On the Ambrosio–Figalli–Trevisan superposition principle for probability solutions to Fokker–Planck–Kolmogorov equations

Vladimir I. Bogachev[a1], **Michael Röckner**[b], **Stanislav V. Shaposhnikov**[a]

[a]*Lomonosov Moscow State University, Russia and National Research University Higher School of Economics, Russian Federation*
[b]*Fakultät für Mathematatik, Universität Bielefeld, D-33501 Bielefeld, Germany*

Abstract We prove a generalization of the known result of Trevisan on the Ambrosio–Figalli–Trevisan superposition principle for probability solutions to the Cauchy problem for the Fokker–Planck–Kolmogorov equation, according to which such a solution is generated by a solution to the corresponding martingale problem. The novelty is that in place of the integrability of the diffusion and drift coefficients A and b with respect to the solution we require the integrability of $(\|A(t,x)\|+|\langle b(t,x),x\rangle|)/(1+|x|^2)$. Therefore, in the case where there are no a priori global integrability conditions the function $\|A(t,x)\|+|\langle b(t,x),x\rangle|$ can be of quadratic growth. Moreover, as a corollary we obtain that under mild conditions on the initial distribution it is sufficient to have the one-sided bound $\langle b(t,x),x\rangle \le C + C|x|^2\log|x|$ along with $\|A(t,x)\| \le C + C|x|^2\log|x|$.

Keywords: Fokker–Planck–Kolmogorov equation, martingale problem, superposition principle

AMS MSC 2010: 60J60; 35Q84

1. Introduction

We study solutions to the Cauchy problem for the Fokker–Planck–Kolmogorov equation

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

Below we write this equation in the short form $\partial_t\mu_t = L^*\mu_t$, where L^* is the formal adjoint operator to the differential operator

$$Lu = a^{ij}\partial_{x_i}\partial_{x_j}u + b^i\partial_{x_i}u,$$

where the usual convention about summation over repeated indices is employed. We assume throughout that the matrix $A(t,x) = (a^{ij}(t,x))_{i,j\le d}$ is symmetric and nonnegative definite and the functions $(t,x)\mapsto a^{ij}(t,x)$ and $(t,x)\mapsto b^i(t,x)$ are Borel measurable on $[0,T]\times\mathbb{R}^d$. By a solution we mean a mapping $t\mapsto\mu_t$ from $[0,T]$ to the space of probability measures $\mathcal{P}(\mathbb{R}^d)$ that is continuous with respect to the weak topology and satisfies the integral equality

$$\int_{\mathbb{R}^d}\varphi\,d\mu_t = \int_{\mathbb{R}^d}\varphi\,d\nu + \int_0^t\int_{\mathbb{R}^d}L\varphi\,d\mu_s\,ds$$

for all $t\in[0,T]$ and all $\varphi\in C_0^\infty(\mathbb{R}^d)$, where it is assumed that a^{ij} and b^i are locally (i.e., on compact sets in $[0,T]\times\mathbb{R}^d$) integrable with respect to the measure $\mu_t\,dt$:

$$a^{ij}, b^i \in L^1_{loc}(\mu_t\,dt).$$

The measure $\mu = \mu_t\,dt$ on $[0,T]\times\mathbb{R}^d$ is defined as usual by the equality

$$\int f\,d\mu = \int_0^T\int_{\mathbb{R}^d} f(t,x)\,\mu_t(dx)\,dt$$

This measure can be identified with the solution, which is also denoted by $\{\mu_t\}$.

Recall that the weak topology on $\mathcal{P}(\mathbb{R}^d)$ is generated by the seminorms

$$\sigma \mapsto \left|\int_{\mathbb{R}^d} f(x)\,\sigma(dx)\right|$$

[1]corresponding author. *E-mail addresses*: vibogach@mail.ru (V. Bogachev), roeckner@math.uni-bielefeld.de (M. Röckner), starticle@mail.ru (S. Shaposhnikov)

on the linear space of all bounded Borel measures, where f is a bounded continuous function on $\mathbb{R}^d$ (see [7]). Recent accounts on the theory of Fokker–Planck–Kolmogorov equations can be found in [11] and [12]. The inner product and norm on $\mathbb{R}^d$ are denoted by $\langle x, y\rangle$ and $|x|$, respectively. The operator norm of a matrix A is denoted by $\|A\|$.

In the case $A = 0$ (the continuity equation) the following superposition principle of Ambrosio [1] is known (see also [2], [4], and [26]). If $\mu = \mu_t\,dt$ with probability measures μ_t on $\mathbb{R}^d$ satisfies the continuity equation

$$\partial_t \mu_t + \operatorname{div}(b\mu_t) = 0$$

and $|b(x,t)|/(1+|x|)$ is μ-integrable, then there exists a nonnegative Borel measure η on the space $\mathbb{R}^d \times C([0,T], \mathbb{R}^d)$ concentrated on the set of pairs (x,ω) such that ω is an absolutely continuous solution of the integral equation

$$\omega(t) = x + \int_0^t b(\omega(s), s)\,ds$$

and, for each function $\varphi \in C_b(\mathbb{R}^d)$ and each $t \in [0,T]$, one has

$$\int \varphi(x)\,\mu_t(dx) = \int \varphi(\omega(t))\,\eta(dx d\omega).$$

In other words, the measure μ_t coincides with the image of η under the evaluation mapping $(x,\omega) \mapsto \omega(t)$. Of course, the integral on the right coincides with the integral against the projection of η on $C([0,T], \mathbb{R}^d)$ (see [1, Remark 3.1] about the connection between the two measures). A discussion of analogous representations for signed solutions and interesting counter-examples can be found in [19].

The case of possibly nonzero A and bounded A and b was first considered by Figalli [17], who proved that every probability solution to the Cauchy problem for the Fokker–Planck–Kolmgorov equation is represented by a martingale measure on the path space. Generalizing this seminal achievement, Trevisan [28] obtained the following important and very general result.

Suppose that a mapping $t \mapsto \mu_t$ from $[0,T]$ to the space of probability measures $\mathcal{P}(\mathbb{R}^d)$ is continuous with respect to the weak topology and satisfies the Cauchy problem (1.1). Suppose also that it satisfies the condition

$$\int_0^T \int_{\mathbb{R}^d} \big[\|A(t,x)\| + |b(t,x)|\big]\,\mu_t(dx)\,dt < \infty. \tag{1.2}$$

Then there exists a Borel probability measure P_ν on the path space

$$\Omega_d := C([0,T], \mathbb{R}^d)$$

of continuous functions $\omega\colon [0,T] \to \mathbb{R}^d$ with its standard sup-norm $\|\omega\| = \sup_t |\omega(t)|$ such that

(i) $P_\nu\big(\omega\colon \omega(0) \in B\big) = \nu(B)$ for all Borel sets $B \subset \mathbb{R}^d$,

(ii) for every function $f \in C_0^\infty(\mathbb{R}^d)$, the function

$$(\omega, t) \mapsto f(\omega(t)) - f(\omega(0)) - \int_0^t Lf(s, \omega(s))\,ds$$

is a martingale with respect to the measure P_ν and the natural filtration $\mathcal{F}_t = \sigma(\omega(s), s \in [0,t])$,

(iii) for every function $f \in C_0^\infty(\mathbb{R}^d)$, there holds the equality

$$\int_{\mathbb{R}^d} f\,d\mu_t = \int_{\Omega_d} f(\omega(t))\,P_\nu(d\omega) \quad \forall\, t \in [0,T].$$

The latter means that μ_t is the law of $\omega(t)$ under P_ν, while (i) means that ν is the law of $\omega(0)$.

In spite of a very general character of condition (1.2), in many simple situations it is not fulfilled. Let us consider a one-dimensional example. Let

$$\varrho \in C^\infty(\mathbb{R}), \quad \varrho > 0, \quad \int \varrho(x)\,dx = 1, \quad b(x) = \varrho'(x)/\varrho(x).$$

Then $\mu_t(dx) = \mu(dx) = \varrho\,dx$ is a stationary solution to the Fokker–Planck–Kolmogorov equation

$$\partial_t \mu = \mu'' - (b\mu)'.$$

In particular, $\mu_t = \mu$ satisfies the Cauchy problem with initial data μ. However, it is easy to find a smooth probability density ϱ such that

$$\int_{\mathbb{R}} |b(x)|\varrho(x)\,dx = \int_{\mathbb{R}} |\varrho'(x)|\,dx = \infty.$$

In this paper we reinforce the aforementioned result by replacing condition (1.2) with a weaker assumption.

Throughout we assume that the coefficients are Borel measurable on $[0, T] \times \mathbb{R}^d$,

$$a^{ij}, b^i \in L^1([0, T] \times U, \mu_t\,dt)$$

for every ball U in $\mathbb{R}^d$, and the following condition is fulfilled:

$$\int_0^T \int_{\mathbb{R}^d} \frac{\|A(t,x)\| + |\langle b(t,x), x\rangle|}{(1+|x|)^2}\,\mu_t(dx)\,dt < \infty. \tag{1.3}$$

It follows from Proposition 2.2 below (see also Example 2.3) that in order to ensure condition (1.3) it suffices that $\log(1+|x|) \in L^1(\nu)$ and

$$\|A(t,x)\| \le C + C|x|^2\log(1+|x|), \quad \langle b(t,x), x\rangle \le C + C|x|^2\log(1+|x|).$$

Obviously, it is also sufficient without any assumptions about ν that

$$\|A(t,x)\| + |\langle b(t,x), x\rangle| \le C + C|x|^2.$$

Our main result is the following theorem (its proof is given in the last section).

Theorem 1.1. *Suppose that $\{\mu_t\}$ is a solution to the Cauchy problem $\partial_t\mu_t = L^*\mu_t$ on $[0, T]$ with $\mu_0 = \nu$ and* (1.3) *is fulfilled. Then there exists a Borel probability measure P_ν on $\Omega_d = C([0, T], \mathbb{R}^d)$ for which all assertions* (i), (ii) *and* (iii) *are true.*

It is important that our theorem assumes no uniqueness of probability solutions to the Cauchy problem, the martingale representation exists for each probability solution satisfying (1.3).

It should be noted that the superposition principle does not work without global assumptions even for smooth coefficients and $A = I$, because it can happen that there are many probability solutions to the Fokker–Planck–Kolmogorov equation (see [12, Section 9.2], while the martingale problem has a unique solution in this case (see [27, Corollary 10.1.2]) and this solution necessarily corresponds to a subprobability solution to the FPK equation due to a blow up.

Note that the integrability of $(1+|x|)^{-2}|\langle b(t,x), x\rangle|$ can hold even in the case where the function $(1+|x|)^{-1}|b(t,x)|$ is not integrable with respect to the solution (see Example 3.2).

Not only is the assumption of integrability of $(1+|x|)^{-2}|\langle b(t,x), x\rangle|$ weaker than the assumption of integrability of $(1+|x|)^{-1}|b(t,x)|$, but it is also simpler to verify. For example, as already noted above, if $\log(1+|x|)$ is ν-integrable, it suffices to have a one-sided bound.

Corollary 1.2. *Let $\log(1+|x|^2) \in L^1(\nu)$ and*

$$\|A(t,x)\| \le C + C|x|^2\log(1+|x|^2), \quad \langle b(t,x), x\rangle \le C + C|x|^2\log(1+|x|^2).$$

Then the hypotheses of the main theorem are fulfilled, hence its conclusion holds.

For the proof, see Example 2.3. Such a bound allows coercive drift coefficients typical in the theory of diffusion processes and is a simple algebraic condition, a verification of which does not use any information about the unknown solution to the Cauchy problem. Note that if no information about the solution $\{\mu_t\}$ is given, then our result applies to A and b of linear growth, more precisely, the function $\|A(t,x)\| + |\langle b(t,x), x\rangle|$ can be of quadratic growth. Note also that the superposition principle holds for nonlinear equations (see Remark 3.4).

2. Auxiliary results

For the proof of the main result we need some auxiliary assertions. The next lemma is a simple consequence of the fact that μ_t is the law of $\omega(t)$ under P_ν, but it will be applied repeatedly below.

Lemma 2.1. *Let $g \geq 0$ be a bounded Borel function on Ω_d and let $f \geq 0$ be a Borel function on $\mathbb{R}^d$ integrable with respect to the measure μ_t for some $t \in (0,T]$. If P_ν is a probability measure on Ω_d with property* (iii) *above, then*

$$\int_{\Omega_d} f(\omega(t)) g(\omega)\, P_\nu(d\omega) \leq \sup_\omega g(\omega) \int_{\mathbb{R}^d} f(x)\, \mu_t(dx). \tag{2.1}$$

The following proposition not only provides an important a priori estimate in the spirit of classical Lyapunov functions (see [12], where a variety of similar results can be found), but also contains an interesting new result: the integrability of $|LV|$ with respect to the solution.

Proposition 2.2. *Suppose that $\{\mu_t\}$ is a solution of the Cauchy problem $\partial_t \mu_t = L^*\mu_t$ with $\mu_0 = \nu$ and there exists a nonnegative function $V \in C^2(\mathbb{R}^d)$ along with a measurable nonnegative function W such that $V \in L^1(\nu)$ and for some numbers $C \geq 0$ and $\tau \in (0,T]$ one has*

$$\lim_{|x|\to+\infty} V(x) = +\infty, \quad LV(t,x) \leq W(t,x) + CV(x), \quad \int_0^\tau \int_{\mathbb{R}^d} W(t,x)\, \mu_t(dx)\, dt < \infty.$$

Then

$$\int_{\mathbb{R}^d} V\, d\mu_t \leq \Bigl(\int_{\mathbb{R}^d} V\, d\nu + \int_0^\tau \int_{\mathbb{R}^d} W\, d\mu_s\, ds\Bigr) e^{Ct} \quad \forall t \in [0,\tau],$$

$$\int_0^\tau \int_{\mathbb{R}^d} |LV|\, d\mu_t\, dt \leq 2e^{C\tau} \Bigl(\int_0^\tau \int_{\mathbb{R}^d} W\, d\mu_s\, ds + \int_{\mathbb{R}^d} V\, d\nu\Bigr).$$

Proof. Let $\zeta_N \in C_b^\infty(\mathbb{R})$ be such that $\zeta_N(t) = t$ if $t \leq N-1$, $\zeta_N(t) = N$ if $t > N+1$ and $0 \leq \zeta_N' \leq 1$, $\zeta_N'' \leq 0$. In the proof we omit indication of $\mathbb{R}^d$ in integration over the whole space. Since

$$L\zeta_N(V) = \zeta_N'(V) L(V) + \zeta_N''(V) |\sqrt{A}\nabla V|^2 \leq \zeta_N'(V) LV,$$

there holds the inequality

$$\int \zeta_N(V)\, d\mu_t \leq \int \zeta_N(V)\, d\nu + \int_0^t \int \zeta_N'(V) LV\, d\mu_s\, ds.$$

Therefore,

$$\int \zeta_N(V)\, d\mu_t \leq \int \zeta_N(V)\, d\nu + \int_0^t \int \zeta_N'(V)(W + CV)\, d\mu_s\, ds$$

and

$$\int \zeta_N(V)\, d\mu_t \leq \int V\, d\nu + \int_0^\tau \int W\, d\mu_s\, ds + C\int_0^t \int \zeta_N(V)\, d\mu_s\, ds.$$

The announced bound on the integral of V against μ_t is obtained with the aid of Gronwall's inequality by passing to the limit as $N \to \infty$.

Next we write the first inequality in the following form:

$$\int \zeta_N(V)\, d\mu_t + \int_0^t \int \zeta_N'(V)(LV)^-\, d\mu_s\, ds \leq \int \zeta_N(V)\, d\nu + \int_0^t \int \zeta_N'(V)(LV)^+\, d\mu_s\, ds,$$

where $(LV)^+ = \max\{LV, 0\}$, $(LV)^- = \max\{-LV, 0\}$ and $LV = (LV)^+ - (LV)^-$.

Since $(LV)^+ \le W + CV$, we have

$$\int \zeta_N(V)\, d\mu_t + \int_0^t \int \zeta_N'(V)(LV)^-\, d\mu_s\, ds \le \int \zeta_N(V)\, d\nu + \int_0^t \int W + CV\, d\mu_s\, ds.$$

On account of the obtained estimate on V we arrive at the inequality

$$\int_0^\tau \int (LV)^-\, d\mu_s\, ds \le e^{C\tau}\Bigl(\int_0^\tau \int W\, d\mu_s\, ds + \int V\, d\nu\Bigr),$$

which yields the announced estimate on the integral of $|LV| = (LV)^+ + (LV)^-$. □

Example 2.3. If $\log(1 + |x|^2) \in L^1(\nu)$ and

$$\|A(t,x)\| \le C + C|x|^2 \log(1 + |x|^2), \quad \langle b(t,x), x\rangle \le C + C|x|^2 \log(1 + |x|^2),$$

then for some number C_1 we have

$$L(\log(1 + |x|^2)) \le C_1 + C_1 \log(1 + |x|^2), \quad \|A(t,x)\|/(1 + |x|^2) \le C_1 + C_1 \log(1 + |x|^2),$$
$$|\langle b(t,x), x\rangle|/(1 + |x|^2) \le |L(\log(1 + |x|^2)| + C_1 + C_1 \log(1 + |x|^2).$$

Hence by Proposition 2.2 the functions $\log(1 + |x|^2)$ and $|L(\log(1 + |x|^2)|$ are integrable on $[0,T] \times \mathbb{R}^d$ with respect to $\mu_t\, dt$ and condition (1.3) is fulfilled.

Proposition 2.4. *Suppose that $V \in C^2(\mathbb{R}^d)$ and $\lim\limits_{|x|\to+\infty} V(x) = +\infty$.*

(i) *There exists a function $\theta \in C^2(\mathbb{R})$ such that $\theta(V) \in L^1(\nu)$ and*

$$\theta \ge 0, \quad \theta(0) = 0, \quad 0 \le \theta'(t) \le 1, \quad \theta'' \le 0, \quad \lim_{t\to+\infty} \theta(t) = +\infty.$$

(ii) *Assume that for some $\tau \in (0,T]$ one has*

$$\int_0^\tau \int_{\mathbb{R}^d} \Bigl(|\sqrt{A}\nabla V|^2 + |LV|\Bigr)\, d\mu_t\, dt < \infty$$

and that θ satisfies all assumptions listed in (i) *and $\theta(V) \in L^1(\nu)$. Then $\theta(V)$ satisfies the following inequality:*

$$\int_0^\tau \int_{\mathbb{R}^d} \Bigl(|\sqrt{A}\nabla\theta(V)|^2 + |L\theta(V)|\Bigr)\, d\mu_t\, dt$$
$$\le 2e^{C\tau}\Bigl(\int \theta(V)\, d\nu + \int_0^\tau \int_{\mathbb{R}^d} \Bigl(|\sqrt{A}\nabla V|^2 + |LV|\Bigr)\, d\mu_t\, dt\Bigr).$$

Proof. In the proof of Proposition 7.1.8 in [12] it was shown that there is a function $\theta \in C^2(\mathbb{R})$ such that $\theta \ge 0$, $\theta(0) = 0$, $0 \le \theta'(t) \le 1$, $\theta'' \le 0$ and $\theta(V) \in L^1(\nu)$. Then

$$|\sqrt{A}\nabla\theta(V)|^2 = |\theta'|^2 |\sqrt{A}\nabla V|^2 \le |\sqrt{A}\nabla V|^2,$$
$$L\theta(V) = \theta''(V)|\sqrt{A}\nabla V|^2 + \theta'(V)LV \le |LV|.$$

Applying Proposition 2.2 with $W = |LV|$ and $C = 0$ we obtain our claim. □

The next assertion enables us to estimate the measure of a ball in the path space with the aid of the function V.

Proposition 2.5. *Let $\tau \in (0,T]$. Suppose that $\{\mu_t\}$ is a solution to the Cauchy problem $\partial_t\mu_t = L^*\mu_t$ on $[0,\tau]$ with $\mu_0 = \nu$ and that there exists a Borel probability measure P_ν on $C([0,\tau], \mathbb{R}^d)$ such that* (i), (ii) *and* (iii) *are fulfilled. Suppose also that there is a nonnegative function $V \in C^2(\mathbb{R}^d)$ with $\lim\limits_{|x|\to\infty} V(x) = +\infty$ such that $V \in L^1(\nu)$ and*

$$\int_0^\tau \int_{\mathbb{R}^d} \Bigl(|\sqrt{A}\nabla V|^2 + |LV|\Bigr)\, d\mu_t\, dt < \infty.$$

Then for every $q > 0$ one has

$$P_\nu\Bigl(\omega\colon \sup_{t\in[0,T]} V(\omega(t)) \ge q\Bigr) \le \frac{2}{q}\Bigl(\int_{\mathbb{R}^d} V\, d\nu + \int_0^\tau \int_{\mathbb{R}^d} \Bigl(|\sqrt{A}\nabla V|^2 + |LV|\Bigr)\, d\mu_s\, ds\Bigr).$$

Proof. Using the function $\zeta_N(V)$ in place of V and the assumption about the integrability of the function LV one can verify that

$$V(\omega(t)) - V(\omega(0)) - \int_0^t LV(s,\omega(s))\,ds$$

is a martingale with the quadratic variation

$$\int_0^t |\sqrt{A}\nabla V(s,\omega(s))|^2\,ds.$$

By Doob's inequality we have

$$P_\nu\Bigl(\omega\colon \sup_{t\in[0,\tau]}\Bigl|V(\omega(t)) - V(\omega(0)) - \int_0^t LV(s,\omega(s))\,ds\Bigr| \ge q\Bigr) \le \frac{1}{q}\int_0^\tau \int_{\mathbb{R}^d} |\sqrt{A}\nabla V|^2\,d\mu_s\,ds.$$

Since

$$\begin{aligned} P_\nu\Bigl(\omega\colon \sup_{t\in[0,\tau]} V(\omega(t)) \ge q\Bigr) &\le P_\nu\Bigl(\omega\colon \sup_{t\in[0,\tau]}\Bigl|V(\omega(t)) - V(\omega(0)) - \int_0^t LV(s,\omega(s))\,ds\Bigr| \ge q/2\Bigr) \\ &\quad + P_\nu\Bigl(\omega\colon \sup_{t\in[0,\tau]}\Bigl|V(\omega(0)) + \int_0^t LV(s,\omega(s))\,ds\Bigr| \ge q/2\Bigr), \end{aligned}$$

we obtain

$$P_\nu\Bigl(\omega\colon \sup_{t\in[0,\tau]} V(\omega(t)) \ge q\Bigr) \le \frac{2}{q}\Bigl(\int_{\mathbb{R}^d} V\,d\nu + \int_0^\tau \int_{\mathbb{R}^d} \Bigl(|\sqrt{A}\nabla V|^2 + |LV|\Bigr)\,d\mu_s\,ds\Bigr),$$

which completes the proof. □

3. Proof of the main result

The proof of the theorem follows the scheme used by Figalli [17] and Trevisan [28]. However, there are some differences: Trevisan's result is not applicable even for smooth coefficients without global integrability of the coefficients. Here we substantionally used some recent results on the uniqueness of probbaility solutions to Fokker–Planck–Kolmogorov equations from [12] and [24]. When reducing the general case to that of smooth coefficients (as also Figalli and Trevisan did), we encounter two problems: 1) it is necessary to control that the solutions with smoothed coefficients converge to the considered solution, which is not automatic due to the lack of uniqueness, 2) for a priori estimates it is necessary to keep condition (1.3) uniformly. The first problem is overcome by using the smoothing involving not only the space variable, but also the time. The second problem is solved with the aid of the equation itself, namely, we estimate the integral of $\langle\beta(t,x),x\rangle$ for the approximating drift β by means of some integral of the diffusion matrix. Note also that before picking a common compact set of measure close to 1 for the corresponding measures on Ω_d we first pick a common ball of measure close to 1. Finally, we verify that for passing to the limit our local integrability conditions on the coefficients are sufficient.

Proof of Theorem 1.1. First we assume that all our hypotheses hold on some larger interval $[0,T_1]$, $T_1 > T$, and at the last step explain how to obtain a representation on $[0,T]$ without that assumption. Set

$$V(x) = \log(1+|x|^2).$$

Take a function θ such that $\theta(\log(3+2|x|^2)) \in L^1(\nu)$ and all conditions from (i) in Proposition 2.4 are fulfilled. According to Proposition 2.2 there is a number N_θ such that

$$\sup_{t\in[0,T_1]} \int \theta(\log(3+2|x|^2))\,\mu_t(dx) \le N_\theta.$$

Note that the coefficients 3 and 2 are only technical things.

I. Justification of the replacement of μ_t by $\mu_t^\delta = \mu_{t+\delta}$.

Passing from μ_t to $\mu_{t+\delta}$ enables us to smmoth solutions not only with respect to x, but also with respect to t.

Suppose that for every $\delta \in (0, T_1 - T)$ there exists a measure P_δ on Ω_d satisfying (i), (ii) and (iii) with the coefficients

$$A_\delta(t,x) = A(t+\delta, x), \quad b_\delta(t,x) = b(t+\delta, x)$$

and the corresponding operator

$$L_\delta f = a_\delta^{ij}\partial_{x_i}\partial_{x_j}u + b_\delta^i \partial_{x_i}u$$

such that $\mu_t^\delta = \mu_{t+\delta}$ solves the Cauchy problem with initial condition $\mu_0^\delta = \mu_\delta$. We show that it is possible to extract from P_δ a weakly convergent sequence P_{δ_n} with $\delta_n \to 0$ such that its limit is a solution to the original martingale problem and gives a representation for the original solution μ_t. We observe that μ_t^δ converges weakly to μ_t as $\delta \to 0$ for every t by the continuity of the mapping $t \mapsto \mu_t$. Omitting again indication of $\mathbb{R}^d$ when integrating over the whole space, we have

$$\int \theta(V)\, d\mu_t^\delta \le \sup_{t\in[0,T_1]} \int \theta(V)\, d\mu_t \le N_\theta.$$

Moreover,

$$\int_0^T \int \Big(|\sqrt{A_\delta}\nabla V|^2 + |L_\delta V|\Big)\, d\mu_t^\delta\, dt \le \int_0^{T_1} \int \Big(|\sqrt{A}\nabla V|^2 + |LV|\Big)\, d\mu_t\, dt \le C_1,$$

where C_1 does not depend on δ. By Proposition 2.4

$$\int_0^T \int \Big(|\sqrt{A_\delta}\nabla \theta(V)|^2 + |L_\delta \theta(V)|\Big) d\mu_t^\delta\, dt \le C(T)(N_\theta + C_1).$$

By Proposition 2.5 applied to the function $\theta(V)$, for every $\varepsilon > 0$, there exists $R > 0$ such that for all $\delta \in (0, T_1 - T)$ we have

$$P_\delta\big(\omega\colon \|\omega\| \le R\big) \ge 1 - \varepsilon.$$

We now need two results from [28]. The first one is Theorem A2. Let $\Theta\colon [0,+\infty) \to [0,+\infty)$ be a lower semicontinuous function and let $\Theta_1, \Theta_2\colon [0,+\infty) \to [0,+\infty)$ be convex functions such that $\Theta_2(2x) \le C\Theta_2(x)$, $\Theta_1(0) = \Theta_2(0) = 0$ and

$$\lim_{x\to+\infty} \Theta(x) = \lim_{x\to+\infty} \frac{\Theta_1(x)}{x} = \lim_{x\to+\infty} \frac{\Theta_2(x)}{x} = +\infty.$$

Then there is a compact function $\Psi\colon C[0,T] \to [0,+\infty]$, i.e., the sets $\{\Psi \le R\}$ are compact for finite R, such that, whenever $\{\alpha_t\}$, $\{\beta_t\}$, $\varphi = \{\varphi_t\}$ are progressively measurable processes on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, P)$ for which

$$M_t := \varphi_t - \int_0^t \beta_s\, ds \quad \text{and} \quad M_t^2 - \int_0^t \alpha_s\, ds$$

are P-a.s. continuous local martingales and $\alpha_t \ge 0$ a.s., one has

$$\mathbb{E}\Psi(\varphi) \le \mathbb{E}\bigg[\Theta(\varphi_0) + \int_0^T [\Theta_1(|\beta_t|) + \Theta_2(\alpha_t)]\, dt\bigg].$$

Next, according to [28, Corollary A5], if $\eta \in \mathcal{P}(C[0,T], \mathbb{R}^d)$ is a solution to the martingale problem associated with an elleiptic operator L (not necessarily our operator), then, for every function $f \in C_0^\infty(\mathbb{R}^d)$ and the marginal distributions η_t for η it holds

$$\int \Psi(f)\, d\eta \le \int \Theta(|f(0,x)|\, \eta_0(dx) + \int_0^T \int \Big[\Theta_1(|Lf|) + \Theta_2(|\sqrt{A}\nabla f|^2|)\Big]\, d\eta_t\, dt,$$

where Ψ, Θ, Θ_1 and Θ_2 are the same as above and $\Psi(f)(\omega) = \Psi(f(\omega(\cdot)))$. Note that the right-hand side is finite due to our hypotheses on A and b and the compactness of support of f in x.

Now take $\psi_R \in C_0^\infty(\mathbb{R}^d)$ with $\psi_R(x) = 1$ if $|x| \le R$, $\psi_R(x) = 0$ if $|x| > 2R$. Let us apply the cited corollary to the function $f_i(x) = x_i\psi_R(x)$ independent of t. Denoting by $\mathbb{E}_{P_\delta}$ the integral with respect to P_δ, we obtain the estimate

$$\mathbb{E}_{P_\delta}\Psi(f_i) \le \int \Theta(x_i\psi_R(x))\,\mu_\delta(dx) \\ + \int_0^T \int \Bigl(\Theta_1(|L_\delta(x_i\psi_R(x))|) + \Theta_2(|\sqrt{A_\delta}\nabla(x_i\psi_R)|^2)\Bigr)\,\mu_t^\delta(dx)\,dt,$$

where for some number C_2 independent of δ the right-hand side is estimated by

$$\sup_x |\Theta(x_i\psi_R(x))| + C_2\int_0^{T_1}\int_{|x|\le 2R}\Bigl(\Theta_1(\|A(t,x)\| + |b(t,x)|) + \Theta_2(\|A(t,x)\|)\Bigr)\,\mu_t(dx)\,dt,$$

which does not depend on δ. As in [28], we consider the compact function

$$\Psi_d(\omega) = \sum_{i=1}^d \Psi(\omega_i)$$

on Ω_d. We have

$$\mathbb{E}_{P_\delta}(I_{\|\omega\|\le R}\Psi_d) = \mathbb{E}_{P_\delta}[I_{\|\omega\|\le R}\Psi(f_i)] \le \sum_{i=1}^d \mathbb{E}_{P_\delta}\Psi(f_i) \le C_3(R),$$

where $C_3(R)$ depends on R, but does not depend on δ. Taking a sufficiently large number M we conclude that for the compact set

$$K = \{\omega\colon \Psi_d(\omega) \le M,\ \|\omega\| \le R\}$$

in Ω_d there holds the estimate

$$P_\delta(K) \ge 1 - 2\varepsilon.$$

Therefore, the family of measures P_δ contains a sequence P_{δ_k} with $\delta_k \to 0$ weakly converging to some probability measure P. Let us verify that the measure P satisfies (i), (ii) and (iii). The first and last properties are obtained in the limit as $\delta_k \to 0$ in the equality

$$\int_{\Omega_d} f(\omega(t))\,P_{\delta_k}(d\omega) = \int f(x)\,\mu_{t+\delta_k}(dx) \quad \forall f \in C_0^\infty(\mathbb{R}^d).$$

For the proof of the second (martingale) property we have to show that for every bounded continuous function

$$g\colon \Omega_d \to \mathbb{R}$$

that is measurable with respect to the σ-algebra $\mathcal{F}_s$ there holds the equality

$$\int_{\Omega_d}\Bigl[f(\omega(t)) - f(\omega(s)) - \int_s^t Lf(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P(d\omega) = 0, \quad t \ge s,$$

where $f \in C_0^\infty(\mathbb{R}^d)$. To this end, it suffices to show that

$$\lim_{\delta_k\to 0}\int_{\Omega_d}\Bigl[\int_s^t L_\delta f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P_\delta(d\omega) = \int_{\Omega_d}\Bigl[\int_s^t Lf(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P(d\omega),$$

because convergence of the integrals of $[f(\omega(t)) - f(\omega(s))]g(\omega)$ is obvious by the continuity of this function on Ω_d. Let $q^{ij}, z^i \in C^\infty([-1,T_1]\times\mathbb{R}^d)$ and

$$\widetilde{L} = q^{ij}\partial_{x_i}\partial_{x_j} + z^i\partial_{x_i}$$

and let $\widetilde{L}_\delta$ be the corresponding operator with the time-shifted coefficients $q^{ij}(t+\delta,x)$ and $z^i(t+\delta,x)$.

It is clear from (2.1) that the difference

$$\int_{\Omega_d}\Bigl[\int_s^t \widetilde{L}f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P_\delta(d\omega) - \int_{\Omega_d}\Bigl[\int_s^t \widetilde{L}f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P(d\omega)$$

tends to zero. Moreover, by Lemma 2.1 the expression

$$\Bigl|\int_{\Omega_d}\Bigl[\int_s^t(\widetilde{L}-L)f(\tau,\omega(\tau))\Bigr]g(\omega)\,P(d\omega)\Bigr|$$

is estimated by

$$\int_0^T\int\Bigl(|a^{ij}-q^{ij}||\partial_{x_i}\partial_{x_j}f|+|b^i-z^i||\partial_{x_i}f|\Bigr)\,d\mu_t\,dt,$$

which can be made arbitrarily small by a suitable choice of q^{ij} and z^i approximating a^{ij} and b^i in L^1 with respect to the measure $\mu_t\,dt$ on $[0,T_1]\times U$, where U is a ball containing the support of f. Since the functions $(\widetilde{L}-\widetilde{L}_\delta)f(t,x)$ converge to zero uniformly on $[0,T]\times\mathbb{R}^d$ as $\delta\to 0$, the expression

$$\int_{\Omega_d}\Bigl[\int_s^t(\widetilde{L}-\widetilde{L}_\delta)f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P_\delta(d\omega)$$

tends to zero as $\delta\to 0$. Finally, we observe that the expression

$$\int_{\Omega_d}\Bigl[\int_s^t(\widetilde{L}_\delta-L_\delta)f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P_\delta(d\omega)$$

is estimated by

$$\int_0^{T_1}\int\Bigl(|a^{ij}-q^{ij}||\partial_{x_i}\partial_{x_j}f|+|b^i-z^i||\partial_{x_i}f|\Bigr)\,d\mu_t\,dt,$$

which can be made arbitrarily small by a suitable choice of q^{ij} and z^i as above. Thus, we have verified (i), (ii), (iii) for P. Therefore, for completing the proof of the theorem it suffices to show that for each fixed $\delta>0$ there exists a suitable measure P_δ for the solution μ_t^δ to the Cauchy problem $\partial_t\mu_t^\delta=L_\delta^*\mu_t^\delta$ with $\mu_0^\delta=\mu_\delta$.

II. Smoothing of the coefficients and verification of the conditions with the Lyapunov function.

Let us fix $\delta\in(0,T_1-T)$. Let $\zeta\in C^\infty([0,+\infty))$, $0\le\zeta\le 1$, $\zeta'\le 0$, $\zeta(t)=1$ if $t<1$ and $\zeta(t)=0$ if $t>2$. Set

$$\eta(t)=\int_t^{+\infty}\zeta(s)\,ds.$$

It is clear that $\eta\ge 0$, $\eta(t)=0$ if $t>2$ and $\eta'(t)=-\zeta(t)$. Let c_1 and c_2 be numbers such that

$$c_1\int_{\mathbb{R}^d}\zeta(|x|^2)\,dx=1,\quad c_2\int_{\mathbb{R}}\zeta(|t|^2)\,dt=1.$$

For every ε with $0<\varepsilon<\min\{\delta/16,1/2\}$ set

$$h_\varepsilon(t,x)=c_1c_2\varepsilon^{-d-1}\zeta(|t|^2/\varepsilon^2)\zeta(|x|^2/\varepsilon^2).$$

Let γ be the standard Gaussian density on $\mathbb{R}^d$. Set

$$\sigma^\varepsilon(t,x)=\varepsilon\gamma(x)+(1-\varepsilon)\int\int h_\varepsilon(t-s,x-y)\mu_s^\delta(dy)\,ds,$$

where the integration is formally taken over all of $\mathbb{R}^{d+1}$. However, we take into account that the function $h_\varepsilon(t-s,x-y)$ vanishes if $s\le-\delta/2$ or $s\ge T+\delta/2$, so that actually the integration in s is taken within the limits $-\delta/2$ and $T+\delta/2$ and for such s the measures μ_s^δ are defined. It is clear that $\sigma^\varepsilon>0$ and

$$\int\sigma^\varepsilon(t,x)\,dx=1.$$

In addition, for every function $f\in C_0^\infty(\mathbb{R}^d)$ and every $t\in[0,T]$ we have

$$\lim_{\varepsilon\to 0}\int f(x)\sigma^\varepsilon(t,x)\,dx=\int f(x)\,\mu_t^\delta(dx).$$

Indeed,

$$\int f(x)\sigma^\varepsilon(t,x)\,dx = \varepsilon\int f\gamma\,dx+$$
$$+(1-\varepsilon)\int c_2\varepsilon^{-1}\zeta(|t-s|^2/\varepsilon^2)\int\Bigl(\int f(x)c_1\varepsilon^{-d}\zeta(|x-y|^2/\varepsilon^2)\,dx\Bigr)\,\mu_s^\delta(dy)\,ds.$$

Since

$$\sup_y\Bigl|\int f(x)c_1\varepsilon^{-d}\zeta(|x-y|^2/\varepsilon^2)\,dx - f(y)\Bigr| \le \varepsilon\sup|\nabla f|c_1\int |x|\zeta(|x|^2)\,dx,$$

it suffices to show that the limit of the expression

$$\int c_2\varepsilon^{-1}\zeta(|t-s|^2/\varepsilon^2)\int f(y)\,\mu_s^\delta(dy)\,ds$$

is equal to the integral of f against μ_t^δ. This follows immediately by the continuity of the function

$$s\mapsto\int f(y)\,\mu_s^\delta(dy).$$

Thus, for every $t\in[0,T]$ the measures $\sigma^\varepsilon(t,x)\,dx$ converge weakly to μ_t^δ.

We recall that $b_\delta(t,x)=b(t+\delta,x)$ and $A_\delta(t,x)=A(t+\delta,x)$. Set

$$\beta_\varepsilon^i(t,x)=\frac{1-\varepsilon}{\sigma^\varepsilon(t,x)}\int\int b_\delta^i(s,y)h_\varepsilon(t-s,x-y)\,\mu_s^\delta(dy)\,ds,$$
$$\alpha_\varepsilon^{ij}(t,x)=\frac{1-\varepsilon}{\sigma^\varepsilon(t,x)}\int\int a_\delta^{ij}(s,y)h_\varepsilon(t-s,x-y)\,\mu_s^\delta(dy)\,ds.$$

Recall that γ is the standard Gaussian density on $\mathbb{R}^d$. We shall deal with the operator

$$\mathcal{L}_\varepsilon u(t,x)=\mathrm{trace}(\alpha_\varepsilon(t,x)D^2u(x))+\langle\beta_\varepsilon(t,x),\nabla u(x)\rangle+\frac{\varepsilon\gamma(x)}{\sigma^\varepsilon(t,x)}\bigl(\Delta u(x)-\langle x,\nabla u(x)\rangle\bigr),$$

which should not be confused with the previously defined L_ε; moreover, $\mathcal{L}_\varepsilon$ depends also on δ, which is now fixed and is not shown in this notation. Set also

$$\mathcal{A}_\varepsilon=\alpha_\varepsilon+\frac{\varepsilon\gamma(x)}{\sigma^\varepsilon(t,x)}I.$$

It is readily seen that σ^ε solves on $[0,T]\times\mathbb{R}^d$ the Cauchy problem

$$\partial_t\sigma^\varepsilon=\mathcal{L}_\varepsilon^*\sigma^\varepsilon,\quad \sigma^\varepsilon(0,x)=\varepsilon\gamma(x)+(1-\varepsilon)\int\int h_\varepsilon(s,x-y)\,\mu_s^\delta(dy)\,ds.$$

We now investigate the expression $\langle\beta_\varepsilon(t,x),x\rangle$. We have

$$\langle\beta_\varepsilon(t,x),x\rangle=\frac{1-\varepsilon}{\sigma^\varepsilon(t,x)}\int\int\langle b_\delta(s,y),y\rangle h_\varepsilon(t-s,x-y)\mu_s^\delta(dy)\,ds$$
$$+\frac{1-\varepsilon}{\sigma^\varepsilon(t,x)}\int\int\langle b_\delta(s,y),x-y\rangle h_\varepsilon(t-s,x-y)\,\mu_s^\delta(dy)\,ds.$$

Let us consider the expression

$$\int\int\langle b_\delta(s,y),x-y\rangle h_\varepsilon(t-s,x-y)\,\mu_s^\delta(dy)\,ds$$
$$=\int c_2\varepsilon^{-1}\zeta(|t-s|^2/\varepsilon)\Bigl(\int\langle b_\delta(s,y),x-y\rangle c_1\varepsilon^{-d}\zeta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\Bigr)\,ds.$$

We observe that $\zeta=-\eta'$ and

$$\langle b_\delta(s,y),x-y\rangle c_1\varepsilon^{-d}\zeta(|x-y|^2/\varepsilon^2)=-2^{-1}c_1\varepsilon^2\Bigl\langle b_\delta(s,y),\nabla_x\bigl(\varepsilon^{-d}\eta(|x-y|^2/\varepsilon^2)\bigr)\Bigr\rangle.$$

Therefore,

$$\int\int \langle b_\delta(s,y), x-y\rangle h_\varepsilon(t-s,x-y)\,\mu_s^\delta(dy)\,ds$$
$$= -2^{-1}c_1c_2\varepsilon^2\partial_{x_i}\biggl(\int\int b_\delta^i(s,y)\varepsilon^{-d-1}\zeta(|t-s|^2/\varepsilon^2)\eta(|x-y|^2/\varepsilon^2)\mu_s^\delta(dy)\,ds\biggr).$$

Recall that

$$\partial_t\mu_t^\delta = \partial_{x_i}\partial_{x_j}(a_\delta^{ij}\mu_t^\delta) - \partial_{x_i}(b_\delta^i\mu_t^\delta)$$

on $(-\delta, T+\delta)\times\mathbb{R}^d$ and for every fix $(t,x)\in[0,T]\times\mathbb{R}^d$ the function

$$\zeta(|t-s|^2/\varepsilon^2)\eta(|x-y|^2/\varepsilon^2)$$

has compact support in $(-\delta, T+\delta)\times\mathbb{R}^d$. Thus, there holds the equality

$$-\partial_{x_i}\biggl(\int\int b_\delta^i(s,y)\varepsilon^{-d-1}\zeta(|t-s|^2/\varepsilon^2)\eta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds\biggr)$$
$$= \partial_t\biggl(\int\int \varepsilon^{-d-1}\zeta(|t-s|^2/\varepsilon^2)\eta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds\biggr)$$
$$-\partial_{x_i}\partial_{x_j}\biggl(\int\int a_\delta^{ij}(s,y)\varepsilon^{-d-1}\zeta(|t-s|^2/\varepsilon^2)\eta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds\biggr).$$

We can estimate the terms in the right-hand as follows:

$$\partial_t\biggl(\int\int \varepsilon^{-d-1}\zeta(|t-s|^2/\varepsilon^2)\eta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds\biggr)$$
$$\le 2\varepsilon^{-d-3}\int\int |t-s||\zeta'(|t-s|^2/\varepsilon^2)|\eta(|x-y|^2/\varepsilon^2)\mu_s^\delta(dy)\,ds,$$

$$-\partial_{x_i}\partial_{x_j}\biggl(\int\int a_\delta^{ij}(s,y)\varepsilon^{-d-1}\zeta(|t-s|^2/\varepsilon^2)\eta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds\biggr)$$
$$= 4\varepsilon^{-d-5}\int\int\langle A_\delta(x-y),(x-y)\rangle\zeta(|t-s|^2/\varepsilon^2)\zeta'(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds$$
$$+2\varepsilon^{-d-3}\int\int(\mathrm{trace}\,A_\delta)\zeta(|t-s|^2/\varepsilon^2)\zeta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds$$
$$\le 2\varepsilon^{-d-3}\int\int(\mathrm{trace}\,A_\delta)\zeta(|t-s|^2/\varepsilon^2)\zeta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds.$$

For obtaining the last inequality we have used that $\zeta'\le 0$ and

$$\langle A_\delta(x-y),(x-y)\rangle \ge 0.$$

We observe that whenever $|x-y|\le 2\varepsilon\le 1$ one has

$$\frac{1}{1+|x|^2}\le\frac{3}{1+|y|^2}.$$

Thus, we have obtained the estimate

$$\frac{\langle\beta_\varepsilon(t,x),x\rangle}{1+|x|^2}\le\frac{3}{\sigma^\varepsilon(t,x)}\int\int\frac{|\langle b_\delta,y\rangle|}{1+|y|^2}h_\varepsilon(t-s,x-y)\,\mu_s^\delta(dy)\,ds$$
$$+\frac{c_1c_2\varepsilon^{-d-1}}{\sigma^\varepsilon(t,x)}\int\int|t-s||\zeta'(|t-s|^2/\varepsilon^2)|\eta(|x-y|^2/\varepsilon^2)\,\mu_s^\delta(dy)\,ds$$
$$+\frac{3c_1c_2\varepsilon^{-d-1}}{\sigma^\varepsilon(t,x)}\int\int\frac{\mathrm{trace}\,A_\delta}{1+|y|^2}\zeta(|t-s|^2/\varepsilon^2)\zeta(|x-y|^2/\varepsilon^2))\,\mu_s^\delta(dy)\,ds.$$

Let us denote the right-hand side of this inequality by $W_1(t,x)$ and observe that $W_1 \geq 0$ and

$$\int_0^T \int W_1(t,x)\sigma^\varepsilon(t,x)\,dx\,dt \leq C_4 \int_{-\delta}^{T+\delta} \int \frac{|\langle b_\delta, y\rangle| + |\mathrm{trace}\, A_\delta|}{1+|y|^2}\,\mu_s^\delta(dy)\,ds \\ + c_1 c_2 \int\int |t|\,|\zeta'(t^2)|\eta(|x|^2)\,dx\,dt,$$

where C_4 does not depend on ε. The function

$$\frac{|\alpha_\varepsilon^{ij}(t,x)|}{1+|x|^2}$$

is estimated by

$$W_2(t,x) := \frac{3}{\sigma^\varepsilon(t,x)} \int\int \frac{|a_\delta^{ij}(s,y)|}{1+|y|^2} h_\varepsilon(t-s, x-y)\,\mu_s^\delta(dy)\,ds.$$

We observe that $W_2 \geq 0$ and

$$\int_0^T \int W_2(t,x)\sigma^\varepsilon(t,x)\,dx\,dt \leq 3 \int_{-\delta}^{T+\delta} \int \frac{|a_\delta^{ij}(s,y)|}{1+|y|^2}\,\mu_s^\delta(dy)\,ds.$$

Set

$$W_3(t,x) = \frac{\varepsilon\gamma(x)}{(1+|x|^2)\sigma^\varepsilon(t,x)}.$$

Note that

$$\int_0^T \int W_3(t,x)\sigma^\varepsilon(t,x)\,d\,dt \leq T \int (1+|x|^2)^{-1}\gamma(x)\,dx.$$

Thus, we arrive at the estimates

$$\mathcal{L}_\varepsilon \log(1+|x|^2) \leq C_5(W_1+W_2+W_3), \quad \Big|\sqrt{\mathcal{A}_\varepsilon}\nabla \log(1+|x|^2)\Big|^2 \leq C_5(W_2+W_3),$$

where C_5 does not depend on ε. Note that for our function $V(x) = \log(1+|x|^2)$ we have

$$\mathcal{L}_\varepsilon\theta(V) = \theta''(V)|\sqrt{\mathcal{A}_\varepsilon}\nabla V|^2 + \theta'(V)\mathcal{L}_\varepsilon V \leq C_5(W_1+W_2+W_3),$$

and

$$|\sqrt{\mathcal{A}_\varepsilon}\nabla\theta(V)|^2 \leq |\sqrt{\mathcal{A}_\varepsilon}\nabla V|^2 \leq C_5(W_2+W_3).$$

Moreover,

$$\int \theta(V(x))\sigma^\varepsilon(0,x)\,dx \leq \varepsilon \int V(x)\gamma(x)\,dx \\ + (1-\varepsilon)\int\int\int h_\varepsilon(s,x-y)V(x)\,\mu_s^\delta(dy)\,dx\,ds.$$

Note that $\log(1+|x|^2) \leq \log(1+2|x-y|^2+2|y|^2) \leq \log(3+2|y|^2)$ if $|x-y| \leq 1$. Recall also that

$$\sup_{t\in[0,T]} \int \theta(\log(3+2|y|^2))\,\mu_t^\delta(dy) \leq N_\theta.$$

Hence

$$\int \theta(V(x))\sigma^\varepsilon(0,x)\,dx \leq \int V(x)\gamma(x)\,dx + N_\theta.$$

Applying Proposition 2.2 we obtain

$$\int_0^T \int \Big(|\mathcal{L}_\varepsilon\theta(V)| + |\sqrt{\mathcal{A}_\varepsilon}\nabla\theta(V)|^2\Big)\sigma^\varepsilon(t,x)\,dx\,dt \leq C_6, \tag{3.1}$$

where C_6 does not depend on ε (recall that $V(x) = \log(1+|x|^2)$).

III. The representation of σ^ε by a solution to the martingale problem with $\mathcal{L}_\varepsilon$.

According to [24, Theorem 3.5] (see also [12, Theorem 9.4.3]), the function σ^ε is a unique subprobability solution to the Cauchy problem. It is shown in [24] and it is very important in our situation that the integrability condition (3.1) just for one probability solution σ^ε implies the uniqueness in the class of all sub-probability solutions.

We show that there exists a solution to the martingale problem with the operator $\mathcal{L}_\varepsilon$ and initial condition $\sigma^\varepsilon(0,x)\,dx$. Let $\varphi_N(x) = \varphi(x/N)$, $0 \le \varphi \le 1$, $\varphi \in C_0^\infty(\mathbb{R}^d)$ and $\varphi(x) = 1$ if $|x| < 1$. Set

$$\mathcal{L}_\varepsilon^N = \varphi_N \mathcal{L}_\varepsilon.$$

According to the Trevisan result (for the case of Dirac's initial condition, see also [27, Theorem 3.2.6]), there exists a solution Q_ε^N to the martingale problem associated with $\mathcal{L}_\varepsilon^N$ and the initial condition $\sigma^\varepsilon(0,x)\,dx$. Let $\varrho_t^N(dx)dt$ be the corresponding probability solution to the Cauchy problem with $\mathcal{L}_\varepsilon^N$.

As in the proof of [24, Theorem 2.5] and [12, Theorem 6.7.3], one can choose a subsequence $\{N_k\}$ such that $\varrho_t^{N_k}(dx)\,dt$ converges weakly to $\varrho_t(dx)\,dt$ on every compact set in $[0,T]\times\mathbb{R}^d$ and $\varrho_t^{N_k}$ converges weakly to ϱ_t on every compact set in $\mathbb{R}^d$ for each $t \in [0,T]$. Moreover, $\varrho_t\,dt$ is a sub-probability solution to the Cauchy problem with $\mathcal{L}_\varepsilon$ and initial condition $\sigma^\varepsilon(0,x)\,dx$. By the cited uniqueness result for the Cauchy problem we have $\varrho_t(dx) = \sigma^\varepsilon(t,x)\,dx$.

Let us prove that the family of measures $Q_\varepsilon^{N_k}$ is compact in the weak topology. Let $q > 1$ and $\zeta_q(t) = t$ if $t < q-1$ and $\zeta_q(t) = q$ if $t > q+1$, $0 \le \zeta_q'(t) \le 1$, $-c \ge \zeta_q'' \le 0$, where c does not depend on q. Set $V_1 = \theta(V)$. Note that

$$Q_\varepsilon^{N_k}\Bigl(\omega\colon \sup_{t\in[0,T]} V_1(\omega(t)) \ge q-1\Bigr) = Q_\varepsilon^{N_k}\Bigl(\omega\colon \sup_{t\in[0,T]} \zeta_q(V_1(\omega(t))) \ge q-1\Bigr)$$

Repeating the arguments from Proposition 2.5 with $\zeta_q(V_1)$ in place of V and taking into account that $\zeta_q'(V_1) = \zeta_q''(V_1) = 0$ if $V_1 > q+1$ we obtain the estimate

$$\begin{aligned} &Q_\varepsilon^{N_k}\Bigl(\omega\colon \sup_{t\in[0,T]} V_1(\omega(t)) \ge q-1\Bigr) \\ &\quad\le \frac{2C_7}{(q-1)}\Bigl(\int V_1(x)\sigma^\varepsilon(0,x)\,dx + \int_0^T \int_{V_1\le q+1} \Bigl(|\mathcal{L}_\varepsilon V_1| + |\sqrt{\mathcal{A}_\varepsilon}\nabla V_1|^2\Bigr)\varrho_t^{N_k}\,dx\,dt\Bigr). \end{aligned}$$

Since $\{x\colon V_1(x) \le q+1\}$ is a compact set, for sufficiently large N_k the last integral is close to

$$\int_0^T \int_{V_1\le q+1} \Bigl(|\mathcal{L}_\varepsilon V_1| + |\sqrt{\mathcal{A}_\varepsilon}\nabla V_1|^2\Bigr)\sigma^\varepsilon(t,x)\,dx\,dt.$$

Thus, for every $\lambda \in (0,1)$ one can take q so large that there exists a number k_0 such that for every $k > k_0$ we have

$$Q_\varepsilon^{N_k}\Bigl(\omega\colon \sup_{t\in[0,T]} V_1(\omega(t)) \ge q-1\Bigr) \le \lambda.$$

It follows that for every $\lambda \in (0,1)$ there exists $R > 0$ such that

$$Q_\varepsilon^{N_k}\Bigl(\omega\colon \|\omega(t)\| \le R\Bigr) \ge 1-\lambda \quad \forall N_k.$$

Repeating the arguments from the first part of the proof and using the functions Θ_1 and Θ_2 appearing from Trevisan's result and de la Vallée Poussin's theorem for $\|A(t,x)\|I_{|x|\le R}$ and $|b(t,x)|I_{|x|\le R}$, we obtain the estimate

$$\begin{aligned} \mathbb{E}_{Q_\varepsilon^{N_k}} \Psi_d(f_i) \le C_8 \sup_x |\Theta(x_i\psi_R(x))| \\ + C_8 T \sup_{x,t}\Bigl(\Theta_1(|\mathcal{L}_\varepsilon(x_i\psi_R(x))|) + \Theta_2(\|\mathcal{A}_\varepsilon(t,x)\|)\Bigr) \end{aligned}$$

with the functions $\Theta(t) = \Theta_1(t) = \Theta_2(t) = t^2$. Here the right-hand side does not depend on N_k. We have

$$\mathbb{E}_{Q_\varepsilon}(I_{\|\omega\|\le R}\Psi_d) \le C_9(R),$$

where $C_9(R)$ depends on R, but does not depend on ε. For any number $\lambda \in (0,1)$ one can take a sufficiently large number M such that for the compact set $K = \{\omega\colon \Psi_d(\omega) \le M,\ \|\omega\| \le R\}$ there holds the estimate

$$Q_\varepsilon^{N_k}(K) \ge 1 - 2\lambda \quad \forall N_k.$$

Therefore, the family of measures $Q_\varepsilon^{N_k}$ contains a sequence $\{Q_\varepsilon^{N_{k_j}}\}$ weakly converging to some probability measure Q_ε. Let us verify that Q_ε corresponds to the solution σ^ε. Clearly, conditions (i) and (iii) are fulfilled. Note that if $f \in C_0^\infty(\mathbb{R}^d)$, then $\varphi_N \mathcal{L}_\varepsilon f = \mathcal{L}_\varepsilon f$ for all sufficiently large N. Hence (ii) follows from weak convergence of $Q_\varepsilon^{N_{k_j}}$.

Thus, for every ε there exists a probability measure Q_ε on Ω_d for which (i), (ii), (iii) are fulfilled with the operator $\mathcal{L}_\varepsilon$ and σ^ε solves the Cauchy problem. Moreover, by Proposition 2.5 for every $\lambda \in (0,1)$ there exists $R > 0$ such that

$$Q_\varepsilon\Bigl(\omega\colon \|\omega\| \le R\Bigr) \ge 1 - \lambda \quad \forall \varepsilon > 0.$$

IV. Verification of the compactness of the family of measures Q_ε and the proof of the fact that the limit is the required measure.

We need the following version of Jensen's inequality. Let Φ be a convex and increasing function on $[0, +\infty)$ with $\Phi(0) = 0$, ν a subprobability measure on some space X, and $f \ge 0$ a measurable function on X. Then

$$\Phi\biggl(\int_X f\, d\nu\biggr) \le \int_X \Phi(f)\, d\nu,$$

which follows by Jensen's inequality and the inequality $\Phi(\alpha t) \le \alpha \Phi(t)$ for $\alpha \in [0,1]$ that follows from the convexity of Φ. Let $\chi_R \ge 0$ be a smooth function such that $\chi_R(x) = 1$ if $|x| \le 2R$ and $\chi_R(x) = 0$ if $|x| \ge 3R$. Note that

$$\frac{\varepsilon\gamma(x)}{\sigma^\varepsilon(t,x)} \le 1.$$

Hence

$$\Phi\Bigl(|\beta_\varepsilon(x,t)|\chi_R(x) + \frac{\varepsilon\gamma(x)|x|}{\sigma^\varepsilon(t,x)}\chi_R(x)\Bigr) \le \frac{1}{2}\Phi(2|\beta_\varepsilon(x,t)|\chi_R(x)) + \frac{1}{2}\Phi(6R).$$

Since

$$\frac{1-\varepsilon}{\sigma^\varepsilon(t,x)} \int\!\!\int h_\varepsilon(t-s, x-y)\, \mu_s^\delta(dy)\, ds \le 1,$$

one has

$$\begin{aligned}
\Phi(2|\beta_\varepsilon(x,t)|\chi_R(x)) &= \Phi\biggl(\frac{1-\varepsilon}{\sigma^\varepsilon(t,x)} \int\!\!\int 2|b_\delta(s,y)|\chi_R(x) h_\varepsilon(t-s, x-y)\, \mu_s^\delta(dy)\, ds\biggr) \\
&\le \frac{1-\varepsilon}{\sigma^\varepsilon(t,x)} \int\!\!\int \Phi(2|b_\delta(s,y)|\chi_R(x)) h_\varepsilon(t-s, x-y)\, \mu_s^\delta(dy)\, ds.
\end{aligned}$$

Therefore, we obtain

$$2\int_0^T \int \Phi\Bigl(|\beta_\varepsilon|\chi_R + \frac{\varepsilon\gamma(x)|x|}{\sigma^\varepsilon(t,x)}\chi_R\Bigr)\, \sigma^\varepsilon(x,t)\, dx\, dt \le \int_0^{T_1} \int \Phi(2|b|\chi_R)\, d\mu_t\, dt + T_1\Phi(6R).$$

In the same way we obtain

$$2\int_0^T \int \Phi(\|\mathcal{A}_\varepsilon\|\chi_R)\, \sigma^\varepsilon(x,t)\, dx\, dt \le \int_0^{T_1} \int \Phi(2\|A\|\chi_R)\, d\mu_t\, dt + \Phi(2).$$

Analogous estimates are fulfilled for

$$\|\mathcal{A}_\varepsilon\| + |\beta_\varepsilon| + \frac{\varepsilon\gamma(x)|x|}{\sigma^\varepsilon(t,x)}.$$

Repeating the arguments from the first part of the proof and again using the functions Θ_1 and Θ_2 appearing from Trevisan's result and de la Vallée Poussin's theorem for $\|A(t,x)\|I_{|x|\le R}$ and $|b(t,x)|I_{|x|\le R}$, we obtain the estimate

$$\mathbb{E}_{Q_\varepsilon}\Psi_d(f_i) \le \int \Theta(x_i\psi_R(x))\,\mu_\delta(dx) \\ + \int\int\Bigl(\Theta_1(|\mathcal{L}_\varepsilon(x_i\psi_R(x))|) + \Theta_2(|\sqrt{\mathcal{A}_\varepsilon}\nabla(x_i\psi_R(x))|^2)\Bigr)\,\sigma^\varepsilon(x,t)\,dx\,dt,$$

where $f_i(x) = x_i\psi_R(x)$ and the right-hand side is estimated by

$$\sup_x |\Theta(x_i\psi_R(x))| \\ + C_{10}\int_0^{T_1}\int_{|x|\le 2R}\Bigl(\Theta_1(\|A(t,x)\| + |b(t,x)|) + \Theta_2(\|A(t,x)\|)\Bigr)\,\mu_t(dx)\,dt \\ + \Theta_1(6R) + \Theta_2(2),$$

which does not depend on ε. Here we apply the above estimates with $\Phi = \Theta_1$ and $\Phi = \Theta_2$ for the measure Q_ε. We have

$$\mathbb{E}_{Q_\varepsilon}(I_{\|\omega\|\le R}\Psi_d) \le C_{11}(R)$$

where $C_{11}(R)$ depends on R, but does not depend on ε. For any number $\lambda \in (0,1)$ one can take a sufficiently large number M such that for the compact set $K = \{\omega\colon \Psi_d(\omega) \le M,\ \|\omega\| \le R\}$ there holds the estimate

$$Q_\varepsilon(K) \ge 1 - 2\lambda \quad \forall\varepsilon > 0.$$

Therefore, the family of measures Q_ε contains a sequence $\{Q_{\varepsilon_k}\}$ with $\varepsilon_k \to 0$ weakly converging to some probability measure P_δ.

Let us verify that the measure P_δ satisfies (i), (ii) and (iii). The first and last properties are obtained in the limit letting $\varepsilon_k \to 0$ in the equality

$$\int_{\Omega_d} f(\omega(t))\,Q_{\varepsilon_k}(d\omega) = \int f(x)\sigma^{\varepsilon_k}(x,t)\,dx \quad \forall f \in C_0^\infty(\mathbb{R}^d).$$

For the proof of (ii) (the martingale property), as above, we have to show that for every bounded continuous function $g\colon \Omega_d \to \mathbb{R}$ that is measurable with respect to the σ-algebra $\mathcal{F}_s$ there holds the equality

$$\int_{\Omega_d}\Bigl[f(\omega(t)) - f(\omega(s)) - \int_s^t L_\delta f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P_\delta(d\omega) = 0, \quad t \ge s,$$

where $f \in C_0^\infty(\mathbb{R}^d)$. To this end, exactly as at the first step, it suffices to show that

$$\lim_{\varepsilon_k\to 0}\int_{\Omega_d}\Bigl[\int_s^t \mathcal{L}_{\varepsilon_k} f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,Q_{\varepsilon_k}(d\omega) = \int_{\Omega_d}\Bigl[\int_s^t L_\delta f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P_\delta(d\omega).$$

Let $q^{ij}, z^i \in C^\infty([-1,T_1]\times\mathbb{R}^d)$ and $\widetilde{L} = q^{ij}\partial_{x_i}\partial_{x_j} + z^i\partial_{x_i}$. It is clear that the difference

$$\int_{\Omega_d}\Bigl[\int_s^t \widetilde{L} f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,Q_{\varepsilon_k}(d\omega) - \int_{\Omega_d}\Bigl[\int_s^t \widetilde{L} f(\tau,\omega(\tau))\,d\tau\Bigr]g(\omega)\,P_\delta(d\omega)$$

tends to zero again by Lemma 2.1. Moreover, the expression

$$\Bigl|\int_{\Omega_d}\Bigl[\int_s^t (\widetilde{L} - L_\delta)f(\tau,\omega(\tau))\Bigr]g(\omega)\,P_\delta(d\omega)\Bigr|$$

is estimated by

$$\int_0^T\int_{\mathbb{R}^d}\Bigl(|a_\delta^{ij} - q^{ij}|\,|\partial_{x_i}\partial_{x_j} f| + |b_\delta^i - z^i|\,|\partial_{x_i} f|\Bigr)\,d\mu_t^\delta\,dt,$$

which can be made arbitrarily small by a suitable choice of q^{ij} and z^i approximating a^{ij}_δ and b^i_δ in L^1 with respect to the measure $\mu^\delta_t\,dt$ on $[0,T_1]\times U$, where U is a ball containing the support of f. Set

$$z^i_\varepsilon(t,x) = \frac{1-\varepsilon}{\sigma^\varepsilon(t,x)}\int\!\!\int z^i(s,y)h_\varepsilon(t-s,x-y)\,\mu^\delta_s(dy)\,ds,$$

$$q^{ij}_\varepsilon(t,x) = \frac{1-\varepsilon}{\sigma^\varepsilon(t,x)}\int\!\!\int q^{ij}(s,y)h_\varepsilon(t-s,x-y)\,\mu^\delta_s(dy)\,ds,$$

$$\widetilde{L}_\varepsilon = q^{ij}_\varepsilon\partial_{x_i}\partial_{x_j} + z^i_\varepsilon\partial_{x_i}.$$

Note that $z^i_\varepsilon - z^i$ and $q^{ij}_\varepsilon - q^{ij}$ converge to zero uniformly on $[0,T_1]\times U$ for every ball U. Since the functions $(\widetilde{L}-\widetilde{L}_{\varepsilon_k})f(t,x)$ converge to zero uniformly on $[0,T]\times\mathbb{R}^d$ as $\varepsilon_k\to 0$, the expression

$$\int_{\Omega_d}\biggl[\int_s^t(\widetilde{L}-\widetilde{L}_{\varepsilon_k})f(\tau,\omega(\tau))\,d\tau\biggr]g(\omega)\,Q_{\varepsilon_k}(d\omega)$$

tends to zero as $\varepsilon_k\to 0$. Let

$$C(f) = d\sup_x|\nabla f(x)| + d\sup_x\|D^2f(x)\|.$$

and $f(x)=0$ if $|x|>r$. Note that

$$\begin{aligned}
|(\widetilde{L}_{\varepsilon_k}-\mathcal{L}_{\varepsilon_k})f(t,x)| &\le \frac{1-\varepsilon_k}{\sigma^{\varepsilon_k}(t,x)}\int\!\!\int\bigl[|a^{ij}_\delta(s,y)-q^{ij}(s,y)||\partial_{x_i}\partial_{x_j}f(x)| \\
&\quad + |b^i_\delta(s,y)-z^i(s,y)||\partial_{x_i}f(x)|\bigr]h_{\varepsilon_k}(t-s,x-y)\,\mu^\delta_s(dy)\,ds \\
&\quad + \frac{\varepsilon_k\gamma}{\sigma^{\varepsilon_k}}(|\Delta f|+|x||\nabla f|) \\
&\le C(f)\frac{1-\varepsilon_k}{\sigma^{\varepsilon_k}(t,x)}\int\!\!\int_{|y|\le r+1}\bigl[|a^{ij}_\delta(s,y)-q^{ij}(s,y)|+|b^i_\delta(s,y)-z^i(s,y)|\bigr] \\
&\quad \times h_{\varepsilon_k}(t-s,x-y)\,\mu^\delta_s(dy)\,ds\,dx + \varepsilon_k C(f)\frac{(1+|x|)\gamma(x)}{\sigma^{\varepsilon_k}}.
\end{aligned}$$

Therefore, the integral

$$\int_{\Omega_d}\biggl[\int_s^t(\widetilde{L}_{\varepsilon_k}-\mathcal{L}_{\varepsilon_k})f(\tau,\omega(\tau))\,d\tau\biggr]g(\omega)\,Q_{\varepsilon_k}(d\omega)$$

is estimated by

$$C(f)\int_0^T\int_{|x|\le r+1}\Bigl(|a^{ij}_\delta-q^{ij}|+|b^i_\delta-z^i|\Bigr)\,d\mu^\delta_t\,dt + \varepsilon_k C(f)\int(1+|x|)\gamma(x)\,dx, \tag{3.2}$$

which can be made arbitrarily small by a suitable choice of q^{ij} and z^i as above. Denoting by $I(L,P)$ the integral

$$\int_{\Omega_d}\biggl[\int_s^t Lf(\tau,\omega(\tau))\,d\tau\biggr]g(\omega)\,P(d\omega),$$

we have

$$\begin{aligned}
I(\mathcal{L}_{\varepsilon_k},P_{\varepsilon_k}) - I(L_\delta,P_\delta) &= (I(\mathcal{L}_{\varepsilon_k},P_{\varepsilon_k}) - I(\widetilde{L}_{\varepsilon_k},P_{\varepsilon_k})) + (I(\widetilde{L}_{\varepsilon_k},P_{\varepsilon_k}) - I(\widetilde{L},P_{\varepsilon_k})) \\
&\quad + (I(\widetilde{L},P_{\varepsilon_k}) - I(\widetilde{L},P_\delta)) + (I(\widetilde{L},P_\delta) - I(L_\delta,P_\delta)).
\end{aligned}$$

Let $\lambda>0$. First we take q^{ij} and z^i such that for the first and forth terms in the right-hand side the integrals with μ^δ_t in the corresponding bounding expressions (3.2) are smaller than λ. Next we take k such that the part with ε_k in (3.2) for the first term and also the second and third terms are smaller than λ. It follows that $I(\mathcal{L}_{\varepsilon_k},P_{\varepsilon_k}) - I(L_\delta,P_\delta)\to 0$. Thus, we have verified (i), (ii), (iii) for P_δ.

V. Extension to the whole interval.

We have constructed martingale representations P_n defined on $C([0,T-1/n],\mathbb{R}^d)$ for every smaller interval $[0,T-1/n]$. It now remains to observe that Trevisan's a priori estimate employed above (that is, [28, Theorem A2 and Corollary A5]) enables one to construct a representation on the whole interval $[0,T]$ on which we have a solution to the Fokker–Planck–Kolmogorov equation. To this end we extend the measures P_n to $C([0,T],\mathbb{R}^d)$ by using the natural extension operator that associates to every function ω on $[0,T-1/n]$ the function that extends it by the constant value $\omega(T-1/n)$ on $(T-1/n,T]$. In addition, the compact function Ψ_d on $C[0,T-1/n],\mathbb{R}^d)$ ensured by Trevisan's result for each P_n (used at Step I) can be chosen in a unified way, namely, by taking such a function on $C([0,T],\mathbb{R}^d)$ and then restricting it to $C([0,T-1/n],\mathbb{R}^d)$ embedded into $C([0,T],\mathbb{R}^d)$ by means of extensions as explained above. It is readily seen from the formulation of the cited results from [28] mentioned above that in this way we obtain a compact function on Ω_d the integrals of which with respect to the extensions of P_n to Ω_d remain uniformly bounded (here it is important, of course, that in our condition (1.3) the integral is taken over all of $[0,T]$). Hence the sequence of extensions of measures P_n contains a weakly convergent subsequence. The limit of this subsequence gives the desired representation. The verification of this is analogous to the previous steps. Of course, the main point is check the martingale property, which is not automatic in case of discontinuous A and b, but follows again my smooth approximations and estimate (2.1). □

The main difficulty with the smoothing of coefficients is due to the necessity to obtain in the limit the solution we consider (but not an arbitrary solution, since there can exist many), in addition, for the approximating solutions we have to keep our Lyapunov-type condition.

Remark 3.1. If the Cauchy problem has a solution on the whole half-line, then a similar reasoning gives a representing martingale measure on the space of paths on $[0,+\infty)$, but here one must be careful, since this space is not separable, so the desired measure is defined not on all Borel sets, but on some smaller σ-field.

Finally, we give an example showing that the integrability of $(1+|x|)^{-2}|\langle b(t,x),x\rangle|$ can hold even in the case where the function $(1+|x|)^{-1}|b(t,x)|$ is not integrable with respect to the solution.

Example 3.2. Let $d=2$ and let (r,φ) be polar coordinates. We construct an example of a stationary solution $\mu=\varrho\,dx$ to the equation with the unit matrix A and the drift coefficient $b=\nabla\varrho/\varrho$ for a suitable function ϱ. We recall that

$$|\nabla\varrho(x)|^2=r^{-2}|\partial_\varphi\varrho|^2+|\partial_r\varrho|^2.$$

Therefore, it suffices to find a smooth nonnegative function ϱ for which

$$\varrho,(1+r)^{-1}|\partial_r\varrho|\in L^1(\mathbb{R}^2),\quad (1+r)^{-2}|\partial_\varphi\varrho|\notin L^1(\mathbb{R}^2).$$

Set

$$\varrho(r,\varphi)=\sum_{n=1}^{\infty}2^{-n}\psi(r-n)(2+\sin(4^n\varphi)),$$

where $\psi\in C_0^\infty((0,1))$ and $\psi\ge 0$ is not identically zero. We observe that for every point $(r,\varphi)\in(0,+\infty)\times[0,2\pi]$ only one term of the series is nonzero. It is clear that $\varrho\in C^\infty(\mathbb{R}^2)$, $\varrho\ge 0$ and

$$\varrho(r,\varphi)+|\partial_r\varrho(r,\varphi)|\le C2^{-r}$$

for some number $C>0$. However,

$$|\partial_\varphi\varrho(r,\varphi)|=\sum_{n=1}^{\infty}2^n\psi(r-n)|\cos(4^n\varphi)|.$$

Since the integral of $|\cos(4^n\varphi)|$ is estimated from below by 2π, we have

$$\int_0^\infty\int_0^{2\pi}r^{-1}|\partial_\varphi\varrho|\,d\varphi\,dr\ge\sum_{n=1}^{\infty}2\pi c_\psi(n+1)^{-1}2^n=+\infty,\quad c_\psi=\int_0^1|\psi(s)|\,ds.$$

Thus, $r^{-2}|\partial_\varphi\varrho| \notin L^1(\mathbb{R}^2)$. In this example b is a gradient. An example without this additional property is even simpler. We observe that the standard Gaussian density γ is a stationary solution to the equation with $A = I$ and $b(x) = -x$, so it remains a solution for the equation with a perturbed drift $-x + v(x)$, where a smooth vector field v is chosen such that $\operatorname{div}(\gamma v) = 0$ and $(x, v(x)) = 0$. For example, we can take v of the form $v(x) = \gamma(x)^{-1}h(|x|^2)Ux$ with an orthogonal operator U such that $(Ux, x) = 0$. Of course, h can be rapidly increasing, so that $|v|\gamma$ will not be integrable.

Remark 3.3. The presented results can be extended with minor technical changes to as follows. Let $V \in C^2([1, +\infty)$, there is $C > 0$ such that

$$\Bigl|\frac{V''(s)}{V''(t)}\Bigr| + \Bigl|\frac{V'(s)}{V'(t)}\Bigr| \le C \quad \text{whenever } |t-s| \le 1,$$

$V \ge 0$, $|V''| + |V'| \le C$ and $\lim\limits_{s\to+\infty} V(s) = +\infty$, i.e., the integral of V' diverges. Then condition (1.3) can be replaced by

$$\int_0^T \int_{\mathbb{R}^d} \Bigl[\Bigl(|V''(1+|x|^2)|(1+|x|^2) + |V'(1+|x|^2)|\Bigr)\|A(t,x)\| \\ + |\langle b(t,x), x\rangle|\,|V'(1+|x|^2)|\Bigr]\,\mu_t(dx)\,dt < \infty.$$

For $V(s) = \log s$ we obtain the original condition (1.3). If $V(s) = \log(1 + \log s)$, then we arrive at the condition

$$\frac{\|A(t,x)\|}{(1+|x|^2)\log(1+|x|^2)},\ \frac{|\langle b(t,x), x\rangle|}{(1+|x|^2)\log(1+|x|^2)} \in L^1(\mu_t\,dt).$$

Remark 3.4. The superposition principle applies not only to linear Fokker–Planck–Kolmogorov equations, but also to nonlinear equations. Let $\{\mu_t\}$ be a solution to the Cauchy problem

$$\partial_t\mu_t = \partial_{x_i}\partial_{x_j}(a^{ij}(t,x,\mu)\mu_t) - \partial_{x_i}(b^i(t,x,\mu)\mu_t), \quad \mu_0 = \nu.$$

For a precise definition of a solution and typical examples of dependence of A and b on the solution μ are given in [12, Chapter 6], [22], [23]. In particular, typical global assumptions are expressed in terms of a Lyapunov function V:

$$L_\mu V \le C(\mu) + C(\mu)V, \quad V \in L^1(\nu).$$

If $V(x) = \log(1 + |x|^2)$, then by Proposition 2.2 the solution $\{\mu_t\}$ satisfies condition (1.3). Given a solution $\{\mu_t\}$, we can regard it as a solution to the linear operator L_μ. Therefore, there exists the corresponding solution P_ν to the martingale problem such that μ_t is the one-dimensional distribution of the measure P_ν on $C([0,T], R^d)$. Hence we can assume that the measure P_ν solves the martingale problem with the operator L_μ that depends on P_ν through μ, i.e., solves the martingale problem corresponding to the stochastic McKean–Vlasov equation (see [18]). Thus, using the superposition principle and solutions to the Fokker–Planck–Kolmogorov equation one can construct solutions to the martingale problem for nonlinear stochastic equations. This approach is applied for constructing probabilistic representations of solutions to PDEs (see, e.g., [6]).

It is worth noting that the superposition principle can be useful for the study of uniqueness problems for Fokker–Planck–Kolmogorov equations with coefficients of low regularity, on this topic see the book [12] and the papers [13], [14], [15], [16], [20], [25], [29]. We also plan to study analogous questions for infinite-dimensional equations in the spirit of [8], [9] and [10].

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

References

[1] Ambrosio L. *Transport equation and Cauchy problem for non-smooth vector fields.* Lecture Notes in Math. 2008. V. 1927. P. 2–41.

[2] Ambrosio, L. Well posedness of ODE's and continuity equations with nonsmooth vector fields, and applications. Rev. Mat. Complut. 30 (2017), no. 3, 427–450.

[3] Ambrosio, L., Gigli, N., Savaré, G. Gradient flows in metric spaces and in the space of probability measures, 2nd ed., Birkhäuser, Basel, 2008.

[4] Ambrosio, L., Trevisan, D. Well-posedness of Lagrangian flows and continuity equations in metric measure spaces, Anal. PDE 7 (2014), no. 5, 1179–1234.

[5] Ambrosio L., Trevisan, D. Lecture notes on the DiPerna–Lions theory in abstract measure spaces. Ann. Fac. Sci. Toulouse Math. (6) 26 (2017), no. 4, 729–766.

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

[7] Bogachev V.I. Weak convergence of measures. Amer. Math. Soc., Rhode Island, Providence, 2018.

[8] Bogachev V.I., Da Prato G., Röckner M. Existence and uniqueness of solutions for Fokker–Planck equations on Hilbert spaces. J. Evol. Equ. 10 (2010), no. 3, 487–509.

[9] Bogachev V.I., Da Prato G., Röckner M. Uniqueness for solutions of Fokker–Planck equations on infinite dimensional spaces. Comm. Partial Differential Equations 36 (2011), no. 6, 925–939.

[10] Bogachev V.I., Da Prato G., Röckner M., Shaposhnikov S.V. An analytic approach to infinite-dimensional continuity and Fokker–Planck–Kolmogorov equations. Annali Scuola Norm. Pisa. V. 14, 983–1023 (2015).

[11] Bogachev V.I., Krylov N.V., Röckner M. Elliptic and parabolic equations for measures. Uspehi Matem. Nauk. 2009. V. 64, N 6. P. 5–116 (in Russian); English transl.: Russian Math. Surveys. 2009. V. 64, N 6. P. 973–1078.

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

[13] Bogachev V.I., Röckner M., Shaposhnikov S.V. On uniqueness problems related to elliptic equations for measures. J. Math. Sci. (New York). 2011. V. 176, N 6. P. 759–773.

[14] Bogachev V.I., Röckner M., Shaposhnikov S.V. On uniqueness problems related to the Fokker–Planck–Kolmogorov equation for measures. J. Math. Sci. (New York), V. 179, N 1, P. 7–47 (2011).

[15] Bogachev V.I., Röckner M., Shaposhnikov S.V. On uniqueness of solutions to the Cauchy problem for degenerate Fokker–Planck–Kolmogorov equations. J. Evol. Equat., V. 13, N 3, P. 577–593 (2013).

[16] Bogachev V.I. , Roeckner M., Shaposhnikov S.V. Uniqueness problems for degenerate Fokker–Planck–Kolmogorov equations. J. Math. Sci. (New York), V. 207, N 2, 147–165 (2015).

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

[18] Funaki T. A certain class of diffusion processes associated with nonlinear parabolic equations. Z. Wahrscheinlichkeitstheor. Verwandte Geb. 67 (1984), 331–348.

[19] Bonicatto P., Gusev N.A. Non-uniqueness of signed measure-valued solutions to the continuity equation in presence of a unique flow. 2018, arXiv:1809.10216.

[20] Le Bris, C., Lions, P.-L. Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients. Comm. Partial Differential Equations 33 (2008), no. 7-9, 1272–1317.

[21] Luo, D. The Itô SDEs and Fokker–Planck equations with Osgood and Sobolev coefficients. Stochastics 90 (2018), no. 3, 379–410.

[22] Manita O.A., Romanov, M.S., Shaposhnikov S.V. On uniqueness of solutions to nonlinear Fokker–Planck–Kolmogorov equations. Nonlinear Anal. 128 (2015), 199–226.

[23] Manita O.A., Shaposhnikov S.V. Nonlinear parabolic equations for measures. Algebra i Analiz 25 (2013), no. 1, 64–93 (in Russian); English transl.: St. Petersburg Math. J. 25 (2014), no. 1, 43–62.

[24] Manita O.A., Shaposhnikov S.V. On the Cauchy problem for Fokker–Planck–Kolmogorov equations with potential terms on arbitrary domains. J. Dynamics Differ. Equ. 28 (2016), 493–518.

[25] Shaposhnikov, S.V. On the uniqueness of the probabilistic solution of the Cauchy problem for the Fokker–Planck–Kolmogorov equation. Teor. Veroyatn. Primen. 56 (2011), no. 1, 77–99 (in Russian); English transl.: Theory Probab. Appl. 56 (2012), no. 1, 96–115.

[26] Stepanov, E., Trevisan, D. Three superposition principles: currents, continuity equations and curves of measures. J. Funct. Anal. 272 (2017), no. 3, 1044–1103.
[27] Stroock, D.V., Varadhan, S.R.S., Multidimensional diffusion processes, Springer-Verlag, Berlin, 2006.
[28] Trevisan, D. Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. Electron. J. Probab. 21 (2016), Paper No. 22, 41 pp.
[29] Zhang, X. Stochastic flows of SDEs with irregular coefficients and stochastic transport equations. Bull. Sci. Math. 134 (2010), no. 4, 340–378.