fh-2r.tex

On local mixing conditions for SDE approximations

S. A. Klokov[*], A. Yu. Veretennikov[†]

July 5, 2011

Abstract

Under minimal assumptions on smoothness of the drift, it is shown that estimates on the rate of beta–mixing for an SDE with an identity diffusion matrix coefficient remain valid for Euler's discretization scheme.

1 Introduction

Consider a stochastic differential equation in $\mathbb{R}^d$,

$$dX_t = F(X_t)\,dt + \sigma dW_t \quad X_0 = x_0 \in \mathbb{R}^d, \tag{1}$$

with a bounded Borel drift $F : \mathbb{R}^d \to \mathbb{R}^d$, non-degenerate constant duffision matrix σ of dimension $d \times d$ and d-dimensional Wiener process W_t. The process (X_t) is a strong solution of the equation (1), which is a homogeneous, strong Markov process and unique in distributions as well as pathwise unique

[*]Omsk branch of S. L. Sobolev Institute of Mathematics, Siberian division of Russian Academy of Sciences, Pevtsova ul. 13, Omsk, 644099 Russia; email: s-klokov @ yandex.ru; the work was supported by the grant of the Foundation of the President of Russia "Scientific School" 3695.2008.1 and the Programme OMN RAS 1.1.

[†]School of Mathematics, University of Leeds, LS2 9JT, Leeds, UK, & Institute of Information Transmission Problems, Moscow, 127994, GSP-4, Russia; email: a.veretennikov @ leeds.ac.uk.

solution (see [12, 23]). Under certain assumptions on the coefficient F, estimates of convergence to equilibrium and β-mixing have been established, in particular, in [8, 20, 22]. The latter estimates may be important in limit theorems and other applications. However, in modelling, often approximation schemes are used,

$$X^h_{(n+1)h} = X^h_{nh} + F(X^h_{nh})h + \xi_{n+1}\sqrt{h}, \qquad X^h_0 = x \in \mathbb{R}^d, \tag{2}$$

where $(\xi_n,\, n \geq 1)$ is a sequence of i.i.d. centered random variables with a common density p, which are *not* necessarily Gaussian. Hence, an important question arises, whether similar *uniform* mixing estimates related to this scheme hold true, independently of the discretization level. Emphasize that in this context “uniform” relates to h, but of course, not to the initial data, likewise for the limiting diffusion, where such estimates are only locally uniform to the initial data. In other words, we would like to be sure that similar mixing bounds hold true for the whole class of the processes defined by the scheme (2) with any h small enough.

In [21, 22, 9], convergence and β-mixing rates have been established for stochastic difference equations of a non-linear auto-regression type,

$$X_{n+1} = g(X_n) + \xi_{n+1}, \qquad X_0 = x_0 \in \mathbb{R}^d. \tag{3}$$

Equations (2) and (3) look similar; in particular, for any $h > 0$ fixed, one could just set $g(x) = x + F(x)h$ and replace ξ_{n+1} in (3) by $\xi_{n+1}\sqrt{h}$, as in (2), so that mixing rates for (2) follow from those for (3). The real question is whether one can get uniform bounds for all h small enough. This question is addressed in this paper. It turns out that for the scheme (2) the answer is positive under rather mild assumptions.

The idea of the method of this paper is (1°) to check recurrence properties, and (2°) verify directly some version of a local mixing condition for the family of Markov chains (cf. [22, 10]), uniformly with respect to $h \leq h_0$ for some $h_0 > 0$. Step (1°) is more or less standard. In step (2°), we use beloved in probability theory Central Limit Theorem and small times, – that is, nh should be small enough, but not tending to zero, – to ensure that the drift term does not spoil local mixing properties too much. Techniques from local theorems, or, in other words, Fourier analysis is used in this step, see [5]. The authors think that this method was not applied earlier in this area and for such purposes.

We consider the case where the process $(X^h_{nh},\, n \geq 0)$ possesses a transition density $p^h_{nh}(x, y)$ given $X^h_0 = x$ with respect to the Lebesgue measure;

notation $(X_t^0,\, t \ge 0)$ is used for solution of the equation (1). Let B be some compact set (usually a closed ball, which will be commented later in due course). The local Markov–Dobrushin's mixing condition may have a few close versions; the one which we are going to check has the following form.

Local Markov–Dobrushin's Condition. There exists some bounded Borel set B of positive Lebesgue measure and values $h_0 > 0$ and $T > 0$ such that for any integer n such that $nh \le T \le (n+1)h$,

$$\inf_{0<h\le h_0} \inf_{x,x'\in B} \int_{\mathbb{R}^d} p^h_{nh}(x,y) \wedge p^h_{nh}(x',y)\, dy =: \rho > 0. \tag{4}$$

Here $a \wedge b = \min(a,b)$. The constant ρ may depend on B, h_0, T, and on the class of processes determined by all assumptions on the equation and on the approximations.

Notice that in some cases it could be useful to have a similar condition for any T greater than some positive value; a statement of such kind is provided in the Corollary 1.

The name of this condition arose from a *global* assumption (i.e. with $B = \mathbb{R}^d$) proposed and used in [3] in the problem of limit theorems for *non-homogeneous* Markov processes, which is now known as Dobrushin's ergodic assumption (or coefficient). For *homogeneous* Markov chains (of course, with a finite state) a similar and also global condition was known long ago in the Ergodic Theorem since A. A. Markov [14], the paper being reprinted in [15]. However, here the name of the condition (4) is proposed for the first time. In coupling, a *construction* based on the same idea – sometimes called Markov's contraction – in the global form was proposed by Vasserstein [19] and, whence, is known in the literature as Vasserstein's coupling.

A little bit different, also local version of the condition (4) was used in mixing problems in [20] and other papers by the same author. In this paper we use a variant that better suits the method.

In the sequel, we will establish the condition (4) for any ball B_R in place of B. In [10], the Malliavin calculus was used to check this condition; naturally, it required some regularity of both the function F and of the density p. Our approach in this paper is different and uses practically no regularity conditions at all. The main assumptions are just that the function F is bounded, and that the density p is strictly positive everywhere and bounded away from zero *on some open set*, possesses a finite third moment, and, for example, is bounded. This, of course, covers a rather general class of densities. Under a stronger assumption that the process has a *locally bounded and* *locally bounded away from zero* transition density there is an easier technique based on "small sets" condition and local version of "condition C" by Doob [4] (apparently proposed by Doeblin), even though the latter technique may provide a slower rate of convergence than ours. However, our assumptions below have another advantage because they are clearly more relaxed. The authors believe that a method that requires weaker assumptions and provides better estimates should be preferred. Also, the assumptions in this paper are considerably weaker than those in [10], although, the potential of this method in the variable diffusion case (not considered in this paper either) is yet unclear. In compare to the famous Doeblin–Doob's condition from [4], our condition (4) looks more suitable because the latter provides a quantitative bound for the rate of convergence, unlike the former one. In any case, Doeblin–Doob's condition does not allow to control any constant in the rate of convergence, nor it defines any class of processes where such rate could be uniform on that class. Finally, although we do not claim here that our estimates below are optimal, the results in [7] and [24] prompt that for a wide class of processes our estimates in a certain sense have nearly optimal order. This provides some reason to drop the question of how our estimates relate to *maximal coupling*, which can be constructed for certain classes of processes. In any case, the authors are not aware of any precise bounds obtained via the latter technique, although apparently such estimates if they existed should have been the best.

The referee recommended to discuss how our method relates to other well-known coupling methods presented, in particular, in [1] or in [16]. The authors are not aware of applications of other methods to the setting under consideration. In principle, *if* one manages to construct a "small set" — that is, a set with respect to which an appropriate recurrence holds true and on which all transition measures are all equivalent with uniformly bounded derivatives with respect to each other and, more than that, both recurrence and the supremum of those derivatives are uniform with respect to the discretization step — then the main results of this paper could have been obtained as a corollary, possibly, with some other constants. The question is, however, exactly about how establish the required properties without assuming stronger conditions. In principle, it follows from a combination of Doob's method in [4] with localization and recurrence that such set exists, however, without any indication about exact value of time when those derivatives become, indeed, bounded. Doob's methods relates upon existence of Lebesgue's derivative of measure, which, apparently, does not admit any effective description, even in the situation of uniform ergodicity. Actually, construction of a small set is always based on bounded derivatives. However, clearly, not all derivatives may be bounded. The method used in this paper does not use this boundedness. Moreover, Nummelin's approach is often correctly called generalized regeneration. Coupling used in this paper is not a generalized regeration and, as it was mentioned above, provides better estimates even if the former were possible.

Briefly, relationship of the methods may be formulated as follows. Generalized regeneration is technically easier, however, it requires some special condition – let us call it local bounded deriative for simplicity – and often does not provide the best estimates. If boundedness of derivatives is not assumed nor established efficiently, then some other method is needed. This is precisely a situation considered in the present paper.

To make the idea of the approach more explicit, we state and prove main results firstly for a simple case $\mathbb{R} = \mathbb{R}^1$ and normally distributed (ξ_n), and then repeat it for the general case. In the latter, we use essentially local theorems techniques based on the Fourier transform, and uniform version of Prokhorov's theorem for densities, due to Statulevicius, Lapinskas and Shervashidze. However, notice that the bound in the Gaussian case, with the Laplace function, is, of course, more precise.

It would be natural to pose a similar question for SDEs with variable diffusions, with an idea to use local theorem approximations such as established in [11]. We do not pursue this goal here, because this technique would inevitably require *some regularity* from the drift, not speaking of diffusion. In this paper we aim to explore *least possible conditions on regularity of the drift*, which seems to be doable for a constant diffusion coefficient. Emphasize that the absence of smoothness condition apparently does not allow application of any technique based on regularity, such as, for example, Malliavin's calculus nor even the method of parametrix. The authors are planning to consider the case of variable diffusion in the future.

The paper is organized as follows. We formulate the main results in the Section 2, along with some corollaries. The latter relate to mixing and convergence rates to equilibrium distributions, and follow easily from the combination of the Theorem 2 and the techniques and results from [20, 22, 8]. The section 3 contains auxiliary lemmae. The proofs of the main theorems and their corollaries and applications are given in the Sections 4-6.

2 Main results and their applications

Assumptions

(A1) The function F is bounded,

$$|F| \le C_1. \tag{5}$$

(A2) There exist an open nonempty set U such that the density p satisfies the following condition:

$$\inf_{x\in U} p(x) \ge c_U = C_2 > 0. \tag{6}$$

Without any loss of generality, we assume in the sequel that U is a ball, $U = B_{r_0}(x_0) \equiv \{x : |x - x_0| < r_0\}$, with some $r_0 > 0$, $x_0 \in \mathbb{R}^d$.

(A3) Random variables $(\xi_n,\, n \ge 1)$ are IID, $\mathbb{E}\xi_1 = 0$, the symmetric covariance matrix $\mathrm{Cov}(\xi_1) = V$ is positive definite and there exists $C_3 < \infty$ such that

$$\mathbb{E}|\xi_1|^3 \le C_3. \tag{7}$$

(A4) There is an integer $n_1 \ge 1$ such that the density q_{n_1} of the normalized sum $n_0^{-1/2} \sum_{k=1}^{n_1} \xi_k$ is bounded,

$$\sup_{x\in\mathbb{R}^d} q_{n_1}(x) \le C_4 < \infty. \tag{8}$$

The assumption (A1) will be required in both of the Theorems 1 and 2; the assumptions (A2)–(A4) will be used only in the Theorem 2, because the Theorem 1 uses a Gaussian distribution (for which (A2)–(A4) are, of course satisfied). Notice that a simple sufficient condition for (A4) is just to assume p *bounded* that is, $n_0 = 1$ in (A4). See, e.g., [2, Theorem 19.1] about some statements equivalent to (A4).

Main results

Let us formulate results about local Markov–Dobrushin's condition for solutions of the equation (2).

Theorem 1 *Let $d = 1$, (X^h_{nh}) satisfy* (2)*, $\xi_j \sim \mathcal{N}(0,1)$, and the assumption* (A1) *hold true. Then, for every $K > 0$ small enough, there exists $c_0 > 0$ which depends only on the values K and C_1, such that for all $0 < nh \le K$,*

$$\int_{\mathbb{R}} p^h_{nh}(x_1, y) \wedge p^h_{nh}(x_2, y)\, dy \ge c_0 \left(1 - 2\Phi_0\left(\frac{|x_2 - x_1|}{\sqrt{nh}}\right)\right), \ x_1, x_2 \in \mathbb{R}, \quad (9)$$

where $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-y^2/2}\, dy$. In particular, for any bounded Borel set $B_0 \subset \mathbb{R}^1$,

$$\inf_{x_1, x_2 \in B_0} \int_{\mathbb{R}} p^h_{nh}(x_1, y) \wedge p^h_{nh}(x_2, y)\, dy \ge C \left(1 - 2\Phi_0\left(\frac{\mathrm{diam}(B_0)}{\sqrt{nh}}\right)\right). \quad (10)$$

The bound (10) prompts an interesting possibility to optimize the rate of convergence or/and mixing by varying the set B_0 treated as a “small set”, that is, a set where mixing occurs, while outside this set we only wait until the process hits B_0 again; this is a standard idea for establishing mixing bounds. So, if we, say, increase B_0, then the mixing rate on this set may decrease, while the recurrence properties may become better, and vice versa. Even more, the same inequality prompts also that in some models mixing may occur everywhere over the state space, although with different rates depending on the location. This could be, possibly, also used for optimization, but the question is open and we do not touch it here.

Theorem 2 *Let (X^h_{nh}) satisfy* (2)*, assumptions* (A1)–(A4) *hold true, and $K_1 \le nh \le K_2$ with some constants $0 < K_1 < K_2 < \infty$. Then, for every K_2 small enough, there exist $c_0, c_1, c_2, n_1 > 0$ such that for any $x_1, x_2 \in \mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} p^h_{nh}(x_1, y) \wedge p^h_{nh}(x_2, y)\, dy \ge c_0(1 - c_1(|x_2 - x_1| + |x_1 - x_2|^2) - c_2 n^{-1/3}), \quad (11)$$

provided that $n \ge n_1$. All these constants c_0, c_1, c_2 and n_1 here depend only on the values C_1, C_2, C_2, C_3, C_4, n_0, and on the set U from the assumptions (A1)–(A4).

Corollary 1 *Under the assumptions of the Theorem* 2*, for* any *$T > 0$ there exists $h_0 > 0$ such that the local Markov–Dobrushin's condition* (4) *holds true.*

Remark. Notice that without conditions that provide uniform recurrence we do not claim that the inequality (4) is uniform with respect to T, nor, say, with respect to $T \ge C$ for any $C > 0$, although the latter may be proved under uniform recurrence and certain mild additional assumptions.

Applications

The first application relates to mixing bound under additional assumptions on recurrence and moments. Some examples shall be found below; however, we emphasize that it is the Theorem 2 about verifying local Markov–Dobrushin's condition that is the main result of the paper. Let us also stress out that the assumptions used for establishing the condition (4) are practically not related to recurrence assumptions. Further applications could be limit theorems, cf. [5]. The Theorem 3 may also be useful in filtering problems, cf. [6].

Let us remind the definition of the β-mixing coefficient for the process X_t in continuous time,

$$\beta_{t,x} = \sup_{s\ge 0} \mathbb{E}_x \operatorname*{var}_{\mathcal{F}^X_{\ge t+s}} \Big(\mathbb{P}_x(B \mid \mathcal{F}^X_{\le s}) - \mathbb{P}_x(B)\Big),$$

where $\mathcal{F}^X_I = \sigma\{X_s : s \in I\}$ and $\mathbb{E}_x$ is expectation given the starting point x, and var is the total variation distance between measures. In the same form the definition above is applied to the discrete version X^h of the process; in such a case notation $\beta^h_{t,x}$ will be used with $t = nh$, $n = 0, 1, 2, \ldots$ Let $\beta_{t,x} \equiv \beta^0_{t,x}$.

Let us remind for better comparison some convergence and mixing bounds established earlier for discrete and continuous time in earlier papers by the authors. Our next aim will be to give similar bounds uniform with respect to h when h is small enough. In the next Proposition, notation $\langle\ \rangle$ stands for the standard inner product in $\mathbb{R}^d$. The value of h in this Proposition is *fixed*, unlike in all previous and next results of the paper.

Proposition 1 ([8, 20, 22]) *Assume that the (discrete) assumption* (4) *holds true and there exist $p \in [0,1]$ and $r > 0$ such that for all x large enough*

$$\langle F(x), x/|x|\rangle \le -r/|x|^p, \quad 0 \le p \le 1,\ r > 0.$$

Then, for each $h \ge 0$, in each of the three cases below, the distributions $\mu^h_{t,x} = \mathcal{L}(X^h_{nh} \mid X^h_0 = x)$ converge to a unique invariant measure μ^h_{inv} in the topology of the total variation norm as $t \to \infty$; the variable t takes values nh, $n = 0, 1, \ldots$ Moreover, the following bounds hold true, with the same convention about time t:

1. If $p = 0$ and there exists $\varepsilon > 0$ such that $\mathbb{E}\exp(\varepsilon|\xi_1|) < \infty$, then for each $c_1 > 0$ small enough there exist $c_0, c_2 > 0$ such that

$$\mathrm{var}(\mu^h_{t,x} - \mu^h_{\mathrm{inv}}) + \beta^h_{t,x} \le c_0 \exp\{c_1|x| - c_2 t\}.$$

2. If $0 < p < 1$ and there exist $\alpha \ge 1 - p$ and $K > 0$ such that for every $0 \le \kappa < K$ the following moment is finite, $\mathbb{E}\exp(\kappa|\xi_1|^\alpha) < \infty$, then for any *$c_1 > 0$ small enough there exist $c_0, c_2 > 0$ such that*

$$\mathrm{var}(\mu^h_{t,x} - \mu^h_{\mathrm{inv}}) + \beta^h_{t,x} \le c_0 \exp\{c_1|x|^{1-p} - c_2 t^{\frac{1-p}{1+p}}\}.$$

3. If $p = 1$, $m > 4$, $\mathbb{E}|\xi_1|^m < \infty$ and $r > r_m := \frac{m-1}{2}\mathbb{E}|\xi_1|^2$, then there exists $c_0 > 0$ such that for any $0 < k < m/2$,

$$\mathrm{var}(\mu^h_{t,x} - \mu^h_{\mathrm{inv}}) + \beta^h_{t,x} \le c_0(1 + |x|^m)(1 + t)^{-k}.$$

Remind that a similar result holds true for a "limiting" process in continuous time, with $h = 0$, see [20, 22]. Let us underline that, generally speaking, all constants in the Proposition 1 may depend on h, while our primary goal in this paper is to obtain certain bounds with no dependence of this sort.

For the next application let us remind some recurrence assertions uniform with respect to h small enough. Let

$$\tau := \inf(t \ge 0 : \ |X^h_t| \le R).$$

Lemma 1 *Let*

$$\langle F(x), x\rangle \le -r|x|^{1-p}, \qquad |x| > R_0, \tag{12}$$

with some $p \in [0, 1]$. Then the following estimates hold true.

1°. *If $p = 0$ and there exists $K > 0$ such that $\mathbb{E}\exp(\kappa|\xi_1|) < \infty$ for any $0 < \kappa < K$, then for every $0 < \kappa < K$ there exist $h_0 > 0$ and $C > 0$ such that*

$$\sup_{h\le h_0}\sup_{t\ge 0} \mathbb{E}_x \exp(\kappa|X^h_t|) \le C\exp(\kappa|x|) < \infty,$$

and for any R large enough and $\alpha > 0$ small enough,

$$\sup_{h\le h_0} \mathbb{E}_x \exp(\alpha\tau) \le C\exp(\kappa|x|).$$

2°. *If* $0 < p < 1$ *and* $\mathbb{E}\exp\{\kappa|\xi_1|^\alpha\} < \infty$ *with some* $(0 < \alpha \le 1 - p$ *and any* $0 < \kappa < K$, *then there exist* $h_0 > 0$ *and* $C > 0$ *such that*

$$\sup_{h \le h_0} \sup_{t \ge 0} \mathbb{E}_x \exp(\kappa |X_t^h|^\alpha) 1(t \le \tau) \le C \exp(\kappa |x|^\alpha),$$

and for any R *large enough and every* $0 < \delta < \alpha/(1+p)$,

$$\sup_{h \le h_0} \mathbb{E}_x \exp(\tau^\delta) \le C \exp(\kappa |x|^\alpha).$$

3°. *If* $p = 1$, $\mathbb{E}|\xi_1|^m < \infty$, $m > 4$, *and* $r > \frac{m-1}{2}\mathbb{E}|\xi_1|^2$, *then there exist* $C > 0$ *and* $h_0 > 0$ *such that for every* R *large enough,*

$$\sup_{h \le h_0} \sup_{t \ge 0} \mathbb{E}_x |X_t^h|^m 1(t \le \tau) \le C(1 + |x|^m),$$

and for every $k < m/2$,

$$\sup_{h \le h_0} \mathbb{E}_x \tau^{k+1} \le C(1 + |x|^m).$$

The condition with the inequality $0 < \kappa < K$ with some $K > 0$ appeared in [10]. The Lemma 1 follows straight away from the auxiliary result in [10] and [22] and an observation that those estimates – see [10, Lemmae 3, 5, 7] and [22, Theorem 3, Lemma 1] – are uniform with respect to $h \le h_0$. Notice that to verify local Markov–Dobrushin's condition, in [10] certain assumptions have been assumed on regularity of coefficients. The assumptions from the Lemma 1 practically do not require any regularity. A "weakened" condition with an indicator function $\mathbf{1}(t \le \tau)$ in 2° and 3° is nearly forced: this condition is *also* sufficient for establishing rate of convergence (see [22]), and at the same time under 3° its verification is much easier than without such indicator.

Now let us apply the above results from the Theorem 2 and Lemma 1 to mixing and convergence bounds for our approximation scheme uniform with respect to small h.

Theorem 3 *Let the family of processes* (X_{nh}^h) *satisfy* (2). *Suppose that the function* F *is bounded, there exist* $R_0 > 0$, $0 \le p \le 1$ *and* $r > 0$ *such that*

$$\langle F(x), x \rangle \le -r|x|^{1-p}, \quad |x| > R_0,$$

and the sequence of i.i.d. random variables (ξ_n) *satisfy the assumptions of the Theorem* 2. *Then, in each of the cases* 1–3 *below,*

(a) *the family of processes* (X^h_{nh}) *satisfies the local Markov–Dobrushin condition* (4) *with some* $h_0 > 0$;

(b) *for each* $0 < h \le h_0$ *with* h_0 *small enough, there exists a unique invariant measure* $\mu^h_{\rm inv}$; *the measures may be different for different values of* h *but all constants and functions in the following estimates can be chosen uniformly in* $h \le h_0$;

(c) *for every* $0 < h \le h_0$, *the marginal distributions* $\mu_{t,x} = \mu^h_{nh,x} = \mathcal{L}(X^h_{nh} \mid X^h_0 = x)$ *converge to the invariant measure* $\mu^h_{\rm inv}$ *at some specific rate on the scale of* t, *and* β*-mixing holds with the same rate, namely:*

1. *If* $p = 0$ *and* $\mathbb{E}\exp\{\kappa|\xi_1|\} < \infty$, $0 < \kappa < K$, *then*

$$\mathrm{var}(\mu^h_{t,x} - \mu^h_{\rm inv}) + \beta^h_{t,x} \le C(x)\exp\{-c(1+t)\},$$

with some $c > 0$ *and* $C(x) = C\exp(\varepsilon|x|)$, *with some* $\varepsilon > 0$.

2. *If* $0 < p < 1$ *and* $\mathbb{E}\exp\{\kappa|\xi_1|^\alpha\} < \infty$ *holds with* $0 < \alpha \le 1 - p$ *and* $0 < \kappa < K$, *then*

$$\mathrm{var}(\mu^h_{t,x} - \mu^h_{\rm inv}) + \beta^h_{t,x} \le C(x)\exp\{-c(1+t)^\delta\},$$

with any $0 < \delta < \frac{\alpha}{1+p}$ *and some constant* $c > 0$ *and a positive function* $C(x)$ *depending on* δ.

3. *If* $p = 1$ *and* $\mathbb{E}|\xi_1|^m < \infty$, $m > 4$, *and* $r > \frac{m-1}{2}\mathbb{E}|\xi_1|^2$, *then*

$$\mathrm{var}(\mu^h_{t,x} - \mu^h_{\rm inv}) + \beta^h_{t,x} \le C(1+|x|^m)(1+t)^{-k},$$

where $C > 0$ *and any* $k < \frac{m}{2}$ *can be used.*

Remark 1 Notice that integration in the condition (4) is over the whole space $\mathbb{R}^d$, while in earlier works [20, 22, 8] such integration was over the set B, i.e.,

$$\inf_{0<h\le h_0}\inf_{x,x'\in B}\int_B p^h_{nh}(x,y)\wedge p^h_{nh}(x',y)\,dy > 0.$$

This difference is not significant and does not affect the proof of the Proposition 1 nor of the Theorem 3, given the moment assumptions. In fact, in all papers cited above, the set of integration could have been replaced by some other set B', bounded or not,

$$\inf_{0<h\le h_0}\inf_{x,x'\in B}\int_{B'} p^h_{nh}(x,y)\wedge p^h_{nh}(x',y)\,dy>0.$$

At the same time, for $B=B_R\equiv(x:\ |x|\le R)$ the latter inequality with $B'=B_{R'}$ and some appropriate $R'<\infty$ follows from (4) by virtue of the assumptions (A1) and (A3). It could be said that it is entirely a convenience of the calculus that determines the choice of the domain of integration in local Markov–Dobrushin's condition.

3 Auxiliary lemmae

Let $p^{(1)}$ and $p^{(2)}$ be two densities. The statements of the Theorems 1 and 2 deal with lower estimates of integrals like $\int p^{(1)}(y)\wedge p^{(2)}(y)\,dy$. We will also work with a "discrete" analogue of this "integral value". Let $P^{(1)}$ and $P^{(2)}$ be two probability measures and (A_j) any finite or countable family of disjoint sets, $A_i\cap A_j=\varnothing$, $i\ne j$. The "discrete" analogue of the integral above is defined as $\sum_j P^{(1)}(A_j)\wedge P^{(2)}(A_j)$. In the Lemmae 2 and 3 we establish lower estimates of the "integral value" via the "discrete" one and vice versa.

Lemma 2 *Let $P^{(1)}$, $P^{(2)}$ and Q be three probability measures on R^d, $q(x)$ the density of Q. Then the convolutions $P^{(1)}*Q$ and $P^{(2)}*Q$ have densities $p^{(1)}$ and $p^{(2)}$, and for every finite or countable family (A_j) of disjoint sets,*

$$\int_{\mathbb{R}^d} p^{(1)}(y)\wedge p^{(2)}(y)\,dy\ge\sum_j\left(P^{(1)}(A_j)\wedge P^{(2)}(A_j)\right)\int_{\mathbb{R}^d}\inf_{x\in A_j}q(y-x)\,dy.$$

Proof Since $p^{(i)}(y)=\int_{\mathbb{R}^d}q(y-x)P^{(i)}(dx)$ and

$$\int_{A_j}q(y-x)P^{(i)}(dx)\ge P^{(i)}(A_j)\inf_{x\in A_j}q(y-x),$$

then

$$\int_{\mathbb{R}^d} p^{(1)}(y)\wedge p^{(2)}(y)\,dy$$

$$
\begin{aligned}
&\geq \int_{\mathbb{R}^d} \left[\sum_j P^{(1)}(A_j) \inf_{x \in A_j} q(y-x) \right] \wedge \left[\sum_j P^{(2)}(A_j) \inf_{x \in A_j} q(y-x) \right] dy \\
&\geq \int_{\mathbb{R}^d} \sum_j (P^{(1)}(A_j) \wedge P^{(2)}(A_j)) \inf_{x \in A_j} q(y-x)\, dy \\
&= \sum_j \left(P^{(1)}(A_j) \wedge P^{(2)}(A_j) \right) \int_{\mathbb{R}^d} \inf_{x \in A_j} q(y-x)\, dy. \qquad \bullet
\end{aligned}
$$

Lemma 3 *Let $Z^{(1)}$ and $Z^{(2)}$ be two d-dimensional random variables with densities $p^{(1)}$ and $p^{(2)}$, $Y^{(1)}$ and $Y^{(2)}$ two d-dimensional bounded random variables, i.e. $\mathbb{P}(|Y^{(i)}| < m/2) = 1$ for some $m > 0$ and $i = 1, 2$. Denote by $P^{(1)}$ and $P^{(2)}$ the distributions of $Z^{(1)} + Y^{(1)}$ and $Z^{(2)} + Y^{(2)}$ respectively. Then for every positive integer k there exists a countable family (A_j) of disjoint sets with diameters at most $km\sqrt{d}$ such that*

$$
\sum_j P^{(1)}(A_j) \wedge P^{(2)}(A_j) \geq \left(1 - \frac{1}{k}\right)^d \int_{\mathbb{R}^d} p^{(1)}(y) \wedge p^{(2)}(y)\, dy.
$$

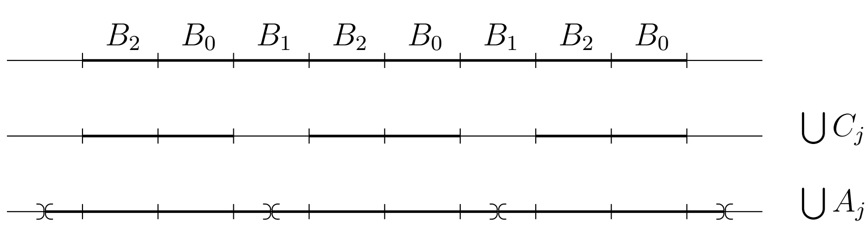
Proof Firstly, we prove the lemma for the case $d = 1$ for the sake of simplicity and in order to present the idea. Consider open intervals

$$
I_i = \{x \in \mathbb{R} : \; im < x < (i+1)m\}, \quad i \in \mathbb{Z},
$$

and define k sets,

$$
B_l = \bigcup_{j \in \mathbb{Z}} I_{jk+l}, \quad l \in \{0, 1, \ldots, k-1\}
$$

as it is shown on the following picture for $k = 3$.



The sets are disjoint and cover $\mathbb{R}$ up to a set of zero Lebesgue measure, so

$$\int_{\mathbb{R}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy = \sum_{l=0}^{k-1} \int_{B_l} p^{(1)}(y) \wedge p^{(2)}(y)\, dy.$$

Find an index l^* such that

$$\int_{B_{l^*}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy \le \frac{1}{k} \int_{\mathbb{R}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy.$$

The set $\mathbb{R} \setminus B_{l^*}$ is a countable union of closed intervals

$$C_j = \{x \in \mathbb{R} :\ (jk + l^* + 1)m \le x \le ((j+1)k + l^*)m\}, \quad j \in \mathbb{Z}.$$

On the picture $l^* = 1$. Now consider sets A_j constructed by "expanding" the sets C_j,

$$A_j = \left\{x \in \mathbb{R} :\ -\frac{m}{2} + (jk + l^* + 1)m < x < ((j+1)k + l^*)m + \frac{m}{2}\right\}, \quad j \in \mathbb{Z}.$$

Since $|Y^{(1)}| < m/2$ and $|Y^{(2)}| < m/2$ with probability 1, then $\{Z^{(i)} \in C_j\} \subset \{Z^{(i)} + Y^{(i)} \in A_j\}$ and

$$P^{(1)}(A_j) \wedge P^{(2)}(A_j) \ge \int_{C_j} p^{(1)}(y) \wedge p^{(2)}(y)\, dy.$$

Therefore,

$$\begin{aligned}
\sum_j P^{(1)}(A_j) \wedge P^{(2)}(A_j) \ &\ge\ \int_{\mathbb{R}\setminus B_{l^*}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy \\
&\ge\ \left(1 - \frac{1}{k}\right) \int_{\mathbb{R}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy.
\end{aligned}$$

Consider the case $d > 1$. Let us start with open sets

$$I_i^1 = \{(x_1, \dots, x_d) \in \mathbb{R}^d :\ im < x_1 < (i+1)m\}, \quad i \in \mathbb{Z},$$

define $B_l^1 = \bigcup_{j\in\mathbb{Z}} I_{jk+l}^1$, $l \in \{0, 1, \dots, k-1\}$, and fix such index l_1^* that

$$\int_{\mathbb{R}^d \setminus B^1_{l_1^*}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy \ge \left(1 - \frac{1}{k}\right) \int_{\mathbb{R}^d} p^{(1)}(y) \wedge p^{(2)}(y)\, dy.$$

Next, consider the sets

$$I_i^2 = \{(x_1, \dots, x_d) \in \mathbb{R}^d \setminus B_{l_1^*} : \ im < x_2 < (i+1)m\}, \quad i \in \mathbb{Z},$$

define $B_l^2 = \bigcup_{j\in\mathbb{Z}} I_{jk+l}^2, \quad l \in \{0, 1, \dots, k-1\}$, and fix such index l_2^* that

$$\int_{\mathbb{R}^d \setminus B_{l_1^*}^1 \setminus B_{l_2^*}^2} p^{(1)}(y) \wedge p^{(2)}(y)\, dy \geq \left(1 - \frac{1}{k}\right)^2 \int_{\mathbb{R}^d} p^{(1)}(y) \wedge p^{(2)}(y)\, dy.$$

Repeat this procedure in the same way $d-2$ times and get a countable union of d-dimensional cubes,

$$\begin{aligned} C_{j_1,\dots,j_d} = \{x \in \mathbb{R}^d \ : \ & (j_1 k + l_1^*)m \leq x_1 \leq ((j_1+1)k + l_1^*)m, \\ & \dots\dots\dots\dots\dots\dots\dots\dots \\ & (j_d k + l_d^*)m \leq x_d \leq ((j_d+1)k + l_d^*)m\}, \ j_1, \dots, j_d \in \mathbb{Z}. \end{aligned}$$

The sets $A_{j_1,\dots,j_d}$ are constructed by "expanding" each cube $C_{j_1,\dots,j_d}$ as it was done for $d = 1$, and the rest of the proof follows the same lines as well. •

Lemma 4 *If* $q(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\{-\dfrac{x^2}{2\sigma^2}\}$ *and* A *is any set of diameter at most* $2a$*, then*

$$\int_{\mathbb{R}} \inf_{x\in A} q(y-x)\, dy \ \ \geq \ \ 1 - 2\Phi_0(a/\sigma) \geq 1 - ca/\sigma, \tag{13}$$

where $\Phi_0(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^x e^{-y^2/2}\, dy$ *and* $c = \sqrt{2/\pi}$*.* •

Proof We have,

$$\begin{aligned} \int_{\mathbb{R}} \inf_{x\in A} q(y-x)\, dy \ &\geq \ \int_{\mathbb{R}} \inf_{|x|\leq a} q(y-x)\, dy = 2\int_0^\infty q(y+a)\, dy \\ &= \ 2\int_a^\infty q(y)\, dy = 1 - 2\Phi_0(a/\sigma). \end{aligned}$$

The second inequality follows due to the estimate $\Phi_0(x) \leq \dfrac{x}{\sqrt{2\pi}}$. •

Lemma 5 *Let* (ξ_n) *be a sequence of* d*-dimensional i.i.d. random variables such that* $\mathbb{E}\xi_n = 0$, $\mathbb{E}|\xi_n|^3 \le C < \infty$, *and such that normalized sums* $s_n = n^{-1/2}\sum_{j=1}^n \xi_j$ *have densities* $q_n(x)$ *bounded for some* $n = n_1$, *and characteristic functions* $f_n(t)$. *Then there exist positive constants* c_1, c_2, *and an integer* n_0 *such that*

$$\int \inf_{|x|\le a} q_n(y-x)\,dy \ge 1 - c_1(a + a^2/2) - c_2 n^{-1/3}, \quad n > n_0. \tag{14}$$

Proof Everywhere in the proof we use the symbol c as a generic positive constant which does not depend on n nor a, and may change from one occurrence to another.

Notice that due to the assumptions, there exists n_1 such that for every $n > n_1$, the function $f_n \in L^1(\mathbb{R}^d)$ and the inversion formula

$$q_n(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} f_n(t)e^{-i\langle t,x\rangle}\,dt$$

holds true (see [2, Theorem 19.1]).

Denote by $\varphi(x)$ the density and by $g(t)$ the characteristic function of the normal distribution with zero mean and covariance matrix $\sigma\sigma^*$,

$$\begin{aligned}
\varphi(x) &= (2\pi\det(\sigma\sigma^*))^{-1/2}\exp\{-|\sigma^{-1}x|^2/2\},\\
g(t) &= \exp\{-|\sigma\, t|^2/2\}.
\end{aligned}$$

For each $N > 0$ we get,

$$\begin{aligned}
\int \inf_{|x|\le a} q_n(y-x)\,dy &\ge \int_{|y|\le N}\inf_{|x|\le a} q_n(y-x)\,dy\\
&\ge \int_{|y|\le N}\inf_{|x|\le a}\Big\{\frac{1}{(2\pi)^d}\int e^{-i\langle t,y-x\rangle}(f_n(t)-g(t))\,dt\Big\}dy\\
&\quad + \int_{|y|\le N}\inf_{|x|\le a}\Big\{\frac{1}{(2\pi)^d}\int e^{-i\langle t,y-x\rangle}g(t)\,dt\Big\}dy\\
&\ge -\int_{|y|\le N}\Big\{\frac{1}{(2\pi)^d}\int |f_n(t)-g(t)|\,dt\Big\}dy\\
&\quad + \int_{|y|\le N}\inf_{|x|\le a}\varphi(y-x)\,dy\\
&\ge -cN^d\int |f_n(t)-g(t)|\,dt\\
&\quad + \int_{|y|\le N}\inf_{|x|\le a}\varphi(y-x)\,dy.
\end{aligned} \tag{15}$$

To estimate the first term in the right hand side of (15) we apply a standard approach for establishing local limit theorems (see, e.g., [5]). Let us fix a small $\varepsilon > 0$, and consider the following four integrals,

$$\begin{aligned} I_1 &= \int_{|t|\le n^{1/6}} |f_n(t) - g(t)|\, dt, \\ I_2 &= \int_{n^{1/6}<|t|\le \varepsilon\sqrt{n}} |f_n(t)|\, dt, \\ I_3 &= \int_{|t|>\varepsilon\sqrt{n}} |f_n(t)|\, dt, \\ I_4 &= \int_{|t|>n^{1/6}} |g(t)|\, dt. \end{aligned}$$

Via Taylor's expansion for the function $f_1(t/\sqrt{n})$ in the region $|t|/\sqrt{n} \le \varepsilon$, we obtain,

$$\begin{aligned} f_n(t) &= f_1^n(t/\sqrt{n}) = \left(1 - \frac{\langle t, \sigma\sigma^* t\rangle}{2n} + O\left(\frac{|t|^3}{n^{3/2}}\right)\right)^n \\ &= \exp\left\{-\frac{\langle t, \sigma\sigma^* t\rangle}{2} + O\left(\frac{|t|^3}{n^{1/2}}\right)\right\} \\ &= g(t)\exp\left\{O\left(\frac{|t|^3}{n^{1/2}}\right)\right\}, \quad n\to\infty. \end{aligned}$$

If $|z| \le M < \infty$, then $|e^z - 1| \le c|z|$ with some $0 < c < \infty$. So, for $|t|^3 \le n^{1/2}$ we have the estimate,

$$I_1 \le \int_{|t|\le n^{1/6}} g(t)\frac{c|t|^3}{n^{1/2}}\, dt \le \frac{c}{n^{1/2}}\int_{\mathbb{R}^d} |t|^3 g(t)\, dt \le \frac{c}{n^{1/2}}. \tag{16}$$

For $n^{1/6} < |t| \le \varepsilon\sqrt{n}$ we get,

$$|f_n(t)| \le \exp\left\{-\frac{\langle t, \sigma\sigma^* t\rangle + \varepsilon O(|t|^2)}{2}\right\} \le \exp\left\{-\frac{\langle t, \sigma\sigma^* t\rangle}{4}\right\},$$

provided $\varepsilon > 0$ is small enough. Here and later the following elementary inequality is used: for any $k > 0$ and $z_0 > 0$ there exists $K > 0$ such that for all $z \ge z_0$

$$\int_{|t|\ge z} \exp\{-k|t|^2\}\, dt \le \exp\{-Kz^2\}. \tag{17}$$

Therefore, by (17)

$$I_2 \le \int_{|t|>n^{1/6}} \exp\left\{-\frac{\langle t,\sigma\sigma^* t\rangle}{4}\right\} dt \le \exp\{-cn^{1/3}\}. \tag{18}$$

The random variables (ξ_n) possess a density, therefore,

$$\sup_{|t|\ge\varepsilon} |f_1(t)| < 1, \tag{19}$$

and

$$I_3 = n^{d/2}\int_{|t|\ge\varepsilon} |f_1(t)|^n\,dt \le n^{d/2}\sup_{|t|\ge\varepsilon}|f_1(t)|^{n-n_0}\int_{\mathbb{R}^d}|f_{n_0}(t)|\,dt \le ce^{-cn}.$$

However, here the right hand side may depend on the distribution of ξ_i, just because the inequality (19) does depend on it, while our aim is a bound uniform over the whole class of distributions described in the assumptions. Hence, we will used the improved exponential bounds by R. Lapinskas [13] and T. Shervashidze [17], which, in turn, extend the one-dimensional estimate due to V. A. Statulevicius [18] and which do not depend on the distribution mentioned above. For the reader's convenience we state here the inequalities from [13, 17] for the characteristic function $f^{n_1}(t)$, $t\in R^d$, where the value n_1 is chosen so that the density p_{n_1} is bounded by some constant C_4 (see (A4)). We use a notation $s := \mathbb{E}|S_{m_0} - \mathbb{E}S_{m_0}|^2$. Then, the inequality from [13] reads,

$$|f^{n_1}(t)| \le 1 - \frac{|t|^2}{96}[(4d)^{d-1}s^{2(d-1)}(\sqrt{d}s|t|+\pi)^2A^2]^{-1}, \tag{20}$$

while the inequality (1.2) from [17] reads,

$$|f^{n_1}(t)| \le 1 - |t|^2[24\times 2^{2d}V_{d-1}^2A^2(2s|t|+\pi)^2s^{2(d-1)}]^{-1}, \tag{21}$$

where $V_{d-1} := \pi^{(d-1)/2}\Gamma(1+(d-1)/2)$. As noticed in [17], the additional assumption of non-degeneracy of the covariance matrix in [13] can be dropped because it is always satisfied.

From either of (21) or (20), we get,

$$\sup_{|t|\ge\varepsilon}|f^{n_1}(t)| := q_\varepsilon < 1, \tag{22}$$

where the value q_ε depends only on s and C_4.

We are now in a position to finish the proof of the Lemma. From (A4) and the Plancherel theorem we get,

$$\int_{\mathbb{R}^d} |f_{n_1}(t)|^2\, dt = (2\pi)^d \int_{\mathbb{R}^d} |q_{n_1}(x)|^2\, dx \le (2\pi)^d C_4. \tag{23}$$

From (22) and (23), for $n \ge 3n_1$, we estimate,

$$\begin{aligned} I_3 &= n^{d/2} \int_{|t|>\varepsilon} |f_1^{n_1}(t)|^{[n/n_1]-2} |f_1^{2n_1}(t)|\, dt \\ &\le \Big(\sup_{|t|>\varepsilon} |f_1^{n_1}(t)| \Big)^{[n/n_1]-2} \int_{|t|>\varepsilon} |f_{n_1}(t)|^2\, dt \le (2\pi)^d C_4 e^{-cn}, \end{aligned} \tag{24}$$

where $[\cdot]$ stands for the integer part.

The integral I_4 is estimated in the same way as I_2 (see (18)),

$$I_4 \le \exp\{-cn^{1/3}\}. \tag{25}$$

The estimates (16), (18)–(25) altogether imply

$$\int_{\mathbb{R}^d} |f_n(t) - g(t)|\, dt \le \frac{c}{n^{1/2}}. \tag{26}$$

Let us estimate the last term in the right hand side of (15). For the standard Gaussian distribution,

$$\inf_{|x|\le a} \varphi(y-x) \ge \varphi(y) e^{-ca|y| - ca^2/2}.$$

Using the inequality $e^{-z} - 1 \ge -z$, $z \in \mathbb{R}$, and the inequality (17), we get,

$$\begin{aligned} &\int_{|y|\le N} \inf_{|x|\le a} \varphi(y-x)\, dy \\ &\qquad \ge \int_{|y|\le N} \varphi(y) e^{-ca|y| - ca^2/2}\, dy \\ &\qquad = 1 - \int_{|y|>N} \varphi(y)\, dy + \int_{|y|\le N} \varphi(y)(e^{-ca|y|-ca^2/2} - 1)\, dy \\ &\qquad \ge 1 - \exp\{-cN^2\} - ca \int_{\mathbb{R}^d} (|y| + a/2) \varphi(y)\, dy. \end{aligned}$$

This estimate together with (26) and (15) shows that

$$\int_{\mathbb{R}^d} \inf_{|x|\le a} q_n(y-x)\,dy \ge 1 - c_1(a + a^2/2) - \exp\{-cN^2\} - cN^d n^{-1/2}.$$

Taking $\varepsilon > 0$ small enough and $N = \varepsilon\sqrt{\ln n}$, we arrive at (14). •

Lemma 6 *Let $b \ge 2$ and $n > b+1$ are two integers. Let*

$$\begin{aligned} r &= \max\{k \in \mathbb{Z} : (b^k - 1)/(b-1) \le n\}, \\ k_j &= 1 + [(5/4)^j], \quad j = 1, \ldots, r, \\ z_j &= b^{r-j}, \quad j = 2, \ldots, r, \\ z_1 &= n - z_2 - \ldots - z_r. \end{aligned}$$

Then

$$\sum_{j=1}^{r} z_j = n, \tag{27}$$

$$\sum_{j=1}^{r-1} \frac{k_j z_j}{\sqrt{z_{j+1}}} \le c'\sqrt{n}, \tag{28}$$

$$\sum_{j=1}^{r-1} \frac{k_j^2 z_j^2}{z_{j+1}} \le c'' n, \tag{29}$$

$$\sum_{j=1}^{r-1} \frac{1}{k_j} \le 4, \tag{30}$$

with some positive constants c', c'' depending only on the value of b.

Proof Since

$$\sum_{j=2}^{r} b^{r-j} = 1 + b + \ldots + b^{r-2} = \frac{b^{r-1} - 1}{b-1} < n,$$

relation (27) follows immediately by definitions of r and z_1. Inequality (30) holds due to the estimate,

$$\sum_{j=1}^{r-1} \frac{1}{k_j} \le \sum_{j=1}^{\infty} \left(\frac{4}{5}\right)^j = 4.$$

For $j \geq 2$,

$$\frac{k_j z_j}{\sqrt{z_{j+1}}} \leq \frac{2 \cdot \left(\frac{5}{4}\right)^j \cdot b^{r-j}}{b^{(r-j-1)/2}} = 2\sqrt{b} \cdot b^{r/2} \cdot \left(\frac{5}{4\sqrt{b}}\right)^j. \tag{31}$$

If $j = 1$, then

$$\frac{k_1 z_1}{\sqrt{z_2}} \leq \frac{2n}{\sqrt{b^{r-2}}} = \frac{2bn}{b^{r/2}}. \tag{32}$$

By definition of r

$$\sqrt{n/2} \leq b^{r/2} \leq \sqrt{bn}, \tag{33}$$

therefore

$$\begin{aligned} \sum_{j=1}^{r-1} \frac{k_j z_j}{\sqrt{z_{j+1}}} &\leq \frac{2bn}{\sqrt{n/2}} + 2\sqrt{b} \cdot \sqrt{bn} \sum_{j=2}^{\infty} \left(\frac{5}{4\sqrt{b}}\right)^j \\ &= \left(2\sqrt{2}b + \frac{25\sqrt{b}}{8\sqrt{b} - 10}\right)\sqrt{n}. \end{aligned}$$

The inequality (28) is proved. Similarly, the estimate

$$\sum_{j=1}^{r-1} \frac{k_j^2 z_j^2}{z_{j+1}} \leq 8b^2 n + 4b^2 n \sum_{j=2}^{\infty} \left(\frac{25}{16b}\right)^j = \left(8b^2 + \frac{2500b}{256b - 400}\right) n$$

proves the inequality (29). •

4 Proof of Theorem 1

Proof The idea is based on the following simple observation: one term $\xi_n\sqrt{h}$ in (2) is more influential than several terms like $F(X^h_{nh})h$ as long as the value of h is small enough, because of the difference in scales ($h = o(\sqrt{h})$ as $h \to 0$). The sums $\xi_{k+1}\sqrt{h} + \ldots + \xi_{k+l}\sqrt{h}$ have Gaussian densities which are easy to calculate. Adding "drift terms" $F(X^h_{(k+1)h})h + \ldots + F(X^h_{(k+l)h})h$ does not change the distributions too much.

Let $b = 2$ and $n > 3$ be arbitrary. By the Lemma 6 compute integers $k_1, \ldots, k_r$ and $z_1, \ldots, z_r$. Consider $r - 1$ random variables

$$U_1 = \sum_{i=0}^{z_1} F(X^h_{ih})h,$$

$$
\begin{aligned}
U_2 &= \sum_{i=z_1+1}^{z_1+z_2} F(X_{ih}^h)h, \\
\cdots &\cdots \cdots\cdots\cdots\cdots\cdots \\
U_{r-1} &= \sum_{i=z_1+\ldots+z_{r-2}+1}^{z_1+\ldots+z_{r-1}} F(X_{ih}^h)h,
\end{aligned}
$$

and r random variables

$$
\begin{aligned}
V_1 &= \sum_{i=1}^{z_1} \xi_i \sqrt{h}, \\
V_2 &= \sum_{i=z_1+1}^{z_1+z_2} \xi_i \sqrt{h}, \\
\cdots &\cdots \cdots\cdots\cdots\cdots\cdots \\
V_r &= \sum_{i=z_1+\ldots+z_{r-1}+1}^{z_1+\ldots+z_r} \xi_i \sqrt{h} = \xi_n \sqrt{h}.
\end{aligned}
$$

It follows from (2) that

$$
X_{nh}^h = x + V_1 + U_1 + V_2 + \ldots + V_{r-1} + U_{r-1} + V_r,
$$

where $|U_j| \le C_1(z_j+1)h < 2C_1 z_j h$ with probability 1, and $V_j \sim \mathcal{N}(0, z_j h)$.

In the Lemma 2 take $P^{(1)} = \delta_x$, $P^{(2)} = \delta_{x'}$, where δ_x denotes the distribution with unit mass concentrated at point x, A_1 the closed interval with endpoints x and x', and let Q be the distribution of V_1 with Gaussian density q. By the virtue of the Lemma 4 where $2a = |x - x'|$ and $\sigma^2 = z_1 h$ we get the estimate

$$
\int_{\mathbb{R}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy \ge \int_{\mathbb{R}} \inf_{x \in A_1} q(y - x)\, dy \ge 1 - 2\Phi_0\left(\frac{|x - x'|}{2\sqrt{z_1 h}}\right),
$$

where $p^{(1)}$ and $p^{(2)}$ are the densities of $x + V_1$ and $x' + V_1$.

In the Lemma 3 take $Z^{(1)} = x + V_1$, $Z^{(2)} = x' + V_1$, $Y^{(1)} = Y^{(2)} = U_1$, $m/2 = 2C_1 z_1 h$ and $k = k_1$. Let us construct a family (A_j) of sets with sizes not more than $k_1 m$ such that

$$
\begin{aligned}
\sum_j P^{(1)}(A_j) \wedge P^{(2)}(A_j) &\ge \left(1 - \frac{1}{k_1}\right) \int_{\mathbb{R}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy \\
&\ge \left(1 - \frac{1}{k_1}\right)\left(1 - 2\Phi_0\left(\frac{|x - x'|}{2\sqrt{z_1 h}}\right)\right), \quad (34)
\end{aligned}
$$

where $P^{(1)}$ and $P^{(2)}$ stand for the distributions of $x+V_1+U_1$ and $x'+V_1+U_1$.

We apply the Lemma 2 with these $P^{(1)}$, $P^{(2)}$, the family (A_j) of disjoint sets and Q and q standing for the distribution and the density of V_2. Together with (34) and the Lemma 4 where $2a = k_1 m = 4C_1 h k_1 z_1$ and $\sigma^2 = z_2 h$ we get,

$$\begin{aligned}\int_{\mathbb{R}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy &\geq \sum_j \left(P^{(1)}(A_j) \wedge P^{(2)}(A_j)\right) \int_{\mathbb{R}} \inf_{x \in A_j} q(y-x)\, dy \\ &\geq \left(1 - 2\Phi_0\left(\frac{2C_1\sqrt{h}\cdot k_1 z_1}{\sqrt{z_2}}\right)\right) \\ &\quad \times \left(1 - \frac{1}{k_1}\right)\left(1 - 2\Phi_0\left(\frac{|x-x'|}{2\sqrt{z_1 h}}\right)\right),\end{aligned}$$

where $p^{(1)}$ and $p^{(2)}$ denote the densities of $x+V_1+U_1+V_2$ and $x'+V_1+U_1+V_2$.

In the Lemma 3 take $Z^{(1)} = x + V_1 + U_1 + V_2$, $Z^{(2)} = x' + V_1 + U_1 + V_2$, $Y^{(1)} = Y^{(2)} = U_2$, $m/2 = 2C_1 z_2 h$ and $k = k_2$. Construct a new family (A_j) of sets with sizes not more than $k_2 m$ such that

$$\begin{aligned}\sum_j P^{(1)}(A_j) \wedge P^{(2)}(A_j) &\geq \left(1 - \frac{1}{k_2}\right) \int_{\mathbb{R}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy \\ &\geq \left(1 - \frac{1}{k_1}\right)\left(1 - 2\Phi_0\left(\frac{|x-x'|}{2\sqrt{z_1 h}}\right)\right) \\ &\times \left(1 - \frac{1}{k_2}\right)\left(1 - 2\Phi_0\left(\frac{2C_1\sqrt{h}\cdot k_1 z_1}{\sqrt{z_2}}\right)\right).\end{aligned}$$

Continue these iterative applications of the Lemmae 2, 4 and 3 until the last summand V_r is added. It results in the inequality,

$$\begin{aligned}\int_{\mathbb{R}} p^{(1)}(y) \wedge p^{(2)}(y)\, dy &\geq \prod_{j=1}^{r-1} \left(1 - 2\Phi_0\left(2C_1\sqrt{h}\cdot \frac{k_j z_j}{\sqrt{z_{j+1}}}\right)\right) \prod_{j=1}^{r-1}\left(1 - \frac{1}{k_j}\right) \\ &\quad \times \left(1 - 2\Phi_0\left(\frac{|x-x'|}{2\sqrt{z_1 h}}\right)\right),\end{aligned}$$

where $p^{(1)}$ and $p^{(2)}$ are the conditional densities of X^h_{nh} given $X^h_0 = x$ and $X^h_0 = x'$ correspondingly, i.e. $p^{(1)}(y) = p^h_{nh}(x,y)$ and $p^{(2)}(y) = p^h_{nh}(x',y)$.

The inequality (30) guarantees that the second product is bounded away from zero. *If the value of nh is small enough,* the estimate (13) implies that the first product is bounded from below by

$$\prod_{j=1}^{r-1}\left(1-\sqrt{2/\pi}\cdot 2C_1\sqrt{h}\cdot\frac{k_j z_j}{\sqrt{z_{j+1}}}\right),$$

and by inequality (28) this product is bounded away from zero as well. Finally, it follows from inequality (33) that $z_1 \geq n/4$ (remind that $b=2$), therefore,

$$\int_{\mathbb{R}} p^h_{nh}(x,y)\wedge p^h_{nh}(x',y)\,dy \geq C\left(1-2\Phi_0\left(\frac{|x-x'|}{\sqrt{nh}}\right)\right),\quad C>0.$$

Hence, we get the desired bound. •

5 Proof of Theorem 2

Proof has the same structure as the one of the Theorem 1 except that the Lemma 5 is used instead of the explicit estimate from the Lemma 4.

Choose an integer b so large that the Lemma 5 is applicable for $n \geq b$, and construct the "blocs" U_j and V_j exactly in the same way as in the proof of the Theorem 1. Again, we have $|U_j| < 2C_1 z_j h$, but V_j are not normally distributed in general.

In the Lemma 2 take $P^{(1)}=\delta_x$, $P^{(2)}=\delta_{x'}$, A_1 the closed parallelepiped with the points x and x' as its opposite vertices and its sides parallel to coordinate system's axes, and let Q be the distribution of V_1 with density q. Together with the Lemma 5 where $2a=2a_0=|x-x'|/\sqrt{z_1 h}$ we get the estimate

$$\int p^{(1)}(y)\wedge p^{(2)}(y)\,dy \geq \int \inf_{x\in A_1} q(y-x)\,dy \geq 1-c_1\left(\frac{|x-x'|}{\sqrt{z_1 h}}+\frac{|x-x'|^2}{2z_1 h}\right)-\frac{c_2}{z_1^{1/3}},$$

where $p^{(1)}$ and $p^{(2)}$ are the densities of $x+V_1$ and $x'+V_1$.

In the Lemma 3 take $Z^{(1)}=x+V_1$, $Z^{(2)}=x'+V_1$, $Y^{(1)}=Y^{(2)}=U_1$, $m/2=2C_1 z_1 h$ and $k=k_1$. Construct a family (A_j) of sets with diameters not more than $k_1 m\sqrt{d}$ such that

$$\sum_j P^{(1)}(A_j)\wedge P^{(2)}(A_j) \ge \left(1-\frac{1}{k_1}\right)^d \int p^{(1)}(y)\wedge p^{(2)}(y)\,dy$$
$$\ge \left(1-\frac{1}{k_1}\right)^d \left(1-c_1\left(\frac{|x-x'|}{\sqrt{z_1 h}}+\frac{|x-x'|^2}{2z_1 h}\right)-\frac{c_2}{z_1^{1/3}}\right), \tag{35}$$

where $P^{(1)}$ and $P^{(2)}$ stand for the distributions of $x+V_1+U_1$ and $x'+V_1+U_1$.

We apply the Lemma 2 with these $P^{(1)}$, $P^{(2)}$, the family (A_j) of disjoint sets and Q and q standing for the distribution and the density of V_2. Together with (35) and the Lemma 5 where $2a = 2a_1 = k_1 m\sqrt{d}/\sqrt{z_2 h} = 4\sqrt{d}C_1\sqrt{h}k_1 z_1/\sqrt{z_2}$ and $a^2/2 = 2dC_1^2 h k_1^2 z_1^2/z_2$, we get,

$$\begin{aligned}
\int p^{(1)}(y)\wedge p^{(2)}(y)\,dy &\ge \sum_j \left(P^{(1)}(A_j)\wedge P^{(2)}(A_j)\right)\int \inf_{x\in A_j} q(y-x)\,dy \\
&\ge \left(1-c_1\left(2\sqrt{d}C_1\sqrt{h}\,\frac{k_1 z_1}{\sqrt{z_2}}+2dC_1^2 h\frac{k_1^2 z_1^2}{z_2}\right)-\frac{c_2}{z_2^{1/3}}\right) \\
&\times \left(1-\frac{1}{k_1}\right)^d \left(1-c_1\left(\frac{|x-x'|}{\sqrt{z_1 h}}+\frac{|x-x'|^2}{2z_1 h}\right)-\frac{c_2}{z_1^{1/3}}\right),
\end{aligned}$$

where $p^{(1)}$ and $p^{(2)}$ denote the densities of $x+V_1+U_1+V_2$ and $x'+V_1+U_1+V_2$.

Continue these iterative applications of the Lemmae 2, 5 and 3 until the summand U_{r-1} is added. It results in the inequality,

$$\begin{aligned}
\sum_j P^{(1)}(A_j)\wedge P^{(2)}(A_j) &\ge \left(1-c_1\left(\frac{|x-x'|}{\sqrt{z_1 h}}+\frac{|x-x'|^2}{2z_1 h}\right)-\frac{c_2}{z_1^{1/3}}\right) \\
&\times \prod_{j=1}^{r-2}\left(1-2c_1\left(\sqrt{d}C_1\sqrt{h}\,\frac{k_j z_j}{\sqrt{z_{j+1}}}+d^2C_1^2 h\,\frac{k_j^2 z_j^2}{z_{j+1}}\right)-\frac{c_2}{z_{j+1}^{1/3}}\right) \\
&\times \prod_{j=1}^{r-1}\left(1-\frac{1}{k_j}\right)^d,
\end{aligned}$$

where $P^{(1)}$ and $P^{(2)}$ are the conditional distributions of $X_{nh}^h - V_r$ given $X_0^h = x$ and $X_0^h = x'$ correspondingly, and (A_j) is a countable family of disjoint sets with diameters at most $2a = 2a_{r-1} = 4\sqrt{d}C_1 h k_{r-1} z_{r-1}$.

The last step is to add $V_r = \xi_n \sqrt{h}$. This term has only one summand and must be treated differently because the Lemma 5 is not applicable here.

Due to the formulae for k_j and z_j from the Lemma 6, we see that

$$\begin{aligned} \frac{2a_{r-1}}{2\sqrt{z_r h}} &\leq 2\sqrt{d} C_1\, k_{r-1} z_{r-1} \sqrt{h} \leq 4\sqrt{d} C_1 b \cdot \left(\frac{5}{4}\right)^{r-1} \sqrt{h} \\ &= 5\sqrt{d} C_1 \left(\frac{5}{4\sqrt{b}}\right)^{r-2} b^{r/2} \sqrt{h} = o(1), \quad n \to \infty, \end{aligned}$$

since $\frac{5}{4\sqrt{b}} < 1$, r grows as n increases, and $b^{r/2}\sqrt{h} \leq \sqrt{bnh}$ is bounded due to the assumptions of the Theorem 2. Therefore, $a' := a/\sqrt{h} \to 0$, $h \to 0$, and, hence, for all n large enough the open set $U \equiv B_{r_0}(x_0)$ contains some ball of radius $2a'$. Denote $U' = \{x : |x - x_0| \leq r_0 - a'\}$. By change of variables,

$$\int \inf_{|x| \leq a} q_{V_r}(y - x)\, dy \geq \int \inf_{|x| \leq a} h^{-d/2} p\left(\frac{y - x}{\sqrt{h}}\right) dy$$

$$\geq \int_{U'} \inf_{|x'| \leq a'} p(z - x')\, dz \geq c_U\, |U'| \geq c_U |U|/2, \qquad h \to 0.$$

Here q_{V_r} denotes the density of the random variable V_r. Denote $c_U|U|/2 = C_U$. Then,

$$\begin{aligned} \int p_{nh}^h(x, y) \wedge p_{nh}^h(x', y)\, dy \geq C_U &\left(1 - c_1\left(\frac{|x - x'|}{\sqrt{z_1 h}} + \frac{|x - x'|^2}{2 z_1 h}\right) - \frac{c_2}{z_1^{1/3}}\right) \\ \times \prod_{j=1}^{r-2} &\left(1 - 2c_1\left(\sqrt{d} C_1 \sqrt{h}\, \frac{k_j z_j}{\sqrt{z_{j+1}}} + d^2 C_1^2 h\, \frac{k_j^2 z_j^2}{z_{j+1}}\right) - \frac{c_2}{z_{j+1}^{1/3}}\right) \\ &\times \prod_{j=1}^{r-1} \left(1 - \frac{1}{k_j}\right)^d. \end{aligned} \tag{36}$$

Inequality (30) implies that

$$\inf_n \prod_{j=1}^{r-1} \left(1 - \frac{1}{k_j}\right)^d \geq c > 0.$$

Due to (28)

$$\sum_{j=1}^{r-1} \frac{k_j z_j}{\sqrt{z_{j+1}}} \sqrt{h} \leq c\sqrt{nh} \leq c\sqrt{T},$$

and

$$\sum_{j=1}^{r-1} \frac{k_j^2 z_j^2}{z_{j+1}} h \le c\,nh \le c\,T,$$

By definition of z_j

$$\sum_{j=2}^{r-1} z_j^{-1/3} \le \sum_{k=1}^{\infty} \left(b^{-1/3}\right)^k = \frac{1}{b^{1/3}-1},$$

which can be made arbitrary small by choosing b large enough. Therefore, if the value of nh is small enough, then the product

$$\prod_{j=1}^{r-2} \left(1 - 2c_1 \left(\sqrt{d} C_1 \sqrt{h}\, \frac{k_j z_j}{\sqrt{z_{j+1}}} + d^2 C_1^2 h\, \frac{k_j^2 z_j^2}{z_{j+1}}\right) - \frac{c_2}{z_{j+1}^{1/3}}\right)$$

is bounded away from zero.

Thus, the two products and the integral term in (36) are bounded away from zero. Hence, there exist constants $c_0, c_1, c_2 > 0$ such that

$$\int p_n^h(x,y) \wedge p_n^h(x',y)\, dy \ge c_0 \left(1 - c_1 \left(\frac{|x-x'|}{\sqrt{T}} + \frac{|x-x'|^2}{2z_1 h}\right) - \frac{c_2}{n^{1/3}}\right).$$

This implies the desired inequality of the theorem. •

Remark 2 Emphasize that the method of the proof intrinsically suits the nature of local Markov–Dobrushin's condition; one might say the proof exploits the form of this condition itself. If one wished to compare to the "small sets" condition, it should be noticed that for the global versions of both conditions, the constant provided by Markov–Dobrushin's condition is always better, although, possibly in a non-strict sense. Indeed, "minorization" in the small sets condition,

$$\inf_x \mathbb{P}_x(X_{nh} \in dy) \ge \varepsilon \nu(dy)$$

with some probability measure ν, directly implies

$$\begin{aligned}\inf_{x,x'} \int_{\mathbb{R}^d} p_{nh}^h(x,y) \wedge p_{nh}^h(x',y)\, dy &= \inf_{x,x'} \int_{\mathbb{R}^d} \frac{\mathbb{P}_x(X_{nh}^h \in dy)}{\nu(dy)} \wedge \frac{\mathbb{P}_{x'}(X_{nh}^h \in dy)}{\nu(dy)}\ \nu(dy) \\ &\ge \int_{\mathbb{R}^d} (\varepsilon \wedge \varepsilon)\, \nu(dy) = \varepsilon.\end{aligned}$$

6 Proof of Theorem 3

Proof follows from the calculus in [8, 22] by virtue of local Markov–Dobrushin's condition (4), which is uniform for $0 < h \le h_0$ and for some $T > 0$, due to the Theorem 2. $\bullet$

7 Proof of Corollary 1

Proof **1.** Let T_0 be small enough (see (38) below). Remind that U is an open ball with a radius r_0 and let

$$R_0 := r_0 \wedge (c_1^{-1}/4) \wedge (c_1^{-1/2}/2), \tag{37}$$

where c_1 is the constant from the inequality (11). Denote by B' the ball with radius $R_0/2$ and the same center as U, and let B'' be a similar ball with the radius $2R_0/3$. Let us choose

$$T_0 \le R_0(\|F\|_B + 1)^{-1}/3 =: \tilde{T}_0. \tag{38}$$

Let h satisfy the bound $h \le h_0$, where h_0 is small enough (see (42) below). Let $M_0 > 0$, and

$$p_{B',T_0} := \inf_{|x|\le M_0} \mathbb{P}(x + \overline{W}_{T_0} \in B'), \tag{39}$$

where $\overline{W}$ is a Brownian motion with a covariance matrix $V = \mathrm{cov}(\xi)$, which is positive definite due to the assumption (A2). Notice that $p_{B',T_0} > 0$ (more than that, see (43) below). We have, due to the Central Limit Theorem, $W^h_{T_0} \Longrightarrow \overline{W}_{T_0}$ (weak convergence), where $\overline{W}_{T_0} \sim \mathcal{N}(0, VT_0)$; moreover, the uniform version of the CLT implies $\sup_{z\in\mathbb{R}^d} |\mathbb{P}(W^h_{T_0} \le z) - \mathbb{P}(\overline{W}_{T_0} \le z)| \to 0,\ h \to 0$. Hence, there exists h_0 small enough such that for every $|x| \le M_0$,

$$\inf_{h\le h_0} \mathbb{P}(x + W^h_{T_0} \in B'') \ge p_{B',T_0}/2 > 0. \tag{40}$$

Then, because of the choice $T_0 \le \tilde{T}_0$, we have,

$$\inf_{|x|\le M_0} \inf_{h\le h_0} \mathbb{P}_x(X^h_{T_0} \in B) \ge p_{B',T_0}/2 > 0. \tag{41}$$

Now, for $T = 2T_0$, the desired inequality (4) with $\rho = c_0 p^2_{B',T_0}/16$ follows from the Markov property, the choice (37) and from the bounds (11) and

(41), if $n \geq n_0$, $n_0 \geq [(4c_2)^3]$, so that $n^{-1/3} \leq [(4c_2)]^{-1}$. Here c_2 is the second constant from (11). Indeed, if X^h and $\tilde X^h$ are two independent copies of our Markov approximation of diffusion, then the integral from (4) is estimated from below as follows,

$$\int P_x(X^h_{2T_0} \in dx') \wedge P_{\tilde x}(\tilde X^h_{2T_0} \in dx')$$

$$\geq P_x(X^h_{T_0} \in B) P_{\tilde x}(\tilde X^h_{T_0} \in B) \inf_{(X^h_{T_0}, \tilde X^h_{T_0}) \in B \times B} \int p(X^h_{T_0}, x') \wedge p(X^h_{T_0}, x')\, dx'$$

$$\geq \left(\frac{p_{B',T_0}}{2}\right)^2 \frac{c_0}{4}.$$

Whence, in particular, it suffices to take

$$h_0 \leq \frac{T_0}{[(4c_2)^3]}. \tag{42}$$

2. Notice that without loss of generality – for T_0 small enough – we may admit that actually

$$\hat p_{B',T_0} := \inf_{T_0 \leq t \leq 2T_0} \inf_{|x| \leq M_0} \mathbb{P}(x + \overline{W}_t \in B') > 0, \tag{43}$$

and, due to that, also

$$\inf_{T_0 \leq t \leq 2T_0} \inf_{h \leq h_0} \mathbb{P}(x + W^h_t \in B'') \geq \hat p_{B',T_0}/2 > 0. \tag{44}$$

Hence, we may conclude that (4) holds true, in fact, for any $T_0 \leq T \leq 2T_0$ with $\rho = c_0 \hat p^2_{B',T_0}/16$, again if $n^{-1/3} \leq [(4c_2)]^{-1}$ and under (42).

3. Now for any $T \geq 2T_0$, define $k := [T/T_0]$. Notice that in this case,

$$T/T_0 - 1 \leq k \leq T/T_0,$$

or,

$$T - T_0 \leq k\, T_0 \leq T \quad \sim \quad k\, T_0 \leq T \leq (k+1)\, T_0.$$

Now the inequality (4) follows from the Markov property and step 2 of this proof with

$$\rho = \frac{c_0}{4}\, (\hat p_{B',\tilde T_0}/2)^{2k}.$$

•

References

[1] *Borovkov, A. A.* Ergodicity and Stability of Stochastic Processes, Wiley, Chichester et al., 1998.

[2] *Bhattacharya, R.N., Ranga Rao, R.* Normal approximation and asymptotic expansions, John Wiley, London et al., 1976.

[3] *Dobrushin, R. L.* Central limit theorem for non-stationary Markov chains, I. Teor. Veroyatnost. i Primenen. 1(1) (1956), 72–89; II. ibid., 1(4) (1956), 365–425.

[4] *Doob, J. L.* Stochastic processes. Wiley, New York, 1953.

[5] *Ibragimov, I. A., Linnik, Yu. V.* Independent and stationary sequences of random variables. Wolters-Noordhoff Publishing, Groningen, 1971, 443 pp.

[6] *Kleptsyna, M. L., Veretennikov, A. Yu.* On Continuous Time Ergodic Filters with Wrong Initial Data. Theory Probab. Appl. 53(2) (2009), 269–300.

[7] *Klokov, S. A.* On lower bounds of mixing rate for one class of Markov processes. (Russian) Teor. Veroyatnost. i Primenen., 2007, 51(3), 600–607.

[8] *Klokov, S. A., Veretennikov, A. Yu.* Subexponential mixing rate for a class of Markov processes (multidimensional case). Engl. transl. Theory Probab. Appl., 49(1) 2004, 1–13.

[9] *Klokov, S. A., Veretennikov, A. Yu.* Sub-exponential mixing rate for a class of Markov processes. Math. Comm. 9, 2004, 9–26.

[10] *Klokov, S. A., Veretennikov, A. Yu.* Mixing and convergence rates for a family of Markov processes apprriximating SDEs. Random Oper. and Stoch. Equ., 14(2), 2006, 103–126.

[11] *Konakov, V., Mammen, M.* Local limit theorems for transition densities of Markov chains converging to diffusions. Probab. Theory Rel. Fields, 117 (2000), 551–587.

[12] *Krylov, N. V.* The selection of a Markov process from a Markov system of processes, and the construction of quasidiffusion processes. (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 37 (1973), 691–708.

[13] *Lapinskas, R.* On local limit theorem and asymptotical expansions in the multidimensional case. Lit. Math. Sb. 11(4), 1971, 817-832.

[14] *Markov, A. A.* Extension of the law of large numbers to dependent quantities (Russian). Izvestiia Fiz.-Matem. Obsch. Kazan Univ (2nd Ser.), 15 (1906), 135–156.

[15] *Markov, A. A.* Selected Works. AN SSSR, Leningrad, 1951.

[16] *Nummelin, E.* General irreducible Markov chains and non-negative operators, Cambridge University Press, Cambridge, 1984.

[17] *Shervashidze, T.* Local limit theorems for sums of random variables and parametric estimators of probability density. Dr. Sc. thesis, N. Muskhelishvili Institute of Computational Mathematics of Georgian Academy of Sciences, Tbilisi, 1999.

[18] *Statuliavicius, V. A.* Limit theorems for densities and asymptotical expansions for distributions of sums of independent random variables. Theor. Veroyatn. Primenen. 10 (1965), 645–659.

[19] *Vaserstein, L. N.* Markov processes on denumerable products of spaces describing large systems automata. Problems Infromation Transmission, 5(3) (1969), 64–72 (in Russian).

[20] *Veretennikov, A. Yu.* Bounds for the mixing rate in the theory of stochastic equations. Theory Probab. Appl. 32, no. 2, 273–281 (1987); translation from Teor. Veroyatn. Primen. 32, no. 2, 299–308 (1987).

[21] *Veretennikov, A. Yu.* Estimates for the mixing rate for Markov processes. (Russian) Litovsk. Mat. Sb. 31(1), 1991, 40–49; Engl. transl.: Lithuanian Math. J. 31(1), 1991, 27–34.

[22] *Veretennikov, A. Yu.* On polynomial mixing and the rate of convergence for stochastic differential and difference equations. (Russian) Teor. Veroyatnost. i Primenen. 44, 1999, no. 2, 312–327; English translation: Theory Probab. Appl. 44, 1999, no. 2, 361–374.

[23] *Veretennikov, A. Yu.* On strong solutions and explicit formulas for solutions of stochastic integral equations. Math. USSR Sb., 39, 1981, 387–403.

[24] *Veretennikov, A. Yu.* On lower bounds for mixing coefficients of Markov diffusions. Kabanov, Yuri (ed.) et al., From stochastic calculus to mathematical finance. The Shiryaev Festschrift. Berlin: Springer, 2006, 623–633.