notice that in Lemma 16 one asks that $2dp/t < \theta$ with θ defined in Hypothesis 3.2, iii) pg 630 in [5]; but as said in Example 1, ii) from the above paper, under our hypothesis we have $\theta = \infty$ so our inequality holds for every $t > 0$). Moreover, taking a look to the proofs of the above results, one can see that the estimates (3.37),(3.38),(3.39) are uniform with respect to $x \in B_R$. Then, using (3.34)

$$|I_1| \leq C\|\phi\|_\infty M^2$$

and the estimate is uniform with respect to $x, y \in B_R$. A similar reasoning gives the same inequality for I_2.

We come now to I_3. We write $\partial_x p_{J,x}(z) = \partial_x \ln p_{J,x}(z) \times p_{J,x}(z)$ so that

$$\begin{aligned} I_3 &= \mathbb{E}(\phi'(F_t^M(x))\psi(F_t^M(x) - y)\partial_x \ln p_{J,x}(\overline{Z}_1, ..., \overline{Z}_J)) \\ &= \mathbb{E}((\phi(F_t^M(x))H_1(F_t^M(x), \psi(F_t^M(x) - y)\partial_x \ln p_{J,x}(\overline{Z}_1, ..., \overline{Z}_J)). \end{aligned}$$

Using (3.34) and (3.36) we obtain

$$|I_3| \leq C\|\phi\|_\infty M^2.$$

□

We will use the following approximation result:

Lemma 3.7. *Let (H_2) holds with $r > d$. For every Lipschitz continuous function f with Lipschitz constant less or equal to one, one has*

$$\left|\mathbb{E}(f(F_t^M(x)) - \mathbb{E}(f(X_t(x))\right| \leq CM^{-(r-d)}. \tag{3.40}$$

where C is a constant which is independent of M.

Proof. We have

$$\left|\mathbb{E}(f(F_t^M(x)) - \mathbb{E}(f(\overline{X}_t^M(x))\right| \leq \sqrt{TU_M}\,\mathbb{E}(|\Delta|) \leq CM^{-(r-d/2)},$$

in which we have used (H_2) in order to estimate U_M in (3.25).

Since the law of $\overline{X}_t^M(x)$ and $X_t^M(x)$ coincide, we use Lemma 4 from [5] and (H_2). So, we obtain

$$\begin{aligned} \left|\mathbb{E}(f(F_t^M(x)) - \mathbb{E}(f(X_t(x))\right| &\leq CM^{-(r-d/2)} + \left|\mathbb{E}(f(X_t^M(x)) - \mathbb{E}(f(X_t(x))\right| \\ &\leq CM^{-(r-d/2)} + C\overline{\gamma}\int_{\{|z|>M\}} \overline{c}(z)dz \\ &\leq CM^{-(r-d)}. \end{aligned}$$

□

We are now able to present our main result.

Theorem 3.8. *Assume Hypotheses (H_i), $i = 0, 1, 2$, hold. Let $q \in \mathbb{N}$ and $p > 1$ be such that $d + 2d(q + 1 + d/p_*) < r$, where r is the constant in (H_2). Then, for every $x \in \mathbb{R}^d$ and $t > 0$ the law of $X_t(x)$ is absolutely continuous with respect to the Lebesgue measure. We denote by $p_t(x, y)$ the density. Moreover, for every $R > 0$, $(x, y) \mapsto p_t(x, y)$ belongs to $W^{q,p}(B_R \times B_R)$ and there exists a constant C (depending on R) such that, for every $M \in \mathbb{N}$ and $\varepsilon > 0$*

$$\|p_t - p_t^M\|_{W^{q,p}(B_R \times B_R)} \leq \frac{C}{M^{r-d-2d(q+1+d/p_*)-\varepsilon}}.$$

Remark 3.9. *If $r > 3d + 2d^2$ then Sobolev embedding theorem ensures that $(x, y) \mapsto p_t(x, y)$ is a continuous function. Moreover for every $x_0 \in R^d$ one may find $y_0 \in R^d$ such that $p_t(x_0, y_0) > 0$ (because $p_t(x_0, y)$ is a probability density, so may not be identically null) and consequently one may find $\delta > 0$ such that*

$$\inf_{|y-y_0|\leq\delta} \inf_{|x-x_0|\leq\delta} p_t(x, y) > 0.$$

This property is crucial in order to use Nummelin's splitting method in order to prove convergence to equilibrium, see e.g. [22] , [34] and [35].

Proof. We will use Theorem 2.11 for the following measures. Given $R > 0$ we denote by $\Psi_R(x)$ a smooth function which verifies $1_{B_R} \leq \Psi_R \leq 1_{B_{R+1}}$ and we define

$$f_{R,M}(x, y) = \Psi_R(x)\Psi_R(y)p_t^M(x, y) \quad \text{and} \quad f_R(x, y) = \Psi_R(x)\Psi_R(y)p_t(x, y).$$

We note that

$$\|p_t - p_t^M\|_{W^{q,p}(B_R \times B_R)} \leq \|f_R - f_{R,M}\|_{W^{q,p}(\mathbb{R}^d \times \mathbb{R}^d)}.$$

We will use Theorem 2.11 to estimate the term in the above r.h.s. Let

$$\mu_{R,M}(dx, dy) = f_{R,M}(x, y)dxdy \quad \text{and} \quad \mu_R(dx, dy) = f_R(x, y)dxdy.$$

For a Lipschitz continuous function with Lipschitz constant ≤ 1, one has

$$\begin{aligned}
&\left|\int g d\mu_R - \int g d\mu_R^M\right| \\
&= \left|\int \Psi_R(x)\Big(\mathbb{E}(g(x, X_t(x))\Psi_R(X_t(x)) - \mathbb{E}(g(x, X_t^M(x))\Psi_R(X_t^M(x))\Big)dx\right| \\
&\leq CM^{-(r-d)},
\end{aligned}$$

in which we have used (3.40). Then, $d_1(\mu_R, \mu_R^M) \leq CM^{-(r-d)}$. By (3.31) we also have

$$\|f_{R,M}\|_{2m+q,2m,p} \leq CM^{d(2m+q+d)}.$$

Now, we fix m and we apply Theorem 2.11 $i)$ with

$$\alpha = \alpha(m) = \frac{r - d}{d(q + 2m + d)}$$

and $\eta(M) = M^{r-d}$. So, we obtain that μ_R is absolutely continuous and if f_R denotes its density, we also get

$$\|f_R - f_{R,M}\|_{W^{q,p}(\mathbb{R}^d \times \mathbb{R}^d)} \leq C\frac{1}{M^{(r-d)\theta}} \quad \text{with} \quad \theta = \frac{1}{\alpha} \wedge \Big(1 - \frac{q + 1 + d/p_*}{\alpha m}\Big).$$

Since $\lim_m m\alpha(m) = \frac{r-d}{2d}$ we obtain

$$(r-d)(1-\frac{q+1+d/p_*}{\alpha m}) \to r-d-2d(q+1+d/p_*)$$

and

$$\frac{r-d}{\alpha} = d(q+2m+d) \to \infty$$

So, taking m sufficiently large we obtain, for each $\varepsilon > 0$

$$\|f_R - f_{R,M}\|_{W^{q,p}(\mathbb{R}^d\times\mathbb{R}^d)} \leq \frac{C}{M^{r-d-2d(q+1+d/p_*)-\varepsilon}}.$$

□

Corollary 3.10. *Suppose that $r \geq 3d+2d^2$ and set $k = \lfloor (r-3d-2d^2)/2d \rfloor$. Then for every $R > 0$ and every $\varepsilon > 0$ there exists a constant $C_{R,\varepsilon} \geq 1$ such that for every multi indexes α, β with $|\alpha| + |\beta| \leq k$*

$$\sup_{|x|\leq R, |y|\leq R} \left| \partial_x^\alpha \partial_y^\beta p_t(x,y) - \partial_x^\alpha \partial_y^\beta p_t^M(x,y) \right| \leq \frac{C_{R,\varepsilon}}{M^{r-d-2d(q+1+d/p_*)-\varepsilon}}.$$

Proof. We take $p > 1$ very close to 1 (so that p_* is very large) and

$$q = \frac{r-d}{2d} - 1 - \frac{d}{p_*}, \qquad k = \left\lfloor q - \frac{d}{p} \right\rfloor = \left\lfloor \frac{r-3d-d^2}{2d} \right\rfloor.$$

Then Sobolev embedding theorem says that for $|\alpha| + |\beta| \leq k$

$$\sup_{|x|\leq R, |y|\leq R} \left| \partial_x^\alpha \partial_y^\beta f(x,y) \right| \leq C_R \|f\|_{W^{q,p}(B_R\times B_R)}$$

and we are done. □

A Hermite expansions and density estimates

The aim of this section is to give the proof of Proposition 2.5. We recall that for $\mu \in \mathcal{M}$ and $\mu_n(x) = f_n(x)dx, n \in \mathbb{N}$,

$$\pi_{q,k,m,\mathbf{e}}(\mu, (\mu_n)_n) = \sum_{n=0}^{\infty} 2^{n(q+k)} \beta_{\mathbf{e}}(2^{nd}) d_k(\mu, \mu_n) + \sum_{n=0}^{\infty} \frac{1}{2^{2nm}} \|f_n\|_{2m+q,2m,\mathbf{e}}.$$

Our proposal for this section is to prove the following

Proposition A.1. *Let $q, k \in \mathbb{N}, m \in \mathbb{N}_*$ and $\mathbf{e} \in \mathcal{E}$. There exists a universal constant C (depending on q, k, m, d and e) such that for every $f, f_n \in C^{2m+q}(\mathbb{R}^d), n \in \mathbb{N}$, one has*

$$\|f\|_{q,\mathbf{e}} \leq C\pi_{q,k,m,\mathbf{e}}(\mu, (\mu_n)_n). \tag{A.1}$$

where $\mu(x) = f(x)dx$ and $\mu_n(x) = f_n(x)dx$.

The proof of Proposition A.1 will follow from the next results and properties of Hermite polynomials, so we postpone it at the end of this section.

We begin with a review of some basic properties of Hermite polynomials and functions. The Hermite polynomials on $\mathbb{R}$ are defined by

$$H_n(t) = (-1)^n e^{t^2} \frac{d^n}{dt} e^{-t^2}, \quad n = 0, 1, ...$$

They are orthogonal with respect to $e^{-t^2} dt$. We denote the L^2 normalized Hermite functions by

$$h_n(t) = (2^n n! \sqrt{\pi})^{-1/2} H_n(t) e^{-t^2/2}$$

and we have

$$\int_{\mathbb{R}} h_n(t) h_m(t) dt = (2^n n! \sqrt{\pi})^{-1} \int_{\mathbb{R}} H_n(t) H_m(t) e^{-t^2} dt = \delta_{n,m}.$$

The Hermite functions form an orthonormal basis in $L^2(\mathbb{R})$. For a multi index $\alpha = (\alpha_1, ..., \alpha_d) \in \mathbb{N}^d$ we define the d-dimensional Hermite function

$$\mathcal{H}_\alpha(x) := \prod_{i=1}^{d} h_{\alpha_i}(x_i), \quad x = (x_1, ..., x_d).$$

The d-dimensional Hermite functions form an orthonormal basis in $L^2(\mathbb{R}^d)$. This corresponds to the chaos decomposition in dimension d (but the notation we gave above is slightly different from the one used in probability; see [26], [29] and [23], where Hermite polynomials are used. One may come back by a renormalization). The Hermite functions are the eigenvectors of the Hermite operator $D = -\Delta + |x|^2$, Δ denoting the Laplace operator, and one has

$$D\mathcal{H}_\alpha = (2|\alpha| + d)\mathcal{H}_\alpha \quad \text{with} \quad |\alpha| = \alpha_1 + ... + \alpha_d. \tag{A.2}$$

We denote $W_n = \text{Span}\{\mathcal{H}_\alpha : |\alpha| = n\}$ and we have $L^2(\mathbb{R}^d) = \oplus_{n=0}^{\infty} W_n$.

For a function $\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and a function $f : \mathbb{R}^d \to \mathbb{R}$ we use the notation

$$\Phi \diamond f(x) = \int_{\mathbb{R}^d} \Phi(x, y) f(y) dy.$$

We denote by J_n the orthogonal projection on W_n and we have

$$J_n v(x) = \bar{\mathcal{H}}_n \diamond v(x) \quad \text{with} \quad \bar{\mathcal{H}}_n(x, y) := \sum_{|\alpha|=n} \mathcal{H}_\alpha(x) \mathcal{H}_\alpha(y). \tag{A.3}$$

Moreover, we consider a function $a : \mathbb{R}_+ \to \mathbb{R}$ whose support is included in $[\frac{1}{4}, 4]$ and we define

$$\bar{\mathcal{H}}_n^a(x, y) = \sum_{j=0}^{\infty} a\Big(\frac{j}{4^n}\Big) \bar{\mathcal{H}}_j(x, y) = \sum_{j=4^{n-1}+1}^{4^{n+1}-1} a\Big(\frac{j}{4^n}\Big) \bar{\mathcal{H}}_j(x, y), \quad x, y \in \mathbb{R}^d,$$

the last equality being a consequence of the support property of the function a.

The following estimate is a crucial point in our approach. It has been proved in [14], [13] and then in [27]. We refer to Corollary 2.3, inequality (2.17), in [27] (we thank to G. Kerkyacharian who signaled us this paper).

Theorem A.2. *Let $a : \mathbb{R}_+ \to \mathbb{R}_+$ be a non negative C^∞ function with the support included in $[\frac{1}{4}, 4]$. We denote $\|a\|_l = \sum_{i=0}^{l} \sup_{t\geq 0} \left|a^{(i)}(t)\right|$. For every multi index α and every $k \in \mathbb{N}$ there exists a constant C_k (depending on k, α, d) such that for every $n \in \mathbb{N}$ and every $x, y \in \mathbb{R}^d$*

$$\left|\frac{\partial^{|\alpha|}}{\partial x^\alpha} \bar{\mathcal{H}}_n^a(x,y)\right| \leq C_k \|a\|_k \frac{2^{n(|\alpha|+d)}}{(1+2^n|x-y|)^k}. \tag{A.4}$$

Following the ideas in [27] we consider a function $a : \mathbb{R}_+ \to \mathbb{R}_+$ of class C_b^∞ with the support included in $[\frac{1}{4}, 4]$ and such that $a(t) + a(4t) = 1$ for $t \in [\frac{1}{4}, 1]$. We may construct a in the following way: we take a function $a : [0,1] \to \mathbb{R}_+$ with $a(t) = 0$ for $t \leq \frac{1}{4}$ and $a(1) = 1$. We may choose a such that $a^{(l)}(\frac{1}{4}) = a^{(l)}(1-) = 0$ for every $l \in \mathbb{N}$. Then we define $a(t) = 1 - a(\frac{t}{4})$ for $t \in [1, 4]$ and $a(t) = 0$ for $t \geq 4$. This is the function we will use in the following. Notice that a has the property:

$$\sum_{n=0}^{\infty} a\Big(\frac{t}{4^n}\Big) = 1 \quad \forall t \geq 1. \tag{A.5}$$

In order to check the above equality we fix n_t such that $4^{n_t - 1} \leq t < 4^{n_t}$ and we notice that $a(\frac{t}{4^n}) = 0$ if $n \notin \{n_t - 1, n_t\}$. So $\sum_{n=0}^{\infty} a(\frac{t}{4^n}) = a(4s) + a(s) = 1$ with $s = t/4^{n_t} \in [\frac{1}{4}, 1]$. In the following we fix a function a and the constants in our estimates will depend on $\|a\|_l$ for some fixed l. Using this function we obtain the following representation formula:

Proposition A.3. *For every $f \in L^2(\mathbb{R}^d)$*

$$f = \sum_{n=0}^{\infty} \bar{\mathcal{H}}_n^a \diamond f$$

the series being convergent in $L^2(\mathbb{R}^d)$.

Proof. We fix N and we denote

$$S_N^a = \sum_{n=1}^{N} \bar{\mathcal{H}}_n^a \diamond f, \quad S_N = \sum_{j=1}^{4^N} \bar{\mathcal{H}}_j \diamond f \quad \textit{and} \quad R_N^a = \sum_{j=4^N+1}^{4^{N+1}} (\bar{\mathcal{H}}_j \diamond f) a\Big(\frac{j}{4^{N+1}}\Big).$$

Let $j \leq 4^{N+1}$. For $n \geq N+2$ one has $a(\frac{j}{4^n}) = 0$. So using (A.5) we obtain $\sum_{n=1}^{N} a(\frac{j}{4^n}) = \sum_{n=1}^{\infty} a(\frac{j}{4^n}) - a(\frac{j}{4^{N+1}}) = 1 - a(\frac{j}{4^{N+1}})$. And for $j \leq 4^N$ one has $a(\frac{j}{4^{N+1}}) = 0$. It follows that

$$\begin{aligned} S_N^a &= \sum_{n=1}^{N} \sum_{j=0}^{\infty} a\Big(\frac{j}{4^n}\Big) \bar{\mathcal{H}}_j \diamond f = \sum_{n=1}^{N} \sum_{j=0}^{4^{N+1}} a\Big(\frac{j}{4^n}\Big) \bar{\mathcal{H}}_j \diamond f \\ &= \sum_{j=0}^{4^{N+1}} (\bar{\mathcal{H}}_j \diamond f) \sum_{n=1}^{N} a\Big(\frac{j}{4^n}\Big) \\ &= \sum_{j=0}^{4^{N+1}} \bar{\mathcal{H}}_j \diamond f - \sum_{j=4^N+1}^{4^{N+1}} (\bar{\mathcal{H}}_j \diamond f) a\Big(\frac{j}{4^{N+1}}\Big) = S_{N+1} - R_N^a. \end{aligned}$$

One has $S_N \to f$ in L^2 and $\|R_N^a\|_2 \leq \|a\|_\infty \sum_{j=4^N+1}^{4^{N+1}} \left\|\bar{\mathcal{H}}_j \diamond f\right\|_2 \to 0$ so the proof is completed. □

We will need the following lemma concerning properties of the Luxembourg norms.

Lemma A.4. *Let $\rho \geq 0$ be a measurable function. Then for every measurable function f one has*

$$\|\rho * f\|_{\mathbf{e}} \leq \|\rho\|_1 \|f\|_{\mathbf{e}} . \tag{A.6}$$

Proof. Let $c = m\|f\|_{\mathbf{e}}$ with $m = \|\rho\|_1 = \int \rho(x-y)dy$. Since $\mathbf{e}$ is convex we obtain

$$\begin{aligned}
\int \mathbf{e}\Big(\frac{1}{c}(\rho * f)(x)\Big)dx &= \int \mathbf{e}\Big(\int \frac{\rho(x-y)}{m} \times \frac{m}{c} f(y)dy\Big)dx \\
&\leq \int dx \int \frac{\rho(x-y)}{m} \times \mathbf{e}\Big(\frac{m}{c} f(y)\Big)dy \\
&= \int \mathbf{e}\Big(\frac{m}{c} f(y)\Big) \int \frac{\rho(x-y)}{m} dxdy = \int \mathbf{e}\Big(\frac{m}{c} f(y)\Big)dy \\
&= \int \mathbf{e}\Big(\frac{1}{\|f\|_{\mathbf{e}}} f(y)\Big)dy \leq 1
\end{aligned}$$

and this means that $\|\rho * f\|_{\mathbf{e}} \leq c = \|\rho\|_1 \|f\|_{\mathbf{e}}$. □

Lemma A.5. *Let $\mathbf{e} \in \mathcal{E}$ and $\rho_{n,p}(z) = (1+2^n|z|)^{-p}$, with $p > d$. There exists a constant C_p depending on p and d such that*

$$\|\rho_{n,p}\|_{\mathbf{e}} \leq \frac{1}{\mathbf{e}^{-1}(\frac{1}{C_p}2^{nd})}. \tag{A.7}$$

In particular, for $p = d+1$ there exists a constant C depending on d and on the doubling constant of $\mathbf{e}$ such that (with $\phi_{\mathbf{e}}$ defined in (2.5))

$$\|\rho_{n,d+1}\|_{\mathbf{e}} \leq \frac{C}{\mathbf{e}^{-1}(2^{nd})} = C2^{-nd}\beta_{\mathbf{e}}(2^{nd}) = C\phi_{\mathbf{e}}(\frac{1}{2^{nd}}). \tag{A.8}$$

Proof. Let $c > 0$. By passing in polar coordinates and by using the change of variable $s = 2^n r$, we obtain

$$\begin{aligned}
\int_{\mathbb{R}^d} \mathbf{e}\Big(\frac{1}{c}\rho_{n,p}(z)\Big)dz &= A_d \int_0^\infty r^{d-1}\mathbf{e}\Big(\frac{1}{c} \times \frac{1}{(1+2^n r)^p}\Big)dr \\
&= 2^{-nd} A_d \int_0^\infty s^{d-1}\mathbf{e}\Big(\frac{1}{c} \times \frac{1}{(1+s)^p}\Big)ds
\end{aligned}$$

where A_d is the surface of the unit sphere in $\mathbb{R}^d$. Using the property (2.1) ii) we upper bound the above term by

$$2^{-nd}\mathbf{e}\Big(\frac{1}{c}\Big)A_d \int_0^\infty s^{d-1} \times \frac{1}{(1+s)^p}ds = C_p 2^{-nd}\mathbf{e}\Big(\frac{1}{c}\Big).$$

In order to prove that $\|\rho_{n,p}\|_{\mathbf{e}} \leq c$ we have to check that $\int_{\mathbb{R}^d} \mathbf{e}(\frac{1}{c}\rho_{n,p}(z))dz \leq 1$. In view of the above inequalities it suffices that $\mathbf{e}(\frac{1}{c}) \leq 2^{nd}/C_p$ that is $c \geq 1/\mathbf{e}^{-1}(2^{nd}/C_p)$. □

Proposition A.6. *Let $\mathbf{e} \in \mathcal{E}$ and $\mathbf{e}_*$ be the conjugate of $\mathbf{e}$. Set α as a multi index.*

i) There exists a universal constant C (depending on α, d and $\mathbf{e}$) such that

$$\begin{aligned}
&a) \qquad \big\|\partial_\alpha \bar{\mathcal{H}}_n^a \diamond f\big\|_{\mathbf{e}} \leq C\|a\|_{d+1} \times 2^{n|\alpha|}\|f\|_{\mathbf{e}}, \\
&b) \qquad \big\|\partial_\alpha \bar{\mathcal{H}}_n^a \diamond f\big\|_{\infty} \leq C\|a\|_{d+1} \times 2^{n|\alpha|}\beta_{\mathbf{e}}(2^{nd})\|f\|_{\mathbf{e}_*}
\end{aligned} \tag{A.9}$$

ii) Let $m \in \mathbb{N}_*$. There exists a universal constant C (depending on α, m, d and $\mathbf{e}$) such that

$$\left\|\bar{\mathcal{H}}_n^a \diamond \partial_\alpha f\right\|_{\mathbf{e}} \leq \frac{C \|a\|_{d+1}^2}{4^{nm}} \|f\|_{2m+|\alpha|,2m,\mathbf{e}} \tag{A.10}$$

iii) Let $k \in \mathbb{N}$. There exists a universal constant C (depending on α, k, d and e) such that

$$\left\|\bar{\mathcal{H}}_n^a \diamond \partial_\alpha (f-g)\right\|_{\mathbf{e}} \leq C \|a\|_{d+1} \times 2^{n(|\alpha|+k)} \beta(2^{nd}) d_k(\mu_f, \mu_g) \tag{A.11}$$

Proof. i) By using (A.4) with $k = d+1$ we get

$$\left|\partial_\alpha \bar{\mathcal{H}}_n^a \diamond f(x)\right| \leq C 2^{n(|\alpha|+d)} \|a\|_{d+1} \int \rho_{n,d+1}(x-y) |f(y)| \, dy. \tag{A.12}$$

Since $\mathbf{e}$ is symmetric, i.e. $\mathbf{e}(|x|) = \mathbf{e}(x)$, one has $\|f\|_{\mathbf{e}} = \||f|\|_{\mathbf{e}}$. Moreover, if $0 \leq f(x) \leq g(x)$ then $\|f\|_{\mathbf{e}} \leq \|g\|_{\mathbf{e}}$. Using these properties in addition to (A.12) and (A.6), we obtain

$$\begin{aligned} \left\|\partial_\alpha \bar{\mathcal{H}}_n^a \diamond f\right\|_{\mathbf{e}} &= \left\|\left|\partial_\alpha \bar{\mathcal{H}}_n^a \diamond f\right|\right\|_{\mathbf{e}} \leq C 2^{n(|\alpha|+d)} \|a\|_{d+1} \|\rho_{n,d+1} * |f|\|_{\mathbf{e}} \\ &\leq C 2^{n(|\alpha|+d)} \|a\|_{d+1} \|\rho_{n,d+1}\|_1 \||f|\|_{\mathbf{e}}. \end{aligned}$$

Using (A.8) with $\mathbf{e}(x) = |x|$ we obtain $\|\rho_{n,d+1}\|_1 \leq C/2^{nd}$. So we conclude that

$$\left\|\partial_\alpha \bar{\mathcal{H}}_n^a \diamond f\right\|_{\mathbf{e}} \leq C \|a\|_{d+1} 2^{n|\alpha|} \||f|\|_{\mathbf{e}}$$

so $a)$ is proved. Again by (A.12)

$$\begin{aligned} \left|\partial_\alpha \bar{\mathcal{H}}_n^a \diamond f(x)\right| &\leq C \|a\|_{d+1} 2^{n(|\alpha|+d)} \int \rho_{n,d+1}(x-y) |f(y)| \, dy \\ &\leq C \|a\|_{d+1} 2^{n(|\alpha|+d)} \|\rho_{n,d+1}\|_{\mathbf{e}} \|f\|_{\mathbf{e}_*}, \end{aligned}$$

the second inequality being a consequence of the Hölder inequality (2.6). Using (A.8), $b)$ is proved as well.

ii) We define the functions $a_m(t) = a(t)t^{-m}$. Since $a(t) = 0$ for $t \leq \frac{1}{4}$ and for $t \geq 4$ we have $\|a_m\|_{d+1} \leq C_{m,d} \|a\|_{d+1}$. Moreover $D\bar{\mathcal{H}}_j \diamond v = (2j+d)\bar{\mathcal{H}}_j \diamond v$ so we obtain

$$\bar{\mathcal{H}}_j \diamond v = \frac{1}{2j}(D-d)\bar{\mathcal{H}}_j \diamond v.$$

We denote $L_{m,\alpha} = (D-d)^m \partial_\alpha$ and we notice that

$$L_{m,\alpha} = \sum_{|\beta| \leq 2m} \sum_{|\gamma| \leq 2m+|\alpha|} c_{\beta,\gamma} x^\beta \partial_\gamma,$$

where $c_{\beta,\gamma}$ are universal constants. It follows that there exists some universal constant C such that

$$\|L_{m,\alpha} f\|_{\mathbf{e}} \leq C \|f\|_{2m+|\alpha|,2m,\mathbf{e}}. \tag{A.13}$$

We take now $v \in L^{\mathbf{e}_*}$ and we write

$$\begin{aligned}
\left\langle v, \bar{\mathcal{H}}_n^a \diamond (\partial_\alpha f)\right\rangle &= \left\langle \bar{\mathcal{H}}_n^a \diamond v, \partial_\alpha f\right\rangle = \sum_{j=0}^{\infty} a\Big(\frac{j}{4^n}\Big) \left\langle \bar{\mathcal{H}}_j \diamond v, \partial_\alpha f\right\rangle \\
&= \sum_{j=1}^{\infty} a\Big(\frac{j}{4^n}\Big) \frac{1}{(2j)^m} \left\langle (D-d)^m \bar{\mathcal{H}}_j \diamond v, \partial_\alpha f\right\rangle \\
&= \frac{1}{2^m} \times \frac{1}{4^{nm}} \sum_{j=1}^{\infty} a_m(\frac{j}{4^n}) \left\langle \bar{\mathcal{H}}_j \diamond v, L_{m,\alpha} f\right\rangle \\
&= \frac{1}{2^m} \times \frac{1}{4^{nm}} \left\langle \bar{\mathcal{H}}_n^{a_m} \diamond v, L_{m,\alpha} f\right\rangle .
\end{aligned}$$

By using the decomposition in Proposition A.3, we write $L_{m,\alpha} f = \sum_{j=0}^{\infty} \bar{\mathcal{H}}_j^a \diamond L_{m,\alpha} f$. For $|j-n| \geq 2$, by the support property of a, one has $a(\frac{k}{4^n})a(\frac{k}{4^j}) = 0$ for every $k \in \mathbb{N}$. One also has $\langle \mathcal{H}_\alpha \diamond v, \mathcal{H}_\beta \diamond L_{m,\alpha} f\rangle = 0$ if $|\alpha| \neq |\beta|$. Then a straightforward decomposition gives $\langle \bar{\mathcal{H}}_n^{a_m} \diamond v, \bar{\mathcal{H}}_j^a \diamond L_{m,\alpha} f\rangle = 0$. So using Hölder's inequality

$$\begin{aligned}
\left|\left\langle v, \bar{\mathcal{H}}_n^a \diamond (\partial_\alpha f)\right\rangle\right| &\leq \frac{1}{2^m} \times \frac{1}{4^{nm}} \sum_{j=n-1}^{n+1} \left|\left\langle \bar{\mathcal{H}}_n^{a_m} \diamond v, \bar{\mathcal{H}}_j^a \diamond L_{m,\alpha} f\right\rangle\right| \\
&\leq \frac{1}{2^m} \times \frac{1}{4^{nm}} \sum_{j=n-1}^{n+1} \left\|\bar{\mathcal{H}}_n^{a_m} \diamond v\right\|_{\mathbf{e}_*} \left\|\bar{\mathcal{H}}_j^a \diamond L_{m,\alpha} f\right\|_{\mathbf{e}} .
\end{aligned}$$

Using point *i)* *a)* with α equal to the void index, we obtain $\left\|\bar{\mathcal{H}}_n^{a_m} \diamond v\right\|_{\mathbf{e}_*} \leq C \left\|a_m\right\|_{d+1} \left\|v\right\|_{\mathbf{e}_*} \leq C \times C_{m,d} \left\|a\right\|_{d+1} \left\|v\right\|_{\mathbf{e}_*}$. Moreover, we have

$$\left\|\bar{\mathcal{H}}_j^a \diamond L_{m,\alpha} f\right\|_{\mathbf{e}} \leq C \left\|a\right\|_{d+1} \left\|L_{m,\alpha} f\right\|_{\mathbf{e}} \leq C \left\|a\right\|_{d+1} \left\|f\right\|_{2m+|\alpha|,2m,\mathbf{e}} ,$$

the last inequality being a consequence of (A.13). We obtain

$$\left|\left\langle v, \bar{\mathcal{H}}_n^a \diamond (\partial_\alpha f)\right\rangle\right| \leq \frac{C \left\|a\right\|_{d+1}^2}{4^{nm}} \left\|v\right\|_{\mathbf{e}_*} \left\|f\right\|_{2m+|\alpha|,2m,\mathbf{e}}$$

and, since $L^{\mathbf{e}}$ is reflexive, (A.10) is proved.

iii) We write

$$\begin{aligned}
\left|\left\langle v, \bar{\mathcal{H}}_n^a \diamond (\partial_\alpha (f-g))\right\rangle\right| &= \left|\left\langle \bar{\mathcal{H}}_n^a \diamond v, \partial_\alpha (f-g)\right\rangle\right| = \left|\left\langle \partial_\alpha \bar{\mathcal{H}}_n^a \diamond v, f-g)\right\rangle\right| \\
&= \left|\int \partial_\alpha \bar{\mathcal{H}}_n^a \diamond v d\mu_f - \int \partial_\alpha \bar{\mathcal{H}}_n^a \diamond v d\mu_g\right|.
\end{aligned}$$

We use the definition of d_k and (A.9) *b)* and we obtain

$$\begin{aligned}
&\left|\int \partial_\alpha \bar{\mathcal{H}}_n^a \diamond v d\mu_f - \int \partial_\alpha \bar{\mathcal{H}}_n^a \diamond v d\mu_g\right| \leq \left\|\partial_\alpha \bar{\mathcal{H}}_n^a \diamond v\right\|_{k,\infty} d_k(\mu_f, \mu_g) \\
&\leq \left\|\bar{\mathcal{H}}_n^a \diamond v\right\|_{k+|\alpha|,\infty} d_k(\mu_f, \mu_g) \leq C \left\|a\right\|_{d+1} 2^{n(k+|\alpha|)} \beta_{\mathbf{e}}(2^{nd}) \left\|v\right\|_{\mathbf{e}_*} d_k(\mu_f, \mu_g)
\end{aligned}$$

which implies (A.11). □

We are now ready for the

Proof of Proposition A.1. Let α with $|\alpha| \leq q$. Using Proposition A.3

$$\partial_\alpha f = \sum_{n=1}^{\infty} \bar{\mathcal{H}}_n^a \diamond \partial_\alpha f = \sum_{n=1}^{\infty} \bar{\mathcal{H}}_n^a \diamond \partial_\alpha (f - f_n) + \sum_{n=1}^{\infty} \bar{\mathcal{H}}_n^a \diamond \partial_\alpha f_n$$

and using (A.11) and (A.10)

$$\begin{aligned} \|\partial_\alpha f\|_{\mathbf{e}} &\leq \sum_{n=1}^{\infty} \left\|\bar{\mathcal{H}}_n^a \diamond \partial_\alpha (f - f_n)\right\|_{\mathbf{e}} + \sum_{n=1}^{\infty} \left\|\bar{\mathcal{H}}_n^a \diamond \partial_\alpha f_n\right\|_{\mathbf{e}} \\ &\leq C \sum_{n=1}^{\infty} 2^{n(|\alpha|+k)} \beta_{\mathbf{e}}(2^{nd}) d_k(\mu_f, \mu_{f_n}) + C \sum_{n=1}^{\infty} \frac{1}{2^{2nm}} \|f_n\|_{2m+|\alpha|,2m,\mathbf{e}} \end{aligned}$$

so (A.1) is proved.

□

B Interpolation spaces

In this section we prove that, in the case of the L^p norms, (that is $\mathbf{e} = \mathbf{e}_p$) the space $\mathcal{S}_{q,k,m,\mathbf{e}_p}$ is an interpolation space between $W_*^{k,\infty}$ (the dual of $W^{k,\infty}$) and $W^{q,2m,p}$. A similar interpretation holds for $\mathbf{e}_{\log}$ but this case is more exotic and we do not enter into details here.

To begin we recall the framework of interpolation spaces. We are given two Banach spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ with $X \subset Y$ (with continuous embedding). We denote $\mathcal{L}(X, X)$ the space of the linear bounded operators from X into itself and we denote by $\|L\|_{X,X}$ the operator norm. A Banach space $(W, \|\cdot\|_W)$ such that $X \subset W \subset Y$ is called an interpolation space for X and Y if $\mathcal{L}(X, X) \cap \mathcal{L}(Y, Y) \subset \mathcal{L}(W, W)$. Let $\gamma \in (0, 1)$. If there exists a constant C such that $\|L\|_{W,W} \leq C \|L\|_{X,X}^{\gamma} \|L\|_{Y,Y}^{1-\gamma}$ for every $L \in \mathcal{L}(X, X) \cap \mathcal{L}(Y, Y)$ then W is an interpolation space of order γ. And if one may take $C = 1$ then W is an exact interpolation space of order γ. There are several methods for constructing interpolation spaces. We focus here on the so called K-method. For $y \in Y$ and $t > 0$ one defines $K(y, t) = \inf_{x \in X}(\|y - x\|_Y + t \|x\|_X)$ and

$$\|y\|_\gamma = \int_0^\infty t^{-\gamma} K(y, t) \frac{dt}{t}, \qquad (X, Y)_\gamma = \{y \in Y : \|y\|_\gamma < \infty\}.$$

Then one proves that $(X, Y)_\gamma$ is an exact interpolation space of order γ. One may also use the following discrete variant of the above norm. Let $\gamma \geq 0$. For $y \in Y$ and for a sequence $x_n \in X, n \in \mathbb{N}$ we define

$$\pi_\gamma(y, (x_n)_n) = \sum_{n=1}^{\infty} 2^{n\gamma} \|y - x_n\|_Y + \frac{1}{2^n} \|x_n\|_X \tag{B.1}$$

and

$$\rho_\gamma^{X,Y}(y) = \inf \pi_\gamma(y, (x_n)_n)$$

with the infimum taken over all the sequences $x_n \in X, n \in \mathbb{N}$. Then a standard result in interpolation theory (the proof is elementary) says that there exists a constant $C > 0$ such that

$$\frac{1}{C} \|y\|_\gamma \leq \rho_\gamma^{X,Y}(y) \leq C \|y\|_\gamma \tag{B.2}$$

so that

$$\mathcal{S}_\gamma(X,Y) =: \{y : \rho_\gamma^{X,Y}(y) < \infty\} = (X,Y)_\gamma$$

Take now $q, k \in \mathbb{N}, m \in \mathbb{N}_*$ and $p > 1$ and set $Y = W_*^{k,\infty}$ and $X = W^{q,2m,p}$. Then with the notation from (2.17) and (2.18)

$$\rho_{q,k,m,\mathbf{e}_p}(\mu) = \rho_\gamma^{X,Y}(\mu) \text{ and } \mathcal{S}_{q,k,m,\mathbf{e}_p} = \mathcal{S}_\gamma(X,Y), \text{ with } \gamma = \frac{q+k+d/p_*}{2m} \tag{B.3}$$

Notice that in the definition of $\mathcal{S}_{q,k,m,\mathbf{e}_p}$ one does not use $\pi_\gamma(y,(x_n)_n)$ but $\pi_\gamma^{(m)}(y,(x_n)_n)$ defined by

$$\begin{aligned}\pi_\gamma^{(m)}(y,(x_n)_n) &= \sum_{n=1}^{\infty} 2^{n(q+k+d/p_*)} \|y-x_n\|_Y + \frac{1}{2^{2mn}} \|x_n\|_X \\ &= \sum_{n=1}^{\infty} 2^{2mn\gamma} \|y-x_n\|_Y + \frac{1}{2^{2mn}} \|x_n\|_X\end{aligned}$$

with $\gamma = \frac{q+k+d/p_*}{2m}$. The fact that one uses 2^{2mn} instead of 2^n has no impact except that it changes the constants in (B.2). So the spaces are the same.

We turn now to a different point. For $p > 1$ and $0 < s < 1$ we denote by $\mathcal{B}^{s,p}$ the Besov space and by $\|f\|_{\mathcal{B}^{s,p}}$ the Besov norm (see Triebel [32] for definitions and notations). Our aim is to give a criterion which guarantees that a function f belongs to $\mathcal{B}^{s,p}$. We will use the classical equality $(W^{1,p}, L^p)_s = \mathcal{B}^{s,p}$.

Lemma B.1. *Let $p > 1$ and $0 < s' < s < 1$. Consider a function $\phi \in C^\infty$ such that $\int_{\mathbb{R}^d} \phi(x)dx = 1$ and let $\phi_\delta(x) = \frac{1}{\delta^d}\phi(\frac{x}{\delta})$ and $\phi_\delta^i(x) = x^i\phi_\delta(x)$. We assume that $f \in L^p$ verifies the following hypothesis: for every $i = 1, ..., d$*

$$\begin{aligned} &i) \qquad \limsup_{\delta\to 0} \delta^{1-s} \|\partial_i(f*\phi_\delta)\|_p < \infty \\ &ii) \qquad \limsup_{\delta\to 0} \delta^{-s} \left\|\partial_i(f*\phi_\delta^i)\right\|_p < \infty. \end{aligned} \tag{B.4}$$

Then $f \in \mathcal{B}^{s',p}$ for every $s' < s$.

Proof. Let $f \in C^1$. We use a Taylor expansion of order one and we obtain

$$\begin{aligned} f(x) - f*\phi_\varepsilon(x) &= \int (f(x) - f(x-y))\phi_\varepsilon(y)dy \\ &= \int_0^1 d\lambda \int \langle \nabla f(x-\lambda y), y\rangle \phi_\varepsilon(y)dy = \int_0^1 d\lambda \int \langle \nabla f(x-z), z\rangle \frac{1}{\lambda}\phi_\varepsilon\Big(\frac{z}{\lambda}\Big)\frac{dz}{\lambda^d} \\ &= \int_0^1 d\lambda \int \langle \nabla f(x-z), z\rangle \phi_{\varepsilon\lambda}(z)\frac{dz}{\lambda} = \sum_{i=1}^d \int_0^1 \partial_i(f*\phi_{\varepsilon\lambda}^i)(x)\frac{d\lambda}{\lambda}. \end{aligned}$$

It follows that

$$\|f - f*\phi_\varepsilon\|_p \le \sum_{i=1}^d \int_0^1 \left\|\partial_i(f*\phi_{\varepsilon\lambda}^i)\right\|_p \frac{d\lambda}{\lambda} \le d\,\varepsilon^s \int_0^1 \frac{d\lambda}{\lambda^{1-s}} = C\varepsilon^s.$$

We also have $\|f * \phi_\varepsilon\|_{W^{1,p}} \leq C(1 + \|f\|_\infty)\varepsilon^{-(1-s)}$ so that

$$K(f,\varepsilon) \leq \|f - f * \phi_\varepsilon\|_p + \varepsilon \|f * \phi_\varepsilon\|_{W^{1,p}} \leq C\varepsilon^s.$$

We conclude that for $s' < s$ we have

$$\int_0^1 \frac{1}{\varepsilon^{s'}} K(f,\varepsilon)\frac{d\varepsilon}{\varepsilon} \leq C \int_0^1 \frac{\varepsilon^s}{\varepsilon^{s'}}\frac{d\varepsilon}{\varepsilon} < \infty$$

so $f \in (W^{1,p}, L^p)_{s'} = \mathcal{B}^{s',p}$. □

C Super kernels

A super kernel $\phi : \mathbb{R}^d \to \mathbb{R}$ is a function which belongs to the Schwartz space $\mathcal{S}$ (infinitely differentiable functions which decrease in a polynomial way to infinity), $\int \phi(x)dx = 1$, and such that for every non null multi index $\alpha = (\alpha_1, ..., \alpha_d) \in \mathbb{N}^d$ one has

$$\int y^\alpha \phi(y)dy = 0 \qquad y^\alpha = \prod_{i=1}^d y_i^{\alpha_i}. \tag{C.1}$$

See [19] Section 3, Remark 1 for the construction of a super kernel. The corresponding ϕ_δ, $\delta \in (0,1)$, is defined by

$$\phi_\delta(y) = \frac{1}{\delta^d}\phi\Big(\frac{y}{\delta}\Big).$$

For a function f we denote $f_\delta = f * \phi_\delta$. We will work with the norms $\|f\|_{k,\infty}$, $\|f\|_{k,1}$ and $\|f\|_{q,l,\mathbf{e}}$ defined in (2.8) and in (2.9). And we have

Lemma C.1. *Let $k, q \in \mathbb{N}$. There exists a universal constant $C_{k,q}$ (depending on $k+q$) such that for every $f \in W^{q,1}$ one has*

$$\|f - f_\delta\|_{W_*^{k,\infty}} \leq C_{k,q} \|f\|_{q,1} \delta^{q+k}. \tag{C.2}$$

Proof. Since $C_b^\infty \subset W^{q,1}$ is dense, we may suppose without loss of generality that $f \in C_b^\infty$. Using Taylor expansion of order $q+k$

$$\begin{aligned} f(x) - f_\delta(x) &= \int (f(x) - f(y))\phi_\delta(x-y)dy \\ &= \int I(x,y)\phi_\delta(x-y)dy + \int R(x,y)\phi_\delta(x-y)dy \end{aligned}$$

with

$$\begin{aligned} I(x,y) &= \sum_{i=1}^{q+k-1} \frac{1}{i!} \sum_{|\alpha|=i} \partial^\alpha f(x)(x-y)^\alpha, \\ R(x,y) &= \frac{1}{(q+k)!} \sum_{|\alpha|=q+k} \int_0^1 \partial^\alpha f(x + \lambda(y-x))(x-y)^\alpha \lambda^{k+q} d\lambda. \end{aligned}$$

Using (C.1) we obtain $\int I(x,y)\phi_\delta(x-y)dy = 0$ and by a change of variable we get

$$\int R(x,y)\phi_\delta(x-y)dy = \frac{1}{(q+k)!}\sum_{|\alpha|=q+k}\int_0^1\int dz\phi_\delta(z)\partial^\alpha f(x+\lambda z)z^\alpha\lambda^{k+q}d\lambda.$$

We consider now $g \in W^{k,\infty}$ and we write

$$\begin{aligned}&\int (f(x)-f_\delta(x))g(x)dx\\ &= \frac{1}{(q+k)!}\sum_{|\alpha|=q+k}\int_0^1 \lambda^{k+q}d\lambda\int dz\phi_\delta(z)z^\alpha\int \partial^\alpha f(x+\lambda z)g(x)dx.\end{aligned}$$

Let us denote $f_a(x) = f(x+a)$. We have $(\partial^\alpha f)(x+a) = (\partial^\alpha f_a)(x)$. Let α with $|\alpha| = \sum_{i=1}^d \alpha_i = q+k$. We split α into two multi indexes β and γ such that $|\beta| = k, |\gamma| = q$ and $\partial^\beta\partial^\gamma = \partial^\alpha$ (this may be done in several ways but any one of them is good for us). Then using integration by parts

$$\begin{aligned}\left|\int \partial^\alpha f(x+\lambda z)g(x)dx\right| &= \left|\int \partial^\beta\partial^\gamma f_{\lambda z}(x)g(x)dx\right|\\ &\le \int |\partial^\gamma f_{\lambda z}(x)|\left|\partial^\beta g(x)\right|dx \le \|g\|_{k,\infty}\int |\partial^\gamma f_{\lambda z}(x)|\,dx\\ &= \|g\|_{k,\infty}\int |\partial^\gamma f(x)|\,dx \le \|g\|_{k,\infty}\|f\|_{q,1}.\end{aligned}$$

For a multi index with $|\alpha| = q+k$ we have

$$\int |\phi_\delta(z)|\,|z^\alpha|\,dz \le \delta^{q+k}\int |\phi(z)|\,|z|^{q+k}\,dz$$

so the proof is completed. □

Remark C.2. *It is clear from the above proof that if $q+k$ is fixed then we do not need to work with a "super" kernel ϕ verifying (C.1) for every α but only with a kernel ϕ_{q+k} of order $q+k$, that is verifying (C.1) for $|\alpha| \le q+k$. The reason to use super kernels (and not a kernel of a given fixed order) is just to avoid to precise each time which is the order of the kernel we need. And this simplifies the already heavy notations.*

Lemma C.3. *i) Let $k, q \in \mathbb{N}, l > d$ and $\mathbf{e} \in \mathcal{E}$. There exists a universal constant $C_{k,q}$ (depending on $q+k$) such that for every $f \in W^{q,l,\mathbf{e}}$ one has*

$$\|f - f_\delta\|_{W_*^{k,\infty}} \le C_{k,q}\|f\|_{q,l,\mathbf{e}}\,\delta^{q+k}. \tag{C.3}$$

ii) Let $l > d, n, q \in \mathbb{N}$, with $n \ge q$, and $\mathbf{e} \in \mathcal{E}$. There exists a universal constant $C_{l,q}$ (depending on l, q, d) such that

$$\|f_\delta\|_{n,l,\mathbf{e}} \le C_{l,q}\|f\|_{q,l,\mathbf{e}}\,\delta^{-(n-q)}. \tag{C.4}$$

Proof. i) Let γ with $|\gamma| \leq q$. We write $\partial^\gamma f(x) = u_l(x) v_\gamma(x)$ with $u_l(x) = (1+|x|)^{-l}$ and $v_\gamma(x) = (1+|x|)^l \partial^\gamma f(x)$. Using Hölder inequality

$$\int |\partial^\gamma f(x)|\, dx \leq C \|u_l\|_{\mathbf{e}_*} \|v_\gamma\|_{\mathbf{e}} \leq C \|u_l\|_{\mathbf{e}_*} \|f\|_{q,l,\mathbf{e}}.$$

By Remark 2.1 $\|u_l\|_{\mathbf{e}_*} < \infty$. This gives $\|f\|_{q,1} \leq C \|f\|_{q,l,\mathbf{e}}$ and (C.3) follows from (C.2).

ii) Let α be a multi index with $|\alpha| = n$ and let β, γ be a splitting of α with $|\beta| = q$ and $|\gamma| = n-q$. Using the triangle inequality, for every y we have $1+|x| \leq (1+|y|)(1+|x-y|)$. Then

$$\begin{aligned} u(x) := (1+|x|)^l |\partial^\alpha f_\delta(x)| &= (1+|x|)^l \left|\partial^\beta f * \partial^\gamma \phi_\delta(x)\right| \\ &\leq \int (1+|x|)^l \left|\partial^\beta f(y)\right| |\partial^\gamma \phi_\delta(x-y)|\, dy \leq \alpha * \beta(x) \end{aligned}$$

with

$$\alpha(y) = (1+|y|)^l \left|\partial^\beta f(y)\right|, \qquad \beta(z) = (1+|z|)^l |\partial^\gamma \phi_\delta(z)|.$$

Using (A.6) we obtain

$$\|u\|_{\mathbf{e}} \leq \|\alpha * \beta\|_{\mathbf{e}} \leq \|\beta\|_1 \|\alpha\|_{\mathbf{e}} \leq \frac{C}{\delta^{n-q}} \|\alpha\|_{\mathbf{e}} = \frac{C}{\delta^{n-q}} \|f_{\beta,l}\|_{\mathbf{e}}.$$

□

References

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

[2] Bally, V. and Caramellino, L. (2014). On the distances between probability density functions. *Electron. J. Probab.* **19**, no. 110, 1-33. MR 3296526

[3] Bally, V. and Caramellino, L. (2012). Regularity of probability laws by using an interpolation method. Preprint `arxiv:1211.0052`

[4] Bally, V. and Caramellino, L. (2013). Regularity of Wiener functionals under a Hörmander type condition of order one. Preprint `arXiv:1307.3942`.

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

[6] Bally, V. and Clément, E. (2011). Integration by parts formulas with respect to jump times and stochastic differential equations. *Stochastic Analysis 2010*, ed. Dan Crisan, Springer. MR 2789077

[7] Bally, V. and Fournier, N. (2011). Regularization properties od the 2D homogeneous Boltzmann equation without cutoff. *Probab. Theory Related Fields* **151**, 659-704. MR 2851696

[8] Bally, V. and Pardoux, E. (1998). Malliavin Calculus for White Noise Driven Parabolic SPDE's. *Potential Anal.* **9**, 27-64. MR 1644120

[9] Bennett, C. and Sharpley, R. (1988). Interpolation of operators. Academic Press INC. MR 0928802

[10] Debussche, A. and Fournier, N. (2013). Existence of densities for stable-like driven SDE's with Holder continuous coefficients. *J. Funct. Anal.* **264**, 1757-1778. MR 3022725

[11] Debussche, A. and Romito, M. (2014). Existence of densities for the 3D Navier–Stokes equations driven by Gaussian noise. *Probab. Theory Related Fields* **158**, 575-596. MR 3176359

[12] De Marco, S. (2011). Smoothness and Asymptotic Estimates of densities for SDEs with locally smooth coefficients and Applications to square-root diffusions. *Ann. Appl. Probab.* **21**, 1282-1321. MR 2857449

[13] Dziubanski, J. (1997). Triebel-Lizorkin spaces associated with Laguerre and Hermite expansions. *Proc. Amer. Math. Soc.* **125**, 3547-3554. MR 1403122

[14] Epperson, Jay (1997). Hermite and Laguerre wave packet expansions. *Studia Math.* **126**, 199-217. MR 1475919

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

[16] Fournier, N. (2008). Smoothness of the law of some one-dimensional jumping SDE's with non constant rate of jump. *Electron. J. Probab.* **13**, 135-156. MR 2375602

[17] Fournier, N. (2015). Finiteness of entropy for the homogeneous Boltzmann equation with measure initial condition. *Ann. Appl. Probab.* **25**, 860-897. MR 3313757

[18] Fournier, N. and Printems, J. (2010). Absolute continuity of some one dimensional processes. *Bernoulli* **16**, 343-360. MR 2668905

[19] Kebaier, A. and Kohatsu-Higa, A. (2008). An optimal control variance reduction method for density estimation. *Stochastic Process. Appl.* **118**, 2143–2180. MR 2474346

[20] Kosmol, P. and Mûller-Wichards, D. (2011). *Optimization in Functional Spaces. With stability considerations in Orlicz spaces.* De Gruyter Series in Nonlinear Analysis and Applications 13. Walter de Gruyter. MR 2760903

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

[22] Löcherbach, E., Loukianova, D. and Loukianov, O. (2011). Polynomial bounds in the ergodic theorem for one-dimensional diffusions and integrability of hitting times. *Ann. Inst. Henri Poincaré Probab. Stat.* **47**, 425-449. MR 2814417

[23] Malliavin, P. (1997). *Stochastic Analysis.* Springer. MR 1450093

[24] Ninomiya, S. and Victoir, N. (2008). Weak approximation of stochastic differential equations and applications to derivative pricing. *Appl. Math. Finance* **15**, 107-121. MR 2409419

[25] Nourdin, I. and Peccati, G. (2012). *Normal approximations with Malliavin calculus. From Stein's method to universality*. Cambridge Tracts in Mathematics, 192. Cambridge University Press. MR 2962301

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

[27] Petrushev, P. and Xu, Y. (2008). Decomposition of spaces of distributions induced by Hermite expansions. *J. Fourier Anal. and Appl.* **14**, 372-414. MR 2399106

[28] Pardoux, E. and Zhang, T. (1993). Absolute continuity for the law of the solution of a parabolic SPDE. *J. Funct. Anal.* **112**, 447-458. MR 1213146

[29] Sanz Solè, M. (2005). *Malliavin Calculus, with Applications to Stochastic Partial Differential Equations*. EPFL Press. Fundamental Sciences, Mathematics. MR 2167213

[30] Talay, D. and Tubaro, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.* **8**, 94-120. MR 1091544

[31] Triebel, H. (1995). *Interpolation theory, function spaces, differential operators*. Second edition. Johann Ambrosius Barth, Heidelberg. MR 1328645

[32] Triebel, H. (2006). *Theory of function spaces III*. Monographs in Mathematics, 100. Birkhäuser Verlag. MR 2250142

[33] Walsh, J.B. (1986). *An introduction to stochastic partial differential equations*. In Ecole d'Eté de Probabilités de Saint Flour XV, Lecture Notes in Math. 1180, Springer; pp 226-437. MR 0876085

[34] Veretennikov, A. Yu. (1997). On polynomial mixing bounds for stochastic differential equations. *Stochastic Process. Appl.* **70**, 115-127. MR 1472961

[35] Veretenikov, A. Yu. and Klokov, S.A. (2005). On the subexponential rate of mixing for Markov processes. *Theory Probab. Appl.* **49**, 110-122(translation from Russian of *Teoriya Veroyatn. Primen.* **49**, 21-35, 2004). MR 2141328