Estimates of distances between solutions of Fokker–Planck–Kolmogorov equations with partially degenerate diffusion matrices

Oxana A. Manita [a], Maxim S. Romanov[a], Stanislav V. Shaposhnikov[a,b] [1]

[a] Department of Mechanics and Mathematics, Moscow State University, Moscow, Russia; [b] National Research University Higher School of Economics, Moscow, Russia.

Using a metric which interpolates between the Kantorovich metric and the total variation norm we estimate the distance between solutions of Fokker–Planck–Kolmogorov equations with degenerate diffusion matrices. Some relations between the degeneracy of a diffusion matrix and the regularity of a drift coefficient are analysed. Applications to nonlinear Fokker–Planck–Kolmogorov equations are given.

Keywords: Fokker–Planck–Kolmogorov equation, Degenerate diffusion matrix.

MSC: 35K10, 35K55, 60J60.

We study the Cauchy problem for the Fokker–Planck–Kolmogorov equation:

$$\partial_t\mu = L^*_{A,b}\mu, \quad \mu|_{t=0} = \mu_0, \tag{1}$$

where μ_0 is a probability measure on $\mathbb{R}^d$ and the operators $L_{A,b}$ and $L^*_{A,b}$ are given by the following expressions

$$L_{A,b}u(t,x) = \sum_{i,j=1}^d a^{ij}(t,x)\partial_{x_i}\partial_{x_j}u(x) + \sum_{i=1}^d b^i(t,x)\partial_{x_i}u(x),$$

$$L^*_{A,b}u(t,x) = \sum_{i,j=1}^d \partial_{x_i}\partial_{x_j}\Big(a^{ij}(t,x)u(x)\Big) - \sum_{i=1}^d \partial_{x_i}\Big(b^i(t,x)u(x)\Big).$$

Further assume that $A(x,t) = (a^{ij}(x,t))_{i,j\le d}$ is a nonnegative symmetric matrix (called the diffusion matrix) with Borel measurable entries, $b(x,t) = (b^i(x,t))_{i=1}^d\colon\ [0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ is a Borel measurable mapping (called the drift coefficient) and a solution $\mu = \mu_t(dx)\,dt$ is given by a family of probability measures μ_t on $\mathbb{R}^d$.

The goal of this paper is to estimate the distance (with respect to a suitable metric) between two solutions $\mu = \mu_t(dx)\,dt$ and $\sigma = \sigma_t(dx)\,dt$ to Fokker–Planck–Kolmogorov equations

$$\partial_t\mu = L^*_{A,b_\mu}\mu \quad \text{and} \quad \partial_t\sigma = L^*_{A,b_\sigma}\sigma$$

with different drifts b_μ and b_σ. The diffusion matrix A is allowed to be fully degenerate. Furthermore we analyse some relations between the degeneracy of the diffusion matrix and the regularity of the drift coefficient. Let us consider two different cases: $A = I$ and $A = 0$. In the first case the estimate

$$\|\mu_t - \sigma_t\|_{TV} \le \|\mu_0 - \sigma_0\|_{TV} + \Big(\int_0^t\int_{\mathbb{R}^d}|b_\mu - b_\sigma|^2\,d\sigma_s\,ds\Big)^{1/2},$$

where $\|\ \|_{TV}$ is the total variation norm, was established in [6, Remark 2.3] for locally bounded coefficients $b^i_\mu, b^i_\sigma \in L^1(\mu+\sigma)$. Note that even the equations with different

[1]Corresponding author

e-mails: oxana.manita@gmail.com (O.A.Manita), mcliz@mail.ru (M.S.Romanov), starticle@mail.ru (S.V.Shaposhnikov).

diffusion matrices were investigated in [6]. In the second case for the Lipschitzian drifts b_μ and b_σ the estimate

$$W(\mu_t, \sigma_t) \le W(\mu_0, \sigma_0) + C\int_0^t \int_{\mathbb{R}^d} |b_\mu - b_\sigma|\, d\sigma_s\, ds,$$

where W is the Kantorovich metric

$$W(\mu^1, \mu^2) = \sup\Bigl\{\int_{\mathbb{R}^d} \varphi\, d(\mu^1 - \mu^2)\colon\ |\varphi| \le 1,\ |\varphi(x) - \varphi(y)| \le |x - y|\Bigr\},$$

can be derived directly from the expressions for the solutions μ_t and σ_t. We stress that the last estimate does not hold for continuous drifts b_μ and b_σ. Moreover the Kantorovich metric cannot be replaced by the total variation norm. The aim of our paper is to study the intermediate case:

$$L_{A,b}u = \sum_{i=1}^{p} \partial_{x_i}^2 u + \sum_{i=1}^{d} b^i \partial_{x_i} u, \quad 0 \le p \le d.$$

In particular we obtain the following estimate. Suppose that b_μ (not b_σ) is a Lipschitz mapping with respect to $(x_{p+1}, \ldots, x_d)$; then the estimate

$$d_p(\mu_t, \sigma_t) \le K d_p(\mu_0, \sigma_0) + K\int_0^t \int_{\mathbb{R}^d} \sum_{i=p+1}^{d} |b^i_\mu - b^i_\sigma|\, d\sigma_s\, ds +$$

$$+K\Bigl(\int_0^t \int_{\mathbb{R}^d} \sum_{i=1}^{p} |b^i_\mu - b^i_\sigma|^2\, d\sigma_s\, ds\Bigr)^{1/2}\Bigl(1+\int_0^t \int_{\mathbb{R}^d} \Bigl[\sum_{i=1}^{p} |b^i_\mu - b^i_\sigma|^2 + \sum_{i=p+1}^{d} |b^i_\mu - b^i_\sigma|\Bigr]\, d\sigma_s\, ds\Bigr)^{1/2}$$

holds true under the condition that $\sum_{i=1}^{p}(|b^i_\mu|^2 + |b^i_\sigma|^2)$ and $\sum_{i=p+1}^{d}(|b^i_\mu| + |b^i_\sigma|)$ are integrable with respect to $\mu + \sigma$. Here the metric d_p is defined by the following way:

$$d_p(\mu^1, \mu^2) = \sup_{\psi} \int_{\mathbb{R}^d} \psi\, d(\mu^1 - \mu^2),$$

where $\psi \in C(\mathbb{R}^d)$, $|\psi| \le 1$ and $|\psi(x+h_p) - \psi(x)| \le |h_p|$ for all $h_p = (0, \ldots, 0, y_{p+1}, \ldots, y_d)$. The main novelty is the investigation of degenerate Fokker–Planck–Kolmogorov equations for measures with nonsmooth unbounded coefficients. In addition we obtain new existence and uniqueness conditions for nonlinear Fokker–Planck–Kolmogorov equations. Since the equations are degenerate the solutions μ and σ do not possess densities with respect to Lebesgue measure. Thus the approach from [6] cannot be applied here and we use the approximative Holmgren method which was developed in [4] and [5]. The main difficulty is to obtain the gradient estimate for a solution of the adjoint equation. The drifts b_μ and b_σ are irregular mappings and we cannot obtain the required estimate by the maximum principle directly. Let us remark that we do not assume that b_μ and b_σ are locally bounded or locally integrable with respect to Lebesgue measure. Thus even in the case $p = d$ our result seems to be new.

Equations with partially degenerate diffusion matrices arise in the Vlasov–Fokker–Planck systems and play a crucial role in physics (see, for instance [14], [7]). The uniqueness of solutions of linear equations with degenerate diffusion matrix is investigated in [3]. The estimates of the total variation and Kantorovich distances between solutions are given in [6] and [10]. In [9] the authors present quantitative stability estimates for the solutions of degenerate Fokker–Planck equations in L^p. The pointwise estimates for the difference of two transition densities of diffusions are given in [8]. In [2], a survey of results about Fokker–Planck–Kolmogorov linear equations is presented. In [12] and [11] the existence and uniqueness of solutions to nonlinear Fokker–Planck–Kolmogorov equations are investigated.

Let us explain precisely our framework.

A Borel measure μ on $[0,T]\times\mathbb{R}^d$ is given by a family of probability measures $(\mu_t)_{t\in[0,T]}$ if $\mu_t\ge 0$, $\mu_t(\mathbb{R}^d)=1$, for every Borel set B the mapping $t\to\mu_t(B)$ is measurable and for every $u\in C_0^\infty((0,T)\times\mathbb{R}^d)$ one has

$$\int_{[0,T]\times\mathbb{R}^d} u\,d\mu=\int_0^T\int_{\mathbb{R}^d}u(x,t)\,\mu_t(dx)\,dt.$$

We write $\mu(dxdt)=\mu_t(dx)\,dt$ or $\mu=\mu_t\,dt$.

We say that a measure $\mu=\mu_t\,dt$ given by a family of probability measures μ_t satisfies the Cauchy problem

$$\partial_t\mu=L^*_{A,b}\mu,\quad \mu|_{t=0}=\mu_0 \tag{2}$$

if $a^{ij},b^i\in L^1([0,T]\times U,\mu)$ for every ball $U\subset\mathbb{R}^d$ and for every function u such that $u(x,t)\equiv 0$ if $|x|\ge R$ for some $R>0$ and $u\in C^{1,2}_{t,x}((0,T)\times\mathbb{R}^d)\bigcap C([0,T]\times\mathbb{R}^d)$ the equality

$$\int_{\mathbb{R}^d}u(x,t)\,\mu_t(dx)=\int_{\mathbb{R}^d}u(x,0)\,d\nu+\int_0^t\int_{\mathbb{R}^d}\big[\partial_t u+L_{A,b}u\big]\,\mu_s(dx)\,ds \tag{3}$$

holds for every $t\in[0,T]$.

Suppose that for a number $\lambda>0$ and some integer $0\le p\le d$ one has

(H1) $\langle A(x,t)\xi,\xi\rangle\ge\lambda\sum_{i=1}^p\xi_i^2$ for all $x,\xi\in\mathbb{R}^d$ and $t\in[0,T]$, where the right hand side equals zero if $p=0$.

Let μ be a bounded Borel measure on $[0,T]\times\mathbb{R}^d$. For $p\ge1$ we denote by μ^p the projection of μ on the first p coordinates $x_1,\dots,x_p$ and t, that is $\mu^p(B)=\mu(B\times\mathbb{R}^{d-p})$ for every Borel set $B\subset[0,T]\times\mathbb{R}^p$.

Proposition 1. *Let $p\ge1$. Suppose that $\mu=\mu_t\,dt$, is a solution of the Cauchy problem* (2) *and μ_t is a family of probability measures on $\mathbb{R}^d$. Suppose also that the diffusion matrix A satisfies the condition* (H1) *and $a^{ij},b^i\in L^1(\mu,[0,T]\times\mathbb{R}^d)$. Then the measure μ^p has a density $\varrho(t,x_1,\dots,x_p)$ with respect to Lebesgue measure on $(0,T)\times\mathbb{R}^d$ and ϱ belongs to $L^{(p+1)/p}_{loc}((0,T)\times\mathbb{R}^p)$.*

Proof. Since a^{ij}, b^i belong to $L^1(\mu,[0,T]\times\mathbb{R}^d)$ we see that the definition of a solution μ holds true for every smooth bounded u that depends only on $x_1,\dots,x_p$ and t. It follows that for every $u\in C_0^\infty((0,T)\times\mathbb{R}^p)$ we have

$$\int_0^T\int_{\mathbb{R}^p}[\partial_t u+\sum_{i,j=1}^p\tilde{a}^{ij}\partial_{x_i}\partial_{x_j}u]\,d\mu^p_t\,dt\le C(\sup|u|+\sup|\nabla_x u|),$$

where $\tilde{a}^{ij}=\mathbb{E}(a^{ij}|\sigma_p)$ and σ_p is generated by $t,x_1,\dots,x_p$. Applying [2, Theorem 6.3.1] we obtain that $(\det\tilde{A})^{1/(p+1)}\cdot\mu^p$ has a density $\varrho\in L^{(p+1)/p}_{loc}((0,T)\times\mathbb{R}^p)$. By (H1) we can find a set $I\subset(0,T)\times\mathbb{R}^p$ such that $\mu^p(I)=1$ and $\langle\tilde{A}(t,x_1,\dots,x_p)\xi,\xi\rangle\ge\lambda|\xi|^2$ for every $(t,x_1,\dots,x_p)\in I$ and every $\xi\in\mathbb{R}^p$. This implies that $\det\tilde{A}(t,x_1,\dots,x_p)\ge\lambda^p>0$ for every $(t,x_1,\dots,x_p)\in I$ and μ^p has a density. $\square$

Suppose also that

(H2) a^{ij} are bounded continuous functions having two bounded continuous spatial derivatives and

$$\sum_{k=p+1}^d(SA(x,t)S)^{kk}\ge\gamma\sum_{k=p+1}^d|\mathrm{tr}(\partial_{x_k}A(x,t)S)|^2$$

for some $\gamma>0$ and every symmetric matrix S.

We stress that according to [15, Lemma 3.2.3] the last inequality holds true if $p=0$. Let us illustrate the case $p\ge1$.

Example 1. Let the diffusion matrix A have the form

$$\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}$$

where R is a symmetric $p \times p$ matrix, $\langle R\xi, \xi\rangle \geq \lambda|\xi|^2$ for every $\xi \in \mathbb{R}^p$ and R depends only on $x_1, \dots, x_p$. It is clear that A satisfies (H1) and (H2).

Example 2. Let A have the following form

$$\begin{pmatrix} R & 0 \\ 0 & Q \end{pmatrix}$$

where $R = (r^{ij})$ is the same as above, $Q = (q^{ij})$ is a symmetric and nonnegative matrix. Let us check that A satisfies (H2). Note that

$$(SAS)^{kk} = \sum_{1\leq ij\leq d} a^{ij} s_{ik} s_{jk} = \sum_{1\leq i,j\leq p} r^{ij} s_{ik} s_{jk} + \sum_{p+1\leq i,j\leq d} q^{ij} s_{ik} s_{jk},$$

where the last term can be represented in the form $(ZQZ)^{kk}$, $Z = (s_{ml})_{p+1\leq m,l\leq d}$. Applying [15, Lemma 3.2.3] we obtain the inequality

$$\mathrm{tr}(ZQZ) \geq \gamma \sum_k |tr(\partial_{x_k} QZ)|^2$$

for some $\gamma > 0$. Since R does not depend on $x_{p+1}, \dots, x_d$ we see that

$$\sum_{k=p+1}^{d} |\mathrm{tr}(\partial_{x_k} A(x,t)S)|^2 = \sum_k |tr(\partial_{x_k} QZ)|^2.$$

It follows that (H2) is fulfilled.

Example 3. Let A have the form

$$\begin{pmatrix} R & Y \\ Y & Q \end{pmatrix}$$

where symmetric and nonnegative matrixes $R = (r^{ij})_{1\leq i,j\leq p}$ and $Q = (q^{ij})_{p+1\leq i,j\leq d}$ do not depend on $x_{p+1}, \dots, x_d$ and A satisfies (H1). Let us prove that A satisfies (H2). The condition (H1) implies that $(SAS)^{kk} \geq \lambda \sum_{i=1}^{p} s_{ik}^2$. Furthermore the inequality

$$8^{-1}|\mathrm{tr}(\partial_{x_k} A(x,t)S)|^2 \leq \sum_{1\leq i\leq p, p+1\leq j\leq d} |\partial_{x_k} y^{ij}|^2 |s_{ij}|^2$$

holds for every $k \geq p+1$. Taking into account that $|\partial_{x_k} y^{ij}|$ are bounded functions we obtain (H2).

Example 4. Assume that the matrix A has the same form as in Example 3, Q depends on $x_1, \dots, x_d$, R does not depend on $x_{p+1}, \dots, x_d$. Assume also that the inequality

$$\langle A\xi, \xi\rangle \geq \lambda \sum_{i=1}^{p} \xi_i^2 + \alpha \sum_{i,j=p+1}^{d} q^{ij} \xi_i \xi_j$$

holds for every $\xi \in \mathbb{R}^d$. Let us show that A satisfies (H2).

Indeed, for every symmetric matrix S we have

$$\sum_{k=p+1}^{d} \sum_{i,j=1}^{d} s_{ki} a^{ij} s_{jk} \geq \lambda \sum_{k=p+1}^{d} \sum_{i=1}^{p} s_{ik}^2 + \alpha \sum_{i,j,k=p+1}^{d} q^{ij} s_{ik} s_{jk}.$$

On the other hand, we obtain the inequality

$$\sum_{k=p+1}^{d}\Big(\sum_{i,j=1}^{d}\partial_{x_k}a^{ij}s_{ji}\Big)^2\le\sum_{k=p+1}^{d}\Big[8\Big(\sum_{1\le i\le p,p+1\le j\le d}\partial_{x_k}y^{ij}s_{ji}\Big)^2+2\Big(\sum_{i,j=p+1}\partial_{x_k}q^{ij}s_{ji}\Big)^2\Big]\le$$

$$\le C\left(\sum_{1\le i\le p,p+1\le j\le d}|\partial_{x_k}y^{ij}|^2|s_{ij}|^2+\sum_{i,j,l=p+1}^{d}q^{ij}s_{il}s_{jl}\right).$$

Thus (H2) is fulfilled.

Example 5. Suppose that $a^{i_0j_0}_{x_k}\neq 0$, for some $i_0,j_0\le p$ and $k>p$; then A does not satisfy (H2). Let $S=(s_{ij})$, $s_{i_0j_0}=s_{j_0i_0}=1$ and $s_{ij}=0$ otherwise. It is easy to prove that

$$\sum_{k=p+1}^{d}(SA(x,t)S)^{kk}=0\quad\text{and}\quad\sum_{k=p+1}^{d}|\mathrm{tr}(\partial_{x_k}A(x,t)S)|^2>0.$$

Recall that

$$d_p(\mu^1,\mu^2)=\sup\Big\{\int_{\mathbb{R}^d}\psi\,d(\mu^1-\mu^2)\colon\ \psi\in C(\mathbb{R}^d),\ |\psi(x)|\le 1,|\psi(x+h_p)-\psi(x)|\le|h_p|\Big\}.$$

Let us formulate our main result.

Theorem 1. *Assume that $\mu=\mu_t\,dt$ and $\sigma=\sigma_t\,dt$ are two solutions of the Cauchy problems* (2) *with the initial conditions μ_0 and σ_0 and with the operators L_{A,b_μ} and L_{A,b_σ}, where A satisfies* (H1), (H2). *Assume also that there exists $\Lambda>0$ such that*

$$|b_\mu(t,x)-b_\mu(t,x+h_p)|\le\Lambda|h_p|\quad\forall\,h_p=(0,\dots,0,y_{p+1},\dots,y_d)$$

and

$$\sum_{i=1}^{p}|b^i_\mu|^2,\quad\sum_{i=1}^{p}|b^i_\sigma|^2,\quad\sum_{i=p+1}^{d}|b^i_\mu|,\quad\sum_{i=p+1}^{d}|b^i_\mu|\quad\textit{belong to}\quad L^1(\mu+\sigma).$$

Then there exists a number $K=K(T,\lambda,\Lambda,\gamma)>0$ such that the estimate

$$d_p(\mu_t,\sigma_t)\le Kd_p(\mu_0,\sigma_0)+K\int_0^t\int_{\mathbb{R}^d}\sum_{i=p+1}^{d}|b^i_\mu-b^i_\sigma|\,d\sigma_s\,ds+$$

$$+K\Big(\sum_{i=1}^{p}\int_0^t\int_{\mathbb{R}^d}|b^i_\mu-b^i_\sigma|^2\,d\sigma_s\,ds\Big)^{1/2}\Big(1+\int_0^t\int_{\mathbb{R}^d}\Big[\sum_{i=1}^{p}|b^i_\mu-b^i_\sigma|^2+\sum_{i=p+1}^{d}|b^i_\mu-b^i_\sigma|\Big]\,d\sigma_s\,ds\Big)^{1/2}.$$

holds for every $t\in[0,T]$.

The proof of Theorem 1 is based on the two lemmas below.

Lemma 1. *Let f be a smooth bounded solution of the Cauchy problem*

$$\partial_t f+\sum_{i,j=1}^{d}a^{ij}\partial_{x_i}\partial_{x_j}f+\sum_{i=1}^{d}h^i\partial_{x_i}f=0,\quad f(T,x)=\psi(x).$$

Suppose that ψ and h are smooth bounded functions having bounded derivatives and

$$|\psi|\le 1,\quad|\partial_{x_k}\psi|\le 1,\quad|\partial_{x_k}h|\le\Lambda,$$

for some $\Lambda>0$ and every $k=p+1,p+2,\dots,d$. Then

$$|f(t,x)|^2+\sum_{k=p+1}^{d}|\partial_{x_k}f(t,x)|^2\le(1+d)e^{2Mt},\quad M=2\lambda^{-1}d^2\Lambda^2+4^{-1}\gamma^{-1}.$$

Proof. It is easily shown that the function $v = (f^2 + \sum_{k=p+1}^d |\partial_{x_k} f|^2)/2$ satisfies the following equation

$$\partial_t v + \sum_{i,j=1}^d a^{ij}\partial_{x_i}\partial_{x_j} v + \sum_{i=1}^d h^i \partial_{x_i} v = Q,$$

where

$$Q = \sum_{1\le i,j\le d} a^{ij}\partial_{x_i} f \partial_{x_j} f + \sum_{k=p+1}^d \sum_{1\le i,j\le d} a^{ij}\partial_{x_i}\partial_{x_k} f \partial_{x_j}\partial_{x_k} f$$
$$- \sum_{k=p+1}^d \sum_{1\le i,j\le d} \partial_{x_k} a^{ij}\partial_{x_i}\partial_{x_j} f \partial_{x_k} f - \sum_{k=p+1}^d \sum_{1\le i\le d} \partial_{x_k} h^i \partial_{x_i} f \partial_{x_k} f.$$

Let $u = (\partial_{x_i} f)_{1\le i\le d}$ and $S = (\partial_{x_i}\partial_{x_j} f)$. The expression Q can be represented in the form

$$Q = \langle Au, u\rangle + \sum_{k=p+1}^d (SAS)^{kk} - \sum_{k=p+1}^d \mathrm{tr}(\partial_{x_k} AS)\partial_{x_k} f - \sum_{k=p+1}^d \sum_{1\le i\le d} \partial_{x_k} h^i \partial_{x_i} f \partial_{x_k} f.$$

Taking into account the estimates

$$\sum_{k=p+1}^d \sum_{1\le i\le d} \partial_{x_k} h^i \partial_{x_i} f \partial_{x_k} f \le d\Lambda |u| \Bigl(\sum_{k=p+1}^d |\partial_{x_k} f|^2\Bigr)^{1/2} \le \lambda |u|^2 + 2\lambda^{-1} d^2 \Lambda^2 v$$

and

$$\sum_{k=p+1}^d \mathrm{tr}(\partial_{x_k} AS)\partial_{x_k} f \le \gamma \sum_{k=p+1}^d |\mathrm{tr}(\partial_{x_k} AS)|^2 + 4^{-1}\gamma^{-1} v,$$

we obtain the inequality

$$Q \ge -(2\lambda^{-1} d^2 \Lambda^2 + 4^{-1}\gamma^{-1})v.$$

Consequently the function v satisfies the inequality

$$\partial_t v + \sum_{i,j=1}^d a^{ij}\partial_{x_i}\partial_{x_j} v + \sum_{i=1}^d h^i \partial_{x_i} v + (2\lambda^{-1} d^2\Lambda^2 + 4^{-1}\gamma^{-1})v \ge 0$$

and the required estimate follows from the maximum principle (see Theorem 3.1.1 [15]). □

Lemma 2. *Let $p \ge 1$ and μ be a bounded nonnegative Borel measure on $[0,T]\times\mathbb{R}^d$. Suppose that the projection μ^p of the measure μ on the first p coordinates $x_1,\dots,x_p$ and t has a density $\varrho \in L^q_{loc}((0,T)\times\mathbb{R}^d)$, where $q > 1$. Suppose also that a measurable function $f \in L^r(\mu)$, where $r \ge 1$, satisfies the following condition:*

(*) *there exists $\Lambda > 0$ such that*

$$|f(t,x) - f(t,x+h_p)| \le \Lambda |h_p| \quad \forall\ h_p = (0,\dots,0,y_{p+1},\dots,y_d).$$

Then there exists a sequence of smooth bounded functions f_n with bounded derivatives such that $\|f - f_n\|_{L^r(\mu)} \to 0$ and

$$|f_n(t,x) - f_n(t,x+h_p)| \le 4\Lambda |h_p| \quad \forall\ h_p = (0,\dots,0,y_{p+1},\dots,y_d).$$

Proof. For simplicity we use the notation $z = (x_1,\dots,x_p)$ and $y = (x_{p+1},\dots,x_d)$.

First let us prove that f can be approximated by the function g such that g satisfies the condition (*), $g(t,z,y) = 0$ if $|z| > R$, $t < \kappa$ or $t > T - \kappa$, and $|g(t,z,y)| \le C$ for some $R > 0$, $\kappa > 0$ and $C > 0$.

Let $I_N(t,z) = 1$ if $|z| < 1/N$, $t \in [N^{-1}, T - N^{-1}]$, and $I_N(t,z) = 0$ otherwise. Let us consider the function $g_N(t,z,y) = I_N(t,z)G_N(f(t,z,y))$, where $G_N(v) = v$ if $|v| \le N$

and $G_N(v) = N\operatorname{sign} v$ if $|v| > N$. Since $|G_N(v_1) - G_N(v_2)| \le |v_1 - v_2|$ the function g_N satisfies (*). By the estimate $|g_N| \le |f|$ and the Lebesgue dominated convergence theorem we have $\|g_N - f\|_{L^r(\mu)} \to 0$ as $N \to \infty$.

Now let us prove that the function g can be approximated by the function η such that η satisfies the condition (*) with 2Λ, $|\eta| \le C$, $\eta(t,z,y) = 0$ if $|z| > R$ or $|y| > R_1$, $t < \kappa$ or $t > T - \kappa$ for some positive numbers R, R_1, κ and C.

Let $\varphi \in C_0^\infty(\mathbb{R}^{d-p})$, $0 \le \varphi \le 1$, $|\nabla\varphi| \le 1$, $\varphi(y) = 1$ if $|y| \le 1$ and $\varphi(y) = 0$ if $|y| > 2$. Let us approximate g by $\eta_M(t,z,y) = \varphi_M(y)g(t,z,y)$, where $\varphi_M(y) = \varphi(y/M)$. Applying the condition (*) we obtain $|g(t,z,y)| \le |g(t,z,0)| + \Lambda|y| \le C + \Lambda|y|$. Let $CM^{-1} < \Lambda$. Then we obtain the estimates

$$|\eta_M(t,z,y) - \eta_M(t,z,y')| \le (CM^{-1} + \Lambda)|y - y'| \le 2\Lambda_f|y - y'|.$$

Moreover, $\|\eta_M - g\|_{L^r(\mu)} \to 0$ as $M \to \infty$.

Finally let us prove that η can be approximated by the required sequence f_n. We can assume that η is a smooth function with respect to y.

Let $\varepsilon > 0$ and $\delta > 0$. Let $\eta(t,z,y) = 0$ if $t < \kappa$ or $t > T - \kappa$ and

$$\eta_\delta(t,z,y) = \int_0^T \int_{\mathbb{R}^p} \omega_\delta(z - v, t - s)\eta(s,v,y)\,dv\,ds,$$

where $\omega_\delta(x,t) = \delta^{-p-1}\omega_1(x/\delta)\omega_2(t/\delta)$ and $\omega_1 \in C_0^\infty(\mathbb{R}^p)$, $\omega_2 \in C_0^\infty(\mathbb{R})$, $0 \le \omega_1 \le 1$, $0 \le \omega_2 \le 1$, $\|\omega_1\|_{L^1} = 1$, $\|\omega_2\|_{L^1} = 1$. There exists a family of Borel sets $\{B_j\}_{j=1}^J$ such that $B_j \subset \mathbb{R}^{d-p}$, $B_j \cap B_i = \emptyset$, $\{y\colon |y| \le R\} \subset \cup_j B_j$ and $\sup_{z,y\in B_j}|z - y| \le \varepsilon$. Let us take a point $y_j \in B_j$. Applying the condition (ii) we obtain

$$\|\eta_\delta - \eta\|^r_{L^r(\mu)} \le \sum_{j=1}^J \int_{([0,T]\times\mathbb{R}^p)\times B_j} |\eta_\delta(t,z,y_j) - \eta(t,z,y_j)|^r\,d\mu + C(r)\Lambda^r\varepsilon^r \le$$

$$\le J\int_{[0,T]\times\mathbb{R}^p} |\eta_\delta(t,z,y_j) - \eta(t,z,y_j)|^r\,d\mu^p + C(r)\Lambda^r\varepsilon^r$$

Since the mapping $(t,z) \to \eta(t,z,y)$ is bounded and $\mu_p = \varrho\,dx\,dt$, where $\varrho \in L^q_{loc}$ and $q > 1$, we can find a number $\delta > 0$ such that $\|\eta_\delta - \eta\|^r_{L^r(\mu)} \le \varepsilon + C(r)\Lambda^r\varepsilon^r$. $\square$

Proof of Theorem 1. Let f be a solution of the Cauchy problem

$$\partial_t f + \sum_{i,j=1}^d a^{ij}\partial_{x_i}\partial_{x_j} f + \sum_{i=1}^d h^i\partial_{x_i} f = 0, \quad f(T,x) = \psi(x),$$

where h and ψ satisfy the conditions of Lemma 1. Substituting u for f in (3) we get for the difference of the solutions $\mu = \mu_t\,dt$ and $\sigma = \sigma_t\,dt$ the following equality

$$\int_{\mathbb{R}^d} \psi\,d(\mu_t - \sigma_t) = \int_{\mathbb{R}^d} f\,d(\mu_0 - \sigma_0) + \int_0^t \int_{\mathbb{R}^d} \langle b_\mu - h, \nabla f\rangle\,d\mu_s\,ds - \int_0^t \int_{\mathbb{R}^d} \langle b_\sigma - h, \nabla f\rangle\,d\sigma_s\,ds.$$

Applying the maximum principle and Lemma 1 we obtain the estimates

$$|f(x,t)| \le 1, \quad \sum_{k=p+1}^d |\partial_{x_k} f(x,t)|^2 \le C_1^2$$

for some $C_1 > 0$. By the definition of d_p we have

$$\int_{\mathbb{R}^d} f\,d(\mu_0 - \sigma_0) \le (1 + C_1)d_p(\mu_0,\sigma_0).$$

Applying the Cauchy inequality we get

$$\int_0^t \int_{\mathbb{R}^d} \langle b_\mu - h, \nabla f\rangle \, d\mu_s \, ds \leq$$
$$\leq \Big(\int_0^t \int_{\mathbb{R}^d} \sum_{i=1}^p |b_\mu^i - h^i|^2 \, d\mu_s \, ds\Big)^{1/2} \Big(\int_0^t \int_{\mathbb{R}^d} \sum_{i=1}^p |\partial_{x_i} f|^2 \, d\mu_s \, ds\Big)^{1/2} +$$
$$+ C_1 \int_0^t \int_{\mathbb{R}^d} \sum_{i=p+1}^d |b_\mu^i - h^i| \, d\mu_s \, ds.$$

Furthermore, we have

$$-\int_0^t \int_{\mathbb{R}^d} \langle b_\sigma - h, \nabla f\rangle \, d\sigma_s \, ds \leq$$
$$\leq \Big(\int_0^t \int_{\mathbb{R}^d} |b_\mu^i - b_\sigma^i|^2 \, d\sigma_s \, ds\Big)^{1/2} \Big(\int_0^t \int_{\mathbb{R}^d} \sum_{i=1}^p |\partial_{x_i} f|^2 \, d\sigma_s \, ds\Big)^{1/2}$$
$$+ C_1 \int_0^t \int_{\mathbb{R}^d} \sum_{i=p+1}^d |b_\mu^i - b_\sigma^i| \, d\sigma_s \, ds - \int_0^t \int_{\mathbb{R}^d} \langle b_\mu - h, \nabla f\rangle \, d\sigma_s \, ds,$$

where the last term is estimated in the following way:

$$-\int_0^t \int_{\mathbb{R}^d} \langle b_\mu - h, \nabla f\rangle \, d\sigma_s \, ds \leq$$
$$\leq \Big(\int_0^t \int_{\mathbb{R}^d} \sum_{i=1}^p |b_\mu^i - h^i|^2 \, d\sigma_s \, ds\Big)^{1/2} \Big(\int_0^t \int_{\mathbb{R}^d} \sum_{i=1}^p |\partial_{x_i} f|^2 \, d\sigma_s \, ds\Big)^{1/2} +$$
$$+ C_1 \int_0^t \int_{\mathbb{R}^d} \sum_{i=p+1}^d |b_\mu^i - h^i| \, d\sigma_s \, ds.$$

Let us estimate the following expression

$$\int_0^t \int_{\mathbb{R}^d} \sum_{k=1}^p |\partial_{x_k} f(x,t)|^2 \, d(\mu_s + \sigma_s) \, ds.$$

Substituting u for f^2 in (3) we obtain

$$\int_{\mathbb{R}^d} \psi^2 \, d\mu_t = \int_{\mathbb{R}^d} f^2 \, d\mu_0 + \int_0^t \int_{\mathbb{R}^d} 2\langle A\nabla f, \nabla f\rangle + 2f\langle b_\mu - h, \nabla f\rangle \, d\mu_s \, ds.$$

Applying the inequalities $|f| \leq 1$ and

$$|\langle b_\mu - h, \nabla f\rangle| \leq \sum_{i=1}^d |b_\mu^i - h^i||\partial_{x_i} f| \leq \frac{\lambda}{2}\sum_{i=1}^p |\partial_{x_i} f|^2 + \frac{1}{2\lambda}\sum_{i=1}^p |b_\mu^i - h|^2 + C_1 \sum_{i=p+1}^d |b_\mu^i - h^i|,$$

we get the estimate

$$\int_0^t \int_{\mathbb{R}^d} \sum_{i=1}^p |\partial_{x_i} f|^2 \, d\mu_s \, ds \leq \frac{1}{\lambda} + R_1(h),$$

where

$$R_1(h) = \int_0^t \int_{\mathbb{R}^d} \lambda^{-2} \sum_{i=1}^p |b_\mu^i - h|^2 + 2C_1\lambda^{-1} \sum_{i=p+1}^d |b_\mu^i - h^i| \, d\mu_s \, ds.$$

By the same argument we obtain the estimate

$$\int_0^t\int_{\mathbb{R}^d}\sum_{i=1}^p |\partial_{x_i}f|^2\,d\sigma_s\,ds\le\frac{1}{\lambda}+R_2(h).$$

where

$$R_2(h)=\int_0^t\int_{\mathbb{R}^d}\lambda^{-2}\sum_{i=1}^p|b^i_\sigma-h|^2+2C_1\lambda^{-1}\sum_{i=p+1}^d|b^i_\sigma-h^i|\,d\sigma_s\,ds.$$

Note that

$$R_2(h)\le Q_1+Q_2(h),$$

where

$$Q_1=\int_0^t\int_{\mathbb{R}^d}2\lambda^{-2}\sum_{i=1}^p|b^i_\mu-b^i_\sigma|^2+2C_1\lambda^{-1}\sum_{i=p+1}^d|b^i_\mu-b^i_\sigma|\,d\sigma_s\,ds,$$

$$Q_2(h)=\int_0^t\int_{\mathbb{R}^d}2\lambda^{-2}\sum_{i=1}^p|b^i_\mu-h^i|^2+2C_1\lambda^{-1}\sum_{i=p+1}^d|b^i_\mu-h^i|\,d\sigma_s\,ds.$$

Applying Lemma 2 (or the standard approximation in the case $p=0$) we find a sequence of smooth vector fields h_n such that

$$\lim_{n\to\infty}\int_0^t\int_{\mathbb{R}^d}2\lambda^{-2}\sum_{i=1}^p|b^i_\mu-h^i_n|^2+\sum_{i=p+1}^d|b^i_\mu-h^i_n|\,d(\mu_s+\sigma_s)\,ds=0.$$

It follows that $R_1(h_n)\to 0$ and $Q_2(h_n)\to 0$ as $n\to\infty$. Substituting h for h_n in the previous estimates and letting $n\to\infty$, we obtain the estimate

$$\int_{\mathbb{R}^d}\psi\,d(\mu_t-\sigma_t)\le(1+C_1)d_p(\mu_0,\sigma_0)+$$
$$+\Bigl(\int_0^t\int_{\mathbb{R}^d}\sum_{i=1}^p|b^i_\mu-b^i_\sigma|^2\,d\sigma_s\,ds\Bigr)^{1/2}\Bigl(\frac{1}{\lambda}+Q_1\Bigr)^{1/2}+C_1\int_0^t\int_{\mathbb{R}^d}\sum_{i=p+1}^d|b^i_\mu-b^i_\sigma|\,d\sigma_s\,ds.$$

This completes the proof. □

We apply the obtained estimates to nonlinear Fokker–Planck–Kolmogorov equations. Suppose that

(NH1) $A=(a^{ij})$ satisfies (H1) and (H2).

Let $\mathcal{M}_\tau$ be the set of all measures $\mu=\mu_t\,dt$ on $[0,\tau]\times\mathbb{R}^d$, where $(\mu_t)_{t\in[0,\tau]}$ is a family of probability measures on $\mathbb{R}^d$. Let $\mathcal{M}_0$ be a subset of $\mathcal{M}_\tau$. Assume that for every $\mu\in\mathcal{M}_0$ we are given Borel measurable functions $b^i(t,x,\mu)$.

Set

$$L_\mu=\sum_{i,j=1}^d a^{ij}(t,x)\partial_{x_i}\partial_{x_j}+\sum_{i=1}^d b^i(t,x,\mu)\partial_{x_i}.$$

We say that $\mu=\mu_t\,dt\in\mathcal{M}_0$ is a solution to the Cauchy problem on $[0,\tau]\times\mathbb{R}^d$

$$\partial_t\mu=L^*_\mu\mu,\quad \mu|_{t=0}=\mu_0, \tag{4}$$

for the nonlinear Fokker–Planck–Kolmogorov equation if μ is a solution to the Cauchy problem (2) on $[0,\tau]\times\mathbb{R}^d$ for the linear Fokker–Planck–Kolmogorov equation with the operator L_μ.

Denote by $\mathcal{P}(\mathbb{R}^d)$ the space of all probability measures on $\mathbb{R}^d$.

Let $V\in C^2(\mathbb{R}^d)$, $V\geq 0$ and $\lim_{|x|\to\infty}V(x)=+\infty$. Let $\alpha>0$. Denote by $B_{\alpha,\tau}(V)$ the set of all mappings $\mu_t\colon[0,\tau]\to\mathcal{P}(\mathbb{R}^d)$ such that $t\to\mu_t(B)$ is a Borel measurable function for every Borel set B and

$$\sup_{t\in[0,\tau]}\int_{\mathbb{R}^d}V(x)\mu_t(dx)\leq\alpha.$$

Proposition 2. *$B_{\alpha,\tau}(V)$ is a complete metric space with respect to the metric*

$$r(\mu,\sigma)=\sup_{t\in[0,\tau]}d_p(\mu_t,\sigma_t).$$

Proof. Note that $\mathcal{P}(\mathbb{R}^d)$ equipped with d_p is a complete metric space. Moreover, for every $\varphi\in C_0^1(\mathbb{R}^d)$ we get the estimate

$$\sup_{t\in[0,\tau]}\Bigl|\int_{\mathbb{R}^d}\varphi\,d(\mu_t-\sigma_t)\Bigr|\leq C(\varphi)r(\mu,\sigma).$$

Assume that $\mu^n\in B_{\alpha,\tau}(V)$ and $r(\mu^n,\mu)\to 0$. It is obvious that $\displaystyle\int_{\mathbb{R}^d}V\,d\mu_t\leq\alpha$. Note that $g_n(t)=\displaystyle\int_{\mathbb{R}^d}\varphi\,d\mu_t^n$ is a Borel measurable function for every n and $\varphi\in C_0^1(\mathbb{R}^d)$. In addition the sequence g_n uniformly converges to $g(t)=\displaystyle\int_{\mathbb{R}^d}\varphi\,d\mu_t$. This yields that g is a Borel measurable for every $\varphi\in C_0^1(\mathbb{R}^d)$. Applying the estimate $\|V\|_{L^1(\mu_t)}\leq\alpha$ we obtain that $g(t)$ is a Borel measurable for every bounded continuous function φ. According to the monotone class theorem (see [1, Theorem 2.12.9]) we have that $g(t)$ is a Borel measurable for every bounded Borel measurable φ. In particular the mapping $t\to\mu_t(B)$ is measurable for every Borel set B. □

Suppose that

(NH2) for every $\alpha>0$ there exists $\Lambda=\Lambda(\alpha)>0$ such that for every $\sigma\in B_{\alpha,T}(V)$ we have

$$|b(t,x,\sigma)-b(t,x+h_p,\sigma)|\leq\Lambda|h_p|\quad\forall\ h_p=(0,\dots,0,y_{p+1},\dots,y_d)$$

and the mappings $x\to b^i(t,x,\sigma)$ are continuous uniformly in $t\in[0,T]$.

Theorem 2. *Suppose that* (NH1) *and* (NH2) *are fulfilled and there exist positive numbers C_1, C_2 and C_3 such that for every $\alpha>0$, $\tau\in(0,T]$ and $\sigma,\mu\in B_{\alpha,\tau}(V)$ we have the estimates*

$$|b(t,x,\sigma)|\leq C_1+C_1\sqrt{V(x)},\quad L_\sigma V(t,x)\leq C_2+C_2V(x),$$
$$|b(t,x,\mu)-b(t,x,\sigma)|\leq C_3(1+\sqrt{V(x)})d_p(\mu_t,\sigma_t)$$

for every $(x,t)\in[0,\tau]\times\mathbb{R}^d$. Then for every probability measure μ_0, such that $V\in L^1(\mu_0)$, there exist $\tau\in(0,T]$ and $\alpha>0$ for which the Cauchy problem (4) *has a unique solution in the space $B_{\alpha,\tau}(V)$.*

Proof. We define the mapping F as follows:

$$\mu=F(\sigma)\Leftrightarrow\partial_t\mu=L_\sigma^*\mu,\quad\mu|_{t=0}=\nu.$$

According to [2, Theorem 6.7.3](see also [13]) and [2, Theorem 9.8.7](see also [3]) F is well-defined on $B_{\alpha,\tau}(V)$. Let $\mu=F(\sigma)$. By [2, Theorem 7.1.1] we get

$$\int_{\mathbb{R}^d}V\,d\mu_t\leq e^{C_2t}+e^{C_2t}\int_{\mathbb{R}^d}V\,d\nu.$$

Setting

$$\alpha=e^{C_2T}+e^{C_2T}\int_{\mathbb{R}^d}V\,d\nu.$$

we have $F\colon B_{\alpha,\tau}(V) \to B_{\alpha,\tau}(V)$ for every $\tau \in (0, T]$. By Theorem 1 we obtain the estimate

$$r(F(\sigma^1), F(\sigma^2)) \le C\tau r(\sigma^1, \sigma^2),$$

where C depends on C_1, C_2, T, $\Lambda(\alpha, T)$ and α. Consequently the mapping F is contractive if $\tau < 1/C$. By the Banach contracting mapping theorem, there exists a unique solution $\mu \in B_{\alpha,\tau}(V)$. $\square$

Acknowledgements.

The authors are grateful to Prof. Vladimir I. Bogachev for fruitful discussions and valuable remarks.

The work was supported by the RFBR grant 15-31-20082, 17-01-00622, the President Grant MD-207.2017.1 and the CRC 701 at Bielefeld University.

References

[1] Bogachev V.I. Measure theory. Springer-Verlag, Berlin, 2007.

[2] Bogachev V.I., Krylov N.V., Röckner M., Shaposhnikov S.V. Fokker–Planck–Kolmogorov equations. Amer. Math. Soc., Providence, Rhode Island, 2015.

[3] Bogachev V.I., Röckner M., Shaposhnikov S.V. Uniqueness problems for degenerate Fokker-Planck-Kolmogorov equations. Journal of Mathematical Sciences (New York), 2015, V. 207, N 2, P. 147–165.

[4] Bogachev V.I., Da Prato G., Röckner M., Shaposhnikov S.V. An analytic approach to infinite-dimensional continuity and Fokker–Planck–Kolmogorov equations. Annali della Scuola Normale Superiore di Pisa, 2015, V. 14, N 3, P. 983-1023.

[5] Bogachev V.I., Da Prato G., Röckner M., Shaposhnikov S.V. On the uniqueness of solutions to continuity equations. Journal of Differential Equations, 2015, V. 259, N 8, P. 3854-3873.

[6] Bogachev V.I., Röckner M., Shaposhnikov S.V. Distances between transition probabilities of diffusions and applications to nonlinear Fokker-Planck-Kolmogorov equations. J. Funct. Anal., 2016, V. 271, P. 1262–1300.

[7] Frank T.D. Nonlinear Fokker-Planck equations. Fundamentals and applications. Springer-Verlag, Berlin, 2005.

[8] Konakov V., Kozhina A., Menozzi S. Stability of densities for perturbed Diffusions and Markov Chains. 2015, arXiv preprint arXiv:1506.08758.

[9] Li H., Luo D. Quantitative stability estimates for solutions of Fokker–Planck equations. 2017, arXiv preprint arXiv:1701.00566.

[10] Manita O.A. Estimates for the Kantorovich distances between solutions to the nonlinear Fokker–Planck–Kolmogorov equation with monotone drift. 2015, arXiv preprint arXiv:1507.04014.

[11] Manita O.A., Romanov M.S., Shaposhnikov S.V. On uniqueness of solutions to nonlinear Fokker–Planck–Kolmogorov equations. Nonlinear Analysis: Theory, Methods and Applications, 2015, V. 128, P. 199–226.

[12] Manita O.A., Shaposhnikov S.V. Nonlinear parabolic equations for measures. St. Petersburg Mathematical Journal, 2014, V. 25, N 1, P. 43-62.

[13] Manita O. A., Shaposhnikov S. V. On the Cauchy problem for Fokker–Planck–Kolmogorov equations with potential terms on arbitrary domains. Journal of Dynamics and Differential Equations, 2016, V. 28, N 2, P. 493-518.

[14] Risken H. The Fokker–Planck Equation. Springer–Verlag, Berlin, 1996.

[15] Stroock D.W., Varadhan S.R.S. Multidimensional diffusion processes. Berlin – New York: Springer-Verlag, 1979.