UDC 519.21

A. M. Kulik (Inst. Math. Nat. Acad. Sci. Ukraine, Kyiv)

MALLIAVIN CALCULUS FOR DIFFERENCE APPROXIMATIONS OF MULTIDIMENSIONAL DIFFUSIONS: TRUNCATED LOCAL LIMIT THEOREM*

ЧИСЛЕННЯ МАЛЛЯВЕНА ДЛЯ РІЗНИЦЕВИХ НАБЛИЖЕНЬ БАГАТОВИМІРНИХ ДИФУЗІЙ: ЛОКАЛЬНА ГРАНИЧНА ТЕОРЕМА ЗІ ЗРІЗАННЯМ

For difference approximations of multidimensional diffusions, the truncated local limit theorem is proved. Under very mild conditions on the distributions of difference terms, this theorem states that the transition probabilities of these approximations, after truncation of some asymptotically negligible terms, possess densities that converge uniformly to the transition probability density for the limiting diffusion and satisfy certain uniform diffusion-type estimates. The proof is based on a new version of the Malliavin calculus for the product of a finite family of measures that may contain non-trivial singular components. Applications to the uniform estimation of mixing and convergence rates for difference approximations of stochastic differential equations and to the convergence of difference approximations of local times of multidimensional diffusions are given.

Для різницевих наближень багатовимірних дифузій доведено локальну граничну теорему зі зрізанням. При дуже слабких умовах на розподіли різницевих членів ця теорема стверджує, що ймовірності переходу таких наближень після видалення певних доданків, якими в асимптотичному сенсі можна знехтувати, мають щільності, які рівномірно прямують до щільності ймовірності переходу граничної дифузії та задовольняють певні рівномірні оцінки дифузійного типу. Доведення базується на новому варіанті числення Маллявена для добутку скінченної сім'ї мір, які можуть містити нетривіальні сингулярні компоненти. Наведено застосування до рівномірного оцінювання коефіцієнта перемішування та швидкості збіжності для різницевих наближень стохастичних диференціальних рівнянь та до збіжності різницевих наближень локальних часів багатовимірних дифузій.

Introduction. Consider a diffusion process X in $\mathbb{R}^d$ defined by the stochastic differential equation

$$X(t) = X(0) + \int_0^t a(X(s))\,ds + \int_0^t b(X(s))\,dW(s), \quad t \in \mathbb{R}^+, \tag{0.1}$$

and a sequence of processes X_n, $n \geq 1$, with their values at the time moments $\frac{k}{n}$, $k \in \mathbb{N}$, defined by a difference relations

$$X_n\left(\frac{k}{n}\right) = X_n\left(\frac{k-1}{n}\right) + a\left(X_n\left(\frac{k-1}{n}\right)\right)\frac{1}{n} + b\left(X_n\left(\frac{k-1}{n}\right)\right)\frac{\xi_k}{\sqrt{n}}, \tag{0.2}$$

and, at all the other time moments, defined in a piece-wise linear way:

*The research was partially supported by the Ministry of Education and Science of Ukraine, project № GP/F26/0106.

$$X_n(t) = X_n\left(\frac{k-1}{n}\right) + (nt-k+1)\left[X_n\left(\frac{k}{n}\right) - X_n\left(\frac{k-1}{n}\right)\right], \quad t \in \left[\frac{k-1}{n}, \frac{k}{n}\right). \tag{0.3}$$

Here and below, W is a Wiener process valued in $\mathbb{R}^d$, $\{\xi_k\}$ is a sequence of i.i.d. random vectors in $\mathbb{R}^d$, that belong to the domain of attraction of the normal law, are centered and have the identity for covariance matrix. Under standard assumptions on the coefficients of the equations (0.1), (0.2) (local Lipschitz condition and linear growth condition), the distributions of the processes X_n in $C(\mathbb{R}^+, \mathbb{R}^d)$ with the given initial value $X_n(0) = x$ converge weakly to the distribution of the process X with $X(0) = x$ [1]. Thus, it is natural to call the sequence $\{X_n\}$ *the difference approximation* for the diffusion X.

Consider the transition probabilities for the processes X, X_n:

$$P_{x,t}(dy) \equiv \mathrm{P}\big(X(t) \in dy \big| X(0) = x\big),$$

$$P^n_{x,t}(dy) \equiv \mathrm{P}\big(X_n(t) \in dy \big| X_n(0) = x\big), \quad t > 0, \quad x \in \mathbb{R}^d.$$

It is well known [2] that if the coefficients a, b are Hölder continuous and bounded and the matrix $b \cdot b^*$ is uniformly non-degenerate, then $P_{x,t}(dy) = p_t(x,y)\,dy$. The function $\big\{p_t(x,y), t \in \mathbb{R}^+, x, y \in \mathbb{R}^d\big\}$ (the *transition probability density* for X) possesses the estimate

$$p_t(x,y) \le C(T)\, t^{-\frac{d}{2}} \exp\left(-\frac{\gamma\|y-x\|^2}{t}\right), \quad t \le T, \quad x, y \in \mathbb{R}^d. \tag{0.4}$$

The general question, that motivates the present paper, is whether any (more or less restrictive) conditions can be imposed on the coefficients a, b and the distribution of ξ_k in order to provide that $P^n_{x,t}(dy) = p^n_t(x,y)\,dy$ for n large enough, the densities p^n possess an estimate analogous to (0.4) and p^n converge to p in an appropriate way. Such a question both is interesting by itself and has its origin in the numerous applications, such as nonparametric estimation problems in time series analysis and diffusion models (see the discussion in the Introduction to [3]), the uniform bounds for the mixing coefficients of the difference approximations to stochastic differential equations (see [4] and Subsection 4.1 below), the difference approximation for local times of multidimensional diffusions (see [5] and Subsection 4.2 below).

In the current paper, we consider the question exposed above in a slightly modified setting. For the distributions P^n, we prove the result that we call *the truncated local limit theorem*. Let us explain this term. We show that the kernel P^n can be decomposed into the sum $P^n = Q^n + R^n$ in such a way that both Q^n and R^n are a non-negative kernels and

(i) for Q^n, its density q^n exists, satisfies an analogue of (0.4) and converges to p;

(ii) for R^n, its total mass can be estimated explicitly and converges to 0.

The kernel Q^n represents *the main term* of the distribution P^n and satisfies the local limit theorem; the kernel R^n represents *the remainder term*, and typically decreases rapidly (see statements (iii) and (iii′) of Theorem 1.1 below). Such kind of a representation appears to be powerful enough to provide non-trivial applications (see Section 2 below). On the other hand, the conditions that we impose on the distribution of ξ_k in order to provide such a decomposition to exist are very mild; in a simplest cases, these conditions have "if and only if" form (see Theorem 1.2 below). Our main tool in the current research is a certain modification of the *Malliavin calculus*.

Let us make a brief overview of the bibliography in the field. Malliavin calculus have not been used widely for studying the properties of the distributions of the processes defined by the difference relations of the type (0.2), (0.3). The only paper in this direction available to the author is [4], where rather restrictive conditions are imposed both on the coefficients (b should be constant) and the distribution of ξ_k (it should possess the density from the class C^d). A powerful group of results is presented in the papers [3, 6], where a modification of the parametrix method in a difference set-up is developed. When applied to the problem formulated above, the results of [3, 6] allow one to prove that p^n converge to p with the best possible rate $O\left(\frac{1}{\sqrt{n}}\right)$. However, conditions imposed on the distribution of ξ_k in [3, 6], are somewhat more restrictive than those used in our approach. For instance, condition (A_2) of [3] requires, in our settings, ξ_k to possess the density of the class $C^4(\mathbb{R}^d)$ (compare with the condition (B_3) in Theorem 1.1 below).

The paper is organized in the following way. In Section 1, we formulate the main theorem of the paper together with its particular version, that is an intermediate between classic Gnedenko's and Prokhorov's local limit theorems. In the same section, we discuss briefly some possible improvements of the main result. In Section 2, two applications are given. In Section 3, the construction of the *partial Malliavin calculus*, that is our main tool, is explained in details. In Section 4, the proofs of the main results are given.

1. The main results. ***1.1. Formulation.*** Let us introduce the notation. We write $\|\cdot\|$ for the Euclidean norm, not indicating explicitly the space this norm is written for. The adjoint matrix for the matrix A is denoted by A^*. The classes of functions, that have k continuous derivatives, and functions, that are continuous and bounded together with their k derivatives, are denoted by C^k and C^k_b, correspondingly. The derivative (the gradient) is denoted by ∇, the partial derivative w.r.t. the variable x_r is denoted by ∂_r. The Lebesgue measure on $\mathbb{R}^d$ is denoted by λ^d. For the measure μ on $\mathfrak{B}(\mathbb{R}^d)$, μ^{ac} denotes its absolutely continuous component w.r.t. λ^d. Any time the kernel P^n is decomposed into a sum $P^n = Q^n + R^n$, we mean that the kernels Q^n, R^n are non-negative; the same convention is used for decompositions of measures, also. In order to simplify notation we consider the processes defined by (0.1), (0.2) and (0.3) for $t \in [0,1]$ only. Of course, all the statements given below have their straightforward analogues on an arbitrary finite time interval $[0,T]$.

Through all the paper, κ is a fixed integer, $\kappa \geq 4$. We denote $\epsilon(\kappa) = \dfrac{\kappa^2 - 3\kappa - 2}{2\kappa + 2}$. We also denote, by μ, the distribution of ξ_1.

Theorem 1.1. *Let the following conditions hold true*:

(B_1) $a \in C_b^{(d+2)^2}(\mathbb{R}^d, \mathbb{R}^d)$, $b \in C_b^{(d+2)^2}(\mathbb{R}^d, \mathbb{R}^{d\times d})$ *and there exists* $\beta = \beta(b) > 0$ *such that*

$$(b(x)b^*(x)v, v)_{\mathbb{R}^d} \geq \beta \|v\|^2, \quad x,\ v \in \mathbb{R}^d;$$

(B_2^κ) $\mathrm{E}\|\xi_1\|^\kappa < +\infty$;

(B_3) *there exist* $\alpha \in (0,1)$ *and bounded open set* $U \subset \mathbb{R}^d$ *such that*

$$\frac{d\mu^{ac}}{d\lambda^d} \geq \frac{\alpha}{\lambda^d(U)} \mathbf{1}_U \quad \lambda^d\text{-a.s.}$$

Then P^n *can be represented in the form* $P^n = Q^n + R^n$ *in such a way that*

(i) $Q^n_{x,t}(dy) = q^n_t(x,y)\,dy$ *and* $q^n \to p$, $n \to +\infty$, *uniformly on the set* $[\delta,1]\times\mathbb{R}^d\times$ $\times\,\mathbb{R}^d$ *for every* $\delta \in (0,1)$;

(ii) *there exist constants* B, C, $\gamma > 0$ *such that, for* $t \in [0,1]$,

$$q^n_t(x,y) \le \begin{cases} Ct^{-\frac{d}{2}} \exp\left(-\dfrac{\gamma\|x-y\|^2}{t}\right), & \|x-y\| \le tBn^{\frac{1}{\kappa+1}}, \\ Ct^{-\frac{d}{2}} \exp\left(-\gamma n^{\frac{1}{\kappa+1}}\|x-y\|\right), & \|x-y\| > tBn^{\frac{1}{\kappa+1}}, \end{cases}$$

in addition, for every $p > 1$ *there exists* $C_p > 0$ *such that, for* $t \in [0,1]$, $x, y \in \mathbb{R}^d$,

$$q^n_t(x,y) \le C_p t^{-\frac{d}{2}} \left(1 + \frac{\|x-y\|^2}{t}\right)^{-p};$$

(iii) *there exist constants* D, $\rho > 0$ *such that* $R^n_{x,t}(\mathbb{R}^d) \le D\big[n^{-\epsilon(\kappa)} + e^{-\rho n t}\big]$, $x \in \mathbb{R}^d$, $t \in [0,1]$.

If the condition (B_2^κ) *is replaced by the stronger condition*

$(\mathrm{B}_2^{\exp})$ $\exists \varkappa > 0$ *such that* $\mathrm{E}\exp[\varkappa\|\xi_k\|^2] < +\infty$,

then the following stronger analogues of (ii), (iii) *hold true*:

(ii′) *there exist constants* C, $\gamma > 0$ *such that*

$$q^n_t(x,y) \le Ct^{-\frac{d}{2}} \exp\left(-\frac{\gamma\|x-y\|^2}{t}\right), \quad t \in [0,1], \quad x,\ y \in \mathbb{R}^d;$$

(iii′) *there exist constants* D, $\rho > 0$ *such that* $R^n_{x,t}(\mathbb{R}^d) \le De^{-\rho n t}$, $x \in \mathbb{R}^d$, $t \in$ $\in [0,1]$.

Let us formulate separately a modification of Theorem 1.1 in the most studied partial case is $a \equiv 0, b \equiv I_{\mathbb{R}^d}$. In this case, $X_n(1)$ is just the normalized sum $n^{-\frac{1}{2}}\sum_{k=1}^n \xi_k$ and the limiting behavior of the distributions of such kind of a sums is given by the Central Limit Theorem. For the densities of the truncated distributions, the following criterium can be derived. We denote by P_n the distribution of $n^{-\frac{1}{2}}\sum_{k=1}^n \xi_k$.

Theorem 1.2. *The following statements are equivalent*:

1. *There exists* $n_0 \in \mathbb{N}$ *such that* $[P_{n_0}]^{ac}$ *is not equal to zero measure.*
2. *There exists a representation of* P_n *in the form* $P_n = Q_n + R_n$, *such that*

(2i) $Q_n(dy) = q_n(y)\,dy$ *and* $\sup_{y\in\mathbb{R}^d}\left|q_n(y) - (2\pi)^{-\frac{d}{2}} e^{-\frac{\|y\|^2}{2}}\right| \to 0$, $n \to \infty$;

(2ii) *there exist constants* D, $\rho > 0$ *such that* $R_n(\mathbb{R}^d) \le De^{-\rho n}$, $n \in \mathbb{N}$.

The well known theorem by Prokhorov states that the given above statement 1 is equivalent to L_1-convergence of the density of $[P_n]^{ac}$ to the standard normal density (see [7], Theorem 4.4.1 for the case $d = 1$). There exist examples showing that, even while $P_1 \ll \lambda^d$, the density of P_n may fail to converge to the standard normal density uniformly (see [7], Ch. 4, § 3 for the example by Kolmogorov and Gnedenko). The criterium of the uniform convergence is given by another well known theorem by Gnedenko: such a convergence holds if and only if there exists $n_0 \in \mathbb{N}$ such that P_{n_0} possesses a *bounded* density (see [7], Theorem 4.3.1 for the case $d = 1$). Theorem 1.2 shows the following curious feature: under condition of the Prokhorov's criterium, some exponentially negligible remainder term can be removed from the total distribution in such a way that, for the truncated distribution, the statement of the Gnedenko's theorem holds. This feature does not seem to be essentially new; one can provide it by using the

Fourier transform technique, that is the standard tool in the proofs of the Prokhorov's and Gnedenko's theorems. We give a simple proof of Theorem 1.2 using the partial Malliavin calculus, developed in Section 3 below. This illustrates that the partial Malliavin calculus is a powerful tool that allow one to provide local limit theorems in a precise (in some cases, an "if and only if") form.

1.2. Some possible improvements. In the present paper, in order to keep exposition reasonably short and transparent, we formulate the main results not in their widest possible generality. In this subsection, we discuss shortly what kind of improvements can be made in the context of our research.

1. Difference relation (0.2) is written w.r.t. uniform partitions $\{0 = t_n^0 < t_n^1 < \ldots\}$, $t_n^k = \dfrac{k}{n}$, $k \in \mathbb{Z}^+, n \in \mathbb{N}$. Without a significant change of the proofs, one can prove analogues of Theorem 1.1 for the processes, defined by the difference relations of the type (0.2) with $\dfrac{\xi_k}{\sqrt{n}}$ replaced by $\xi_k\sqrt{t_n^k - t_n^{k-1}}$ and partitions $\{t_n^k\}$ satisfying condition

$$\exists c,\ C,\ d,\ D > 0\colon\ \lim\inf_{n\to+\infty}\frac{1}{n}\#\left\{k\middle|t_n^k \le t, (t_n^k - t_n^{k-1}) \in \left[\frac{c}{n}, \frac{C}{n}\right]\right\} \ge td,$$

$$\lim\sup_{n\to+\infty}\frac{1}{n}\#\left\{k\middle|t_n^k \le t, (t_n^k - t_n^{k-1}) \in \left[\frac{c}{n}, \frac{C}{n}\right]\right\} \le tD, \quad t \in (0,1]. \tag{1.1}$$

2. One can, without a significant change of the proofs, replace the sequence of i.i.d. random vectors $\{\xi_k\}$ in (0.2) by a triangular array $\{\xi_{n,k}, k \le n\}$ of independent random vectors, possibly not identically distributed, having zero mean and identity for the covariance matrix. Under such a modification, condition (B_2^κ) should be replaced by $\sup_{n,k} \mathrm{E}\|\xi_{n,k}\|^\kappa < +\infty$, and condition ($B_3$) by

(B_3') $\exists \alpha,\ r > 0,\ x_n \in \mathbb{R}^d\colon\ \dfrac{d[\mu_{n,k}]^{ac}}{d\lambda^d} \ge \alpha \mathbf{1}_{B(x_n,r)}\ \lambda^d$-a.s.,

here $\mu_{n,k}$ denotes the distribution of $\xi_{n,k}$, $B(x,r)$ denotes the open ball in $\mathbb{R}^d$ with the centrum x and radius r. Also, the phase space for $\xi_{n,k}$ may be equal $\mathbb{R}^m$ with $m \ge d$ (note that the case $m < d$ is excluded by the condition (B_1)).

3. Under an appropriate regularity conditions on a, b, Malliavin's representation, analogous to (3.25), can be written for the derivatives of the truncated density of an arbitrary order with respect to both x and y. Thus, after some standard technical steps, one can obtain the following estimate, that generalize statement (ii′) of Theorem 1.1: for a given $k, l \in \mathbb{N}$,

$$\frac{\partial^{k+l}}{\partial x^k \partial y^l} q_t^n(x,y) \le C_{k+l} t^{-\frac{d+k+l}{2}} \exp\left(-\frac{\gamma\|y-x\|^2}{t}\right)$$

under (B_1), ($B_2^{\exp}$), (B_3) and $a \in C_b^{(d+k+l+1)^2}(\mathbb{R}^d, \mathbb{R}^d)$, $b \in C_b^{(d+k+l+1)^2}(\mathbb{R}^d, \mathbb{R}^{d\times d})$.

4. Theorem 3.1 provides the truncated limit theorem without essential restrictions on the structure of the functionals. For instance, one can apply this theorem in order to obtain a truncated local limit theorem for difference approximations of integral functionals, etc.

5. Like the Malliavin calculus for (continuous time) diffusion processes, the partial Malliavin calcucus, developed in Section 3, can be applied when the diffusion matrix is not uniformly elliptic, but locally elliptic, only. However, the changes that should be done in the proof are significant; in particular, Theorem 3.1 is not powerful enough

to cover this case. Thus we postpone the detailed investigation of this case (and more generally, the case of coefficients satisfying an analogue of Hörmander condition) to some further research.

6. In the present paper, we concentrate on the individual estimates for the densities q_t^n and do not deal with the convergence rate in the statement (i) of Theorem 1.1. The (seemingly) possible way to establish such a rate is to write the Malliavin's representation, analogous to (3.25), for the limiting density p and then construct both the functionals $X_n(t) \equiv f_n$ and $X(t) = f$ and the corresponding *weights* Υ^{f_n} and Υ^f, involved into the Malliavin's representation, on the same probability space with a controlled L_2-distance between (f_n, Υ^{f_n}) and (f, Υ^f). Since the question about the estimates in the *strong invariance principle* for the pair (f_n, Υ^{f_n}) is far from being trivial, we postpone the detailed investigation of the rate of convergence in Theorem 1.1 to some further research. We remark that the modification of the parametrix method, developed in [3, 6], provides, under more restrictive conditions on the distribution of $\{\xi_k\}$, the best possible converence rate $O\left(\frac{1}{\sqrt{n}}\right)$.

2. Applications. In this section, we formulate two applications of Theorem 1.1. The proofs are given in Section 4.

2.1. Mixing and convergence rates for difference approximations to stochastic differential equations. Under condition (B_1) and some recurrence conditions, the process X is *ergodic*, i.e., possesses a unique invariant distribution μ_{inv} (see [8, 9]). Moreover, an explicit estimates for the *β-mixing coefficients* and for the rate of convergence of $P_{x,t} \equiv \mathrm{P}(X(t) \in \cdot | X(0) = x)$ to μ_{inv} in total variation norm are also available.

The processes X_n, restricted to $\frac{1}{n}\mathbb{Z}_+$, are a Markov chains. The following natural question takes its origins in a numerical applications: can the mentioned above estimates for the mixing and convergence rate be made uniform over the class $\{X_n, n \geq 1, X\}$? This question is studied in the recent paper [4], see more discussion therein.

In this subsection, we use the truncated local limit theorem (Theorem 1.1) in order to establish the required uniform estimates. Denote, by $\|\cdot\|_{\mathrm{var}}$, the total variation norm. Recall that the β-mixing coefficient for X is defined by

$$\beta_x(t) \equiv \sup_{s \in \mathbb{R}^+} \mathrm{E}\, \|\mathrm{P}(\cdot | \mathcal{F}_0^s, X(0) = x) - \mathrm{P}(\cdot | X(0) = x)\|_{\mathrm{var}, \mathcal{F}_{t+s}^{\infty}}, \quad t \in \mathbb{R}^+,$$

where $\mathcal{F}_a^b \equiv \sigma(X(s), s \in [a, b])$, $\mathrm{P}(\cdot | \mathcal{F}_0^s, X(0) = x)$ denotes the conditional distribution of the process X with $X(0) = x$ w.r.t. $\mathcal{F}_0^s$, and

$$\|\varkappa\|_{\mathrm{var}, \mathcal{G}} \stackrel{\mathrm{df}}{=} \sup_{\substack{B_1 \cap B_2 = \varnothing,\ B_1 \cup B_2 = C(\mathbb{R}^+, \mathbb{R}^m) \\ B_1, B_2 \in \mathcal{G}}} \big[\varkappa(B_1) - \varkappa(B_2)\big].$$

The β-mixing coefficient $\{\beta_x^n(t), t \in \frac{1}{n}\mathbb{Z}_+\}$ for X_n is defined analogously.

Theorem 2.1. *Let conditions* (B_1) *and* (B_3) *hold true. Suppose also that*

(B_4) *there exists* $R_0 > 0$ *and* $r > 0$ *such that*

$$(a(x), x)_{\mathbb{R}^d} \leq -r\|x\|, \quad \|x\| \geq R_0;$$

(B_5) *there exists* $\varkappa > 0$: $\mathrm{E}\, \exp[\varkappa\|\xi\|] < +\infty$.

Then, for every process X, X_n, $n \geq 1$, there exists unique invariant distribution μ_{inv}, μ^n_{inv}. Moreover, there exist $n_0 \in \mathbb{N}$, a function $C\colon \mathbb{R}^d \to \mathbb{R}^+$ and a constant $c > 0$ such that

$$\|P^n_{x,t} - \mu^n_{\mathrm{inv}}\|_{\mathrm{var}} \leq C(x)e^{-ct}, \quad t \in \frac{1}{n}\mathbb{Z}_+, n \geq n_0,$$

$$\|P_{x,t} - \mu_{\mathrm{inv}}\|_{\mathrm{var}} \leq C(x)e^{-ct}, \quad t \in \mathbb{R}^+,$$

$$\beta^n_x(t) \leq C(x)e^{-ct}, \quad t \in \frac{1}{n}\mathbb{Z}_+, \quad n \geq n_0, \quad \beta_x(t) \leq C(x)e^{-ct}, \quad t \in \mathbb{R}^+.$$

***Remarks.* 2.1.** The statement of Theorem 2.1 is analogous to the one of Theorem 1 [4]. The main improvement is that the conditions (D_1)–(D_3) of Theorem 1 [4] are replaced by (seemingly, the mildest possible) condition (B_3). In addition, Theorem 2.1, unlike Theorem 1 [4], admits non-constant diffusion coefficients b.

2.2. The mixing and convergence rates established in Theorem 2.1 are called an *exponential* ones. If the recurrence condition (B_4) is replaced by a weaker ones, then the *subexponential* or *polynomial* rates can be established (see Theorem 1 [4], cases 2 and 3). We do not give an explicit formulation here in order to shorten the exposition.

2.2. Difference approximation for local times of multidimensional diffusions. Consider a *W-measure* μ on $\mathbb{R}^d$, that is, by definition [10] (Chapter 8), a σ-finite measure satisfying the condition

$$\sup_{x\in\mathbb{R}^m} \int\limits_{\|y-x\|\leq 1} w_d(\|y-x\|)\mu(dy) < +\infty \quad \text{with} \quad w_d(r) = \begin{cases} r, & d = 1, \\ \max(-\ln r, 1), & d = 2, \\ r^{2-d}, & d > 2. \end{cases} \tag{2.1}$$

Every such a measure generates a *W-functional* [10] (Chapter 6) of a Wiener process W on $\mathbb{R}^d$,

$$\varphi^{s,t} = \varphi^{s,t}(W) = \int\limits_s^t \frac{d\mu}{d\lambda^d}(W(r))\, dr, \quad 0 \leq s \leq t. \tag{2.2}$$

For singular μ, equality (2.2) is a formal notation, that can be substantiated via an approximative procedure with μ approximated by an absolutely continuous measures [10] (Chapter 8). The functional φ is naturally interpreted as the *local time* for the Wiener process, correspondent to the measure μ.

Next, let the process X be defined by (0.1) and satisfy (0.4), that means that the asymptotic behavior of its transition probability density as $t \to 0+$ is similar to the one of the transition probability density for the Wiener process. Then the estimates, analogous to those given in [10] (Chapter 8) provide that the W-functional of the process X

$$\varphi^{s,t} = \varphi^{s,t}(X) = \int\limits_s^t \frac{d\mu}{d\lambda^d}(X(r))\, dr, \quad 0 \leq s \leq t, \tag{2.3}$$

is well defined. We interpret this functional as the *local time* for the diffusion process X, correspondent to the measure μ.

At last, let the sequence X_n, $n \in \mathbb{N}$ of difference approximations for X be defined by (0.2), (0.3). Consider a sequence of the functionals $\varphi_n(X_n)$ of the processes X_n of the form

$$\varphi_n^{s,t} = \varphi_n^{s,t}(X_n) \stackrel{\text{df}}{=} \frac{1}{n} \sum_{k:\, s \le \frac{k}{n} < t} F_n\left(X_n\left(\frac{k}{n}\right)\right), \quad 0 \le s < t. \tag{2.4}$$

In Theorem 2.2 below, we establish sufficient conditions for the joint distributions of (φ_n, X_n) to converge weakly to the joint distribution of (φ, X). Thus, it is natural to say that the functionals φ_n defined by (2.4) provide the difference approximation for the local time φ defined by (2.3). For the further discussion and references concerning this problem, we refer the reader to the recent paper [5].

We fix $x \in \mathbb{R}^d$ and suppose that $X_n(0) = X(0) = x$. We denote $\mathbb{T} = \{(s,t)\colon 0 \le \le s \le t\}$. In order to shorten exposition, we suppose μ to be finite and to have a compact support. Together with the functionals φ_n that are discontinuous w.r.t. variables s, t, we consider the "random broken line" processes

$$\psi_n^{s,t} = \varphi_n^{\frac{j-1}{n}, \frac{k-1}{n}} - (ns - j + 1)\varphi_n^{\frac{j-1}{n}, \frac{j}{n}} + (nt - k + 1)\varphi_n^{\frac{k-1}{n}, \frac{k}{n}},$$
$$s \in \left[\frac{j-1}{n}, \frac{j}{n}\right), \quad t \in \left[\frac{k-1}{n}, \frac{k}{n}\right).$$

Theorem 2.2. *Let conditions* (B_1), (B_2^6), (B_3) *hold true. Suppose also that*

(B_6) $F_n(x) \ge 0$, $x \in \mathbb{R}^d$, $n \ge 1$ *and* $\frac{1}{n} \sup_{x \in \mathbb{R}^d} F_n(x) \to 0$, $n \to \infty$;

(B_7) *measures* $\mu_n(dx) \equiv F_n(x)\lambda^d(dx)$ *weakly converge to* μ;

(B_8) $\lim_{\delta \downarrow 0} \limsup_{n \to +\infty} \sup_{x \in \mathbb{R}^d} \int_{\|y-x\| \le \delta} w_d(\|y-x\|)\mu_n(dy) \to 0.$

Then $(X_n, \psi_n(X_n)) \Rightarrow (X, \varphi(X))$ *in a sense of weak convergence in* $C(\mathbb{R}^+, \mathbb{R}^d) \times \times C(\mathbb{T}, \mathbb{R}^+)$.

Remarks. **2.3.** The statement of Theorem 2.2 is analogous to the one of Theorem 2.1 [5]. The main improvement is that the condition A3) of Theorem 2.1 [5] is replaced by (seemingly, the mildest possible) condition (B_3).

2.4. Once Theorem 2.2 is proved, one can use the standard truncation procedure in order to replace the moment condition (B_2^6) by the Lyapunov type condition "$\exists\, \delta > > 0\colon \mathrm{E}\|\xi_k\|^{2+\delta} < +\infty$" (e.g. [11], Section 5).

2.5. For examples and a discussion on the relation between conditions (2.1) and (B_8), we refer the reader to [5].

3. Partial Malliavin calculus on a space with a product measure. For every given $n \in \mathbb{N}$ and $t \in [0,1]$, the value $X_n(t)$ is a functional of $\xi_1, \ldots, \xi_n$ and thus can be interpreted as a functional on the space $(\mathbb{R}^d)^n$ with the product measure μ^n. However, under the conditions of Theorem 1.1, μ may contain a singular component and therefore it may fail to have logarithmic derivative. Thus, in general, one can not write the integration-by-parts formula on the probability space $\left((\mathbb{R}^d)^n, (\mathfrak{B}(\mathbb{R}^d))^{\otimes n}, \mu^n\right)$. We overcome this difficulty by using the following trick. Under condition (B_3), the measure μ can be decomposed into a sum

$$\mu = \alpha \pi_U + (1-\alpha)\nu, \tag{3.1}$$

where π_U is the uniform distribution on U. One can write (on an appropriate probability space) the representation for $\{\xi_k\}$ corresponding to (3.1):

$$\xi_k = \varepsilon_k \eta_k + (1-\varepsilon_k)\zeta_k, \tag{3.2}$$

where $\eta_k \sim \pi_U$, $\zeta_k \sim \nu$, and the distribution $\varkappa$ of ε_k is equal to Bernoulli distribution with $\varkappa\{1\} = \alpha$. This representation allows one to consider the family $\xi_1, \dots, \xi_n$ (and, therefore, the process X_n) as a functional on the following probability space:

$$\Omega = (\mathbb{R}^d \times \{0,1\} \times \mathbb{R}^d)^n, \qquad \mathcal{F} = (\mathfrak{B}(\mathbb{R}^d) \otimes 2^{\{0,1\}} \otimes \mathfrak{B}(\mathbb{R}^d))^n, \qquad \mathrm{P} = (\pi_U \times \varkappa \times \nu)^n. \tag{3.3}$$

Now, the measure π_U has a logarithmic derivative w.r.t. a properly chosen vector field, and some kind of an integration-by-parts formula can be written on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$ (see Subsection 3.1 below). The Malliavin-type calculus, associated to this formula, is our main tool in the proof of Theorems 1.1, 1.2. We call this calculus a *partial* one because the stochastic derivative, this calculus is based on, is defined w.r.t. a proper group of variables, while the other variables play the role of interfering terms. In this section, we give the main constructions of the partial Malliavin calculus, associated to the representation (3.2).

3.1. Integration-by-parts formula. Derivative and divergence. Sobolev classes. Denote $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$,

$$\Omega_1 = \Omega_3 = (\mathbb{R}^d)^n, \quad \Omega_2 = \{0,1\}^n.$$

We write a point $\omega \in \Omega$ in the form $\omega = (\eta, \varepsilon, \zeta)$, where

$$\eta = (\eta_1, \dots, \eta_n) \in (\mathbb{R}^d)^n, \qquad \varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \{0,1\}^n,$$

$$\zeta = (\zeta_1, \dots, \zeta_n) \in (\mathbb{R}^d)^n$$

and $\eta_k = (\eta_{k1}, \dots, \eta_{kd})$, $\zeta_k = (\zeta_{k1}, \dots, \zeta_{kd})$. In this notation, the random variables $\eta_k, \varepsilon_k, \zeta_k$ are defined just as the coordinate functionals:

$$\eta_k(\omega) = \eta_k, \quad \varepsilon_k(\omega) = \varepsilon_k, \quad \zeta_k(\omega) = \zeta_k, \quad \omega = (\eta, \varepsilon, \zeta) \in \Omega.$$

Denote by $\mathcal{C}$ the set of bounded measurable functions f on Ω such that, for every $(\varepsilon, \zeta) \in \Omega_2 \times \Omega_3$, the function $f(\cdot, \varepsilon, \zeta)$ belongs to the class $C^\infty(\mathbb{R}^d)$ and

$$\operatorname*{ess\,sup}_{\eta,\varepsilon,\zeta} \left\| [\nabla_\eta]^j f(\eta, \varepsilon, \zeta) \right\| < +\infty, \quad j \in \mathbb{N},$$

where ∇_η denotes the gradient w.r.t. variable η.

For $f \in \mathcal{C}$ and $k = 1, \dots, n$, $r = 1, \dots, d$, denote by $\partial_{kr} f$ the derivative of f w.r.t. the variable η_{kr}. Also, denote by H the space $\mathbb{R}^{d\times n}$ considered as a (finite-dimensional) Hilbert space with the usual Euclid norm, and by $\{e_{kr}, r = 1, \dots, d, k = 1, \dots, n\}$ the canonical basis in it: all coordinates of the vector e_{kr} are equal to zero except the coordinate with the index kr being equal to one. For a given functions $\psi\colon \mathbb{R}^d \to \mathbb{R}$ and $\theta_n\colon (\mathbb{R}^d)^n \to [0,1]$, define the stochastic gradient D by the formula

$$[Df](\eta,\varepsilon,\zeta)=\theta_n(\zeta)\sum_{k,r}\psi(\eta_k)[\partial_{kr}f](\eta,\varepsilon,\zeta)e_{kr},\quad f\in\mathcal{C}. \tag{3.4}$$

This definition can be naturally extended to the functionals taking their values in a finite-dimensional Hilbert space Y (actually, in any separable Hilbert space, but we do not need such a generality in our further construction). Given an orthonormal basis $\{y_l\}$ in Y, denote by $\mathcal{C}^Y$ the set of the functions of the type

$$y=\sum_l f_l y_l,\quad \{f_l\}\subset\mathcal{C},$$

and put for such a function

$$Dy=\sum_l [Df_l]\otimes y_l.$$

It is easy to see that the definitions of the class $\mathcal{C}^Y$ and the derivative D do not depend on the choice of the basis $\{y_l\}$. By the construction, D satisfies the chain rule: for any two spaces Y,Z and for any $f_1,\ldots,f_m\in\mathcal{C}^Y$, $F\in C^\infty(Y^m,Z)$, $m\geq 1$,

$$F(f_1,\ldots,f_m)\in\mathcal{C}^Z\quad\text{and}\quad D\big[F(f_1,\ldots,f_m)\big]=\sum_{j=1}^m[\partial_j F](f_1,\ldots,f_m)Df_j. \tag{3.5}$$

We denote $D^0f=f, D^1f=Df$. The higher derivatives D^j, $j>1$, are defined iteratively: $D^j=\underbrace{D\cdot\ldots\cdot D}_{j}$ (note that the first operator in this product acts on the elements of $\mathcal{C}^Y$ while the last one acts on the elements of $\mathcal{C}^{H^{\otimes(j-1)}\otimes Y}$).

Everywhere below, we suppose that U is an open ball $B(z,r)$ (this obviously does not restrict generality). We define the function ψ in (3.4) by $\psi(x)=r^2-\|x-z\|^2$. Due to this choice, $\psi\in C^\infty(\mathbb{R}^d)$ and $\psi=0$ on ∂U. These properties of ψ imply the following integration by parts formula:

$$\int_U[\partial_r f](x)\psi(x)\,dx=-\int_U[\partial_r\psi](x)f(x)\,dx,\qquad f\in C^1(\mathbb{R}^d),\quad r=1,\ldots,d.$$

As a corollary of this formula, we obtain the following statement.

Proposition 3.1. *For every $h\in H$ and every $f\in\mathcal{C}$, the following integration-by-parts formula holds true:*

$$\mathrm{E}(Df,h)_H=-\mathrm{E}(\rho,h)_H f,\qquad \rho\equiv\theta_n(\zeta)\sum_{k,r}[\partial_r\psi](\eta_k)e_{kr}. \tag{3.6}$$

The formula (3.6) allows one to introduce, in a standard way, the divergence operator corresponding to the derivative D. For $g\in\mathcal{C}^{H\otimes Y}$, put

$$\delta(g)=-\sum_{k,r,l}\big[(\rho,e_{kr})g_{krl}+(Dg_{krl},e_{kr})_H\big]y_l,\quad g_{krl}=(g,e_{kr}\otimes y_l)_{H\otimes Y}. \tag{3.7}$$

By the choice of the function ψ, $\delta(g)\in\mathcal{C}^Y$ as soon as $g\in\mathcal{C}^{H\otimes Y}$. The chain rule (3.5) and the integration-by-parts formula (3.6) imply that the operators D and δ are mutually adjoint in a sense of the following *duality formula*:

$$\mathrm{E}(Df, g)_H = \mathrm{E} f \delta(g), \qquad f \in \mathcal{C}^Y, \quad g \in \mathcal{C}^{Y,H}. \tag{3.8}$$

Since, for every $p \geq 1$ and every Y, $\mathcal{C}^Y$ is dense in $L_p(\Omega, \mathrm{P}, Y)$, the duality formula (3.8) provides that, for any $p \geq 1$ and any Y, the operators D, δ are closable as densely defined unbounded operators

$$D\colon L_p(\Omega, \mathrm{P}, Y) \to L_p(\Omega, \mathrm{P}, H \otimes Y), \qquad \delta\colon L_p(\Omega, \mathrm{P}, H \otimes Y) \to L_p(\Omega, \mathrm{P}, Y).$$

Definition 3.1. *The Sobolev class $W_p^m(Y)$, $p \geq 1$, $m \in \mathbb{Z}_+$, is the completion of the class $\mathcal{C}^Y$ w.r.t. the norm*

$$\|f\|_{p,m} \equiv \left[\sum_{j=0}^{m} \mathrm{E}\|D^j f\|^p_{H^{\otimes j} \otimes Y}\right]^{\frac{1}{p}} < +\infty.$$

Since D is closable in L_p sense, there exists the canonical embedding of $W_p^m(Y)$ into $L_p(\Omega, \mathrm{P}, Y)$.

We denote $W_\infty^\infty(Y) = \bigcap_{m,p} W_p^m(Y)$. If $Y = \mathbb{R}$ then we denote the corresponding Sobolev spaces simply by W_p^m.

3.2. Algebraic relations for derivative and divergence. Moment estimates. Let us introduce some notation. We denote by C a constant such that its value can be calculated explicitly, but this calculation is omitted. The value of C may vary from line to line. If the value of the constant C depends on some parameters, say m, d, then we write $\mathsf{C}(m, d)$. The latter notation indicates that the value of the constant does not depend on other parameters (for instance, n). If, in a sequel, the constant C is referred to, then we endow it with the lower index like $\mathsf{C}_1, \mathsf{C}_2$, etc. We use standard notation $\{\delta_{jk}, j, k \in \mathbb{N}\}$ for the Kronecker's symbol.

For an $H \otimes H \otimes Y$-valued element K, we denote by K^* the element such that

$$(K^*, h \otimes g \otimes y)_{H \otimes H \otimes Y} = (K, g \otimes h \otimes y)_{H \otimes H \otimes Y}, \quad h, g \in H, y \in Y.$$

For an $X \otimes Y$-valued element g_1 and $X \otimes Z$-valued element g_2, we denote by $(g_1, g_2)_X$ the $Y \otimes Z$-valued element

$$(g_1, g_2)_X \equiv \sum_{l_1, l_2, l_3} (g_1, x_{l_1} \otimes y_{l_2})_{X \otimes Y} (g_2, x_{l_1} \otimes z_{l_3})_{X \otimes Z} [y_{l_2} \otimes z_{l_3}],$$

here $\{x_l\}$, $\{y_l\}$, $\{z_l\}$ are orthonormal bases in X, Y and Z, correspondingly. We also denote for an $Y \otimes X$-valued element g_1 and $Z \otimes X$-valued element g_2

$$(g_1, g_2)_X \equiv \sum_{l_1, l_2, l_3} (g_1, y_{l_2} \otimes x_{l_1})_{Y \otimes X} (g_2, z_{l_3} \otimes x_{l_1})_{Z \otimes X} [y_{l_2} \otimes z_{l_3}].$$

Although the same notation $(\cdot, \cdot)_X$ is used for two slightly different objects, it does not cause misunderstanding further.

Consider the $\mathcal{L}(H)$-valued random element (i.e., *random operator* in H) B, defined by the relations

$$\Big(Be_{k_1r_1}, e_{k_2r_2}\Big)_H = -\Big(D(\rho, e_{k_1r_1})_H, e_{k_2r_2}\Big)_H = -\theta_n^2(\zeta)\delta_{k_1k_2}[\partial_{r_1}\partial_{r_2}\psi](\eta_{k_1})\psi(\eta_{k_1}), \tag{3.9}$$

$k_{1,2} = 1, \ldots, n$, $r_{1,2} = 1, \ldots, d$. We define the action of B on $H \otimes Y$-valued element g by

$$Bg = \sum_{k,r,l} (g, e_{k,r} \otimes y_l)_{H\otimes Y} \Big[[Be_{k,r}] \otimes y_l\Big].$$

Using representation (3.7), one can deduce the commutation relations for the operators D, δ, analogous to those for the stochastic derivative and integral for the Wiener process (the proof is straightforward and omitted; for the Wiener case, see [12], § 1.2).

Proposition 3.2. I. *If $f \in \mathcal{C}$, $g \in \mathcal{C}^{H\otimes Y}$, then $f \cdot g \in \mathcal{C}^{H\otimes Y}$ and*

$$\delta(f \cdot g) = f \cdot \delta(g) - (Df, g)_H.$$

II. *If $g \in \mathcal{C}^{H\otimes Y}$, then*

$$D\big[\delta(g)\big] = Bg + \delta\big([Dg]^*\big).$$

III. *If $g_1, g_2 \in \mathcal{C}^H$, then*

$$(D[\delta(g_1)], g_2)_H = (Bg_1, g_2)_H + \delta\Big((Dg_1, g_2)_H\Big) + ([Dg_1]^*, Dg_2)_{H\otimes H}.$$

The main result of this subsection is given by the following lemma.

Lemma 3.1. *Let $m, l \in \mathbb{N}$, $g \in W_{2m}^{2m+l-1}(H)$. Then there exists $\delta(g) \in W_{2m}^l$ and*

$$\|\delta(g)\|_{2m,l} \le \mathsf{C}(m, l, d, \psi)\|g\|_{2m,2m+l-1}. \tag{3.10}$$

***Remarks.* 3.1.** On the Wiener space, the typical way to prove estimates of the type (3.10) is to use Meyer's inequalities for the generator $L = \delta D$ of the Ornstein–Uhlenbeck semigroup (see, for instance, [12], § 2.4). Moreover, on the Wiener space, (3.10) can be made more precise: the similar inequality holds with $2m + l - 1$ replaced by $l + 1$. In our settings, it is not clear whether the operator $\delta \cdot D$ provides the analogues of Meyer's inequalities, since it does not have the specific structural properties of the Ornstein–Uhlenbeck generator (such as Mehler's formula, hypercontractivity of the associated semigroup, etc.). Thus we prove (3.10) straightforwardly by using an iterative integration-by-parts procedure.

3.2. Throughout the exposition, the function ψ is fixed together with the set $U = B(x, z)$. However, when the constant C depends on the values of ψ or its derivatives, we indicate it explicitly in the notation for C.

In order to prove Lemma 3.1, we need some auxiliary statements and notation. For $g \in \mathcal{C}^Y$ and $m \in \mathbb{Z}_+$, we define the random variable $|g|_m$ by

$$|g|_m \equiv \left(\sum_{j=0}^m \|D^j g\|^2_{H^{\otimes j}\otimes Y}\right)^{\frac{1}{2}}.$$

Lemma 3.2. *If $g \in \mathcal{C}^{H\otimes Y}$ then $Bg \in \mathcal{C}^{H\otimes Y}$ and, for every $m \in \mathbb{Z}_+$,*

$$|Bg|_m \le \mathsf{C}(m, d, \psi)|g|_m.$$

Proof. Write Bg in the coordinate form:

$$Bg = \sum_{k,r^1,r^2,l} (g, e_{kr^1} \otimes y_l)_{H\otimes Y} b_{k,r^1,r^2}[e_{kr^2} \otimes y_l],$$

where $b_{k,r^1,r^2} = (Be_{kr^1}, e_{kr^2})_H$ (recall that $(Be_{k^1r^1}, e_{k^2r^2})_H = 0$ as soon as $k^1 \neq k^2$). Write the Leibnitz formula for the higher derivatives:

$$D^m(Bg) =$$

$$= \sum_{k,r^1,r^2,l} \sum_{\Theta\in 2^{\{1,\ldots,m\}}} \sum_{k_1,r_1,\ldots,k_m,r_m} \left(D^{\#\Theta}(g, e_{kr^1} \otimes y_l)_{H\otimes Y}, \bigotimes_{i\in\Theta} e_{k_ir_i} \right)_{H^{\otimes\#\Theta}} \times$$

$$\times \left(D^{m-\#\Theta} b_{kr^1r^2}, \bigotimes_{i\notin\Theta} e_{k_ir_i} \right)_{H^{\otimes(m-\#\Theta)}} \left[\bigotimes_{i=1}^m e_{k_ir_i} \otimes e_{kr^2} \otimes y_l \right], \tag{3.11}$$

where $\#\Theta_m$ denotes the number of elements in the set Θ. We write

$$S_\Theta = \sum_{k,r^1,r^2,l,k_1,r_1,\ldots,k_m,r_m} \left(D^{\#\Theta}(g, e_{kr^1} \otimes y_l)_{H\otimes Y}, \bigotimes_{i\in\Theta} e_{k_ir_i} \right)_{H^{\otimes\#\Theta}} \times$$

$$\times \left(D^{m-\#\Theta} b_{kr^1r^2}, \bigotimes_{i\notin\Theta} e_{k_ir_i} \right)_{H^{\otimes(m-\#\Theta)}} \left[\bigotimes_{i=1}^m e_{k_ir_i} \otimes e_{kr^2} \otimes y_l \right]$$

and estimate $\|S_\Theta\|_{H^{\otimes(m+1)}\otimes Y}$. The function ψ belongs to C^∞ and is bounded together with all its derivatives on U. Thus, one can deduce from the representation (3.9) and formula (3.4) that

$$\left\| D^M b_{kr^1r^2} \right\|_{H^{\otimes M}} \leq \mathsf{C}(M, d, \psi), \quad M \in \mathbb{N}.$$

In addition, due to (3.9),

$$\left(D^{m-\#\Theta} b_{kr^1r^2}, \bigotimes_{i\notin\Theta} e_{k_ir_i} \right)_{H^{\otimes(m-\#\Theta)}} = 0$$

as soon as $k_i \neq k$ for some $i \notin \Theta$. Using these facts, we deduce that

$$\sum_{k_i\in\{1,\ldots,n\},r_i\in\{1,\ldots,d\},i\notin\Theta} \left(D^{m-\#\Theta} b_{kr^1r^2}, \bigotimes_{i\notin\Theta} e_{k_ir_i} \right)^2_{H^{\otimes(m-\#\Theta)}} \leq \mathsf{C}(m, d, \psi).$$

Thus

$$\|S_\Theta\|^2_{H^{\otimes(m+1)}\otimes Y} \leq \mathsf{C}(m, d, \psi) \times$$

$$\times \sum_{k,r^1,r^2,l,k_i\in\{1,\ldots,n\},r_i\in\{1,\ldots,d\},i\in\Theta} \left(D^{\#\Theta}(g, e_{kr^1} \otimes y_l)_{H\otimes Y}, \bigotimes_{i\in\Theta} e_{k_ir_i} \right)^2_{H^{\otimes\#\Theta}} \leq$$

$$\leq \mathsf{C}(m,d,\psi)|g|^2_{\#\Theta} \leq \mathsf{C}(m,d,\psi)|g|^2_m.$$

Taking the sum over $\Theta \in 2^{\{1,\ldots,m\}}$ and using the Cauchy inequality, we obtain the required statement.

The lemma is proved.

Proposition 3.3. I. *Let* $g_1 \in \mathcal{C}^{X\otimes Y}$, $g_2 \in \mathcal{C}^{X\otimes Z}$, *then* $(g_1,g_2)_X \in \mathcal{C}^{Y\otimes Z}$ *and*

$$\big|(g_1,g_2)_X\big|_m \leq \mathsf{C}(m)|g_1|_m|g_2|_m, \quad m \geq 0.$$

II. *Let* $g \in \mathcal{C}^Y$ *and* $A \in \mathcal{L}(Y,Z)$, *then* $Ag \in \mathcal{C}^Z$ *and* $|Ag|_m \leq \|A\||g|_m$, $m \geq 0$.

The second statement is a straightforward corollary of the chain rule (3.5). The first one can be proved using the Leibnitz formula; the proof is totally analogous to the one of Lemma 3.2, and thus we omit the detailed exposition.

***Remark* 3.3.** Taking $X=\mathbb{R}$, we obtain that, for $g_1 \in \mathcal{C}^Y$, $g_2 \in \mathcal{C}^Z$, $g_1\otimes g_2 \in \mathcal{C}^{Y\otimes Z}$ with $|g_1\otimes g_2|_m \leq \mathsf{C}(m)|g_1|_m|g_2|_m$, $m \geq 0$.

Using iteratively statement II of Proposition 3.2, we obtain that, for $g \in \mathcal{C}^H$ and $m \geq 1$, the derivative $D^m[\delta(g)]$ can be expressed in the form

$$D^m\big[\delta(g)\big] = F_m(g) + \delta\big(G_m(g)\big),$$

where $F_m(g) \in \mathcal{C}^{H^{\otimes m}}$, $G_m(g) \in \mathcal{C}^{H^{\otimes(m+1)}}$ are defined via the iterative procedure

$$F_0(g)=0, \qquad G_0(g)=g, \qquad G_{i+1}(g)=\big[DG_i(g)\big]^*,$$

$$F_{i+1}(g) = DF_i(g) + BG_i(g), \quad i \geq 0.$$

The mapping $K \mapsto K^*$ is an isometry in $H\otimes H\otimes Y$, thus statement II of Proposition 3.3 provides that

$$\big|G_m(g)\big|_j \leq |g|_{m+j}, \quad m,\ j \geq 0. \tag{3.12}$$

Using Lemma 3.2, we deduce that

$$\big|F_m(g)\big|_j \leq \mathsf{C}(m,j,d,\psi)|g|_{m+j}, \quad m,\ j\geq 0. \tag{3.13}$$

Thus, in order to prove inequality (3.10) for $g \in \mathcal{C}^H$, it is sufficient to prove that

$$\mathrm{E}\,\big\|\delta(g)\big\|^{2m}_Y \leq \mathsf{C}(m,d,\psi)\mathrm{E}|g|^{2m}_{2m-1} \tag{3.14}$$

for any $m\geq 1$, $g \in \mathcal{C}^{H\otimes Y}$ and arbitrary Hilbert space Y. In order to prove estimate (3.14) we embed it into a larger family of estimates. Consider the following objects.

1. Numbers $k_0,\ldots,k_v \in \mathbb{Z}_+$, $v \leq 2m$, such that $k_0+\ldots+k_v = 2m$. Denote $I_j = [k_0+\ldots+k_{j-1}+1, k_0+\ldots+k_j]\cap\mathbb{N}$, $j=1,\ldots,v$, $I_0=[1,k_0]\cap\mathbb{N}$ (if $k_0=0$ then $I_0=\varnothing$).

2. Function $\sigma\colon \{1,\ldots,2m\} \to \{1,\ldots,m\}$ such that $\#\sigma^{-1}(\{i\})=2$ and $\#\Big[I_j\cap$ $\cap\,\sigma^{-1}(\{i\})\Big] \leq 1$ for every $i=1,\ldots,m$ and $j=0,\ldots,v$.

Lemma 3.3. *Let $f_0 \in \mathcal{C}^{Y^{\otimes k_0}}$, $g_j \in \mathcal{C}^{H\otimes Y^{\otimes k_j}}$. Then*

$$\mathrm{E}\sum_{l_1,\dots,l_m}\left(g_0,\bigotimes_{i\in I_0} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_0}}\prod_{j=1}^{v}\delta\left(\left(g_j,\bigotimes_{i\in I_j} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_j}}\right)\le$$

$$\le \mathsf{C}(v,d,\psi)\mathrm{E}\left[|g_0|_v\prod_{j=1}^{v}|g_j|_{v-1}\right]. \tag{3.15}$$

***Remark* 3.4.** The left-hand side of (3.14) can be rewritten to the form

$$\mathrm{E}\sum_{l_1,\dots,l_m}\delta\big((g,y_{l_1})_Y\big)\delta\big((g,y_{l_1})_Y\big)\delta\big((g,y_{l_2})_Y\big)\delta\big((g,y_{l_2})_Y\big)\dots$$

$$\dots\delta\big((g,y_{l_m})_Y\big)\delta\big((g,y_{l_m})_Y\big). \tag{3.16}$$

If $v=2m$, $k_0=0$, $k_1=\ldots=k_{2m}=1$, $g_0=1, g_1=\ldots=g_{2m}=g\in\mathcal{C}^{H\otimes Y}$, then the left-hand side of (3.15) coincides with the expression written in (3.16). Thus Lemma 3.3 implies estimate (3.14).

Proof of the lemma. We use induction by v. For $v=0$, conditions imposed on σ can be satisfied if $m=0$, only (i.e., if g_0 is a function valued in $\mathbb{R}$). Thus, for $v=0$, (3.15) is trivial since $g_0\le|g_0|\equiv|g_0|_0$. For $v=1$, conditions imposed on σ imply that $I_0=\{1,\dots,m\}$, $I_1=\{m+1,\dots,2m\}$ and the function σ, restricted to either I_0 or I_1, is bijective. Thus the left-hand side of (3.15) can be rewritten to the form

$$\mathrm{E}\sum_{l_1,\dots,l_m}\left(g_0,\bigotimes_{i=1}^{m} y_{l_i}\right)_{Y^{\otimes m}}\delta\left(\left(g_j,\bigotimes_{i=1}^{m} y_{l_{\pi(i)}}\right)_{Y^{\otimes m}}\right),$$

where π is some permutation of $\{1,\dots,m\}$. Using duality formula (3.8), we rewrite this as

$$\mathrm{E}\big(Dg_0,A_\pi g\big)_{H\otimes Y^{\otimes m}},$$

where the operator $A_\pi\in\mathcal{L}(H\otimes Y^{\otimes m})$ is defined by

$$A[h\otimes y_{l_1}\otimes\ldots\otimes y_{l_m}]=h\otimes y_{l_{\pi(1)}}\otimes\ldots\otimes y_{l_{\pi(m)}}. \tag{3.17}$$

One can easily see that A_π is an isometry operator, and thus Proposition 3.3 provides that (3.15) holds true for $v=1$ with $\mathsf{C}(1,d,\psi)=1$.

Suppose that, for some $V\ge 2$, (3.15) holds true for all $v\le V-1$. Let us prove that (3.15) holds for $v=V$, also. For every $l_1,\dots,l_m$, take

$$g=\left(g_1,\bigotimes_{i\in I_1} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_1}},$$

$$f=\left(g_0,\bigotimes_{i\in I_0} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_0}}\prod_{j=2}^{V}\delta\left(\left(g_j,\bigotimes_{i\in I_j} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_j}}\right)$$

and apply duality formula (3.8). Then the left-hand side of (3.15) transforms to the form

$$\mathrm{E}\sum_{l_1,\ldots,l_m}\left(D\left(g_0,\bigotimes_{i\in I_0}y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_0}},\ \left(g_1,\bigotimes_{i\in I_1}y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_1}}\right)_H\times$$

$$\times\prod_{j=2}^{V}\delta\left(\left(g_j,\bigotimes_{i\in I_j}y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_j}}\right)+$$

$$+\sum_{r=2}^{V}\mathrm{E}\sum_{l_1,\ldots,l_m}\left(g_0,\bigotimes_{i\in I_0}y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_0}}\times$$

$$\times\left(D\left(\delta\left(\left(g_r,\bigotimes_{i\in I_r}y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_r}}\right)\right),\ \left(g_1,\bigotimes_{i\in I_1}y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_1}}\right)_H\times$$

$$\times\prod_{j\in\{2,\ldots,V\}\setminus\{r\}}\delta\left(\left(g_j,\bigotimes_{i\in I_j}y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_j}}\right). \tag{3.18}$$

Let us estimate every summand in (3.18) separately. The idea is that every such summand can be written as

$$\mathrm{E}\sum_{l_1,\ldots,l_{\tilde m}}\left(\tilde g_0,\bigotimes_{i\in\tilde I_0}y_{l_{\tilde\sigma(i)}}\right)_{Y^{\otimes\tilde k_0}}\prod_{j=1}^{v}\delta\left(\left(\tilde g_j,\bigotimes_{i\in\tilde I_j}y_{l_{\tilde\sigma(i)}}\right)_{Y^{\otimes\tilde k_j}}\right) \tag{3.19}$$

with $v=V-1$ or $v=V-2$ and some new $\tilde m$, $\tilde k_0,\ldots,k_{\tilde m}v$, $\tilde g_0$, $\tilde g_v$, $\tilde\sigma$, and thus the inductive supposition can be applied.

Consider the first summand. Denote by J the set of indices $r\in\{1,\ldots,m\}$ such that $\sigma^{-1}\big(\{r\}\big)\subset I_0\cup I_1$. In order to shorten notation, we suppose that $J=\{1,\ldots,\#J\}$ (this does not restrict generality since one can make an appropriate permutation of the set $\{1,\ldots,m\}$ in order to provide such a property). Take permutations $\pi_0\colon I_0\to I_0$ and $\pi_1\colon I_1\to I_1$ such that

$$[\sigma\circ\pi_0](i)=i,\quad i\in\{1,\ldots,\#J\},$$

$$[\sigma\circ\pi_1](i)=i-k_0,\quad i\in\{k_0+1,\ldots,k_0+\#J\}.$$

Then the first summand in (3.18) can be rewritten to the form

$$\mathrm{E}\sum_{l_{\#J+1},\ldots,l_m}\left[\sum_{l_1,\ldots,l_{\#J}}\left(D\left(A_{\pi_0}g_0,\bigotimes_{i=1}^{k_0}y_{l_{\sigma(\pi_0(i))}}\right)_{Y^{\otimes k_0}},\right.\right.$$

$$\left.\left.\left(A_{\pi_1}g_1,\bigotimes_{i=k_0+1}^{k_0+k_1}y_{l_{\sigma(\pi_1(i))}}\right)_{Y^{\otimes k_1}}\right)_H\right]\prod_{j=2}^{V}\delta\left(\left(g_j,\bigotimes_{i\in I_j}y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_j}}\right),$$

where the operators A_{π_0} and A_{π_1} are defined analogously to (3.17). Denote

$$\tilde{g}_0 = (DA_{\pi_0}g_0, A_{\pi_1}g_1)_{H\otimes Y^{\otimes(\#J)}},$$

then

$$\left[\sum_{l_1,\ldots,l_{\#J}} \left(D\left(A_{\pi_0}g_0, \bigotimes_{i=1}^{k_0} y_{l_{\sigma(\pi_0(i))}}\right)_{Y^{\otimes k_0}}, \left(A_{\pi_1}g_1, \bigotimes_{i=k_0+1}^{k_0+k_1} y_{l_{\sigma(\pi_1(i))}}\right)_{Y^{\otimes k_1}}\right)_H\right] =$$

$$= \left(\tilde{g}_0, \left[\bigotimes_{i=\#J+1}^{k_0} y_{l_{\sigma(\pi_0(i))}}\right] \otimes \left[\bigotimes_{i=k_0+\#J+1}^{k_1} y_{l_{\sigma(\pi_1(i))}}\right]\right)_{Y^{k_0+k_1-2\#J}}.$$

Put $\tilde{m} = m - \#J$, $\tilde{k}_0 = k_0 + k_1 - 2\#J$, $\tilde{k}_1 = k_2, \ldots, \tilde{k}_{V-1} = k_V$ and let $\{\tilde{I}_0, \tilde{I}_1, \ldots$ $\ldots, \tilde{I}_{V-1}\}$ be the partition of $\{1, \ldots, 2\tilde{m}\}$ corresponding to the family $\{\tilde{k}_0, \ldots, \tilde{k}_{V-1}\}$. Put $\tilde{g}_1 = g_2, \ldots \tilde{g}_{V-1} = g_V$ ($\tilde{g}_0$ is already defined). At last, define function $\tilde{\sigma}$ by

$$\tilde{\sigma}(i) = \begin{cases} \sigma(\pi_0(i + \#J)), & i = 1, \ldots, k_0 - \#J, \\ \sigma(\pi_1(i + 2\#J)), & i = k_0 - \#J + 1, \ldots, k_0 + k_1 - 2\#J, \\ \sigma(i + 2\#J), & i = k_0 + k_1, \ldots, 2\tilde{m}. \end{cases}$$

Under such a notation, the first summand in (3.18) has exactly the form (3.19) with $v = V - 1$, and the inductive supposition provides that this summand is dominated by the term

$$\mathsf{C}(V-1, d, \psi)\mathrm{E}\left[|\tilde{g}_0|_{V-1} \prod_{j=1}^{V-1} |\tilde{g}_j|_{V-2}\right].$$

Since A_{π_0}, A_{π_1} are isometric operators, we can apply Proposition 3.3 and obtain that

$$|\tilde{g}_0|_{V-1} = \left|(DA_{\pi_0}g_0, A_{\pi_1}g_1)_{H\otimes Y^{\otimes(\#J)}}\right|_{V-1} \le$$

$$\le \mathsf{C}(V-1)|DA_{\pi_0}g_0|_{V-1}|A_{\pi_1}g_1|_{V-1} \le \mathsf{C}(V-1)|g_0|_V|g_1|_{V-1}.$$

For every $j = 1, \ldots, V-1$, $|\tilde{g}_j|_{V-2} = |g_{j+1}|_{V-2} \le |g_{j+1}|_{V-1}$. Thus, under the inductive supposition, the first summand in (3.18) is dominated by the expression given in the right-hand side of (3.15).

All the $V-1$ summands in the second sum in (3.18) have the same form and can be estimated similarly; let us make such an estimation for $r = 2$. Using Proposition 3.2, we rewrite this summand to the form

$$\mathrm{E}\sum_{l_1,\ldots,l_m} \left(g_0, \bigotimes_{i\in I_0} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_0}} \times$$

$$\times \left(B\left(\left(g_2, \bigotimes_{i\in I_r} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_r}}\right), \left(g_1, \bigotimes_{i\in I_1} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_1}}\right)_H \times$$

$$\times \prod_{j=3}^{V} \delta\left(\left(g_j, \bigotimes_{i\in I_j} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_j}}\right) +$$

$$+ \mathrm{E} \sum_{l_1,\ldots,l_m} \left(g_0, \bigotimes_{i\in I_0} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_0}} \delta\left(\left(D\left(\left(g_2, \bigotimes_{i\in I_r} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_r}}\right),\right.\right.$$

$$\left.\left.\left(g_1, \bigotimes_{i\in I_1} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_1}}\right)_H\right) \prod_{j=3}^{V} \delta\left(\left(g_j, \bigotimes_{i\in I_j} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_j}}\right) +$$

$$+ \mathrm{E} \sum_{l_1,\ldots,l_m} \left(g_0, \bigotimes_{i\in I_0} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_0}} \left(\left[D\left(\left(g_2, \bigotimes_{i\in I_r} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_r}}\right)\right]^*,\right.$$

$$\left.\left(Dg_1, \bigotimes_{i\in I_1} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_1}}\right)_{H\otimes H} \prod_{j=3}^{V} \delta\left(\left(g_j, \bigotimes_{i\in I_j} y_{l_{\sigma(i)}}\right)_{Y^{\otimes k_j}}\right). \tag{3.20}$$

Let us show that, after an appropriate rearrangement of the indices i, every summand in (3.20) can be rewritten to the form (3.19). Such a rearrangement can be organized in the way, totally analogous to the one used before while the first summand in (3.18) was estimated. Therefore, in order to shorten exposition, we do not write here an explicit form for the permutations of the indices, used in such a rearrangement. The first summand in (3.20) has the form (3.19) with $v = V-2$, $\tilde{g}_j = g_{j+2}$, $j = 1,\ldots,V-2$,

$$\tilde{g}_0 = \Big(A_{\pi_1}[Bg_2], A_{\pi_2}[g_1 \otimes g_0]\Big)_{H\otimes Y^{\otimes \# J^1}},$$

where J^1 is the set of such $i \in \{1,\ldots,m\}$ that $\sigma^{-1}(\{i\}) \subset I_0 \cup I_1 \cup I_2$ (we do not write here an explicit expressions neither for the permutations π_1, π_2 nor for $\tilde{k}_0,\ldots,\tilde{k}_{V-2}$, $\tilde{\sigma}$). Under inductive supposition, this summand is estimated by

$$\mathsf{C}(V-2,d,\psi)\,\mathrm{E}|\tilde{g}_0|_{V-2}\prod_{j=3}^{V}|g_j|_{V-3} \le$$

$$\le \mathsf{C}(V,d,\psi)\,\mathrm{E}|Bg_2|_{V-2}|g_1|_{V-2}|g_0|_{V-2}\prod_{j=3}^{V}|g_j|_{V-3} \le$$

$$\le \mathsf{C}(V,d,\psi)\mathrm{E}|g_2|_{V-2}|g_1|_{V-2}|g_0|_{V-2}\prod_{j=3}^{V}|g_j|_{V-3} \le \mathsf{C}(V,d,\psi)\mathrm{E}|g_0|_V\prod_{j=1}^{V}|g_j|_{V-1},$$

here we used Proposition 3.3 and Lemma 3.2. The second summand in (3.20) has the form (3.19) with $v = V-1$, $\tilde{g}_j = g_{j+1}$, $j = 2,\ldots,V-1$, $\tilde{g}_0 = g_0$,

$$\tilde{g}_1 = \Big(D[A_{\pi_1} g_2], A_{\pi_2} g_1\Big)_{H\otimes Y^{\otimes \# J^2}},$$

where J^2 is the set of such $i \in \{1, \ldots, m\}$ that $\sigma^{-1}(\{i\}) \subset I_1 \cup I_2$. This summand again is estimated by

$$\mathsf{C}(V-1, d, \psi)\, \mathrm{E}|g_0|_{V-1} |\tilde{g}_1|_{V-2} \prod_{j=3}^{V} |g_j|_{V-2} \le$$

$$\le \mathsf{C}(V, d, \psi)\, \mathrm{E}|g_0|_{V-1} ||Dg_2|_{V-2} |g_1|_{V-2} |g_0|_{V-2} \prod_{j=3}^{V} |g_j|_{V-3} \le$$

$$\le \mathsf{C}(V, d, \psi)\, \mathrm{E}|g_0|_V \prod_{j=1}^{V} |g_j|_{V-1}.$$

At last, the third summand in (3.20) has the form (3.19) with $v = V-2$, $\tilde{g}_j = g_{j+2}$, $j = 1, \ldots, V-2$,

$$\tilde{g}_0 = \Big(A_{\pi_1} g_0, ([DA_{\pi_2} g_2]^*,\ DA\pi_3 g_2)_{H\otimes H}\Big)_{Y^{\otimes \# J^1}},$$

and again is estimated by

$$\mathsf{C}(V-2, d, \psi)\, \mathrm{E}|\tilde{g}_0|_{V-2} \prod_{j=3}^{V} |g_j|_{V-3} \le$$

$$\le \mathsf{C}(V, d, \psi)\, \mathrm{E}|g_0|_{V-2} |Dg_2|_{V-2} |Dg_1|_{V-2} \prod_{j=3}^{V} |g_j|_{V-3} \le$$

$$\le \mathsf{C}(V, d, \psi)\, \mathrm{E}|g_0|_V \prod_{j=1}^{V} |g_j|_{V-1}.$$

The estimates given above show that (3.15) holds for $v = V$ as soon as it holds for $v = V-2$ and $v = V-1$. We have already proved that (3.15) holds for $v = 0, 1$. Thus, (3.15) holds for every v.

The lemma is proved.

***Proof of Lemma* 3.1.** We have already proved (3.10) to hold for every $g \in \mathcal{C}^H$. Now, let $g \in W_{2m}^{2m+l-1}(H)$. Consider $\{g_n\} \subset \mathcal{C}^H$ such that $g_n \to g$ in $W_{2m}^{2m+l-1}(H)$ (recall that $\mathcal{C}^H$ is dense in any $W_p^k(H)$ by definition). By (3.10), for any $k = 0, \ldots, l$,

$$\big\|D^k \delta(g_n) - D^k \delta(g_N)\big\|_{L_{2m}(\Omega, \mathrm{P}, H^{\otimes k})} \le$$

$$\le \mathsf{C}(m, l, d, \psi) \|g_n - g_N\|_{2m, 2m+l-1} \to 0, \quad n,\ N \to +\infty.$$

Thus there exist $F_k \in L_{2m}(\Omega, \mathrm{P}, H^{\otimes k})$, $k = 0, \ldots, l$, such that

$$\left\|D^k\delta(g_n)-F_k\right\|_{L_{2m}(\Omega,\mathrm{P},H^{\otimes k})}\to 0,\qquad n\to+\infty,\quad k=0,\ldots,l.$$

Since operator δ is closed, $F_0=\delta(g)$. Using that operator D is closed, one can verify inductively that $F_k=DF_{k-1}$, $k=1,\ldots,l$. This means that $\delta(g)\in W^l_{2m}$ with $D^k\delta(g)=F_k$, $k=0,\ldots,l$. At last, using (3.10) we get

$$\|\delta(g)\|^{2m}_{2m,l}=\mathrm{E}\sum_{k=0}^{l}\|F_k\|^{2m}_{H^{\otimes k}}\le\limsup_n\mathrm{E}\sum_{k=0}^{l}\left\|D^k\delta(g_n)\right\|^{2m}_{H^{\otimes k}}=$$

$$=\limsup_n\|\delta(g_n)\|^{2m}_{2m,l}\le\mathsf{C}(m,l,d,\psi)\limsup_n\|g_n\|^{2m}_{2m,3m-1}=$$

$$=\mathsf{C}(m,d,\psi)\|g\|^{2m}_{2m,2m+l-1}.$$

The lemma is proved.

3.3. Malliavin's representation for the densities of the truncated distributions of smooth functionals. The typical result in the Malliavin calculus on the Wiener space is that, when the components $f_1,\ldots,f_d$ of a random vector $f=(f_1,\ldots,f_d)$ are smooth enough and the Malliavin matrix $\sigma^f=\left\{(Df_i,Df_j)_H\right\}^d_{i,j=1}$ is non-degenerate in a sense that

$$[\det\sigma^f]^{-1}\in\bigcap_{p\ge1}L_p(\Omega,\mathcal{F},\mathrm{P}),\tag{3.21}$$

the distribution of f has a smooth density (see, for instance, [12], § 3.2). Such kind of a result is useless in the framework, introduced in Subsection 3.1, since there does not exist any functional f satisfying (3.21): if $\varepsilon_1=\ldots=\varepsilon_n=0$ then $Df=0$ for every $f\in\mathcal{C}$. In order to overcome this difficulty we use the following truncation procedure: we consider, instead of P, a new (non-probability) measure $\mathrm{P}_\Xi(\cdot)=\mathrm{P}(\cdot\cap\Xi)$ with some set $\Xi\in\sigma(\varepsilon,\zeta)$. If this set is chosen in such a way that (3.21) holds true with P replaced by P_Ξ then the Malliavin's calculus can be applied in order to investigate the law of f w.r.t. P_Ξ. In this subsection, we give the Malliavin's representation for the density of this law. All principal steps in our consideration are analogous to those in the standard Malliavin calculus on the Wiener space (see, for instance, [12], Chapter 3). Therefore, we sketch the proofs only.

Let $f_1,\ldots,f_d\in\mathcal{C}$ be fixed, consider the Malliavin matrix $\sigma^f=(\sigma^f_{ij})^d_{i,j=1}$,

$$\sigma^f_{ij}=(Df_i,Df_j)_H=\sum_{k,r}\psi(\eta_k)\big[\partial_{kr}f_i(\eta,\varepsilon,\zeta)\big]\big[\partial_{kr}f_j(\eta,\varepsilon,\zeta)\big].$$

Consider a set $\Xi\in\sigma(\varepsilon,\zeta)$ such that $\Xi\subset\{\det\sigma^f>0\}$ and

$$\mathrm{E}\mathbf{1}_\Xi[\det\sigma^f]^{-p}<\infty,\quad p\ge1.\tag{3.22}$$

Then $\mathbf{1}_\Xi\in\mathcal{C}$ and $D\mathbf{1}_\Xi=0$. Put

$$\varrho^{f,\Xi}(\omega)=\begin{cases}[\sigma^f(\omega)]^{-1}, & \omega\in\Xi,\\ 0, & \omega\notin\Xi.\end{cases}$$

Proposition 3.4. $\varrho^{f,\Xi} \in W^\infty_\infty(\mathbb{R}^{d\times d})$ *and*

$$(D\varrho^{f,\Xi}, h)_H = -\varrho^{f,\Xi}(D\sigma^f, h)_H \varrho^{f,\Xi}, \quad h \in H. \tag{3.23}$$

Sketch of the proof. It is enough to prove that $\varrho^{f,\Xi} \in \bigcap_{p\geq 1} W^1_p(\mathbb{R}^{d\times d})$ and (3.23) holds true. Suppose that $\sigma^f \geq cI_{\mathbb{R}^d}$ with some $c > 0$. Then one can easily see that $\varrho^{f,\Xi} \in \mathcal{C}$ and (3.23) follows from the well known formula for the derivative of the inverse matrix,

$$\frac{d}{dt}\big[A(t)\big]^{-1} = -\big[A(t)\big]^{-1}\left[\frac{d}{dt}A(t)\right]\big[A(t)\big]^{-1}.$$

In the general case, consider the matrix-valued functions $\sigma^{f,c} = \sigma^f + cI_{\mathbb{R}^d}$ and $\varrho^{f,\Xi,c} = \mathbb{I}_\Xi \cdot [\sigma^{f,c}]^{-1}$, $c > 0$. Condition (3.22) provides that $\varrho^{f,\Xi,c} \to \varrho^{f,\Xi}, c \to 0+$ in any L_p. It is already proved that (3.23) holds true for the functionals indexed by c. Thus, passing to the limit as $c \to 0+$, we obtain the required statement.

Denote $\vartheta^{f,\Xi}_i = \sum_{k=1}^d \varrho^{f,\Xi}_{ki} Df_k$, $i = 1,\ldots,d$. Also denote, by E_Ξ, the expectation w.r.t. P_Ξ.

Proposition 3.5. *For every* $i = 1,\ldots,d$, $\upsilon \in W^\infty_\infty$ *and every* $F \in C^\infty_b(\mathbb{R}^d)$,

$$\mathrm{E}_\Xi[\partial_i F](f_1,\ldots,f_d)\upsilon = \mathrm{E}_\Xi F(f_1,\ldots,f_d)\delta\big(\upsilon\vartheta^{f,\Xi}_i\big). \tag{3.24}$$

Sketch of the proof. It follows from Propositions 3.3, 3.4 and Lemma 3.1 that $\upsilon\cdot\vartheta^f_i \in \mathrm{Dom}(\delta)$. Since $\Xi \in \sigma(\varepsilon,\zeta)$, the function $\mathbb{I}_\Xi$ belongs to $\mathcal{C}$ and has its stochastic derivative equal to 0. Proposition 3.2 provides that $\delta(\mathbb{I}_\Xi g) = \mathbb{I}_\Xi\delta(g)$, $g \in \mathrm{Dom}(\delta)$. Therefore

$$\mathrm{E}[\partial_i F](f_1,\ldots,f_d)\mathbb{I}_\Xi\upsilon = \sum_{j=1}^d \mathrm{E}[\partial_j F](f_1,\ldots,f_d)\mathbb{I}_\Xi\upsilon[\sigma^f\varrho^{f,\Xi}]_{ij} =$$

$$= \sum_{j=1}^d\sum_{k=1}^d \mathrm{E}[\partial_j F](f_1,\ldots,f_d)\mathbb{I}_\Xi\upsilon\varrho^{f,\Xi}_{ki}\sigma^f_{jk} =$$

$$= \mathrm{E}\sum_{k=1}^d \mathbb{I}_\Xi\upsilon\varrho^{f,\Xi}_{ki}\left(\sum_{j=1}^d[\partial_j F](f_1,\ldots,f_d)Df_j, Df_k\right)_H =$$

$$= \mathrm{E}\left(D[F(f_1,\ldots,f_d)], \mathbb{I}_\Xi\upsilon\vartheta^{f,\Xi}_i\right)_H =$$

$$= \mathrm{E}F(f_1,\ldots,f_d)\delta\left(\mathbb{I}_\Xi\upsilon\vartheta^f_i\right) = \mathrm{E}F(f_1,\ldots,f_d)\mathbb{I}_\Xi\delta\left(\upsilon\vartheta^f_i\right),$$

that provides (3.24).

Put

$$\upsilon^{f,\Xi}_1 = 1, \qquad \upsilon^{f,\Xi}_{l+1} = \delta\left(\upsilon^{f,\Xi}_l\vartheta^{f,\Xi}_l\right), \quad l = 1,\ldots,d,$$

$$\Upsilon^{f,\Xi} = \upsilon^{f,\Xi}_{d+1}, \qquad \Upsilon^{f,\Xi}_i = \delta\left(\Upsilon^{f,\Xi}\vartheta^{f,\Xi}_i\right), \quad i = 1,\ldots,d.$$

Write P_{Ξ}^{f} for the distribution of f w.r.t. P_{Ξ}. For $\alpha_1, \ldots, \alpha_d \in \{0,1\}$, denote

$$\mathbf{1}_{\alpha_1\ldots\alpha_d}(x) = \mathbf{1}_{(-1)^{\alpha_1}x_1\geq 0,\ldots,(-1)^{\alpha_d}x_d\geq 0}, \quad x \in \mathbb{R}^d.$$

Proposition 3.6. *The distribution* P_{Ξ}^{f} *has a density* p_{Ξ}^{f}, *bounded together with all its derivatives* $\partial_i p_{\Xi}^{f}, i = 1, \ldots, d$. *For any* $\alpha_1, \ldots, \alpha_d \in \{0,1\}$,

$$p_{\Xi}^{f}(y) = (-1)^{\alpha_1+\ldots+\alpha_d}\mathrm{E}\mathbf{1}_{\alpha_1\ldots\alpha_d}(f-y)\Upsilon^{f,\Xi}, \tag{3.25}$$

$$\partial_i p_{\Xi}^{f}(y) = (-1)^{\alpha_1+\ldots+\alpha_d+\alpha_i+1}\mathrm{E}\mathbf{1}_{\alpha_1\ldots\alpha_d}(f-y)\Upsilon_i^{f,\Xi}, \tag{3.26}$$

$$y \in \mathbb{R}^d.$$

Sketch of the proof. Applying iteratively (3.24) one can deduce that, for every $F \in C_b^{\infty}(\mathbb{R}^d)$,

$$\mathrm{E}_{\Xi}[\partial_1 \ldots \partial_d F](f_1, \ldots, f_d) = \mathrm{E}F(f_1, \ldots, f_d)\Upsilon^{f,\Xi}, \tag{3.27}$$

$$\mathrm{E}_{\Xi}[\partial_1 \ldots \partial_d \partial_i F](f_1, \ldots, f_d) = \mathrm{E}F(f_1, \ldots, f_d)\Upsilon_i^{f,\Xi}, \quad i = 1, \ldots, d. \tag{3.28}$$

Now, the informal way to get representation (3.25) is to apply (3.27) to $F = \mathbf{1}_{\alpha_1\ldots\alpha_d}$:

$$\begin{aligned} p_{\Xi}^{f}(y) &= (-1)^{\alpha_1+\ldots+\alpha_d+d}\partial_1 \ldots \partial_d \mathrm{E}_{\Xi}\mathbf{1}_{\alpha_1\ldots\alpha_d}(f-y) = \\ &= (-1)^{\alpha_1+\ldots+\alpha_d}\mathrm{E}_{\Xi}[\partial_1 \ldots \partial_d \mathbf{1}_{\alpha_1\ldots\alpha_d}](f-y) = \\ &= (-1)^{\alpha_1+\ldots+\alpha_d}\mathrm{E}\mathbf{1}_{\alpha_1\ldots\alpha_d}(f-y)\Upsilon^{f,\Xi}. \end{aligned} \tag{3.29}$$

In order to justify (3.29) one should consider smooth approximations F_n for the function $F = \mathbf{1}_{\alpha_1\ldots\alpha_d}$ and use Fubini theorem (we omit detailed exposition here, referring the reader, for instance, to [12], § 3.1, 3.2). Similarly, (3.26) is provided by (3.28) and the formula

$$\partial_i p_{\Xi}^{f}(y) = (-1)^{\alpha_1+\ldots+\alpha_d+1}\mathrm{E}_{\Xi}[\partial_1 \ldots \partial_d \partial_i \mathbf{1}_{\alpha_1\ldots\alpha_d}](f-y).$$

3.4. Estimates for the densities of the truncated distributions of smooth functionals. Proposition 3.6 immediately provides the following family of estimates for the density p_{Ξ}^{f} of the truncated distribution of f.

***Corollary* 3.1.** *For any* $y \in \mathbb{R}^d$,

$$p_{\Xi}^{f}(y) \leq \|\Upsilon^{f,\Xi}\|_{L_2} \min_{\alpha_1,\ldots,\alpha_d\in\{0,1\}} \mathrm{P}_{\Xi}^{\frac{1}{2}}\Big((-1)^{\alpha_1}f_1 \geq (-1)^{\alpha_1}y_1, \ldots$$

$$\ldots, (-1)^{\alpha_d}f_d \geq (-1)^{\alpha_d}y_d\Big) \leq \|\Upsilon^{f,\Xi}\|_{L_2},$$

$$\partial_i p_{\Xi}^{f}(y) \leq \|\Upsilon_i^{f,\Xi}\|_{L_2} \min_{\alpha_1,\ldots,\alpha_d\in\{0,1\}} \mathrm{P}_{\Xi}^{\frac{1}{2}}\Big((-1)^{\alpha_1}f_1 \geq (-1)^{\alpha_1}y_1, \ldots$$

$$\ldots, (-1)^{\alpha_d}f_d \geq (-1)^{\alpha_d}y_d\Big) \leq \|\Upsilon_i^{f,\Xi}\|_{L_2},$$

$$i = 1, \ldots, d.$$

In particular, p^f_Ξ *satisfies Lipschitz condition with the constant* $L=\sum_{i=1}^d \|\Upsilon^{f,\Xi}_i\|_{L_2}$.

In this subsection we give explicit estimates for $\|\Upsilon^{f,\Xi}\|_{L_2}$, $\|\Upsilon^{f,\Xi}_i\|_{L_2}$, $i=1,\ldots,d$. Our estimates somewhat differ from the standard Malliavin-type ones. In our considerations, we operate with the matrix $[\sigma^f]^{-1}$ straightforwardly and do not use (unlike in the standard Malliavin's approach) representation of this matrix via the Cramer's formula $[\sigma^f]^{-1}=[\det\sigma^f]^{-1}\Sigma^f$ (Σ^f denotes the cofactor matrix for σ^f). This is caused by our goal to prove, together with existence of the density, an explicit estimates for it like the estimate (ii) of Theorem 1.1.

Let us give an iterative description of the family $\{\upsilon^{f,\Xi}_l\}$ involved into construction of $\Upsilon^{f,\Xi}$, $\Upsilon^{f,\Xi}_i$. We introduce two families of operators acting on W^∞_∞:

$$I_i\colon \varphi\mapsto\delta(\varphi Df_i),\qquad J_{ijk}\colon\varphi\mapsto\varphi(D\sigma^f_{jk},Df_i)_H,\quad i,j,k=1,\ldots,d.$$

We call any operator $I_1,\ldots,I_d$ an *operator of the type* I, and any operator from the set $\{J_{ijk},\ i,j,k=1,\ldots,d\}$ an *operator of the type* J. We denote by $\mathcal{K}(m,M)$ the class of all functions that can be obtained from $\varphi\equiv 1$ by applying, in arbitrary order, of m operators of the type I and M operators of the type J.

Proposition 3.7. *For any* $l=1,\ldots,d+1$, *there exist constant* $\mathsf{C}(d,l)\in\mathbb{N}$ *such that* $\upsilon^{f,\Xi}_l$ *is a sum of at most* $\mathsf{C}(d,l)$ *summands of the type*

$$\varphi\prod_{k=1}^r\varrho^{f,\Xi}_{i_kj_k},\tag{3.30}$$

where $i_k,\ j_k=1,\ldots,d$ *are arbitrary and* φ *belongs to some class* $\mathcal{K}(m,M)$ *with* $m+M=l-1$ *and* $r=M+l-1$.

Proof. We use induction by l. For $l=1$, the statement is trivial since $\upsilon^{f,\Xi}_1=1\in\in\mathcal{K}(0,0)$. Suppose the statement of the Lemma to hold true for some $l\le d$. Let us prove this statement for $l+1$. Due to the inductive supposition, $\upsilon^{f,\Xi}_l$ is a sum of at most $\mathsf{C}(d,l)$ summands of the type

$$\delta\Big(\varphi\prod_{k=1}^r\varrho^{f,\Xi}_{i_kj_k}\vartheta^{f,\Xi}_l\Big),$$

$\varphi\in\mathcal{K}(m,M)$, $m+M=l$, $M+l=r$. We have

$$\delta\left(\varphi\prod_{k=1}^r\varrho^{f,\Xi}_{i_kj_k}\vartheta^{f,\Xi}_l\right)=\sum_{q=1}^d\delta\left(\varphi\left[\prod_{k=1}^r\varrho^{f,\Xi}_{i_kj_k}\right]\varrho^{f,\Xi}_{ql}Df_l\right).$$

Thus, $\upsilon^{f,\Xi}_l$ is a sum of at most $d\mathsf{C}(d,l)$ summands of the type

$$\delta\left(\varphi\prod_{k=1}^{r+1}\varrho^{f,\Xi}_{i_kj_k}Df_l\right),$$

$\varphi\in\mathcal{K}(m,M)$, $m+M=l$, $M+l=r$. Due to Propositions 3.2 and 3.4,

$$\delta\left(\varphi\prod_{k=1}^{r+1}\varrho^{f,\Xi}_{i_kj_k}Df_l\right)=\delta(\varphi Df_l)\prod_{k=1}^{r+1}\varrho^{f,\Xi}_{i_kj_k}-\left(D\prod_{k=1}^{r+1}\varrho^{f,\Xi}_{i_kj_k},\ \varphi Df_l\right)_H=$$

$$=\delta(\varphi Df_l)\prod_{k=1}^{r+1}\varrho^{f,\Xi}_{i_kj_k}-\sum_{q=1}^{l+1}\varphi\left[\prod_{k\le r+1,\ k\ne q}\varrho^{f,\Xi}_{i_kj_k}\right]\left(\varrho^{f,\Xi}(D\sigma^f,Df_l)\varrho^{f,\Xi}\right)_{i_qj_q}. \quad (3.31)$$

Since $\varphi\in\mathcal{K}(m,M)$, $\delta(\varphi Df_l)\in\mathcal{K}(m+1,M)$. Thus, the first term in the right-hand side of (3.31) has the form (3.30). Every summand in the sum in the right-hand side of (3.31) is a sum of d^2 terms of the type

$$\varphi(D\sigma^f_{\tilde i\tilde j},Df_l)_H\left[\prod_{k=1}^{r+2}\varrho^{f,\Xi}_{\tilde i_k\tilde j_k}\right].$$

Every such a term has the form (3.30), since $\varphi(D\sigma^f_{\tilde i\tilde j},Df_l)_H\in\mathcal{K}(m,M+1)$ for $\varphi\in$ $\in\mathcal{K}(m,M)$. Therefore, the statement of the Lemma holds true for $l+1$, also, with $\mathsf{C}(d,l+1)=\mathsf{C}(d,l)\big[1+d^2(l+1)\big]$.

The proposition is proved.

Recall that $\Upsilon^{f,\Xi}=\upsilon^{f,\Xi_{d+1}}$, and thus Proposition 3.7 provides that $\Upsilon^{f,\Xi}$ is a sum of not more than $\mathsf{C}(d)$ summands of the type (3.30) with $\varphi\in\mathcal{K}(d-M,M)$ and $r=M+d$ (M may vary from 0 to d). For every such a summand,

$$\left\|\varphi\prod_{k=1}^{M+d}\varrho^{f,\Xi}_{i_kj_k}\right\|_{L_2}\le\|\varphi\|_{L_4}\left[\mathrm{E}\|\varrho^{f,\Xi}\|^{4(M+d)}_{\mathcal{M}}\right]^{\frac14}, \quad (3.32)$$

where $\|A\|_{\mathcal{M}}=\max_{i,j=1,\dots,d}|A_{ij}|$, $A\in\mathbb{R}^{d\times d}$. Thus, in order to estimate $\|\Upsilon^{f,\Xi}\|_{L_2}$, it is sufficient to estimate $\max_{\varphi\in\mathcal{K}(d-M,M)}\|\varphi\|_{L_4}$. Denote $\alpha_i=Df_i$, $\beta_{ijk}=$ $=(D\sigma^f_{jk},Df_i)_H$.

Proposition 3.8. *For every* $\varphi\in\mathcal{K}(d-M,M)$,

$$\|\varphi\|_{L_4}\le\mathsf{C}(d,\psi)\Big(\max_i\|\alpha_i\|_{2(d+1)(d+2),(d+1)^2-1}\Big)^{d-M}\times$$

$$\times\Big(\max_{ijk}\|\beta_{ijk}\|_{2(d+1)(d+2),(d+1)^2-1}\Big)^{M}. \quad (3.33)$$

Proof. By Lemma 3.1 and Proposition 3.3, for any $i=1,\dots,d$, $\varphi\in W^\infty_\infty$, $m\ge0$

$$\big\|\delta(\varphi\alpha_i)\big\|_{2m+2,m^2-1}\le\mathsf{C}(m,d,\psi)\|\varphi\alpha_i\|_{2m+2,(m+1)^2-1}\le$$

$$\le\mathsf{C}(m,d,\psi)\|\varphi\|_{2m+4,(m+1)^2-1}\|\alpha_i\|_{2(m+1)(m+2),(m+1)^2-1}. \quad (3.34)$$

By Proposition 3.3, for any $i,j,k=1,\dots,d$, $\varphi\in W^\infty_\infty$, $m\ge0$

$$\|\varphi\beta_{ijk}\|_{2m+2,m^2-1}\le\mathsf{C}(m)\|\varphi\|_{2m+4,m^2-1}\beta_{ijk}\|_{2(m+1)(m+2),m^2-1}\le$$

$$\le\mathsf{C}(m,d,\psi)\|\varphi\|_{2m+4,(m+1)^2-1}\|\beta_{ijk}\|_{2(m+1)(m+2),(m+1)^2-1}. \quad (3.35)$$

Recall that $\|\cdot\|_{L_4}=\|\cdot\|_{4,0}$ and $\varphi\in\mathcal{K}(d-M,M)$ is obtained from 1 by applying (in some order) of $d-M$ operators of the type I and M operators of the type J. Thus, in order to obtain (3.33), one should consequently put $m = d, d-1, \ldots, 1$ and apply either inequality (3.34) or inequality (3.35) depending on what type of the operator (I or J) was applied at this position in the construction of the function φ.

The proposition is proved.

Recall that $\sigma^f_{jk} = (Df_j, Df_k)_H$. By Proposition 3.3,

$$|\beta_{ijk}|_m \leq \mathsf{C}(m)|Df_i|_m|D\sigma^f_{jk}|_m \leq \tilde{\mathsf{C}}(m)|Df_i|_m|Df_j|_{m+1}|Df_k|_{m+1}.$$

Thus, Proposition 3.8 provides the following estimate for $\|\Upsilon^{f,\Xi}\|_{L_2}$. Denote

$$N_d(f) = \max_{i=1,\ldots,d}\ \max_{m=1,\ldots,(d+1)^2}\left[\mathrm{E}\|D^m f_i\|^{2(d+1)(d+2)}_{H^{\otimes m}}\right]^{\frac{1}{2(d+1)(d+2)}}.$$

***Corollary* 3.2.** *There exists a constant L_d, dependent on d and ψ only, such that*

$$\|\Upsilon^{f,\Xi}\|_{L_2} \leq \mathsf{L}_d \sum_{M=0}^{d}\Big[N_d(f)\Big]^{d+2M}\Big[\mathrm{E}\|\varrho^{f,\Xi}\|^{4(M+d)}_{\mathcal{M}}\Big]^{\frac{1}{4}}. \tag{3.36}$$

For $\|\Upsilon^{f,\Xi}_i\|_{L_2}$, the following estimate holds true (we omit the proof since it is totally analogous to the proof of (3.36) given above).

Proposition 3.9. *For any $i = 1, \ldots, d$,*

$$\|\Upsilon^{f,\Xi}_i\|_{L_2} \leq \mathsf{L}_{d+1} \sum_{M=0}^{d+1}\Big[N_{d+1}(f)\Big]^{d+2M+1}\Big[\mathrm{E}\|\varrho^{f,\Xi}\|^{4(M+d+1)}_{\mathcal{M}}\Big]^{\frac{1}{4}}. \tag{3.37}$$

At the end of this section, we formulate a general local limit theorem. This theorem is a straightforward corollary of the representation given by Proposition 3.6 and the estimates (3.36) and (3.37). Consider the sequence of probability spaces $\{(\Omega^n, \mathcal{F}^n, \mathrm{P}^n), n \geq \geq 1\}$ of the type (3.3) with the given measures $\pi_U, \varkappa, \nu$. Let the function ψ and the functions θ_n be fixed. Denote $H_n = \mathbb{R}^{d\times n}$ and consider the derivative, gradient and Sobolev spaces constructed in Subsection 3.1. For a sequence of random vectors $f^n\colon \Omega^n \to \mathbb{R}^d$ and a sequence of sets $\{\Xi_n \in \sigma(\varepsilon, \zeta)\}$ denote by P^{n,f^n} the distribution of f^n w.r.t. P^n and by $\mathrm{P}^{n,f^n}_{\Xi_n}$ the distribution of f^n w.r.t. $\mathrm{P}_{\Xi_n} \equiv \mathrm{P}^n(\cdot \cap \Xi_n)$. Denote

$$\mathsf{K}_d(f, \Xi) = \mathsf{L}_d \sum_{M=0}^{d}\Big[N_d(f)\Big]^{d+2M}\Big[\mathrm{E}\|\varrho^{f,\Xi}\|^{4(M+d)}_{\mathcal{M}}\Big]^{\frac{1}{4}}.$$

Theorem 3.1. *Suppose that $\{f^n\}$ and $\{\Xi_n\}$ satisfy condition*

(C_1) $f^n \in W^{(d+2)^2}_{2(d+2)(d+3)}(\mathbb{R}^d)$ *and* $\sup_n \mathsf{K}_{d+1}(f^n, \Xi_n) < +\infty$.

Then $\mathrm{P}^{n,f^n}_{\Xi_n}$ possess a densities $p^{f^n}_{\Xi_n}$. Moreover,

(a) $p^{f^n}_{\Xi_n}(y) \leq \mathsf{K}_d(f^n, \Xi_n)\mathrm{P}^{\frac{1}{2}}_{\Xi_n}(\|f^n\| \geq \|y\|)$;

(b) $p^{f^n}_{\Xi_n}$ *satisfy Lipschitz condition with the common constant equal to* $d\sup_n \mathsf{K}_{d+1}(f^n, \Xi_n)$.

If, additionally,

(C_2) f^n *converge in distribution to some random vector* f;

(C_3) $\mathrm{P}^n(\Xi_n) \to 1, n \to +\infty$;

then the distribution of the vector f possess a density p^f and

(c) $\sup_{y\in\mathbb{R}^d} |p^{f^n}_{\Xi_n}(y) - p^f(y)| \to 0$, $n \to +\infty$.

Proof. Although the statements of Corollaries 3.1 and 3.2 are formulated for $f \in \in \mathcal{C}^{\mathbb{R}^d}$, they can be extended to $f \in W^{(d+2)^2}_{2(d+2)(d+3)}(\mathbb{R}^d)$ by a standard approximation procedure. For any $y \in \mathbb{R}^d$, there exist a choice of the signs $\alpha_1, \ldots, \alpha_d$ such that

$$\mathrm{P}_{\Xi_n}\Big((-1)^{\alpha_1} f_1^n \geq (-1)^{\alpha_1} y_1, \ldots, (-1)^{\alpha_d} f_d^n \geq (-1)^{\alpha_d} y_d\Big) \leq \mathrm{P}_{\Xi_n}(\|f^n\| \geq \|y\|).$$

Therefore, statements (a) and (b) follow immediately from Corollaries 3.1 and 3.2. Statement (b) provides that the sequence $\{p^{f^n}_{\Xi_n}\}$ has a compact closure in the space $C(\mathbb{R}^d)$ with the topology of uniform convergence on a compacts. This together with the conditions (C_2), (C_3) provides (c).

The theorem is proved.

4. Proofs of Theorems 1.1 – 2.2. ***4.1. Proof of Theorem* 1.1.** We reduce the proof of Theorem 1.1 to the verification of the conditions of Theorem 3.1 and explicit estimation of the expression in the right-hand side of (a). We use, without additional discussion, notation introduced in Section 3.

Denote $f^n_{x,t} = X_n(t) - x$, where the processes X_n are defined by (0.2), (0.3) with the initial value $X_n(0) = x \in \mathbb{R}^d$. When it does not cause misunderstanding, we omit the indices x, t and write f^n for $f^n_{x,t}$.

We conduct the proof in several steps. First, we give explicit expressions for the derivatives of the functionals f^n. Next, we estimate the moments of these derivatives (this allows us to estimate $N_{d+1}(f^n)$). Then, on the properly chosen Ξ_n, we estimate the inverse matrix for the Malliavin matrix σ^{f^n} (this allows us to estimate $\mathsf{K}_{d+1}(f^n, \Xi_n)$). At last, we estimate the tail probabilities $\mathrm{P}_{\Xi_n}(\|f^n\| \geq \|y\|)$ in order to provide the estimates given in the statement (ii) of Theorem 1.1.

Everywhere below we suppose conditions (B_1), (B_2^κ), (B_3) of Theorem 1.1 to hold true. We prove (i) – (iii) in details and give a brief sketch of changes that should be made in order to prove (ii′), (iii′). We put

$$\theta_n(\zeta) = 1\!\!\mathrm{I}_{\max_{k\leq n}\|\zeta_k\|\leq n^{\varsigma}}, \qquad \varsigma = \frac{\kappa - 1}{2\kappa + 2}. \tag{4.1}$$

In order to make notation more convenient, we rewrite (0.2) to the form

$$X_n\left(\frac{k}{n}\right) = X_n\left(\frac{k-1}{n}\right) + a\left(X_n\left(\frac{k-1}{n}\right)\right)\frac{1}{n} + \sum_{r=1}^{d} b_r\left(X_n\left(\frac{k-1}{n}\right)\right)\frac{\xi_{kr}}{\sqrt{n}}, \tag{4.2}$$

here $\xi_{k1}, \ldots, \xi_{kd}$ are the components of the vector ξ_k and $b_1, \ldots, b_d$ are the columns of the matrix b.

Lemma 4.1. *For every* $t \in [0,1]$, $X_n(t) \in \bigcap_{p>1} W_p^{(d+2)^2}(\mathbb{R}^d)$. *Derivatives* $Y_n(t) = DX_n(t), t \in [0,1]$, *satisfy relations*

$$Y_n(0) = 0,$$

$$Y_n\left(\frac{k}{n}\right) = Y_n\left(\frac{k-1}{n}\right) + \nabla a\left(X_n\left(\frac{k-1}{n}\right)\right) Y_n\left(\frac{k-1}{n}\right)\frac{1}{n} +$$

$$+\sum_{r=1}^{d}\left[\nabla b_r\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right)\frac{\xi_{kr}}{\sqrt{n}}+\right.$$

$$\left.+\frac{\theta_n(\zeta)\mathbf{1}_{\varepsilon_k=1}\psi(\eta_k)}{\sqrt{n}}b_r\left(X_n\left(\frac{k-1}{n}\right)\right)\otimes e_{kr}\right], \tag{4.3}$$

$$Y_n(t)=Y_n\left(\frac{k-1}{n}\right)+(nt-k+1)\left[Y_n\left(\frac{k}{n}\right)-Y_n\left(\frac{k-1}{n}\right)\right], \tag{4.4}$$

$$t\in\left[\frac{k-1}{n},\frac{k}{n}\right),\quad k=1,\ldots,n.$$

Sketch of the proof. The proof is quite standard, and thus we just outline its main steps. Using induction by k, one can easily verify that, for every j, k, r, there exists $\partial_{jr}X_n\left(\frac{k}{n}\right)=\left(Y_n\left(\frac{k}{n}\right),e_{jr}\right)_{H_n}$ with Y_n defined by (4.3). One can see that $Y_n\equiv 0$ as soon as $\max_{k\le n}\|\zeta_k\|>n^{\varsigma}$ and, therefore, $\operatorname{ess\,sup}\left\|Y_n\left(\frac{k}{n}\right)\right\|<+\infty$ for every $k\le n$. Iterating these considerations, one can verify that $\operatorname{ess\,sup}\left\|\nabla_\eta^m X_n\left(\frac{k}{n}\right)\right\|<+\infty$ for every $k\le n$, $m\le(d+2)^2$, that means that $X_n\left(\frac{k}{n}\right)\in\bigcap_{p>1}W_p^{(d+2)^2}(\mathbb{R}^d)$ with $DX_n\left(\frac{k}{n}\right)=Y_n\left(\frac{k}{n}\right)$, that gives the statement of the Lemma for $t=\frac{k}{n}$. For arbitrary $t\in[0,1]$, this statement holds by linearity.

Denote $\mu_\kappa(\xi)=\mathrm{E}\|\xi_1\|^\kappa$.

Lemma 4.2. *For every* $p\ge 1$, $m\in\mathbb{N}$, *there exist constant* $\mathsf{C}(a,b,d,U,\mu_\kappa(\xi),m,p)$ *such that*

$$\mathrm{E}\|D^mX_n(t)\|^p_{H^{\otimes m}\otimes\mathbb{R}^d}\le\mathsf{C}(a,b,d,U,\mu_\kappa(\xi),m,p)t^{\frac{p}{2}},\quad t\in\left[\frac{1}{n},1\right]. \tag{4.5}$$

Proof. Consider first the case $m=1$. It is enough to prove (4.5) for $t=\frac{k}{n}$, $k=1,\ldots,n$ and $p=2q$, $q\in\mathbb{N}$. We have

$$\left\|Y_n\left(\frac{k}{n}\right)\right\|^2_{H\otimes\mathbb{R}^d}=\left\|Y_n\left(\frac{k-1}{n}\right)\right\|^2_{H\otimes\mathbb{R}^d}+$$

$$+\frac{1}{n}\left(Y_n\left(\frac{k-1}{n}\right),\ \nabla a\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right)\right)_{H\otimes\mathbb{R}^d}+$$

$$+\sum_{r=1}^{d}\left(Y_n\left(\frac{k-1}{n}\right),\ \nabla b_r\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right)\right)_{H\otimes\mathbb{R}^d}\frac{\xi_{kr}}{\sqrt{n}}+$$

$$+\frac{1}{n^2}\left\|\nabla a\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right)\right\|^2_{H\otimes\mathbb{R}^d}+$$

$$+\sum_{r=1}^{d}\left(\nabla a\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right),\right.$$

$$\left.\nabla b_r\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right)\right)_{H\otimes\mathbb{R}^d}\frac{\xi_{kr}}{n^{\frac{3}{2}}}+$$

$$+\sum_{r_1,r_2=1}^{d}\left(\nabla b_{r_1}\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right),\right.$$

$$\left.\nabla b_{r_2}\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right)\right)_{H\otimes\mathbb{R}^d}\frac{\xi_{kr_1}\xi_{kr_2}}{n}+$$

$$+\frac{\theta_n(\zeta)\mathbf{1}_{\varepsilon_k=1}\psi^2(\eta_k)}{n}\sum_{r=1}^{d}\left\|b_r\left(X_n\left(\frac{k-1}{n}\right)\right)\right\|^2_{\mathbb{R}^d}, \tag{4.6}$$

here we have used the fact that $Y_n\left(\frac{k-1}{n}\right)$, $\nabla a\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right)$, $\nabla b_r\left(X_n\left(\frac{k-1}{n}\right)\right)Y_n\left(\frac{k-1}{n}\right)$, $r=1,\ldots,d$, belong to the subspace generated by the vectors of the type $v\otimes e_{jr}$, $v\in\mathbb{R}^d$, $j<k$, $r=1,\ldots,d$, and $b_r\left(X_n\left(\frac{k-1}{n}\right)\right)\otimes$ $\otimes e_{kr}$, $r=1,\ldots,d$, are orthogonal to this subspace.

Denote $\left\|Y_n\left(\frac{k}{n}\right)\right\|^2_{H\otimes\mathbb{R}^d}=\Upsilon_k$. Recall that $Y_n\left(\frac{k}{n}\right)=0$, $k=1,\ldots,n$, as soon as there exist $j=1,\ldots,d$ such that $\|\zeta_j\|>n^{\varsigma}$. Since coefficients a, b are bounded together with their derivatives, we can rewrite (4.6) as

$$\Upsilon_k=\left[\Upsilon_{k-1}+\Theta_{k-1}\frac{1}{n}+\sum_{r=1}^{d}\Lambda_{k-1,r}\frac{\xi_{kr}}{\sqrt{n}}\mathbf{1}_{\|\zeta_k\|\le n^{\varsigma}}+\right.$$

$$\left.+\sum_{r_1,r_2=1}^{d}\Delta_{k-1,r_1,r_2}\frac{\xi_{kr_1}\xi_{kr_2}}{n}\mathbf{1}_{\|\zeta_k\|\le n^{\varsigma}}\right]\mathbf{1}_{\max_{j<k}\|\zeta_j\|\le n^{\varsigma}},\quad k=1,\ldots,n,$$

with $\mathcal{F}_{k-1}\equiv\sigma(\eta_1,\varepsilon_1,\zeta_1,\ldots,\eta_{k-1},\varepsilon_{k-1},\zeta_{k-1})$ —- measurable Θ_{k-1}, Λ_{k-1,r_1,r_2} such that

$$|\Theta_{k-1}|\le \mathsf{C}(a,b,d)\left[1+\Upsilon_{k-1}\right],\qquad |\Lambda_{k-1,r_1,r_2}|\le \mathsf{C}(a,b,d)\Upsilon_{k-1}. \tag{4.7}$$

Since $\Upsilon_k\ge 0$, we have

$$\mathrm{E}\Upsilon_k^q\le \mathrm{E}\left(\Upsilon_{k-1}+\Theta_{k-1}\frac{1}{n}\right)^q+\sum_{i=0}^{q-1}\frac{q!}{i!(q-i)!}\mathrm{E}\left(\Upsilon_{k-1}+\Theta_{k-1}\frac{1}{n}\right)^i\times$$

$$\times\left[\sum_{r=1}^{d}\Lambda_{k-1,r}\frac{\xi_{kr}}{\sqrt{n}}\mathbf{1}_{\|\zeta_k\|\le n^{\varsigma}}+\sum_{r_1,r_2=1}^{d}\Delta_{k-1,r_1,r_2}\frac{\xi_{kr_1}\xi_{kr_2}}{n}\mathbf{1}_{\|\zeta_k\|\le n^{\varsigma}}\right]^{q-i}. \tag{4.8}$$

We have $\xi_{kr} = \varepsilon_k\eta_{kr} + (1-\varepsilon_k)\zeta_{kr}$ and the set U of the possible values of $\eta_k = (\eta_{k1},\ldots,\eta_{kd})$ is bounded. In addition, $|\zeta_{kr}|\mathbf{1}_{\|\zeta_k\|\le n^\varsigma} \le n^\varsigma$. Therefore,

$$\mathrm{E}\left|\frac{\xi_{kr}}{\sqrt{n}}\right|^l \mathbf{1}_{\|\zeta_k\|\le n^\varsigma} \le \frac{(\mathsf{C}(U)n^\varsigma)^l}{n^{\frac{l}{2}}} = \mathsf{C}^l(U)n^{l(\varsigma-\frac12)} \le \frac{\mathsf{C}^l(U)}{n}, \quad l\ge\kappa+1. \tag{4.9}$$

Since $\mathrm{E}\|\xi_k\|^\kappa < +\infty$,

$$\mathrm{E}\left|\frac{\xi_{kr}}{\sqrt{n}}\right|^l \mathbf{1}_{\|\zeta_k\|\le n^\varsigma} \le \frac{\mathrm{E}|\xi_{kr}|^l}{n^{\frac{l}{2}}} \le \frac{\mathrm{E}(\|\xi_k\|^\kappa\vee 1)}{n}, \quad l=2,\ldots,\kappa. \tag{4.10}$$

Similarly,

$$\left|\mathrm{E}\frac{\xi_{kr_1}\xi_{kr_2}}{n}\mathbf{1}_{\|\zeta_k\|\le n^\varsigma}\right| \le \frac1n\mathrm{E}\|\xi_k\|^2 \le \frac{\mathrm{E}(\|\xi_k\|^\kappa\vee 1)}{n}. \tag{4.11}$$

At last,

$$\left|\mathrm{E}\frac{\xi_{kr}}{\sqrt{n}}\mathbf{1}_{\|\zeta_k\|\le n^\varsigma}\right| = \left|\mathrm{E}\frac{\xi_{kr}}{\sqrt{n}}\mathbf{1}_{\|\zeta_k\|> n^\varsigma}\right| \le n^{-\frac12}\Big[\mathrm{E}|\xi_{kr}|^\kappa\Big]^{\frac1\kappa}\Big[\mathrm{P}(\|\zeta_{kr}\|\ge n^\varsigma)\Big]^{\frac{\kappa-1}{\kappa}} \le$$

$$\le \mathsf{C}(U,\mu_\kappa(\xi))n^{-\frac12}\Big[n^{-\varsigma\kappa}\Big]^{\frac{\kappa-1}{\kappa}} \le \frac{\mathsf{C}(U,\mu_\kappa(\xi))}{n} \tag{4.12}$$

(recall that $\mathrm{E}\xi_{kr}=0$). The triple $(\eta_k,\varepsilon_k,\zeta_k)$ is independent of $\mathcal{F}_{k-1}$. Thus, taking in (4.8) conditional expectation w.r.t. $\mathcal{F}_{k-1}$ and taking into account inequalities (4.7), we obtain an estimate

$$\mathrm{E}\Upsilon_k^q \le \mathrm{E}\left(\Upsilon_{k-1}+\Theta_{k-1}\frac1n\right)^q + \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),q)\frac{1+\mathrm{E}\Upsilon_{k-1}^q}{n} \le$$

$$\le \left(1+\frac{\mathsf{C}_1(a,b,d,U,\mu_\kappa(\xi),q)}{n}\right)\mathrm{E}\Upsilon_{k-1}^q +$$

$$+\ \mathsf{C}_2(a,b,d,U,\mu_\kappa(\xi),q)\sum_{l=0}^{q-1}\left(1+\frac1n\right)^l\frac{1}{n^{q-l}}\mathrm{E}\Upsilon_{k-1}^l. \tag{4.13}$$

Let us show that (4.13) provide the family of estimates

$$\mathrm{E}\Upsilon_k^q \le \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),q)\left(\frac kn\right)^q, \qquad k=1,\ldots,n, \quad q\in\mathbb{N} \tag{4.14}$$

(note that (4.14) is exactly (4.5) with $m=1$ and $p=2q$). We use induction by q. For $q=1$, (4.13) implies that

$$\mathrm{E}\Upsilon_k \le \frac{\mathsf{C}_2(a,b,d,U,\mu_\kappa(\xi),1)}{n} +$$

$$+\ \frac{\mathsf{C}_2(a,b,d,U,\mu_\kappa(\xi),1)}{n}\left(1+\frac{\mathsf{C}_1(a,b,d,U,\mu_\kappa(\xi),1)}{n}\right)+\ldots$$

$$\ldots+\frac{\mathsf{C}_2(a,b,d,U,\mu_\kappa(\xi),1)}{n}\left(1+\frac{\mathsf{C}_1(a,b,d,U,\mu_\kappa(\xi),1)}{n}\right)^{k-1} \le$$

$$\leq \frac{k}{n}\mathsf{C}_2(a,b,d,U,\mu_\kappa(\xi),1)e^{\mathsf{C}_1(a,b,d,U,\mu_\kappa(\xi),1)}.$$

Similarly, if (4.14) holds true for all $q \leq Q-1$, then (4.13) implies that

$$\mathrm{E}\Upsilon_k^Q \leq \mathsf{C}_2(a,b,d,U,\mu_\kappa(\xi),Q)e^{\mathsf{C}_1(a,b,d,U,\mu_\kappa(\xi),Q)}\times$$

$$\times\sum_{l=0}^{Q-1}\frac{2^l}{n^{Q-l}}\mathsf{C}(a,b,d,U,\mu_\kappa(\xi),l)\left(\frac{k}{n}\right)^l \leq$$

$$\leq \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),Q)\left(\frac{k}{n}\right)^Q,$$

that proves (4.14) for $q = Q$. This proves the statement of the lemma for $m = 1$. For arbitrary m, the proof is analogous: one should write difference relations for the higher derivatives of X_n, analogous to (4.3), and then again use the moment estimates of the same type with the given above. This step does not differ principally from the one for SDE's driven by a Wiener process (see, for instance [13], Chapter V, § 8), and thus we omit its detailed exposition here.

The lemma is proved.

***Corollary* 4.1.** *The following estimate holds*:

$$N_j(X_n(t)) \leq \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),j)\sqrt{t}, \qquad t\in\left[\frac{1}{n},1\right], \quad j\in\mathbb{N}.$$

Let us proceed with the investigation of the Malliavin's matrix σ^{f^n} for $f^n = X_n(t)$. Denote by $\mathcal{E}^n_{i,j}$, $0 \leq i \leq j \leq n$, the difference analogue of the stochastic exponent for (4.3), i.e., the family of $\mathbb{R}^{d\times d}$-valued variables satisfying the relations

$$\mathcal{E}^n_{i,i} = I_{\mathbb{R}^d},$$

$$\mathcal{E}^n_{i,j} = \mathcal{E}^n_{i,j-1} + \nabla a\left(X_n\left(\frac{j-1}{n}\right)\right)\mathcal{E}^n_{i,j-1}\frac{1}{n} + \tag{4.15}$$

$$+\sum_{r=1}^d\left[\nabla b_r\left(X_n\left(\frac{j-1}{n}\right)\right)\mathcal{E}^n_{i,j-1}\frac{\xi_{jr}}{\sqrt{n}}\right], \quad j=i,\ldots,n.$$

Then one can easily obtain the representation for $Y_n(\cdot)$,

$$Y_n\left(\frac{k}{n}\right) = \theta_n(\zeta)\sum_{j=1}^k\sum_{r=1}^d\frac{\mathbf{1}_{\varepsilon_j=1}\psi(\eta_k)}{\sqrt{n}}\left[\mathcal{E}^n_{j,k}b_r\left(X_n\left(\frac{j-1}{n}\right)\right)\right]\otimes e_{jr}. \tag{4.16}$$

Denote $f^{n,k} = X\left(\dfrac{k}{n}\right)$ and $\sigma_{n,k} = \sigma^{f^{n,k}}$. By (4.16), we have

$$\sigma_{n,k} = \frac{\theta_n(\zeta)}{n}\sum_{j=1}^k\sum_{r=1}^d[\psi^2(\eta_j)\mathbf{1}_{\varepsilon_j=1}]\times$$

$$\times\left[\mathcal{E}^n_{j,k}b_r\left(X_n\left(\frac{j-1}{n}\right)\right)\right]\otimes\left[\mathcal{E}^n_{j,k}b_r\left(X_n\left(\frac{j-1}{n}\right)\right)\right] =$$

$$= \frac{\theta_n(\zeta)}{n} \sum_{j=1}^{k} [\psi^2(\eta_j) \mathbf{1}_{\varepsilon_j=1}] \left[\mathcal{E}_{j,k}^n b \left(X_n \left(\frac{j-1}{n} \right) \right) \right] \left[\mathcal{E}_{j,k}^n b \left(X_n \left(\frac{j-1}{n} \right) \right) \right]^* =$$

$$= \frac{\theta_n(\zeta)}{n} \sum_{j=1}^{k} [\psi^2(\eta_j) \mathbf{1}_{\varepsilon_j=1}] \mathcal{E}_{j,k}^n [bb^*] \left(X_n \left(\frac{j-1}{n} \right) \right) [\mathcal{E}_{j,k}^n]^*.$$

Together with the family $\{\mathcal{E}_{i,j}^n\}$, we consider the family $\{\tilde{\mathcal{E}}_{i,j}^n\}$ defined by

$$\tilde{\mathcal{E}}_{i,i}^n = I_{\mathbb{R}^d}, \qquad \tilde{\mathcal{E}}_{i,j}^n = \tilde{\mathcal{E}}_{i,j-1}^n + \nabla a \left(X_n \left(\frac{j-1}{n} \right) \right) \tilde{\mathcal{E}}_{i,j-1}^n \frac{1}{n} +$$

$$+ \sum_{r=1}^{d} \left[\nabla b_r \left(X_n \left(\frac{j-1}{n} \right) \right) \tilde{\mathcal{E}}_{i,j-1}^n \frac{\xi_{jr}^n}{\sqrt{n}} \right], \quad j = i, \ldots, n, \tag{4.17}$$

where $\xi_i^n = \xi_i \mathbf{1}_{\|\zeta_i\| \le n^{\varsigma}}$, $i = 1, \ldots, n$. By the construction, $\tilde{\mathcal{E}}_{i,j}^n = \mathcal{E}_{i,j}^n$ on the set $\{\theta_n(\zeta) = 1\}$, therefore

$$\sigma_{n,k} = \frac{\theta_n(\zeta)}{n} \sum_{j=1}^{k} [\psi^2(\eta_j) \mathbf{1}_{\varepsilon_j=1}] \tilde{\mathcal{E}}_{j,k}^n [bb^*] \left(X_n \left(\frac{j-1}{n} \right) \right) [\tilde{\mathcal{E}}_{j,k}^n]^*.$$

We have

$$\tilde{\mathcal{E}}_{i,j}^n = \prod_{l=j}^{i+1} \left[I_{\mathbb{R}^d} + \nabla a \left(X_n \left(\frac{l-1}{n} \right) \right) \frac{1}{n} + \sum_{r=1}^{d} \nabla b_r \left(X_n \left(\frac{l-1}{n} \right) \right) \frac{\xi_{lr}^n}{\sqrt{n}} \right].$$

Since ∇a, ∇b are bounded and $|\xi_{jr}^n| \mathbf{1}_{\|\zeta_j\| \le n^{\varsigma}} \le \max(\max_{x \in U} \|x\|, n^{\varsigma})$, there exists $n_0 = n_0(a, b, d, U, \varsigma)$ such that

$$\left\| \nabla a \left(X_n \left(\frac{l-1}{n} \right) \right) \frac{1}{n} + \sum_{r=1}^{d} \nabla b_r \left(X_n \left(\frac{l-1}{n} \right) \right) \frac{\xi_{lr}^n}{\sqrt{n}} \right\| \mathbf{1}_{\|\zeta_j\| \le n^{\varsigma}} \le \frac{1}{2}, \quad n \ge n_0.$$

Then $\tilde{\mathcal{E}}_{i,j}^n$ is invertible and

$$[\tilde{\mathcal{E}}_{i,j}^n]^{-1} = \prod_{l=i+1}^{j} \left[I_{\mathbb{R}^d} + \nabla a \left(X_n \left(\frac{l-1}{n} \right) \right) \frac{1}{n} + \sum_{r=1}^{d} \nabla b_r \left(X_n \left(\frac{l-1}{n} \right) \right) \frac{\xi_{lr}^n}{\sqrt{n}} \right]^{-1}.$$

Thus, on the set $\{\theta_n(\zeta) = 1\} \backslash \{\varepsilon_1 = \ldots = \varepsilon_n = 0\}$, the matrix $\sigma_{n,k}$ is invertible and

$$\|\sigma_{n,k}^{-1}\| = \left[\inf_{\|v\|=1} (\sigma_{n,k} v, v) \right]^{-1} \le$$

$$\le \beta^{-1}(b) \left[\max_{i \le j \le n} \left\| \left[\tilde{\mathcal{E}}_{i,j}^n \right]^{-1} \right\| \right]^2 \left\{ \frac{1}{n} \sum_{j=1}^{k} \psi^2(\eta_j) \mathbf{1}_{\varepsilon_j=1} \right\}^{-1}. \tag{4.18}$$

Lemma 4.3. *For every $p \ge 1$,*

$$\mathrm{E}\left[\max_{i\le j\le n}\big\|\big[\tilde{\mathcal{E}}^n_{i,j}\big]^{-1}\big\|\right]^p\le \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),p),\quad n\in\mathbb{N}.$$

Proof. Since $\big[\tilde{\mathcal{E}}^n_{i,j}\big]^{-1}=\tilde{\mathcal{E}}^n_{0,i}\big[\tilde{\mathcal{E}}^n_{0,j}\big]^{-1}$, it is enough to prove that

$$\mathrm{E}[\max_{i\le n}\|\tilde{\mathcal{E}}^n_{0,i}\|]^p\le \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),p), \tag{4.19}$$

$$\mathrm{E}\max_{j\le n}\big|\det\tilde{\mathcal{E}}^n_{0,j}\big|^{-p}\le \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),p). \tag{4.20}$$

Let us prove inequality

$$\mathrm{E}\left[\max_{i\le n}\|\tilde{\mathcal{E}}^n_{0,i}\|_2\right]^p\le \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),p) \tag{4.21}$$

with $\|A\|_2\equiv\sqrt{\sum_{lr}A^2_{lr}}$; this will provide inequality (4.19). We deduce from (4.17) that $Z^n_i\equiv\|\tilde{\mathcal{E}}^n_{0,i}\|^2_2$ satisfy relations analogous to (4.6), i.e.,

$$Z^n_i=Z^n_{i-1}+V^{1,n}_{i-1}\frac{1}{n}+\sum_r V^{2,n}_{i-1}\frac{\xi^n_{ir}}{\sqrt{n}}+\sum_{r_1,r_2}V^{3,n}_{i-1}\frac{\xi^n_{ir_1}\xi^n_{ir_1}}{n},\quad i=1,\ldots,n, \tag{4.22}$$

with an $\{\mathcal{F}_i\}$-adapted sequences $V^{1,n}_i$, $V^{2,n}_{i,\cdot}$, $V^{3,n}_{i,\cdot,\cdot}$ such that

$$\begin{gathered}\big|V^{1,n}_i\big|\le \mathsf{C}(a,b,d,U)(1+Z^n_i),\qquad \big|V^{2,n}_{i,r}\big|\le \mathsf{C}(a,b,d,U)Z^n_i,\\ \big|V^{3,n}_{i,r_1,r_2}\big|\le \mathsf{C}(a,b,d,U)Z^n_i.\end{gathered} \tag{4.23}$$

Then the moment estimates analogous to those made in the proof of Lemma 4.2 provide that

$$\max_{i\le n}(\mathrm{E}Z^n_i)^{\frac{p}{2}}\le \mathsf{C}(a,b,d,U,\mu_\kappa(\xi),p). \tag{4.24}$$

Denote $A^n_j=Z^n_0+\sum_{i=1}^j\Delta A^n_i$, $M^n_j=\sum_{i=1}^j\Delta M^n_i$ with

$$\Delta M^n_i=\sum_r V^{2,n}_{i-1}\frac{\xi^n_{ir}-\mathrm{E}\xi^n_{ir}}{\sqrt{n}}+\sum_{r_1,r_2}V^{3,n}_{i-1}\frac{\xi^n_{ir_1}\xi^n_{ir_2}-\mathrm{E}\xi^n_{ir_1}\xi^n_{ir_2}}{n}$$

and

$$\Delta A^n_i=Z^n_i-Z^n_{i-1}-M^n_i=V^{1,n}_{i-1}\frac{1}{n}+\sum_r V^{2,n}_{i-1}\frac{\mathrm{E}\xi^n_{ir}}{\sqrt{n}}+\sum_{r_1,r_2}V^{3,n}_{i-1}\frac{\mathrm{E}\xi^n_{ir_1}\xi^n_{ir_2}}{n}.$$

Then $Z_i=A_i+M_i$. By (4.12) and (4.23),

$$|\Delta A^n_i|\le\frac{\mathsf{C}(a,b,U,\mu_\kappa(\xi))}{n}(1+Z_i),$$

and therefore (4.24) provides that

$$\mathrm{E}\max_{j\le n}|A_j^n|^{\frac{r}{2}} \le \mathsf{C}(a,b,U,\mu_\kappa(\xi),p).$$

Similarly, Burkholder inequality together with (4.23) and (4.9) – (4.12) provides that

$$\mathrm{E}\max_{j\le n}|M_j^n|^{\frac{p}{2}} \le \mathsf{C}(a,b,U,\mu_\kappa(\xi),p),$$

that proves (4.21), and therefore (4.19).

On the set $\left\{\|A\|_2\le \frac{1}{2}\right\}\subset\mathbb{R}^{d\times d}$, the function $\Phi\colon A\mapsto \big(\det[I_{\mathbb{R}^d}+A]\big)^{-1}$ can be represented in the form

$$\Phi(A)=1+Q(A)+\vartheta(A),$$

where Q is a polynomial of A with $\deg Q\le\kappa$ and $|\vartheta(A)|\le \mathsf{C}\|A\|_2^{\kappa+1}$. We have

$$\left\|\nabla a\left(X_n\left(\frac{l-1}{n}\right)\right)\frac{1}{n}+\sum_{r=1}^d \nabla b_r\left(X_n\left(\frac{l-1}{n}\right)\right)\frac{\xi_{lr}^n}{\sqrt{n}}\right\|_2^{\kappa+1}\le \frac{\mathsf{C}(a,b,d,U)}{n}.$$

Therefore,

$$(\det\tilde{\mathcal{E}}_{0,j}^n)^{-1}=(\det\tilde{\mathcal{E}}_{0,j-1}^n)^{-1}\left[1+Q_{j-1}^n\left(\frac{\xi_{j1}^n}{\sqrt{n}},\ldots,\frac{\xi_{jd}^n}{\sqrt{n}}\right)+\vartheta_i^n\right],$$

where $|\vartheta_i^n|\le\dfrac{\mathsf{C}(a,b,d,U)}{n}$, ϑ_i^n is $\mathcal{F}_j$-measurable, Q_{j-1}^n is a polynomial with $\deg Q_{j-1}^n\le$ $\le\kappa$ and its coefficients are $\mathcal{F}_{j-1}$-measurable and bounded by some constant depending on the coefficients a,b. Repeating the arguments used in the proof of (4.19) we obtain (4.20).

The lemma is proved.

Lemma 4.4. *For every* $p\in\mathbb{N}$, $c>0$,

$$\mathrm{E}\left\{\sum_{j=1}^k\psi^2(\eta_j)\mathbf{1}_{\varepsilon_j=1}\right\}^{-p}\mathbf{1}_{\sum_{j=1}^k\varepsilon_j\ge ck}\le \mathsf{C}(c,\psi,p)k^{-p},\quad k\ge\frac{2p+1}{c}.$$

Remark **4.1.** For arbitrary $\psi\in C^\infty(\mathbb{R}^d)$ with $\psi=0$ on ∂U, the given above statement may fail. It is crucial for ψ to have non-zero normal derivative at (some part of) the boundary in order to provide (4.25) below to hold true.

Proof. Since η and ε are independent,

$$\mathrm{E}\left\{\sum_{j=1}^k\psi^2(\eta_j)\mathbf{1}_{\varepsilon_j=1}\right\}^{-p}\mathbf{1}_{\sum_{j=1}^k\varepsilon_j\ge ck}\le \mathrm{E}\left\{\sum_{j=1}^{]ck[}\psi^2(\eta_j)\right\}^{-p},$$

where $]x[\overset{\mathrm{df}}{=}\min\{n\in\mathbb{Z}\mid n\ge x\}$. By the construction of the function ψ,

$$\mathrm{P}(\psi^2(\eta_j)\le z)\sim \mathsf{C}(\psi)\sqrt{z},\quad z\to 0+. \tag{4.25}$$

Therefore

$$\mathrm{P}(\psi^2(\eta_1)+\ldots+\psi^2(\eta_l)\le z)\sim \mathsf{C}(\psi,l)z^{\frac{l}{2}},\quad z\to 0+,$$

and

$$\mathrm{E}(\psi^2(\eta_1)+\ldots+\psi^2(\eta_l))^{-p}<+\infty \tag{4.26}$$

as soon as $l>2p$. We put $q=2p+1$, $N=\left[\frac{]ck[}{q}\right]$ and divide the set $\{1,\ldots,]ck[\}$ on the blocks

$$\{1,\ldots,q\},\ \{q+1,\ldots,2q\},\ldots,\{(N-1)q+1,\ldots,Nq\},\ \{Nq+1,\ldots,]ck[\}$$

(the last block may be empty). We denote

$$\vartheta_i=\sum_{j=(i-1)q+1}^{iq}\psi^2(\eta_j),\quad i=1,\ldots,N.$$

We have

$$\mathrm{E}\left\{\sum_{j=1}^{]ck[}\psi^2(\eta_j)\right\}^{-p}\le \mathrm{E}\left(\sum_{i=1}^{N}\vartheta_i\right)^{-p}=N^{-p}\mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}\vartheta_i\right)^{-p}.$$

The function $x\mapsto x^{-p}$ is convex on $\mathbb{R}^+$, and therefore

$$\mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}\vartheta_i\right)^{-p}\le \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}\vartheta_i^{-p}\right)=\mathrm{E}\vartheta_1^{-p}<+\infty$$

(the last inequality follows from (4.26)). If $k\ge\frac{2p+1}{c}$, then $\frac{]ck[}{q}\ge 1$ and therefore $N=\left[\frac{]ck[}{q}\right]\ge\frac{]ck[}{2q}\ge\frac{c}{4p+2}k$. Thus

$$\mathrm{E}\left\{\sum_{j=1}^{]ck[}\psi^2(\eta_j)\right\}^{-p}\le \mathsf{C}(\psi,p)\left(\frac{c}{4p+2}\right)^{-p}k^{-p}.$$

The lemma is proved.

Inequality (4.18), Lemmas 4.3 and 4.4 provide the following estimate. For a given $c>0$ and $t\in[0,1]$, we put $\tilde{\Xi}_n=\{\theta_n(\zeta)=1\}\bigcap\left\{\sum_{j=1}^{[tn]}\varepsilon_j\ge c[tn]\right\}$.

***Corollary* 4.2.** *For $p\in\mathbb{N}$ and $[tn]>\frac{2p+1}{c}$,*

$$\mathrm{E}\big\|\varrho^{f^n,\tilde{\Xi}_n}\big\|^p_{\mathcal{M}}\le \mathsf{C}(a,b,d,U,\psi,\mu_\kappa(\xi),p)t^{-p}.$$

At last, let us give an estimates for the tail probabilities for f^n. The following lemma is completely analogous to Lemma 4.2; the proof is omitted.

Lemma 4.5. *For every $p \ge 1$, there exist constant $\mathsf{C}(a, b, d, U, \mu_\kappa(\xi), p)$ such that*

$$\mathrm{E}\|X_n(t) - X(0)\|^p_{\mathbb{R}^d} \mathbf{1}_{\{\theta_n(\zeta)=1\}} \le \mathsf{C}(a, b, d, U, \mu_\kappa(\xi), p) t^{\frac{p}{2}}, \quad t \in \left[\frac{1}{n}, 1\right].$$

***Corollary* 4.3.** *For every $p \ge 1$, there exists constant C_p, dependent on a, b, d, U, $\mu_\kappa(\xi)$, p, such that*

$$\mathrm{P}\big(\|X_n(t) - X_n(0)\| \ge y, \ \ \theta_n(\zeta) = 1\big) \le C_p \left(1 + \frac{\|y\|^2}{t}\right)^{-p}, \ \ n \in \mathbb{N}, \ \ t \in \left[\frac{1}{n}, 1\right].$$

***Remark* 4.2.** For $\frac{\|y\|^2}{t}$ small, the latter inequality is trivial since $\mathrm{P}(\cdot) \le 1$. For $\frac{\|y\|^2}{t}$ large, it comes from Chebyshev's inequality.

Lemma 4.6. *There exist constants C_1, C_2, C_3, dependent on a, b, U, d, $\mu_\kappa(\xi)$, such that, for every $\lambda \in \mathbb{R}^d$ with $\|\lambda\| \le \mathsf{C}_1 n^{\frac{1}{\kappa+1}}$,*

$$\mathrm{E} e^{(\lambda, X_n(t) - X_n(0))} \mathbf{1}_{\{\theta_n(\zeta)=1\}} \le \mathsf{C}_2 e^{\mathsf{C}_3 t \|\lambda\|^2}, \qquad t \in \left[\frac{1}{n}, 1\right], \quad n \in \mathbb{N}. \tag{4.27}$$

Proof. For a given λ, denote $Z_n(t) = e^{(\lambda, X_n(t) - X_n(0))}$. We have $Z_n(0) = 1$. On the other hand,

$$\left|\frac{\xi_{kr}}{\sqrt{n}}\right| \le \frac{\max_{y \in U} \|y\|}{\sqrt{n}} + n^{\varsigma} n^{-\frac{1}{2}} = \frac{\max_{y \in U} \|y\|}{\sqrt{n}} + n^{-\frac{1}{\kappa+1}}$$

on the set $\{\theta_n(\zeta) = 1\}$. Thus there exists a constant C_4 such that, for $\|\lambda\| \le \mathsf{C}_1 n^{\frac{1}{\kappa+1}}$,

$$\left| \frac{1}{n}\left(\lambda, a\left(X_n\left(\frac{k-1}{n}\right)\right)\right) + \sum_{r=1}^d \frac{\xi_{kr}}{\sqrt{n}} \left(\lambda, b_r\left(X_n\left(\frac{k-1}{n}\right)\right)\right)\right| \le \mathsf{C}_4$$

on the set $\{\theta_n(\zeta) = 1\}$. Using the elementary inequality $e^x \le 1 + x + \mathsf{C}x^2$, $|x| \le \mathsf{C}_4$, we obtain that, on the same set,

$$Z_n\left(\frac{k}{n}\right) = Z_n\left(\frac{k-1}{n}\right) \exp\left[\frac{1}{n}\left(\lambda, a\left(X_n\left(\frac{k-1}{n}\right)\right)\right) + \right.$$

$$\left. + \sum_{r=1}^d \frac{\xi_{kr}}{\sqrt{n}} \left(\lambda, b_r\left(X_n\left(\frac{k-1}{n}\right)\right)\right)\right] \le$$

$$\le Z_n\left(\frac{k-1}{n}\right) \left[1 + \Theta_{k-1} \frac{\|\lambda\| + \|\lambda\|^2}{n} + \sum_{r=1}^d \Lambda_{k-1,r} \frac{\xi_{kr}(\|\lambda\| + \|\lambda\|^2)}{\sqrt{n}} + \right.$$

$$\left. + \sum_{r_1, r_2=1}^d \Delta_{k-1, r_1, r_2} \frac{\xi_{kr_1}\xi_{kr_2}\|\lambda\|^2}{n}\right] \tag{4.28}$$

with an $\mathcal{F}_{k-1}$-measurable coefficients Θ_{k-1}, $\Lambda_{k-1,r}$, Δ_{k-1,r_1,r_2}, bounded by some constant C. An arguments, analogous to those used in the proof of Lemma 4.2, provide that (4.28) implies the estimate

$$\mathrm{E}Z_n\left(\frac{k}{n}\right) \le \left(1 + \mathsf{C}\frac{\|\lambda\| + \|\lambda\|^2}{n}\right)^k \le \exp\left[2\mathsf{C}\frac{k}{n}(1+\lambda^2)\right], \quad k,\ n \in \mathbb{N}. \tag{4.29}$$

This is exactly (4.27) for $t = \dfrac{k}{n}$. For $t \in \left(\dfrac{k}{n}, \dfrac{k+1}{n}\right)$, $Z_n(t)$ is a linear combination of $Z_n\left(\dfrac{k}{n}\right)$ and $Z_n\left(\dfrac{k+1}{n}\right)$. Therefore, (4.27) follows from (4.29) and relation $\dfrac{k+1}{n} \le \le 2\dfrac{k}{n} \le 2t$ (recall that $t \ge \frac{1}{n}$ and thus $k \ \ge \ 1$).

The lemma is proved.

***Corollary* 4.4.** *There exist constants* C_5, C_6, C_7, *dependent on* a, b, U, d, $\mu_\kappa(\xi)$, *such that*

$$\mathrm{P}\Big(\|X_n(t) - X_n(0)\| \ge y, \theta_n(\zeta) = 1\Big) \le \mathsf{C}_6 e^{-\mathsf{C}_7\frac{y^2}{t}},$$

$$y \in (0, \mathsf{C}_5 t n^{\frac{1}{\kappa+1}}), \quad n \in \mathbb{N}, \quad t \in \left[\frac{1}{n}, 1\right],$$

and

$$\mathrm{P}\Big(\|X_n(t) - X_n(0)\| \ge y, \theta_n(\zeta) = 1\Big) \le \mathsf{C}_6 e^{-\mathsf{C}_7 n^{\frac{1}{\kappa+1}} y},$$

$$y \ge \mathsf{C}_5 t n^{\frac{1}{\kappa+1}}, \quad n \in \mathbb{N}, \quad t \in \left[\frac{1}{n}, 1\right].$$

Proof. It is enough to verify that, for any coordinate $(X_n)_j$ of the process X_n, $j = 1, \ldots, d$, there exist constants $\tilde{\mathsf{C}}_5$, $\tilde{\mathsf{C}}_6$, $\tilde{\mathsf{C}}_7$, $\tilde{\mathsf{C}}_8$ such that

$$\mathrm{P}\Big(\pm((X_n)_j(t) - (X_n)_j(0)) \ge y, \theta_n(\zeta) = 1\Big) \le \tilde{\mathsf{C}}_6 e^{-\tilde{\mathsf{C}}_7\frac{y^2}{t}}, \tag{4.30}$$

$$y \in (0, \tilde{\mathsf{C}}_5 t n^{\frac{1}{\kappa+1}}), \quad n \in \mathbb{N}, \quad t \in \left[\frac{1}{n}, 1\right],$$

and

$$\mathrm{P}\Big(\pm((X_n)_j(t) - (X_n)_j(0)) \ge y, \theta_n(\zeta) = 1\Big) \le \tilde{\mathsf{C}}_6 e^{-\tilde{\mathsf{C}}_8 n^{\frac{1}{\kappa+1}} y}, \tag{4.31}$$

$$y \ge \tilde{\mathsf{C}}_5 t n^{\frac{1}{\kappa+1}}, \quad n \in \mathbb{N}, \quad t \in \left[\frac{1}{n}, 1\right].$$

Inequality (4.30) with $\tilde{\mathsf{C}}_5 = 2\mathsf{C}_1\mathsf{C}_3$ and $\tilde{\mathsf{C}}_7 = [2\mathsf{C}_3]^{-1}$ follows from (4.27) with $\lambda = = \left(\pm\dfrac{y}{2\mathsf{C}_3 t}\right)e_j$, where e_j is the j-th coordinate vector in $\mathbb{R}^d$. Inequality (4.31) with the same $\tilde{\mathsf{C}}_5$ and $\tilde{\mathsf{C}}_8 = \dfrac{\mathsf{C}_1}{2}$ follows from (4.27) with $\lambda = \big(\pm\mathsf{C}_1 n^{\frac{1}{\kappa+1}}\big)e_j$.

***Proof of Theorem* 1.1.** We take $p = 8(d+1)$ and fix some $c \in (0, \alpha)$ (α is given in condition (B_3)). We write $n_* = n_0(a, b, U, \varsigma)$ (see the notation before Lemma 4.3) and put

$$\Xi_n^t = \begin{cases} \{\theta_n(\zeta) = 1\} \bigcap \left\{ \sum_{j=1}^{[tn]} \varepsilon_j \geq c[tn] \right\}, & n \geq n_*, \quad [tn] > \dfrac{2p+1}{c}, \\ \varnothing, & \text{otherwise}, \end{cases}$$

$$Q_{x,t}^n(dy) = \mathrm{P}(X_n(x,t) \in dy, \Xi_n^t) = \mathrm{P}_{\Xi_n^t}(f_{x,t}^n \in dy - x),$$

$$R_{x,t}^n(dy) = \mathrm{P}(X_n(x,t) \in dy, \Omega \backslash \Xi_n^t).$$

Corollaries 4.1 and 4.2 provide condition (C_1) of the Theorem 3.1. By the statement (a) of this theorem,

$$Q_{x,t}^n(dy) = q_{x,t}^n(y)\,dy \quad \text{with} \quad q_{x,t}^n \leq \mathsf{K}_d(f^n, \Xi_n)\mathrm{P}^{\frac{1}{2}}(\|f^n\| \geq \|y\|, \theta_n(\zeta) = 1). \tag{4.32}$$

Moreover, Corollaries 4.1 and 4.2 provide an explicit estimate for $\mathsf{K}_d(f^n, \Xi_n)$. Namely, for some constant C dependent on a, b, c, d, $\mu_\kappa(\xi)$, U, ψ,

$$\mathsf{K}_d(f^n, \Xi_n) \leq \mathsf{C} \sum_{M=0}^{d} [\sqrt{t}]^{d+2M} [t^{-4(M+d)}]^{\frac{1}{4}} = (d+1)\mathsf{C}\, t^{-\frac{d}{2}}. \tag{4.33}$$

Thus the statement (a) of Theorem 3.1 and Corollaries 4.4, 4.3 provide statement (ii) of Theorem 1.1.

By Chebyshev inequality,

$$\mathrm{P}(\theta_n(\zeta) = 0) \leq n n^{-\kappa \cdot \varsigma} = n^{-\epsilon(\kappa)}.$$

Take $\lambda = \ln\left(\dfrac{\alpha(1-c)}{c(1-\alpha)}\right) > 0$. By Chebyshev inequality, we get, after some simple calculations,

$$\mathrm{P}\left(\sum_{j=1}^{k} \varepsilon_j < ck\right) \leq \frac{\mathrm{E}e^{-\lambda \sum_{j=1}^{k} \varepsilon_j}}{e^{-\lambda ck}} = [\Psi(\alpha, c)]^k, \quad k \in \mathbb{N}, \tag{4.34}$$

with $\Psi(\alpha, c) = \left(\dfrac{1-\alpha}{1-c}\right)^{1-c} \left(\dfrac{\alpha}{c}\right)^{c}$. One can verify that $\Psi(\alpha, c) < 1$ for $0 < c < \alpha < 1$. Thus, we can conclude that, for $\rho = -\dfrac{1}{2} \ln \Psi(\alpha, c) > 0$,

$$\mathrm{P}(\Omega \backslash \Xi_n^t) \leq n^{-\epsilon(\kappa)} + e^{-\rho n t} \quad \text{when} \quad n \geq n_*, \quad [tn] > \frac{2p+1}{c}$$

$\left(\text{we have used here that } [nt] \geq \dfrac{nt}{2} \text{ for } t \geq \dfrac{1}{n}\right)$. Thus, for all $t > 0$,

$$\mathrm{P}(\Omega \backslash \Xi_n^t) \leq D\big[n^{-\epsilon(\kappa)} + e^{-\rho n t}\big] \tag{4.35}$$

with the constant D dependent on n_*, p, c, ρ. This provides statement (iii) of Theorem 1.1.

We have shown that if $x \in \mathbb{R}^d$, $t > 0$ are fixed then the functions $f^n = f_{x,t}^n$ and the sets $\Xi_n = \Xi_n^t$ satisfy all the conditions of Theorem 3.1. This means that

$q^n_{x,t}(y) \to p_{x,t}(y)$ uniformly w.r.t. $y \in \mathbb{R}^d$. In order to show that this convergence holds uniformly w.r.t. $t \in [\delta, 1]$, $x,\ y \in \mathbb{R}^d$, we need to show that, for every sequences $\{t_n\} \subset [\delta, 1]$, $\{x_n\}$, $\{y_n\} \subset \mathbb{R}^d$,

$$q^n_{x_n,t_n}(y_n) - p_{x_n,t_n}(y_n) \to 0, \quad n \to \infty. \tag{4.36}$$

We can suppose that $\{t_n\}$ converges to some $t \in [\delta, 1]$. The functions a, b are bounded together with their derivatives up to the second order, and therefore the sequences of the functions

$$a_n(\cdot) = a(\cdot + x_n), \qquad b_n(\cdot) = b(\cdot + x_n)$$

are pre-compact in $C^1(\mathbb{R}^d, \mathbb{R}^d)$ and $C^1(\mathbb{R}^d, \mathbb{R}^{d\times d})$, correspondingly. We can suppose that

$$a_n \to \tilde{a} \text{ in } C^1(\mathbb{R}^d, \mathbb{R}^d), \quad b_n \to \tilde{b} \quad \text{in} \quad C^1(\mathbb{R}^d, \mathbb{R}^{d\times d}).$$

Consider the processes Z_n defined by the relations of the type (0.2), (0.3) with $Z_n(0) = 0$ and the coefficients a, b replaced by a_n, b_n. Also, consider the processes Z^n defined by the stochastic differential equations of the type (0.1) with $Z^n(0) = 0$ and the coefficients a, b replaced by a_n, b_n. At last, consider the process Z defined by the stochastic differential equations of the type (0.1) with $Z(0) = 0$ and the coefficients a, b replaced by $\tilde{a}$, $\tilde{b}$. Denote $f_n = Z_n(t_n)$, $\Xi_n = \Xi^{t_n}_n$. It is easy to verify that Z_n converge weakly in $C([0, 1], \mathbb{R}^d)$ to Z (see, for instance, Proposition 5.1 [14]). Thus, for the sequences f_n, Ξ_n, all the conditions of Theorem 3.1 hold true with $f = Z(t)$. This means that f possesses a distribution density p^f and

$$\sup_y \left| p^{f_n}_{\Xi_n}(y) - p^f(y) \right| \to 0. \tag{4.37}$$

Similarly, one can show that, for $f^n = Z^n(t_n)$, the distribution density p^{f^n} exists and

$$\sup_y \left| p^{f^n}(y) - p^f(y) \right| \to 0. \tag{4.38}$$

Now, (4.36) is provided by the relations (4.37), (4.38) and

$$q_{x_n,t_n}(y_n) = p^{f_n}_{\Xi_n}(y_n - x_n), \qquad p_{x_n,t_n}(y_n) = p^{f^n}(y_n - x_n).$$

This proves statement (i) of Theorem 1.1.

The proof of (ii′) and (iii′) can be conducted analogously, with an appropriate changes of the truncation procedure and corresponding estimates. Under ($\mathrm{B}_2^{\exp}$), we put, instead of (4.2),

$$\theta_n(\zeta) = \mathbf{1}_{\max_{k\le n} \|\zeta_k\| \le \delta\sqrt{n}}.$$

By the Chebyshev's inequality,

$$\mathrm{P}(\theta_n(\zeta) = 0) \le n e^{-\varkappa(\delta\sqrt{n})^2} \le e^{-\tilde{\rho} n}$$

with an appropriate $\tilde{\rho} > 0$. This and the estimate (4.34) provide statement (iii′).

Under ($B_2^{\exp}$) and the truncation level given above, the estimates (4.10) – (4.12) have their (simpler) analogues, and thus the statement of Lemma 4.2 holds. The constant δ in the definition of the truncation level can be made small enough to provide inequality

$$\left\|\nabla a\left(X_n\left(\frac{l-1}{n}\right)\right)\frac{1}{n}+\sum_{r=1}^{d}\nabla b_r\left(X_n\left(\frac{l-1}{n}\right)\right)\frac{\xi_{lr}^n}{\sqrt{n}}\right\|\mathbf{1}_{\|\zeta_j\|\le\delta\sqrt{n}}\le\frac{1}{2},$$

$$n\ge n_0,$$

to hold with some n_0 dependent on a, b, U. Then the statement of Lemma 4.3 holds true, also. Lemma 4.4 does not depend on the truncation procedure. Thus the Corollaries 4.1, 4.2 hold true and provide the principal estimate (4.33).

Under condition ($B_2^{\exp}$),

$$\mathrm{E}\left[\exp\left[\frac{1}{n}\left(\lambda, a\left(X_n\left(\frac{k-1}{n}\right)\right)\right)+\right.\right.$$

$$\left.\left.+\sum_{r=1}^{d}\frac{\xi_{kr}}{\sqrt{n}}\left(\lambda, b_r\left(X_n\left(\frac{k-1}{n}\right)\right)\right)\right]\Big|\mathcal{F}_{\frac{k-1}{n}}\right]\le$$

$$\le\tilde{C}_1e^{\tilde{C}s_2\frac{\|\lambda\|^2}{n}}\quad\text{a.s.}$$

for every $\lambda\in\mathbb{R}^d$ with some constants $\tilde{C}_1$, $\tilde{C}_2$ dependent on a, b, d, $\varkappa$, $\mathrm{E}\,e^{\varkappa\|\xi_k\|^2}$. Then the arguments analogous to those made in the proof of Lemma 4.2 provide that the estimate (4.27) holds true for every $\lambda\in\mathbb{R}^d$. Consequently, the first inequality in Corollary 4.4 holds true for every $y>0$. This inequality, the estimate (4.33) and Theorem 3.1 provide (iii′).

This completes the proof of Theorem 1.1.

***4.2. Proof of Theorem* 1.2.** The implication $2\Rightarrow1$ is obvious. Let us first prove 2 under additional supposition (B_3). We put, in the notation of Section 3, $\theta_n(\zeta)\equiv1$, $f^n=\frac{1}{\sqrt{n}}\sum_{k=1}^n\xi_k$,

$$\Xi_n=\begin{cases}\left\{\sum_{j=1}^{n}\varepsilon_j\ge cn\right\}, & n\ge n_*,\\ \varnothing, & \text{otherwise},\end{cases}$$

with n_*, c that will be defined later. Then

$$Df^n=\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\sum_{r=1}^{d}\psi(\eta_k)\mathbf{1}_{\varepsilon_k=1}b_r\otimes e_{kr},$$

where b_r stands for the r-th coordinate vector in $\mathbb{R}^d$ (the proof is straightforward and omitted). Since ψ is bounded on U together with all its derivatives, this provides the estimates analogous to those given in Lemma 4.2. The statement of Lemma 4.3 is trivial now, since $\mathcal{E}_{i,j}=I_{\mathbb{R}^d}$ (the identity matrix in $\mathbb{R}^d$) for every i,j. Now, take $p=8(d+1)$, $c\in(0,\alpha)$, $n_*>\frac{2p+1}{c}$. Using Lemma 4.4, we obtain the estimate (4.33) with the constant C dependent on c, d, ψ. The estimate (4.34) provides that $\mathrm{P}(\Xi_n) \le De^{-\rho n}$ for an appropriate $D, \rho > 0$. Now the statement 2 follows from Theorem 3.1.

Let us replace the additional supposition (B$_3$) by the condition 1. It is enough to prove the statement 2 for $n \in m\mathbb{N}$ with some given $m \in \mathbb{N}$. We take $m = 2n_0$ with n_0 given in the statement 1 and have

$$\frac{d[P_m]^{ac}}{d\lambda^d} \ge \frac{d[P_{n_0}]^{ac}}{d\lambda^d} * \frac{d[P_{n_0}]^{ac}}{d\lambda^d}.$$

The function $\frac{d[P_{n_0}]^{ac}}{d\lambda^d} * \frac{d[P_{n_0}]^{ac}}{d\lambda^d}$ is continuous since λ^d-almost all points of $\mathbb{R}^d$ are a Lebesgue points (i.e., a points of λ^d-almost continuity) for any function $g \in L_1(\mathbb{R}^d)$. In addition, $\frac{d[P_{n_0}]^{ac}}{d\lambda^d} * \frac{d[P_{n_0}]^{ac}}{d\lambda^d}$ is not an identical zero due to the statement 1. Thus there exist $\tilde{\alpha} > 0$ and an open set $U \subset \mathbb{R}^d$ such that $\frac{d[P_{n_0}]^{ac}}{d\lambda^d} * \frac{d[P_{n_0}]^{ac}}{d\lambda^d} \ge \tilde{\alpha}\mathbf{1}_U$. Therefore the distribution of $\xi_1 + \ldots + \xi_m$ satisfies (B$_3$). Using what we have proved before, we deduce that the statement 2 holds for $n \in m\mathbb{N}$, and therefore for $n \in \mathbb{N}$.

This completes the proof of Theorem 1.2.

***4.3. Sketch of the proof of Theorem* 2.1.** We will show that, under conditions of Theorem 2.1, the following *uniform local Doeblin condition* holds true. For two measures μ_1, μ_2, denote

$$[\mu_1 \wedge \mu_2](dy) = \min\left[\frac{d\mu_1}{d(\mu_1+\mu_2)}(y),\ \frac{d\mu_2}{d(\mu_1+\mu_2)}(y)\right](\mu_1+\mu_2)(dy).$$

Proposition 4.1. *For every ball B there exists $n_B \in \mathbb{N}$, $T_B, \gamma_B > 0$ such that*

$$\Big[P_{x,T_B} \wedge P_{x',T_B}\Big](\mathbb{R}^d) \ge \gamma_B, \quad x,\ x' \in B, \tag{4.39}$$

and, for every $n \ge n_B$, there exists $T_B^n \in \frac{1}{n}\mathbb{Z}_+$, $T_B^n \le T_B$ such that

$$\Big[P^n_{x,T_B^n} \wedge P^n_{x',T_B^n}\Big](\mathbb{R}^d) \ge \gamma_B, \quad x,\ x' \in B, \quad n \ge n_B. \tag{4.40}$$

Once Proposition 4.1 is proved, one can finish the proof of Theorem 2.1 following the proof of Theorem 1 in [4] literally. We omit this part of the discussion and prove Proposition 4.1, only.

***Proof of Proposition* 4.1.** Since (B$_5$) implies (B$_2^\kappa$) for any κ, we can apply Theorem 1.1. One can easily see that

$$\Big[P^n_{x,t} \wedge P^n_{x',t}\Big](\mathbb{R}^d) \ge \int_{\mathbb{R}^d} \min[q^n_{x,t}(y), q^n_{x',t}(y)]\,dy.$$

Thus, for any sequence $t_n \to t > 0$ we have, by the statement (i),

$$\lim\inf_{n\to+\infty}\inf_{x,x'\in B}\Big[P^n_{x,t} \wedge P^n_{x',t}\Big](\mathbb{R}^d) \ge \inf_{x,x'\in B}\int_{\mathbb{R}^d} \min\big[p_{x,t}(y), p_{x',t}(y)\big]\,dy.$$

On the other hand (see [2]), under condition (B$_1$) the function

$$\mathbb{R}^d \times (0, +\infty) \times \mathbb{R}^d \ni (x, t, y) \mapsto p_{x,t}(y)$$

is continuous and strictly positive at every point. Therefore, for a given B,

$$\gamma_B \stackrel{\mathrm{df}}{=} \frac{1}{2} \inf_{x,x' \in B} \int_{\mathbb{R}^d} \min \left[p_{x,t}(y), p_{x',t}(y) \right] dy > 0,$$

and (4.39), (4.40) hold true for $T_B = T_B^n = 1$ and sufficiently large n_B.

The proposition is proved.

***4.4. Sketch of the proof of Theorem* 2.2.** It was already mentioned in Subsection 2.1 that Theorem 2.2 is analogous to Theorem 2.1 [5]. We refer the reader to the paper [5] for the detailed proof. Here, we expose a principal estimate only, demonstrating that, in this proof, the truncated local limit theorem can be used efficiently instead of the usual one, that was used in [5].

Theorem 2.1 [5] is derived from the general theorem on convergence in distribution of a sequence of additive functionals of Markov chains, given in the paper [11] (Theorem 1). *The characteristics* of the functionals $\varphi(X)$, $\varphi_n(X_n)$ are defined by the relations

$$f^t(x) \stackrel{\mathrm{df}}{=} \mathrm{E}\big[\varphi^{0,t}(X_n) \mid X(0) = x\big],$$

$$f_n^{s,t}(x) \stackrel{\mathrm{df}}{=} \mathrm{E}\big[\varphi_n^{s,t}(X_n) \mid X_n(s) = x\big], \quad s = \frac{i}{n}, \quad i \in \mathbb{Z}_+, \quad t > s, \quad x \in \mathbb{R}^d.$$

The first relation is due to [10], Chapter 6. The second relation was introduced in [11] by an analogy with the first one.

The key condition of Theorem 1 [11] is

$$\sup_{x \in \mathbb{R}^d,\ s = \frac{i}{n},\ t \in (s,T)} \left| f_n^{s,t}(x) - f^{t-s}(x) \right| \to 0, \quad n \to \infty. \tag{4.41}$$

Here, we need to verify this condition only, since, for all the other conditions, the proof from [5] can be used literally. We have

$$f_n^{s,t}(x) = \frac{1}{n} F_n(x) + \frac{1}{n} \sum_{k \in \mathbb{N}, \frac{k}{n} < t-s} \int_{\mathbb{R}^m} F_n(y) P^n_{x, \frac{k}{n}}(dy) = f_n^{0,t-s}(x), \quad s \le t, \ x \in \mathbb{R}^m.$$

We use the decomposition $P^n = Q^n + R^n$ from Theorem 1.1 and write

$$f_n^{0,t}(x) = \frac{1}{n} F_n(x) + \frac{1}{n} \sum_{k \in \mathbb{N}, \frac{k}{n} < t} \int_{\mathbb{R}^m} F_n(y) R^n_{x, \frac{k}{n}}(dy) + \frac{1}{n} \sum_{k \in \mathbb{N}, \frac{k}{n} < t} \int_{\mathbb{R}^m} F_n(y) q^n_{x, \frac{k}{n}}(y) dy.$$

The statement (ii) of Theorem 1.1 and the estimates, analogous to the estimates (4.2) – (4.10) from [5], imply that

$$\sup_{x \in \mathbb{R}^d, t \le T} \left| \frac{1}{n} \sum_{k \in \mathbb{N}, \frac{k}{n} < t} \int_{\mathbb{R}^m} F_n(y) q^n_{x, \frac{k}{n}}(y)\, dy - f^t(x) \right| \to 0, \quad n \to \infty.$$

On the other hand, $\epsilon(\kappa) = \dfrac{8}{7} > 1$ for $\kappa = 6$. Thus, by condition (B_6) and the statement (iii) of Theorem 1.1,

$$\mathrm{E}\left[\frac{1}{n}F_n(x)+\frac{1}{n}\sum_{k\in\mathbb{N},\,\frac{k}{n}<t}\int_{\mathbb{R}^m}F_n(y)R^n_{x,\frac{k}{n}}(dy)\right]\le$$

$$\le n^{-1}\sup_{x'}F_n(x')\left[1+\sum_{k<tn}R^n_{x,\frac{k}{n}}(\mathbb{R}^d)\right]\le$$

$$\le n^{-1}\sup_{x'}F_n(x')\left[1+Dn^{-\frac{8}{7}}nt+D\sum_{k\in\mathbb{N}}e^{-\gamma k}\right]\to 0,\quad n\to\infty,$$

uniformly for $x\in\mathbb{R}^d$, $t\le T$ for any $T\in\mathbb{R}^+$. This proves (4.41).

1. *Skorokhod A. V.* Asymptotic methods in the theory of stochastic differential equations. – Kiev: Naukova dumka, 1987. – 328 p.
2. *Il'in A. M., Kalashnikov A. S., Oleinik O. A.* Linear second order parabolic equations // Uspekhi Mat. Nauk. – 1962. – **17**, № 3. – P. 3 – 143.
3. *Konakov V., Mammen E.* Local limit theorems for transition densities of Markov chains converging to diffusions // Probab.Theory and Relat. Fields. – 2000. – **117**. – P. 551 – 587.
4. *Klokov S. A., Veretennikov A. Yu.* Mixing and convergence rates for a family of Markov processes approximating SDEs // Random Oper. and Stochast. Equat. – 2006. – **14**, № 2. – P. 103 – 126.
5. *Kulik A. M.* Difference approximation for local times of multidimensional diffusions. – Preprint is available at arxiv:math.PR/0702175.
6. *Konakov V.* Small time asymptotics in local limit theorems for Markov chains converging to diffusions. – 2006. – arxiv:math. PR/0602429.
7. *Ibragimov I. A., Linnik V. Yu.* Independent and stationary connected variables. – Moscow: Nauka, 1965. – 524 p.
8. *Veretennikov A. Yu.* On estimates of mixing rate for stochastic equations // Teor. Veroyatnost. i Prim. — 1987. – **32**. – P. 299 – 308.
9. *Veretennikov A. Yu.* On polynomial mixing and rate of convergence for stochastic differential and difference equations // Teor. Veroyatnost. i Prim. – 1999. – **44**. – P. 312 – 327.
10. *Dynkin E. B.* Markov processes. – Moscow: Fizmatgiz, 1963. – 860 p.
11. *Kartashov Yu. N., Kulik A. M.* Invariance principle for additive functionals of Markov chains. – Preprint is available at arXiv:0704.0508v1.
12. *Nualart D.* Analysis on Wiener space and anticipating stochastic calculus // Lect. Notes Math. – 1998. – **1690**. – P. 123 – 227.
13. *Ikeda N., Watanabe S.* Stochastic differential equations and diffusion processes. – Amsterdam: North-Holland, 1981. – 456 p.
14. *Kurtz T. G., Protter Ph.* Weak limit theorems for stochastic integrals and SDE's // Ann. Probab. – 1991. – **19**, № 3. – P. 1035 – 1070.

Received 10.01.08