On convergence to stationary distributions for solutions of nonlinear Fokker–Planck–Kolmogorov equations

Vladimir I. Bogachev, Michael Röckner, Stanislav V. Shaposhnikov

Abstract. We obtain conditions under which solutions to a nonlinear Fokker–Planck–Kolmogorov equation with the diffusion matrix depending on the solution converge to the stationary solution.

Keywords: Fokker–Planck–Kolmogorov equation, stationary distribution, convergence in variation

MSC 2010: 60J75, 47G20, 60G52

1. Introduction

This paper is devoted to the study of the behavior, as $t \to \infty$, of the probability solution μ_t to the Cauchy problem for the nonlinear Fokker–Planck–Kolmogorov equation

$$\partial_t \mu_t = \partial_{x_i}\partial_{x_j}\big(a^{ij}(x,\mu_t)\mu_t\big) - \partial_{x_i}\big(b^i(x,\mu_t)\mu_t\big), \quad \mu_0 = \nu \tag{1.1}$$

with respect to probability measures on $\mathbb{R}^d$, where ν is a given initial distribution.

Let

$$L_\mu = a^{ij}(x,\mu)\partial_{x_i}\partial_{x_j} + b^i(x,\mu)\partial_{x_i}$$

with summation over repeated indices, where the coefficients are defined for all $x \in \mathbb{R}^d$ and probability measures μ from some set of probability measures on $\mathbb{R}^d$.

A family of probability measures $\{\mu_t\}$ on $\mathbb{R}^d$ is a solution to the Cauchy problem (1.1) on $[0,T]$ if, for every function $\varphi \in C_0^\infty(\mathbb{R}^d)$, there holds the equality

$$\int \varphi\, d\mu_t = \int \varphi\, d\nu + \int_0^t \int L_{\mu_s}\varphi\, d\mu_s\, ds$$

for almost all $t \in [0,T]$. Here we assume that the coefficients are Borel functions integrable with respect to the measure $\mu_s(dx)\,ds$ on bounded sets. A stationary solution is a probability measure μ on $\mathbb{R}^d$ satisfying the equation

$$\partial_{x_i}\partial_{x_j}\big(a^{ij}(x,\mu)\mu\big) - \partial_{x_i}\big(b^i(x,\mu)\mu\big) = 0, \tag{1.2}$$

which is understood in the sense of the integral equality

$$\int L_\mu \varphi\, d\mu = 0$$

for every function $\varphi \in C_0^\infty(\mathbb{R}^d)$. Below for brevity we write equations (1.1) and (1.2) in the form

$$\partial_t \mu_t = L^*_{\mu_t}\mu_t \quad \text{and} \quad L^*_\mu \mu = 0.$$

Equations of this form arise in physical, biological and probabilistic models (see [1]). In this paper we consider a nonlocal type of nonlinearity, where the coefficients depend on the measure, but not on the pointwise values of its density. Typical examples are equations with coefficients of the form

$$\int K(x,y)\,\mu_t(dy).$$

Nonlinear Fokker–Planck–Kolmogorov equations with such nonlinearities arise naturally in the following problem. Let us consider the system of differential equations

$$\dot{x}^i = N^{-1}\sum_{j=1}^{N} K(x^i, x^j),$$

where $K(x,y)$ is some function of two variables. It is readily verified that the measure

$$\mu_t^N = N^{-1}\sum_{j=1}^{N} \delta_{x_t^j}$$

satisfies the Vlasov equation

$$\partial_t \mu_t^N = -\partial_x (b(x,\mu_t^N)\mu_t^N), \quad b(x,\mu) = \int K(x,y)\,\mu(dy),$$

which is a particular case of the Fokker–Planck–Kolmogorov equation. Under broad conditions on K one can show that, as $N\to\infty$, the measures μ_t^N converge weakly to a measure μ_t that also satisfies the Vlasov equation. There is an extensive literature on such equations, we mention in the first place the pioneering paper of Dobrushin [2] and also the papers [3], [4], [5]. In case of functions $K(x,y)$ of a special form Vlasov equations can be regarded as a particular case of the theory of gradient flows (see [6], [7]). If in place of the system of deterministic equations we consider the system of stochastic equations

$$dx_t^i = N^{-1}\sum_{j=1}^{N} K(x_t^i, x_t^j)dt + \sqrt{2}dw_t^i,$$

where w_t^i are independent Wiener processes, then one can show that the one-dimensional distributions $\mu_t^{N,i}$ (i.e., $\mu_t^{N,i}(B) = P(x_i\in B)$) converge weakly to some measure μ_t as $N\to\infty$ and this measure satisfies the McKean–Vlasov equation

$$\partial_t \mu_t = \Delta\mu - \partial_x (b(x,\mu_t)\mu_t), \quad b(x,\mu) = \int K(x,y)\,\mu(dy),$$

which is also a particular case of the Fokker–Planck–Kolmogorov equation. There are fundamental works [8], [9], [10], [11] on such equations. In this paper, as compared to the above examples, we study a more general situation where the diffusion matrix depends on the solution.

Note that there exists a deep connection between Fokker–Planck–Kolmogorov equations and the corresponding stochastic equations, which is expressed in the form of the so-called superposition principle (see [12], [13], [14], [15]). This principle asserts that under broad conditions on the coefficients of the linear Fokker–Planck–Kolmogorov equation for every solution $\{\mu_t\}$ there exists a probability measure P on $C[0,T]$ such that

$$\int \varphi\,d\mu_t = \int_{C[0,T]} \varphi(x(t))\,dP \quad \forall\varphi\in C_0^\infty(\mathbb{R}^d)$$

and the process

$$\varphi(x(t)) - \varphi(x(0)) - \int_0^t [\mathrm{tr}(AD^2\varphi) + \langle b,\nabla\varphi\rangle](x(s))\,ds$$

is a martingale with respect to the measure P and the filtration $\sigma(x(s), s\in[0,t])$. Thus, to every solution of the Fokker–Planck–Kolmogorov equation there corresponds a solution to the martingale problem, moreover, an analogous assertion is also true for nonlinear equations, which finds applications in deriving and justifying probability representations of solutions to partial differential equations (see [15], [16]). The problems of existence and uniqueness of solutions to stochastic McKean–Vlasov equations under very weak restrictions on their coefficients are discussed in the papers [17], [18], [19], [20]. The last paper gives a generalization of the known method of Lyapunov functions to the nonlinear case.

The problems of existence and uniqueness of probability solutions to the Cauchy problem (1.1) and to the stationary equation (1.2) are discussed in the papers [21], [22], [23], [24], [25]. Infinite-dimensional nonlinear Fokker–Planck–Kolmogorov equations are discussed in [26], [27]. A survey of the theory of linear Fokker–Planck–Kolmogorov equations is given in the book [28].

In this paper we study convergence to the stationary distribution and show that in the case where the nonlinear Fokker–Planck–Kolmogorov equation is in a sense a small perturbation of the linear equation, there holds the estimate

$$\|W^k(\mu_t - \overline{\mu})\|_{TV} \le q_1 e^{-q_2 t}$$

with

$$W(x) = 1 + |x|,$$

where $\{\mu_t\}$ is the solution to the Cauchy problem (1.1), $\overline{\mu}$ is the solution to the stationary equation (1.2), $\|\cdot\|_{TV}$ denotes the total variation norm on the space of measures and $W^k\sigma$ is the measure with density W^k with respect to σ.

The problems of convergence to stationary distributions of solutions to nonlinear Fokker–Planck–Kolmogorov equation are considered in many papers, among which we note [29], [30], [31], and [32]. Most of the known results give merely weak convergence or convergence in the Kantorovich metric. Convergence in variation is concerned only in the recent papers [33], [34] and [35], in which, however, only the drift coefficient b depends on μ, but the diffusion matrix A does not depend on μ. Dependence of the diffusion matrix on the solution considerably complicates this study, which is seen even in the proof of uniqueness of solutions to the Cauchy problem. For instance, an example is given in [25] in which the drift coefficient b is zero and the diffusion matrix $A(t, x, \mu)$ is Lipschitz in μ, but the Cauchy problem has several solutions.

In case of locally bounded coefficients and a nondegenerate diffusion matrix the solution $\{\mu_t\}$ to the Fokker–Planck–Kolmogorov equation for almost all $t > 0$ has a density $\varrho(x, t)$ with respect to Lebesgue measure and this density possesses a version continuous in all variables (see [28]). A similar assertion is true for solutions to the stationary equation. Thus, in the case of a nondegenerate diffusion matrix it is naturally to consider convergence in variation (in $L^1(\mathbb{R}^d)$), but not in the Kantorovich metric.

2. Main results

Let $k \ge 1$. Let $\mathscr{P}^k(\mathbb{R}^d)$ denote the set of probability measures μ on $\mathbb{R}^d$ with finite kth moment, i.e., $|x|^k \in L^1(\mu)$. Let $\mathscr{P}^k_a(\mathbb{R}^d)$ denote the subset of $\mathscr{P}^k(\mathbb{R}^d)$ consisting of absolutely continuous measures. The space $\mathscr{P}^k_a(\mathbb{R}^d)$ with the metric $\|W^k(\mu - \sigma)\|_{TV}$ is complete.

Suppose that for every measure $\mu \in \mathscr{P}^k_a(\mathbb{R}^d)$ and every $\varepsilon \in [0, 1]$ we are given Borel functions $x \mapsto a^{ij}_\varepsilon(x, \mu)$ and $x \mapsto b^i_\varepsilon(x, \mu)$. Let the following conditions be fulfilled.

(H1) The matrix $A_\varepsilon(x,\mu)=(a^{ij}_\varepsilon(x,\mu))$ is symmetric and positive definite, moreover, there exist constants $\Lambda>0$ and $\alpha>0$, independent of ε and μ, such that

$$|A_\varepsilon(x,\mu)-A_\varepsilon(y,\mu)|\le \Lambda|x-y|,\quad \alpha^{-1}I\le A_\varepsilon(x,\mu)\le \alpha I.$$

(H2) There exist constants $\beta_0>0$, $\beta_1>0$, $\beta_2>0$, $\beta_3>0$ and $m\ge 0$, independent of ε and μ, such that

$$|b_\varepsilon(x,\mu)|\le \beta_0(1+|x|)^m\|W^k\mu\|_{TV},\quad \langle b_\varepsilon(x,\mu),x\rangle\le \beta_1-\beta_2|x|^{2k}+\varepsilon\beta_3\|W^k\mu\|^2_{TV}.$$

(H3) There exists a constant $N>0$, independent of ε, μ and σ, such that

$$|A_\varepsilon(x,\mu)-A_\varepsilon(x,\sigma)|\le \varepsilon N\|W^k(\mu-\sigma)\|_{TV}$$

and for all $1\le i,j\le d$

$$|\partial_{x_j}a^{ij}_\varepsilon(x,\mu)-\partial_{x_i}a^{ij}_\varepsilon(x,\sigma)|+|b^j_\varepsilon(x,\mu)-b^j_\varepsilon(x,\sigma)|\le \varepsilon N(1+|x|)^m\|W^k(\mu-\sigma)\|_{TV}.$$

Example 2.1. Let

$$A_\varepsilon(x,\mu)=I+\varepsilon\int Q(x,y)\,\mu(dy),\quad b_\varepsilon(x,\mu)=-x+\varepsilon\int K(x,y)\,\mu(dy).$$

Suppose that Q is a symmetric matrix and

$$0\le Q(x,y)\le C_1I,\quad |Q(x,y)-Q(z,y)|\le C_2|x-z|,$$
$$\langle K(x,y),x\rangle\le C_3(1+|y|)(1+|x|),\quad |K(x,y)|\le C_4(1+|x|)^m(1+|y|).$$

Then conditions (H1), (H2), (H3) are fulfilled.

Throughout we consider the solution $\{\mu_t^\varepsilon\}$ to the Cauchy problem (1.1) in which the coefficients of the Fokker–Planck–Kolmogorov equation are precisely A_ε and b_ε, but for notational simplicity we omit the index ε. Similarly we do in the case of a stationary solution and a stationary Fokker–Planck–Kolmogorov equation.

The following theorem is the main result of this paper.

Theorem 2.2. *Let conditions* (H1), (H2), (H3) *be fulfilled and* $r\in(0,\beta_2/2\alpha k)$. *Suppose that* $\exp(r|x|^{2k})\in L^1(\nu)$. *Then there exists* $\varepsilon_0\in(0,1]$ *along with positive numbers* q_1, q_2 *such that for all* $\varepsilon\in[0,\varepsilon_0]$ *there holds the estimate*

$$\|W^k(\mu_t^\varepsilon-\overline{\mu}^\varepsilon)\|_{TV}\le q_1e^{-q_2t},$$

where $\{\mu_t^\varepsilon\}$ *is the solution to the Cauchy problem* (1.1) *with initial condition* ν *and* $\overline{\mu}^\varepsilon$ *is the solution to the stationary equation* (1.2). *The constants* ε_0, q_1 *and* q_2 *depend only on* d, ν, r *and the constants from conditions* (H1), (H2), (H3).

We observe that the existence of a solution $\{\mu_t\}$ on $[0,T]$ for every $T>0$ follows from [25, Theorem 3]. Sufficient conditions for uniqueness are obtained in the papers [23] and [25]. In [25, Theorem 5] under conditions close to (H1), (H2) and (H3), uniqueness is proved under the additional very essential assumption that the initial distribution ν has a density ϱ_0 with respect to Lebesgue measure and $\varrho_0\ln\varrho_0\in L^1(\mathbb{R}^d)$. In this paper we do not impose this restriction on the initial distribution. However, it will be shown below (see Corollary 3.5) that $\varrho_\varepsilon(x,t)\ln\varrho_\varepsilon(x,t)\in L^1(\mathbb{R}^d)$ for all $t>1$, where $\varrho_\varepsilon(x,t)$ is the density of the solution μ_t^ε. An assertion analogous to [25, Theorem 5], is discussed in Remark 3.8. Thus, we can use the uniqueness condition for $t>1$, which is sufficient for extending the solution to $[0,+\infty)$. Hence under our conditions there exists a solution $\{\mu_t^\varepsilon\}$ to the Cauchy problem (1.1) on the whole half-line $[0,+\infty)$. Note that for the existence of solutions we do not need that ε be small.

We now show that for ε sufficiently small there exists a stationary solution from the class $\mathscr{P}^k_a(\mathbb{R}^d)$ introduced above.

Proposition 2.3. *Suppose that conditions* (H1), (H2), (H3) *are fulfilled. Then there exists* $\varepsilon_0 > 0$ *such that for all* $\varepsilon \in [0, \varepsilon_0]$ *the corresponding stationary equation has a solution* μ^ε *in the class* $\mathscr{P}_a^k(\mathbb{R}^d)$.

Proof. Let $\sigma \in \mathscr{P}_a^k(\mathbb{R}^d)$. According to [28, Corollary 2.4.2 and Theorem 4.1.6], there exists a unique solution μ to the linear stationary equation $L_\sigma^*\mu = 0$. The solution μ possesses a density with respect to Lebesgue measure. Note that

$$L_\sigma|x|^2 \le 2\alpha d + 2\beta_1 - 2\beta_2|x|^{2k} + 2\varepsilon\beta_3\|W^k\sigma\|_{TV}^2.$$

According to [28, Theorem 2.3.2], there holds the estimate

$$\int |x|^{2k}\,\mu(dx) \le \alpha d\beta_2^{-1} + \beta_1\beta_2^{-1} + \varepsilon\beta_3\beta_2^{-1}\|W^k\sigma\|_{TV}^2.$$

Suppose that $\|W^{2k}\sigma\|_{TV} \le R$. Then $\|W^k\sigma\|_{TV}^2 \le R$ and

$$\|W^{2k}\mu\|_{TV} \le C_1 + \varepsilon C_2 R,$$

where C_1, C_2 depend only on α, β_1, β_2, β_3 and k. Let $R = 2C_1$ and $\varepsilon \le 1/(2C_2)$. Then we have $C_1 + \varepsilon C_2 R \le R$. We denote by $\mathscr{P}_R$ the set of probability measures $\sigma \in \mathscr{P}_a^k(\mathbb{R}^d)$ satisfying the inequality

$$\|W^{2k}\sigma\|_{TV} \le R.$$

Thus, we have the well-defined mapping T that to every measure $\sigma \in \mathscr{P}_R$ associates the solution $\mu \in \mathscr{P}_R$. We observe that for every measure $\sigma \in \mathscr{P}_R$ we have

$$|b_\varepsilon(x,\sigma)| \le \beta_0\sqrt{R}(1+|x|)^m, \quad \langle b_\varepsilon(x,\sigma), x\rangle \le \beta_1 + \beta_3 R - \beta_2|x|^{2k}.$$

Moreover, applying again [28, Theorem 2.3.2], one can show that for $\mu = T(\sigma)$ and any $p \ge 1$ there holds the estimate

$$\int (1+|x|)^p\,\mu(dx) \le C(p),$$

in which the constant $C(p)$ depends only on p, d, α, β_1, β_2, β_3 and R. Let $\mu_1 = T(\sigma_1)$ and $\mu_2 = T(\sigma_2)$. Denote by ϱ_{μ_2} the density of the measure μ_2. By [36, Theorem 3.1] with $V(x) = (1+|x|)^k$ one has the estimate

$$\|W^k(\mu_1-\mu_2)\|_{TV}^2 \le C_3\Bigl(\int |A(\mu_1)^{-1/2}\Phi|^2\,d\mu_2\Bigr)\Bigl(\int (1+|x|^{2k})(1+|A(\mu_1)^{-1/2}\Phi|^2)\,d\mu_2\Bigr), \tag{2.1}$$

where

$$\Phi = \frac{\bigl(A(\mu_1) - A(\mu_2)\bigr)\nabla\varrho_{\mu_2}}{\varrho_{\mu_2}} - \bigl(h(\mu_2) - h(\mu_1)\bigr),$$

$$h^i(\mu) = b^i(\mu) - \partial_{x_j}a^{ij}(\mu).$$

Here for shortening the notation we omit indication of x in the coefficients. According to [28, Corollary 3.4.7], for every $p \ge 1$ there exists a constant C_2, depending on p, R and the constants from conditions (H1), (H2), (H3), for which

$$\int \Bigl|\frac{\nabla\varrho_{\mu_2}}{\varrho_{\mu_2}}\Bigr|^p\,d\mu_2 \le C_4.$$

We observe that

$$(1+|A(\mu_1)^{-1/2}\Phi|^2)^2 \le C_5\Bigl|\frac{\nabla\varrho_{\mu_2}}{\varrho_{\mu_2}}\Bigr|^4 + C_6(1+|x|)^{4m}.$$

Therefore,

$$\int(1+|x|^{2k})(1+|A(\mu_1)^{-1/2}\Phi|^2)\,d\mu_2$$
$$\le\Bigl(\int(1+|x|^{2k})^2\,d\mu_2\Bigr)^{1/2}\Bigl(\int(1+|A(\mu_1)^{-1/2}\Phi|^2)^2\,d\mu_2\Bigr)^{1/2}\le C_7.$$

From the estimate

$$|A(\mu_1)^{-1/2}\Phi|^2\le 2\alpha\Bigl|\frac{\nabla\varrho_{\mu_2}}{\varrho_{\mu_2}}\Bigr|^2|A(\mu_1)-A(\mu_2)|^2+2\alpha|h(\mu_2)-h(\mu_1)|^2$$

and condition (H3) we obtain the estimate

$$|A(\mu_1)^{-1/2}\Phi|^2\le 2\alpha N^2\varepsilon^2\|W^k(\sigma_1-\sigma_2)\|^2_{TV}\Bigl(\Bigl|\frac{\nabla\varrho_{\mu_2}}{\varrho_{\mu_2}}\Bigr|^2+(1+|x|)^{2m}\Bigr),$$

whence

$$\int|A(\mu_1)^{-1/2}\Phi|^2\,d\mu_2\le\varepsilon^2C_8\|W^k(\sigma_1-\sigma_2)\|^2_{TV}.$$

Thus, we have estimated the expressions in the right-hand side of (2.1) and now we can write down the estimate for the distance between μ_1 and μ_2 through the distance between σ_1 and σ_2:

$$\|W^k(\mu_1-\mu_2)\|_{TV}\le\varepsilon C_9\|W^k(\sigma_1-\sigma_2)\|_{TV},$$

where C_9 does not depend on ε. It is clear that for sufficiently small $\varepsilon>0$ the mapping T is contracting. Hence there exists a unique fixed point $\mu=T(\mu)$, which is the desired solution to the stationary equation. □

The uniqueness of a stationary solution in the class $\mathscr{P}^k_a(\mathbb{R}^d)$ follows immediately from Theorem 2.2. Note that under somewhat different assumptions the existence and uniqueness theorem was obtained in the papers [21], [37].

3. Auxiliary results and proofs

For the proof of the main result we need several auxiliary assertions, in which we substantially use various facts from the theory of linear Fokker–Planck–Kolmogorov equations. In particular, we apply a priori estimates with a Lyapunov function, upper bounds on the density of the solution, estimates for the logarithmic gradient and entropy, and estimates for distances between solutions.

We recall that $\mu^\varepsilon_t=\varrho_\varepsilon(x,t)\,dx$ is the solution to the Cauchy problem (1.1) and that we assume throughout that the coefficients of the Fokker–Planck–Kolmogorov equation satisfy conditions (H1), (H2), (H3). For shortening the notation we omit the index ε and write simply μ_t and $\varrho(x,t)$.

Below we always assume that the constants N_i can depend only on r, d, ν and the constants from conditions (H1), (H2), (H3), but do not depend on t, τ, T, ε, μ_t.

The next lemma is based on estimates with a Lyapunov function.

Lemma 3.1. *There holds the estimate*

$$\int\exp(r|x|^{2k})\,\mu_t(dx)\le N_1\quad\forall t>0.$$

Proof. Note that the constants C_i, C'_i and M_i used below depend only on d, r, k, α and β_i. First we obtain a bound on $\|W^{2k}\mu_t\|_{TV}$. There holds the inequality

$$L_{\mu_t}|x|^{2k}\le C_1-C_2|x|^{4k-2}+\varepsilon C_3|x|^{2k-2}\|W^k\mu_t\|^2_{TV}.$$

Note that $\|W^k\mu_t\|^2_{TV} \le \|W^{2k}\mu_t\|_{TV}$. In addition, by Young's inequality

$$|x|^{2k-2}|y|^{2k} \le \frac{2k-2}{4k-2}|x|^{4k-2} + \frac{2k}{4k-2}|y|^{4k-2}.$$

Therefore,

$$L_{\mu_t}|x|^{2k} \le C_1' - C_2'|x|^{4k-2} + \varepsilon C_3' \int |y|^{4k-2}\,\mu_t(dy).$$

It is readily seen that

$$\int \Bigl(C_1' - C_2'|x|^{4k-2} + \varepsilon C_3' \int |y|^{4k-2}\,\mu_t(dy)\Bigr)\,\mu_t(dx) = C_1' - (C_2' - \varepsilon C_3')\int |x|^{4k-2}\,\mu_t(dx).$$

Let $C_4' \in (0, C_2')$ and let ε be so small that $C_4' + \varepsilon C_3' < C_2'$. Repeating the reasoning from the proof of [28, Theorem 7.1.1] and taking into account that $-|x|^{4k-2} \le 1 - |x|^{2k}$, we obtain the inequality

$$\int e^{C_4' t}|x|^{2k}\,\mu_t(dx) \le \int |x|^{2k}\,\nu(dx) + C_5' e^{C_4' t}.$$

Thus, $\|(1+|y|)^{2k}\mu_t\| \le R$, where

$$R = 4^k + 4^k \int |x|^{2k}\,\nu(dx) + 4^k C_5'.$$

Therefore, $\|(1+|x|)^k\mu_t\|^2 \le R$ and

$$\langle b(x,\mu_t), x\rangle \le \beta_1' - \beta_2|x|^{2k}, \quad \beta_1' = \beta_1 + \beta_3 R.$$

We now obtain a bound on the integral of $\exp(r|x|^{2k})$ for $r < \beta_2/2\alpha k$. There holds the inequality

$$L_{\mu_t}\exp(r|x|^{2k}) \le M_1 - M_2\exp(r|x|^{2k}).$$

Let $M_3 \in (0, M_2)$ and $V(x) = \exp(r|x|^{2k})$. Then

$$\partial_t(e^{M_3 t}V) + L_{\mu_t}(e^{M_3 t}V) \le M_1 e^{M_3 t}.$$

By [28, Theorem 7.1.1] we have

$$\int e^{M_3 t} V\,d\mu_t \le \int V\,d\nu + M_1 M_3^{-1} e^{M_3 t}.$$

Therefore,

$$\int V\,d\mu_t \le e^{-M_3 t}\int V\,d\nu + M_1 M_3^{-1} \le \int V\,d\nu + M_1 M_3^{-1} = N_1,$$

which completes the proof. □

Remark 3.2. Below we always assume that $\varepsilon \le \varepsilon_1$. In the proof of Lemma 3.1 we obtained the estimate $\|W^k\mu_t\|^2_{TV} \le R$. Therefore, condition (H2) can be rewritten as follows:

$$|b(x,\mu_t)| \le \beta_0'(1+|x|)^m, \quad \langle b(x,\mu_t), x\rangle \le \beta_1' - \beta_2|x|^{2k},$$

where $\beta_0' = \beta_0\sqrt{R}$ and $\beta_1' = \beta_1 + \beta_3 R$. We shall use condition (H2) in this form.

Remark 3.3. Let $\mu \in \mathscr{P}_a^k(\mathbb{R}^d)$ and $\tau > 0$. Let us consider a probability solution $\{\sigma_t\}$ to the Cauchy problem

$$\partial_t\sigma_t = L_\mu^*\sigma_t, \quad \sigma_\tau = \mu_\tau, \quad t > \tau.$$

By Lemma 3.1

$$\int \exp(r|x|^{2k})\,d\mu_\tau \le N_1.$$

In addition, as in the proof of Lemma 3.1, one can obtain the inequality

$$L_\mu \exp(r|x|^{2k}) \le M_1(\mu) - M_2(\mu)\exp(r|x|^{2k}),$$

whence one can derive the bound

$$\int \exp(r|x|^{2k})\,\sigma_t(dx) \le \int \exp(r|x|^{2k})\,\mu_\tau(dx) + M_3(\mu) = N_1 + M_3(\mu) = N_1'(\mu),$$

where the constants $M_i(\mu)$ and $N_1'(\mu)$ depend on d, r, k, the constants from conditions (H1) and (H2) and the measure μ through the expression $\|W^k\mu\|_{TV}$.

In the following lemma we establish an upper bound on the solution density.

Lemma 3.4. *For all $t > 1$ one has $\varrho(x,t) \le N_2$.*

Proof. We observe that the measures $\mu_{t-1+s} = \varrho(x, t-1+s)\,dx$ satisfy the Fokker–Planck–Kolmogorov equation with the operators $L_{\mu_{t-1+s}}$ on the interval $(0,2)$. Let γ be a number larger than $(d+2)/2$. By [28, Corollary 7.2.3], in which we set $\Theta = 1/2$, one has the inequality

$$\varrho(x, t-1+s) \le C_1 s^{-(d+2)/2}(1+s^{2\gamma}) \int_{s/2}^{s} \int (1+|x|^{2\gamma m})\,\mu_{t-1+\tau}(dx)\,d\tau,$$

where C_1 depends only on d, γ, β_0' and the constants from conditions (H1) and (H2). By the estimate

$$|x|^{2\gamma m} \le C_2 \exp(r|x|^{2k}),$$

where C_2 depends only on γ, m, r, k, the integral of $|x|^{2\gamma m}$ with respect to the measure $\mu_{t-1+\tau}$ is estimated from above by a constant C_2N_1. Therefore, for $s = 1$ we obtain the estimate $\varrho(x,t) \le C_1(1 + C_2N_1)$. □

From Lemma 3.4 one immediately deduces an estimate for the entropy, which we apply below.

Corollary 3.5. *For all $t > 1$ one has*

$$\int \varrho(x,t) \ln \varrho(x,t)\,dx \le \ln N_2.$$

In the following assertion we obtain an estimate on the logarithmic derivative of the solution.

Lemma 3.6. *Let $\tau > 1$ and $T > 0$. Then*

$$\int_\tau^{\tau+T} \int \frac{|\nabla \varrho(x,t)|^2}{\varrho(x,t)}\,dx\,dt \le N_3(1+T).$$

Proof. The measure $\mu_{\tau+t} = \varrho(x, \tau+t)\,dx$ satisfies on the interval $[0,T]$ the Fokker–Planck–Kolmogorov equation with the operator $L_{\mu_{\tau+t}}$ and at $t = 0$ it equals the measure μ_τ. By [28, Theorem 7.4.1] (see also the remark after the cited theorem) one has the estimate

$$\int_0^T \int \frac{|\nabla \varrho(\tau+t,x)|^2}{\varrho(\tau+t,x)}\,dx\,dt \le C_1 + \alpha \int_0^T \int |b(x,\mu_{\tau+t})|^2\,\varrho(x,\tau+t)\,dx\,dt$$
$$+ 2\alpha \int \varrho(x,\tau)\ln\varrho(x,\tau)\,dx + 2\alpha(d+1)\int \ln(\max|x|,1)\varrho(T+\tau,x)\,dx,$$

where C_1 depends only on α, d, Λ. The integral of $\varrho(x,\tau)\ln\varrho(x,\tau)$ is estimated by a constant $\ln N_2$ according to the previous assertion. In addition, the function $|b(x,\mu_{\tau+t})|^2$ is estimated by the expression $(\beta_0')^2(1+|x|)^{2m}$, hence, it is estimated by the expression $\exp(r|x|^{2k})$ with some constant factor. The function $\ln(\max|x|,1)$

is also estimated by the expression $\exp(r|x|^{2k})$ with some constant factor. Thus, the integral of $|b(x,\mu_{\tau+t})|^2$ is estimated by the quantity C_2TN_1 and the integral of $\ln(\max|x|,1)$ is estimated by the quantity C_2N_1, where C_2 depends only on β_0', m, r, k. Therefore,

$$\int_0^T\int\frac{|\nabla\varrho(\tau+t,x)|^2}{\varrho(\tau+t,x)}\,dx\,dt\le C_1+\alpha\ln N_2+2\alpha(d+1)C_2N_1+\alpha C_2N_2T,$$

which completes the proof. □

The next assertion is a corollary of estimates obtained in the paper [24] and plays the key role in justification of convergence to the stationary distribution.

Lemma 3.7. *Let $\tau>1$, $T>0$ and let μ be some probability measure in $\mathscr{P}_a^k(\mathbb{R}^d)$. Suppose that a family of measures σ_t from $\mathscr{P}_a^k(\mathbb{R}^d)$ for $t\in[\tau,\tau+T]$ is a solution to the Cauchy problem*

$$\partial_t\sigma_t=L_\mu^*\sigma_t,\quad \sigma_\tau=\mu_\tau.$$

Then

$$\|W^k(\mu_t-\sigma_t)\|_{TV}\le\varepsilon N_4(\mu)\Bigl(\int_\tau^t G(s)\|W^k(\mu_s-\mu)\|_{TV}^2\,ds\Bigr)^{1/2},$$

where G is a positive function such that

$$\int_\tau^{\tau+T}G(t)\,dt\le N_3T+N_5.$$

Proof. Applying [24, Corollary 1.2], we obtain the estimate

$$\|W^k(\mu_t-\sigma_t)\|_{TV}^2\le 4^kr^{-1}(1+\ln\eta(t))\int_\tau^t\int|A(\mu)^{-1/2}\Phi|^2\,d\mu_s\,ds,$$

where

$$\Phi=\frac{(A(\mu)-A(\mu_t))\nabla\varrho}{\varrho}-(h(\mu_t)-h(\mu)),$$
$$h^i(\mu)=b^i(\mu)-\partial_{x_j}a^{ij}(\mu),$$
$$\eta(t)=\int\exp(4^{-k}r(1+|x|)^{2k})\,\sigma_t(dx)\le e^r\int\exp(r|x|^{2k})\,\sigma_t(dx).$$

Here for shortening the notation we omit indication of x in the coefficients. We recall that by conditions (H1) and (H3) the following inequalities are fulfilled:

$$A(\mu)^{-1/2}\le\alpha^{1/2}I,\quad |A(\mu)-A(\mu_t)|\le\varepsilon N\|W^k(\mu-\mu_t)\|_{TV},$$
$$|h(\mu)-h(\mu_t)|\le\varepsilon N(1+|x|)^m\|W^k(\mu-\mu_t)\|_{TV}.$$

Set

$$G(t)=\int_{\mathbb{R}^d}\frac{|\nabla\varrho(t,x)|^2}{\varrho(t,x)}\,dx+\int_{\mathbb{R}^d}(1+|x|)^{2m}\varrho(x,t)\,dx.$$

Then

$$\int|A(\mu)^{-1/2}\Phi|^2\varrho\,dx\le\varepsilon^2\alpha N^2G(t)\|W^k(\mu-\mu_t)\|_{TV}^2.$$

We now observe that by Lemma 3.1, Remark 3.3 and Lemma 3.6 we have

$$\eta(t)\le e^rN_1'(\mu),\qquad \int_\tau^{\tau+T}G(t)\,dt\le N_3T+N_5.$$

We finally obtain

$$\|W^k(\mu_t-\sigma_t)\|_{TV}^2\le\varepsilon^2\alpha N^24^kr^{-1}(1+r+\ln N_1'(\mu))\int_\tau^tG(s)\|W^k(\mu-\mu_s)\|_{TV}^2\,ds,$$

as required. □

Remark 3.8. Let us show that with the aid of [24, Corollary 1.2] one can derive the uniqueness of solutions of the Cauchy problem $\partial_t\mu_t = L^*_{\mu_t}\mu_t$, $\mu_0 = \nu$ on a sufficiently small interval $[0,\tau]$, provided that the initial distribution ν has a density ϱ_0 with respect to Lebesgue measure such that $\varrho_0 \ln \varrho_0 \in L^1(\mathbb{R}^d)$. The following reasoning actually repeats the proof of Theorem 5 from [25]. Similarly to the proof of Lemma 3.7 one can apply [24, Corollary 1.2] and [28, Theorem 7.4.1], which for two solutions μ_t^1 and μ_t^2 to the Cauchy problem yields the estimate

$$\sup_t \|W^k(\mu_t^1 - \mu_t^2)\|^2_{TV} \le C_1 \int_0^\tau G(s)\,ds \sup_t \|W^k(\mu_t^1 - \mu_t^2)\|^2_{TV},$$

where C_1 does not depend on τ and the function $G(s)$ is integrable on $[0,\tau]$. Let τ be so small that

$$\int_0^\tau G(s)\,ds < 1/2C_1.$$

Then $\sup_t \|W^k(\mu_t^1 - \mu_t^2)\|_{TV} = 0$. Note that without the assumption of finite entropy in the general case a solution can be nonunique (see [25, Example 3]).

The next assertion is a particular case of [35, Lemma 3.5] (see also [38]).

Lemma 3.9. *Let $\overline{\mu} \in \mathscr{P}^k_a(\mathbb{R}^d)$ be a stationary solution and let $\{\sigma_t\}$ be the solution to the Cauchy problem $\partial_t\sigma_t = L^*_{\overline{\mu}}\sigma_t$, $\sigma_0 = \eta \in \mathscr{P}^k(\mathbb{R}^d)$. Then there exist numbers $\lambda_1 > 0$ and $\lambda_2 > 0$ such that*

$$\|W^k(\sigma_t - \overline{\mu})\|_{TV} \le \lambda_1 e^{-\lambda_2 t}\|W^k(\eta - \overline{\mu})\|_{TV}.$$

In [35] (as well as in [38]) this assertion was proved with the aid of the known Harris theorem (see [39]). A drawback of applying the Harris theorem is that the constants λ_1 and λ_2 depend on the coefficients of the operator L_μ in a complicated way. There are other methods of justification of such convergence, in particular, the method of coupling (see [34]). However, these methods usually require greater regularity of the coefficients than is necessary for applying the Harris theorem. Note that if the constants λ_1 and λ_2 are known, then it is not very difficult to control dependence of the constants q_1 and q_2 on the coefficients of the equation.

We are now able to prove Theorem 2.2. The idea of the proof is this. It suffices to show that for some $T > 0$ there exists $q \in (0,1)$ such that for every $\tau > 1$ one has the estimate

$$\|W^k(\mu_{\tau+T} - \overline{\mu})\|_{TV} \le q\|W^k(\mu_\tau - \overline{\mu})\|_{TV}.$$

For the proof of this estimate we compare on the interval $[\tau, \tau+T]$ the solution $\{\mu_t\}$ and the solution $\{\sigma_t\}$ to the Cauchy problem $\partial_t\sigma_t = L^*_{\overline{\mu}}\sigma_t$ with $\sigma_\tau = \mu_\tau$ and use that for σ_t convergence to $\overline{\mu}$ is already known.

Proof of Theorem 2.2. Let $\tau > 1$ and $T > 0$. Let $\{\sigma_t\}$ be the solution to the Cauchy problem

$$\partial_t\sigma_t = L^*_{\overline{\mu}}\sigma_t, \quad \sigma_\tau = \mu_\tau.$$

We have

$$\|W^k(\mu_t - \overline{\mu})\|_{TV} \le \|W^k(\sigma_t - \overline{\mu})\|_{TV} + \|W^k(\mu_t - \sigma_t)\|_{TV}.$$

Then by Lemma 3.7 and Lemma 3.9 there holds the inequality

$$\|W^k(\mu_t - \overline{\mu})\|_{TV} \le \lambda_1 e^{-\lambda_2(t-\tau)}\|W^k(\mu_\tau - \overline{\mu})\|_{TV} + \varepsilon N_4\Bigl(\int_\tau^t G(s)\|W^k(\mu_s - \overline{\mu})\|^2_{TV}\,ds\Bigr)^{1/2}. \quad (3.1)$$

We pick $T > 1$ such that

$$\lambda_1 e^{-\lambda_2 T} < 1/2.$$

From the inequalities $\varepsilon \le 1$ and $e^{-\lambda_2(t-\tau)} \le 1$ we obtain

$$\|W^k(\mu_t - \overline{\mu})\|_{TV}^2 \le 2\lambda_1^2 \|W^k(\mu_\tau - \overline{\mu})\|_{TV}^2 + 2N_4^2 \int_\tau^t G(s)\|W^k(\mu_s - \overline{\mu})\|_{TV}^2\, ds.$$

By Gronwall's inequality

$$\|W^k(\mu_t - \overline{\mu})\|_{TV}^2 \le 2\lambda_1^2 \|W^k(\mu_\tau - \overline{\mu})\|_{TV}^2 \exp\Bigl(2N_4^2 \int_\tau^{\tau+T} G(t)\, dt\Bigr).$$

Applying the estimate for the integral of the function G, we obtain

$$\|W^k(\mu_t - \overline{\mu})\|_{TV}^2 \le 2\lambda_1^2 \|W^k(\mu_\tau - \overline{\mu})\|_{TV}^2 \exp\Bigl(2N_4^2(N_3T + N_5)\Bigr).$$

Substituting this estimate into (3.1) and estimating the integral of the function G again, for $t = \tau + T$ we arrive at the inequality

$$\|W^k(\mu_{\tau+T} - \overline{\mu})\|_{TV} \le q\|W^k(\mu_\tau - \overline{\mu})\|_{TV},$$

where

$$q = \frac{1}{2} + \varepsilon N_4 \lambda_1 2^{1/2}(N_3T + N_5)^{1/2} \exp\Bigl(N_4^2(N_3T + N_5)\Bigr).$$

Let us pick ε so small that $q < 1$. By the fact that τ was arbitrary we obtain

$$\|W^k(\mu_{nT} - \overline{\mu})\|_{TV} \le q^n\|W^k(\mu_T - \overline{\mu})\|_{TV}.$$

It remains to observe that on $[0, T]$ the expression $\|W^k(\mu_t - \overline{\mu})\|_{TV}$ is bounded. We note that the constants q_1 and q_2 from the formulation of the theorem depend on $\|W^k\overline{\mu}\|_{TV}$. However, by the uniqueness of $\overline{\mu}$, which follows from the established convergence of μ_t to $\overline{\mu}$, we can assume that we consider the stationary solution constructed in Proposition 2.3. For this sttaionary solution the expression $\|W^k\overline{\mu}\|_{TV}$ is estimated by a quantity depending only on the constants from conditions (H1), (H2) and (H3). □

Remark 3.10. It is easy to see in the proof of Lemma 3.1 that, taking t sufficiently large, we can assume that the constant N_1 does not depend on the initial distribution ν. This implies that for τ sufficiently large all constants N_i appearing in the estimates of the solution μ_t on the interval $[\tau, \tau + T]$ do not depend on ν. Therefore, the number q from the proof of Theorem 2.2 does not depend on ν, hence the constant q_2 from the formulation of the theorem does not depend on the measure ν.

Remark 3.11. The assumption that ε is small is essential. Technically it is used in Lemma 3.1 for obtaining a priori estimates with Lyapunov functions and in the proof of Theorem 2.2, when we ensure the smallness of the factor q in the key estimate. It is important to bear in mind that without this assumption a stationary solution need not exist. Let us consider a simple example. Let $d = 1$, $A = I$ and $b(x, \mu) = -x + 1 + \varepsilon B(\mu)$, where

$$B(\mu) = \int_{\mathbb{R}} y\, \mu(dy).$$

It is readily verified that conditions (H1), (H2) and (H3) are fulfilled. However, for $\varepsilon = 1$ the stationary equation has no solutions, since for a stationary solution $\overline{\mu}$ the equality $(1 - \varepsilon)B(\overline{\mu}) = 1$ must hold.

In conclusion we formulate several problems, investigation of which is of interest in connection with the study of convergence of solutions of nonlinear Fokker–Planck–Kolmogorov equations to the stationary distribution.

1. For comparing the solution to the linear equation and the solution to the nonlinear equation an important role is played by the estimate of the integral of the

square of the logarithmic derivative of the density of one of the solutions. It is of great interest to obtain various estimates on the logarithmic derivative of the density of the solution in the case where the coefficients of the Fokker–Planck–Kolmogorov equation have a considerable growth at infinity.

2. We have established above convergence to the stationary solution in the case where the stationary solution is unique. It is important and interesting to study the case where there exist several stationary solutions. This case is not sufficienty studied (see a discussion of this problem in [35]).

3. It has been shown above that convergence holds in the case of a sufficiently small parameter ε. This actually means that the nonlinear perturbation is so small that it does not spoil the linear equation, whose solutions converge to the stationary solution. It would be interesting to obtain results on convergence in the case where the nonlinear part is not small and convergence is due to the nonlinearity.

This work is supported by the RFBR Grants 17-01-00662, 18-31-20008, the CRC 1283 at Bielefeld University, the DFG Grant RO 1195/12-1, and the Simons fundation.

References

[1] T.D. Frank, *Nonlinear Fokker–Planck equations, fundamentals and applications*, Springer-Verlag, Berlin, 2005.

[2] R.L. Dobrushin, “Vlasov equations”, *Funkt. Anal. i Pril.* **13**, No. 2, 48–58 (1979) (in Russian); English transl.: *Funct. Anal. Appl.* **13**, No. 2, 115–123 (1979).

[3] R.J. DiPerna, P.L. Lions, “Ordinary differential equations, transport theory and Sobolev spaces”, *Invent. Math.* **98**, 511–547 (1989).

[4] V.V. Kozlov, “The generalized Vlasov kinetic equations”, *Uspekhi Mat. Nauk* **63**, No. 4, 93–130 (2008) (in Russian); English transl.: *Russian Math. Surveys* **63**, No. 4, 691–726 (2008).

[5] V.V. Kozlov, “Kinetic Vlasov equation, dynamics of continua and turbulence”, *Nonlinear Dynamics* **6**, No. 3, 489–512 (2010).

[6] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Math. ETH Zürich, Birkhäuser, Basel (2005).

[7] J.A. Carrillo, M. Difrancesco, A. Figalli, T. Laurent, D. Slepcev, “Global-in-time weak measure solutions and finite-time aggregation for non-local interaction equations”, *Duke Math. J.* **156**, 229–271 (2011).

[8] M. Kac, “Foundations of kinetic theory”, In: Proc. 3rd Berkeley Sympos. Math. Statist. Probability, V. 3, pp. 171–197 (1956).

[9] H.P. McKean, “A class of Markov processes associated with nonlinear parabolic equations”, *Proc. Natl. Acad. Sci. USA* **56**, 1907–1911 (1966).

[10] H.P. McKean, “Propagation of chaos for a class of non-linear parabolic equations”, In: Lecture Series in Differential Equations, session 7, pp. 177–194, Catholic Univ. (1967).

[11] T. Funaki, “A certain class of duffusion processes associated with nonlinear parabolic equations”, *Z. Wahrscheinlichkeitstheorie verw. Geb.* **67**, 331–348 (1984).

[12] L. Ambrosio, “Transport equation and Cauchy problem for non-smooth vector fields”, Lecture Notes in Math. V. **1927**, 2–41 (2008).

[13] A. Figalli, “Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients”, *J. Funct. Anal.* **254**, No. 1, 109–153 (2008).

[14] D. Trevisan, “Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients”, *Electron. J. Probab.* **21**, Paper No. 22, 41 pp. (2016).

[15] V.I. Bogachev, M. Röckner, S.V. Shaposhnikov, “On the Ambrosio–Figalli–Trevisan superposition principle for probability solutions to Fokker–Planck–Kolmogorov equations”, ArXiv:1903.10834 (2019).

[16] V. Barbu, M. Röckner, “Probabilistic representation for solutions to nonlinear Fokker–Planck equations”, *SIAM J. Math. Anal.* **50**, No. 4, 4246–4260 (2018).

[17] Yu.S. Mishura, A.Yu. Veretennikov, “Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations”, ArXiv:1603.02212 (2016).

[18] S. Mehri, W. Stannat, “Weak solutions to Vlasov–McKean equations under Lyapunov-type conditions”, *Stochastics and Dynamics* (2019) DOI: 10.1142/S0219493719500424

[19] F.-Y. Wang, "Distribution dependent SDEs for Landau type equations", *Stoch. Processes Appl.* **128**, No. 2, 595–621 (2018).
[20] W. Hammersley, D. Siska, L. Szpruch, "McKean–Vlasov SDEs under measure dependent Lyapunov conditions", arXiv:1802.03974 (2018).
[21] L.G. Tonoyan, "Nonlinear elliptic equations for measures", *Dokl. Akad. Nauk* **439**, No. 2, 174–177 (2011) (in Russian); English transl.: *Dokl. Math.* **84**, No. 1, 558–561 (2011).
[22] O.A. Manita, S.V. Shaposhnikov, "Nonlinear parabolic equations for measures", *St. Petersburg Math. J.* **25**, No. 1, 43–62 (2014).
[23] O.A. Manita, M.S. Romanov, S.V. Shaposhnikov, "On uniqueness of solutions to nonlinear Fokker–Planck–Kolmogorov equations", *Nonlin. Anal.* **128**, 199–226 (2015).
[24] V.I. Bogachev, M. Röckner, S.V. Shaposhnikov, "Distances between transition probabilities of diffusions and applications to nonlinear Fokker–Planck–Kolmogorov equations", *J. Funct. Anal.* **271**, 1262–1300 (2016).
[25] S.V. Shaposhnikov, "Nonlinear Fokker–Planck–Kolmogorov equations for measures", In: Stochastic Partial Differential Equations and Related Fields, pp. 367–379, Springer Proc. Math. Stat., **229**, Springer, Cham, 2018.
[26] V.I. Bogachev, G. Da Prato, M. Röckner, S.V. Shaposhnikov, "Nonlinear evolution equations for measures on infinite dimensional spaces", In: Stochastic Partial Differential Equations and Applications. Quaderni di Matematica, Dipartimento di Matematica Seconda Universita di Napoli Napoli, **25**, 51–64 (2010).
[27] O.A. Manita, "Nonlinear Fokker–Planck–Kolmogorov equations in Hilbert spaces", *J. Math. Sci. (New York)* **216**, No. 1, 120–135 (2016).
[28] V.I. Bogachev, N.V. Krylov, M. Röckner, S.V. Shaposhnikov, *Fokker–Planck–Kolmogorov equations*, Amer. Math. Soc., Providence, Rhode Island, 2015.
[29] N.U. Ahmed, X. Ding, "On invariant measures of nonlinear Markov processes", *J. Appl. Math. Stoch. Anal.* **6**, No. 4, 385–406 (1993).
[30] Y. Tamura, "On asymptotic behaviors of the solution of a nonlinear diffusion equation", *J. Fac. Sci. Univ. Tokyo, Sect. 1a, Math.* **31**, 195–221 (1984).
[31] Y. Tamura, "Free energy and the convergence of distributions of diffusion processes of McKean type", *J. Fac. Sci. Univ. Tokyo, Sect. 1a, Math.* **34**, 443–484 (1987).
[32] A.Yu. Veretennikov, "On ergodic measures for McKean–Vlasov stochastic equations", In: Niederreiter H., Talay D. (eds.) Monte Carlo and Quasi-Monte Carlo Methods, pp. 471–486. Springer, Berlin, 2004.
[33] O.A. Butkovsky, "On ergodic properties of nonlinear Markov chains and stochastic McKean–Vlasov equations", *Theory Probab. Appl.* **58**, 661–674 (2014).
[34] A. Eberle, A. Guillin, R. Zimmer, "Quantitative Harris type theorems for diffusions and McKean–Vlasov processes", ArXiv:1606.06012 (2016).
[35] V.I. Bogachev, M. Röckner, S.V. Shaposhnikov, "Convergence in variation of solutions of nonlinear Fokker–Planck–Kolmogorov equations to stationary measures", *J. Funct. Anal.* (2019).
[36] V.I. Bogachev, M. Röckner, S.V. Shaposhnikov, "The Poisson equation and estimates for distances between stationary distributions of diffusions", *J. Math. Sci. (New York)* **232**, No. 3, 254–282 (2018).
[37] V.I. Bogachev, A.I. Kirillov, S.V. Shaposhnikov, "Distances between stationary distributions of diffusions and solvability of nonlinear Fokker–Planck–Kolmogorov equations", *Teor. Veroyatn. Primen.* **62**, No. 1, 16–43 (2017) (in Russian); English transl.: *Theory Probab. Appl.* **62**, No. 1, 12–34 (2018).
[38] M. Ji, Z. Shen, Y. Yi, "Convergence to equilibrium in Fokker–Planck equations", *J. Dynamics Differ. Equ.* (2018). Doi: 10.1007/s10884-018-9705-8.
[39] M. Hairer, J.C. Mattingly, "Yet another look at Harris' ergodic theorem for Markov chains", In: Seminar on Stochastic Analysis, Random Fields and Applications VI, 2011, Progress in Probab., V. 63, pp. 109–117.

V.I. Bogachev: Lomonosov Moscow State University, 119991 Moscow, Russia; National Research University Higher School of Economics, Moscow, Russia

M. Röckner: Bielefeld University, Bielefeld, 33501, Germany

S.V. Shaposhnikov: Lomonosov Moscow State University, 119991 Moscow, Russia; National Research University Higher School of Economics, Moscow, Russia