FORM-BOUNDEDNESS AND SDES WITH SINGULAR DRIFT

DAMIR KINZEBULATOV

Abstract. We survey and refine recent results on weak and strong well-posedness of stochastic differential equations with singular drift satisfying some minimal assumptions.

Contents

1. Introduction 2
2. Notations 7
3. Form-bounded drifts. Semigroup in L^2 9
4. Sharp solvability 12
4.1. Critical magnitude of drift singularities 15
5. Three approaches to constructing Feller semigroup for $-\Delta + b \cdot \nabla$ 17
6. Basic result on weak well-posedness of SDEs with singular drift 20
7. Proof of Theorem 6.1 26
8. Proof of Theorem 6.2 29
9. Time-inhomogeneous form-bounded drifts and Feller theory via iterations 32
10. SDEs with time-inhomogeneous form-bounded drifts 40
11. "Form-bounded" diffusion coefficients 42
12. Stochastic transport equation and strong solutions to SDEs 46
13. Strong well-posedness via Röckner-Zhao's approach 49
14. More singular than form-bounded. Semigroup in $\mathcal{W}^{\frac{1}{2},2}$ 51
15. Weakly form-bounded drifts and SDEs 56
16. Time-inhomogeneous drifts in Morrey class 62
17. SDEs driven by α-stable process 68

Appendix A. Proof of Lemma 6.1 74

Appendix B. Some examples of form-bounded vector fields 77

Appendix C. Smooth approximations of form-bounded (-type) vector fields 78

Appendix D. Trotter's approximation theorem 81

References 81

Université Laval, Département de mathématiques et de statistique, Québec, QC, Canada

E-mail address: `damir.kinzebulatov@mat.ulaval.ca`.

2010 *Mathematics Subject Classification.* 60H10, 47D07 (primary), 35J75 (secondary).

Key words and phrases. Form-boundedness, stochastic differential equations, weak solutions, singular drifts, Feller semigroups, Morrey class, weak well-posedness, Brownian motion, α-stable process.

November 7, 2023.

The research of the author is supported by the NSERC (grant RGPIN-2017-05567).

1. Introduction

Various applications in physics and technology dictate the need to work with stochastic differential equations

$$dX_t = -b(X_t)dt + \sqrt{2}dW_t, \quad X_0 = x \in \mathbb{R}^d, \tag{1.1}$$

having an irregular, locally unbounded drift $b : \mathbb{R}^d \to \mathbb{R}^d$. Here $\{W_t\}_{t\geq 0}$ is a d-dimensional Brownian motion in $\mathbb{R}^d$ defined on some complete filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbf{P})$. This naturally leads to the problem of finding the least restrictive assumptions on b that ensure well-posedness of (1.1), in one sense or another. More specifically, one asks: what integral characteristics of b determine whether there exists a unique solution of (1.1)? The same question arises when one considers more general SDEs, also dictated by applications:

$$dX_t = -b(t, X_t)dt + \sqrt{2}dW_t, \tag{1.2}$$

with drift $b : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ that can also be singular in time, and

$$dX_t = -b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{1.3}$$

with diffusion coefficients $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ that can be discontinuous. Regarding SDE (1.2), one illustrative example is the "passive tracer model" that describes the motion of a small particle in a turbulent flow, i.e. (1.2) with the velocity field b obtained by solving three-dimensional Navier-Stokes equations [MK].

The paper deals with weak and strong well-posedness of SDEs (1.1)-(1.3), for every initial point $x \in \mathbb{R}^d$. Recall that a weak solution of (1.1)-(1.3) is a pair of continuous processes $\{(X_t, W_t)\}_{t\geq 0}$ defined on some complete probability space, such that $\{W_t\}$ is a Brownian motion and the identity in (1.1)-(1.3) holds a.s. for all $t \geq 0$. In turn, a strong solution of (1.1)-(1.3) is a continuous process X_t that is adapted to the natural filtration of the Brownian motion $\{W_t\}$, and such that the identity in (1.1)-(1.3) holds a.s. for $t \geq 0$. That is, X_t is a strong solution if it only depends on time and the driving process $\{W_s\}_{0\leq s\leq t}$.

The question of what local singularities of b are admissible, so that SDEs (1.1)-(1.3) are weakly or strongly well-posed, was thoroughly studied in the literature. Below we give a brief outline of the results on multidimensional SDEs with the focus on the singularities of the drift. We will keep the chronological order of appearance of preprints, where applicable. However, we will be somewhat loose with the terminology by including in "well-posedness" uniqueness results of varying strength (in general, within some large classes of solutions).

Veretennikov [V] was the first who established, using Zvonkin's method [Zv], strong well-posedness of (1.2) when $|b|$ is bounded measurable. Portenko [P1] considered drift b in the sub-critical Ladyzhenskaya-Prodi-Serrin class

$$|b| \in L^l(\mathbb{R}, L^p(\mathbb{R}^d)), \quad p \geq d, l \geq 2, \quad \frac{d}{p} + \frac{2}{l} < 1 \tag{1.4}$$

and proved existence of weak solution to SDE (1.2) and its uniqueness in law. Krylov-Röckner further established, using Yamada-Watanabe theorem, that for such b the SDE (1.2) is, in fact, strongly well-posed. A number of important results for SDEs with drift satisfying (1.4) were established next by X. Zhang [Z1, Z2, Z3]. In the period between the papers of Portenko and Krylov-Röckner, Bass-Chen [BC] proved existence and uniqueness in law of weak solutions of (1.1) for time-homogeneous drift $b = b(x)$ in the Kato class of vector fields, with arbitrarily small δ, cf. (14.3). The Kato class of vector fields contains $\{|b| \in L^p(\mathbb{R}^d), p > d\}$ as well as some vector fields with entries not even in $L^{1+\varepsilon}_{\rm loc}(\mathbb{R}^d)$, $\varepsilon > 0$. However, Kato class does not contain $\{|b| \in L^d(\mathbb{R}^d)\}$. Speaking of time-homogeneous drifts, the fact that $p = d$ is the optimal exponent on the scale of Lebesgue spaces can be seen already from rescaling the parabolic equation. In [BFGM], Beck-Flandoli-Gubinelli-Maurelli developed an approach to establishing strong well-posedness of (1.2) with drift b in the critical Ladyzhenskaya-Prodi-Serrin (LPS) class

$$|b| \in L^l(\mathbb{R}, L^p(\mathbb{R}^d)), \quad p \geq d, l \geq 2, \quad \frac{d}{p} + \frac{2}{l} \leq 1. \tag{1.5}$$

for a.e. initial point $x \in \mathbb{R}^d$, via stochastic transport and stochastic continuity equations. They also discussed the following attracting drift:

$$b(x) = \sqrt{\delta}\frac{d-2}{2}\mathbf{1}_{|x|<1}|x|^{-2}x \tag{1.6}$$

(note that $|b|$ misses $L^d(\mathbb{R}^d)$ by a logarithmic factor) and provided a detailed proof of the fact that for

$$\delta > 4(\frac{d}{d-2})^2, \tag{1.7}$$

i.e. when the attraction to the origin by the drift is strong enough, then SDE (1.1) with initial point $x = 0$ does not have a weak solution. In [KiS1], Semënov and the author showed that the constructed in [Ki1] Feller generator $-\Delta + b \cdot \nabla$ for b satisfying weak form-boundedness condition (1.11), see below, determines, for every initial point $x \in \mathbb{R}^d$, a weak solution to SDE (1.1) that is unique among weak solutions that can be constructed via approximation. To the best of the author's knowledge, this was the first result on weak well-posedness of (1.1) that included both $|b| \in L^d(\mathbb{R}^d)$ and the model vector field (1.6) with δ small. It also included the elliptic Morrey class $M_{1+\varepsilon}$, see below, and the Kato class considered by Bass-Chen. The construction of the Feller generator with such b used in an essential manner some inequalities for symmetric Markov generators, and hence required time-homogeneity of the drift. Returning to time-inhomogeneous drifts, we note that almost at the same time Jin [J] proved weak well-posedness of (1.2) with time-inhomogeneous Kato class drifts, Wei-Lv-Wu [WLW] and Nam [N] obtained results on weak well-posedness of (1.2), for every $x \in \mathbb{R}^d$, for time-inhomogeneous vector fields b that can be more singular than the ones in (1.4). Nevertheless, their results excluded $b = b(x)$ with $|b| \in L^d(\mathbb{R}^d)$. In [XXZZh], Xia-Xie-Zhang-Zhao established, among other results, weak well-posedness for every initial point of SDE (1.1) with b having entries in $C_b(\mathbb{R}, L^d(\mathbb{R}^d))$. In [Kr1, Kr2, Kr3], Krylov proved weak and strong well-posedness of SDEs (1.1)-(1.3) with $|b|$ in $L^d(\mathbb{R}^d)$ and beyond, e.g. in a large Morrey class (in fact, below we use an argument from these papers to prove some gradient bounds). In [YZ], S. Yang-T. Zhang proved strong well-posedness of (1.2) for time-inhomogeneous drifts b with $|b|^2$ "almost" in the Kato class of potentials, cf. (3.9) (which, to make the comparison clear at least in the time-homogeneous case, is smaller than the Kato class of vector fields in [BC]; still, the class considered by Yang-Zhang contains some drifts with $|b| \notin L^{2+\varepsilon}_{\rm loc}$, $\varepsilon > 0$). In [RZh1], Röckner-Zhao established weak well-posedness of (1.1), for any initial point $x \in \mathbb{R}^d$, for drifts in $L^\infty(\mathbb{R}, L^{d,w}(\mathbb{R}^d))$, plus the drifts in the critical LPS class. Here $L^{d,w}(\mathbb{R}^d)$ denotes the weak L^d class that contains vector fields in $L^d(\mathbb{R}^d)$, as well as more singular vector fields, such as (1.6). In [RZh2], the authors obtained strong well-posedness of (1.1), for any initial point, with b in the critical LPS class. In [KiM1], Madou and the author established weak well-posedness of (1.2), for every initial point, for b in the class of time-inhomogeneous form-bounded drifts (containing (1.8) below). This class contains $L^\infty(\mathbb{R}, L^{d,w}(\mathbb{R}^d))$ and the critical LPS class, as well as some drifts that are not even in $L^\infty(\mathbb{R}, L^{2+\varepsilon}(\mathbb{R}^d))$ for a given $\varepsilon > 0$. In [KiS4], Semënov and the author proved existence of a weak solution to SDE (1.2) with time-inhomogeneous form-bounded drift, see (1.8), covering the entire range of admissible form-bounds $\delta < 4$, cf. (1.7). The critical value $\delta = 4$ was recently attained at the PDE level in [Ki6]. In [Ki5], the author established weak well-posedness of SDE (1.2) and proved Feller property for time-inhomogeneous drifts in essentially the largest possible parabolic Morrey class, which contains the class of time-inhomogeneous form-bounded drifts, the time-homogeneous Morrey drifts in $M_{1+\varepsilon}$, and allows to include drifts b having critical behaviour both in time and in space, e.g.

$$|b(t,x)| \leq \frac{c_1}{|x|} + \frac{c_2}{\sqrt{t}}, \quad t > 0, \quad x \in \mathbb{R}^d.$$

More recently, Krylov [Kr5] established weak well-posedness of SDE (1.3) with discontinuous diffusion coefficients in the VMO class and time-inhomogeneous drifts in a large parabolic Morrey class; restricted to time-homogeneous drifts his assumptions read as $|b| \in M_q$, $q > \frac{d}{2}$. This result was further refined by Krylov in [Kr6], see also [Kr7] regarding regularity theory of parabolic equations with VMO diffusion coefficients and drift and potential in Morrey classes.

Below we survey and refine recent results on weak and strong well-posedness of SDEs (1.1)-(1.3) with $|b|$ satisfying some minimal constraints, as in [KiS1, KiM1, KiS4, Ki5] mentioned above. For instance, regarding SDE (1.1), our assumption on the order of singularities of $|b|$ is basically that $-\Delta + b \cdot \nabla$ must generate a quasi contraction strongly continuous semigroup in L^2. That is, we will be assuming that $|b|$ is form-bounded:

$$|b|^2 \leq \delta(-\Delta) + c \quad \text{(in the sense of quadratic forms)} \tag{1.8}$$

for some constants δ and c. See Definition 3.1 below. This translates, by means of the Cauchy-Schwarz inequality, into the assumption of smallness of $b \cdot \nabla$ with respect to $-\Delta$. A broad elementary sufficient condition for (1.8) is the scaling-invariant Morrey class $M_{2+\varepsilon}$, i.e.

$$\|b\|_{M_{2+\varepsilon}} := \sup_{r>0, x\in\mathbb{R}^d} r\biggl(\frac{1}{|B_r|}\int_{B_r(x)} |b|^{2+\varepsilon} dx\biggr)^{\frac{1}{2+\varepsilon}} < \infty \tag{1.9}$$

where $B_r(x)$ is the ball of radius r centered at x. Here ε can be taken arbitrarily small. One has $\delta = C\|b\|_{M_{2+\varepsilon}}$ for appropriate constant C. The class $M_{2+\varepsilon}$ contains all $|b|$ in L^d or, more generally, in the weak L^d class (we recall its definition in Section 2).

Regarding the relationship between operator $-\Delta + b \cdot \nabla$ and SDE (1.1), one expects that for $b = b(x)$ the function

$$v(t,x) := \mathbf{E}_{X_0=x}[f(X_t)], \tag{1.10}$$

solves Cauchy problem

$$(\partial_t - \Delta + b \cdot \nabla)v = 0, \quad v|_{t=0} = f.$$

The intimate relationship between parabolic equations and SDEs is a consequence of the fact that both describe the same physical process of diffusion.

In Section 15, we discuss results on weak well-posedness of (1.1) under more general assumption on the drift than its form-boundedness. Namely, our hypothesis on the order of singularities of $|b|$ will be that $-\Delta + b \cdot \nabla$ generates a quasi contraction strongly continuous semigroup in the Bessel space $\mathcal{W}^{1/2,2}$, i.e. we will require

$$|b| \leq \delta(\lambda - \Delta)^{\frac{1}{2}} \quad \text{(in the sense of quadratic forms).} \tag{1.11}$$

Such vector fields b are called weakly form-bounded. This class of weakly form-bounded vector fields contains scaling-invariant Morrey class $M_{1+\varepsilon}$. It also contains the Kato class of vector fields.

One of our goals will be to bootstrap the semigroup in L^2 or in $\mathcal{W}^{\frac{1}{2},2}$ to a strongly continuous semigroup in C_∞, the space of continuous functions on $\mathbb{R}^d$ vanishing at infinity, endowed with the sup-norm (that is, to a Feller semigroup). This will come at the cost of restricting admissible values of constant δ (in terms of the Morrey class, this means that the Morrey norm cannot be too large). We emphasize that while the classes (1.8), (1.11) determine the order of singularities of drift b, the value of δ measures the magnitude of its singularities. We are particularly interested in the optimal assumptions on δ.

The requirement that there should be a properly defined operator behind (1.1) is reasonable, since it gives a reasonably complete solution theory of the corresponding parabolic equation. That being said, there are situations where one does not want to insist on a strong link between parabolic equations and stochastic processes in order to treat, in one sense or another, more singular drifts. See e.g. [W] which considers solutions of martingale problem with test functions cutting out the singular set of the drift, or [NU, Za1] which deal with elliptic or parabolic equations with supercritical (in the sense of scaling) drift where one no longer has Hölder continuity of solution. We are interested, on the contrary, in finding the maximal singularities of the drift that still give a more or less classical theory of parabolic equations and SDEs, including the possibility to consider SDEs with arbitrarily fixed initial point, e.g. in the singular set of the drift.

One natural question is: why not restrict attention to the Morrey class of drifts (1.9), a broad subclass of (1.8) defined in elementary terms (and which also allows to use e.g. harmonic-analytic arguments that are not available for form-bounded or weakly form-bounded drifts, cf. Remark 16.1)? First, there is a refinement of the Morrey class (1.9) due to Chang-Wilson-Wolff, which is still contained in the class of form-bounded vector fields (see Appendix B), and there is no reason to believe that their result itself cannot be refined, in elementary terms, even further. Second, and more importantly, by assuming e.g. form-boundedness of b (1.8) we impose a condition on b that is ultimate in some precise sense, i.e. the existence of quasi contraction strongly continuous semigroup in L^2. The latter is, arguably, an extra hypothesis on the diffusion process, but it is explicitly spelled out. Also, at least in the case of Schrödinger operators, see below, form-boundedness is a physical assumption on the potential.

We impose similar assumptions on the time-inhomogeneous drift in SDE (1.2). We will deal with SDE (1.3) with diffusion coefficients σ that can have some critical discontinuities, but only those that are allowed by the singularities of the drift b, which is our main focus.

In Section 17 we discuss weak well-posedness of SDE

$$X_t = x - \int_0^t b(X_s)ds + Z_t, \quad t \geq 0, \quad x \in \mathbb{R}^d, \tag{1.12}$$

driven by rotationally symmetric α-stable process Z_t, $1 < \alpha < 2$, with drift $b : \mathbb{R}^d \to \mathbb{R}^d$ satisfying

$$|b| \leq \delta(\lambda - \Delta)^{\frac{\alpha-1}{2}} \quad \text{(in the sense of quadratic forms)}, \tag{1.13}$$

e.g.

$$|b|^{\frac{1}{\alpha-1}} \in M_{1+\varepsilon}$$

for some $\varepsilon > 0$. There is a rich literature devoted to equation (1.12), which also covers the range $0 < \alpha \leq 1$ (see Remark 17.1). In the case $1 < \alpha < 2$, which allows to deal with locally unbounded drifts, earlier results on weak well-posedness of (1.12) include $|b| \in L^\infty$ due to Komatsu [Ko], $|b| \in L^p$, $p > \frac{d}{\alpha-1}$ due to Portenko and Podolynny-Portenko [P2, PP], and, more generally, $|b|$ in appropriate Kato class of vector fields, see Chen-Kim-Song [CKS], Kim-Song [KSo], Chen-Wang [CW]. All of these classes are contained in (1.13).

See also Priola [Pr] and X. Zhang [Z5] regarding strong well-posedness of (1.12) and its generalizations.

We discuss conditions on b stated in terms of $|b|$. The latter allows to include measure-valued drifts, see Remark 15.5. However, distributional drifts lie outside of the scope of this paper (regarding distributional drifts, see [FIR, CM, CJM, PZ, ZZh2] and references therein). We also do not discuss here many interesting results that require additional structure of b such as existence of non-positive divergence $\operatorname{div} b$ or b of the form $b = \nabla V$ for appropriate potential V; we only refer to Bresch-Jabin-Wang [BJW], Fournier-Jourdain [FJ], Röckner-Zhao [RZh1] and references therein.

The drifts that we consider in this paper in general destroy both the upper and the lower Gaussian bounds on the heat kernel of the corresponding to (1.1) and (1.2) parabolic equations.

In this paper we are interested in local singularities of the drift, although our drifts can still be unbounded at infinity along some subsets (e.g. $b(x) = \sum_{k=1}^\infty c_k |x - a_k|^{-2}(x - a_k)$, $x \in \mathbb{R}^d$, where $\sum_{k=1}^\infty |c_k|^{\frac{1}{2}} < \infty$ and $a_k \in \mathbb{R}^d$, $|a_k| \to \infty$).

Throughout the paper, dimension $d \geq 3$. Dimension $d = 2$ does not present an obstacle for our methods, however, in our opinion it deserves a separate study. The exposition of the results does not follow the chronological order of their appearance (on arXiv), but proceeds from weaker singularities of the drift to stronger singularities.

Structure of the paper. In Section 3 we introduce the class of time-homogeneous form-bounded vector fields.

In Section 4 we discuss the results on weak solvability of SDE (1.1) and on solution theory of the corresponding parabolic equation under essentially sharp assumption on the value of form-bound δ of the drift.

In Section 5 we describe two existing approaches to constructing a Feller semigroup associated with $-\Delta + b \cdot \nabla$, and introduce a new one.

In Sections 6, 7 and 8 we prove a detailed weak well-posedness result for (1.1) with form-bounded drift, however, at expense of requiring smaller δ.

In Sections 9 and 10 we discuss results on weak well-posedness of SDE (1.2) having time-inhomogeneous form-bounded drift. Their proofs use a different (iteration) technique.

In Section 11 we discuss an extension of the previous results to some discontinuous diffusion coefficients.

In Sections 12 and 13 we discuss strong well-posedness of SDEs (1.1) and (1.2), via stochastic transport equation and via relative compactness criterion for random fields on the Wiener-Sobolev space (Röckner-Zhao's approach).

In Sections 14 and 15 we substantially enlarge the class of admissible in SDE (1.1) time-homogeneous drifts (i.e. to weakly form-bounded drifts), but at expense of δ that now needs to be smaller than in Section 6.

In Section 16 we consider again time-inhomogeneous drifts and strengthen and simplify all aspects of the results from Sections 9 and 10 except their assumptions on δ. We reach, in particular, critical singularities of drift in the time variable. Compared to Section 15 we, however, restrict somewhat the class of the spatial singularities of the drift to essentially the largest possible Morrey class.

In Section 17 we discuss analogues of the weak well-posedness results from Section 14 for the SDE (1.12) driven by symmetric α-stable process.

2. Notations

1. $\mathbb{R}_+ := [0,\infty[$. In what follows, $B_r(x)$ is the open ball of radius r centered at $x \in \mathbb{R}^d$. Put

$$\nabla_i f := \partial_{x_i} f,$$

where $f = f(x)$ or $f = f(t,x)$, $x = (x_1, \dots, x_d)$.

For $\alpha, \beta \in \mathbb{R}$, set

$$\alpha \vee \beta := \max\{\alpha, \beta\}, \quad \alpha \wedge \beta := \min\{\alpha, \beta\}.$$

Let $\mathcal{B}(X,Y)$ denote the space of bounded linear operators between Banach spaces $X \to Y$, endowed with the operator norm $\|\cdot\|_{X\to Y}$. $\mathcal{B}(X) := \mathcal{B}(X,X)$.

We write $T = s\text{-}X\text{-}\lim_n T_n$ for T, $T_n \in \mathcal{B}(X,Y)$ if

$$\lim_n \|Tf - T_n f\|_Y = 0 \quad \text{for every } f \in X.$$

Put

$$L^p = L^p(\mathbb{R}^d), \quad W^{1,p} = W^{1,p}(\mathbb{R}^d).$$

Set

$$\|\cdot\|_p := \|\cdot\|_{L^p}$$

and

$$\|\cdot\|_{p\to q} := \|\cdot\|_{L^p \to L^q}.$$

Let $\mathcal{W}^{\alpha,p}$, $\alpha > 0$, denote the Bessel potential space on $\mathbb{R}^d$ endowed with norm $\|u\|_{p,\alpha} := \|g\|_p$, $u = (1-\Delta)^{-\frac{\alpha}{2}} g$, $g \in L^p$. Let $\mathcal{W}^{-\alpha,p'}$, $p' = p/(p-1)$ denote the anti-dual of $\mathcal{W}^{\alpha,p}$.

For a given vector field b and $1 \leq p < \infty$, put

$$b^{\frac{1}{p}} := b|b|^{-1+\frac{1}{p}}.$$

Put

$$\langle f, g \rangle = \langle f\bar{g} \rangle := \int_{\mathbb{R}^d} f\bar{g}\,dx$$

(some of our functions will be complex-valued).

C_c (C_c^∞) denotes the space of continuous (smooth) functions on $\mathbb{R}^d$ having compact support.

$C_\infty := \{f \in C(\mathbb{R}^d) \mid \lim_{x\to\infty} f(x) = 0\}$ (with the sup-norm).

$\mathcal{S}$ is the L. Schwartz' space of test functions.

We denote by $\restriction$ the restriction of an operator (or a function) to a subspace (a subset).

Given linear operators A, B, we write $B \supset A$ if B is an extension of A.

Let

$$\big[A \upharpoonright D(A) \cap L^p\big]^{\rm clos}_{L^p \to L^p}$$

denote the closure of operator A as an operator $L^p \to L^p$ (if it exists).

2. Fix $0 < T < \infty$. Let $D([0,T],\mathbb{R}^d)$, the space of right-continuous functions having left limits, be endowed with the filtration $\mathcal{B}'_t = \sigma(\omega_r \mid 0 \leq r \leq t)$, where ω_t is the coordinate process on $D([0,T],\mathbb{R}^d)$.

We will also need the canonical space $(C([0,T],\mathbb{R}^d), \mathcal{B}_t = \sigma(\omega_r \mid 0 \leq r \leq t))$, where ω_t is the coordinate process on $C([0,T],\mathbb{R}^d)$.

Recall that a probability measure $\mathbb{P}_x$, $x \in \mathbb{R}^d$ on $(C([0,T],\mathbb{R}^d), \mathcal{B}_t)$ is called a martingale solution to SDE

$$dX_t = -b(t,X_t)dt + \sqrt{2}dW_t, \quad X_0 = x \tag{2.1}$$

if

1) $\mathbb{P}_x[\omega_0 = x] = 1$;

2)

$$\mathbb{E}_x \int_0^T |b(r,\omega_r)|dr < \infty;$$

3) for every $f \in C_c^2(\mathbb{R}^d)$ the process

$$t \mapsto f(\omega_t) - f(x) + \int_0^t (-\Delta f + b \cdot \nabla f)(\omega_r)dt$$

is a $\mathcal{B}_t$-martingale under $\mathbb{P}_x$.

A martingale solution $\mathbb{P}_x$ of (2.1) is called a weak solution if, upon completing filtration $\mathcal{B}_t$ with respect to $\mathbb{P}_x$ (to, say, $\hat{\mathcal{B}}_t$), there exists a Brownian motion $\{W_t\}$ on $\big(C([0,T],\mathbb{R}^d), \hat{\mathcal{B}}_t, \mathbb{P}_x\big)$ such that

$$\omega_t = x - \int_0^t b(r,\omega_r)dr + \sqrt{2}W_t, \quad 0 \leq t \leq T \quad \mathbb{P}_x\text{-a.s.}$$

3. A function $h : \mathbb{R}^d \to \mathbb{R}$ is said to be in the weak L^d class (denoted by $L^{d,w}$) if

$$\|h\|_{d,w} := \sup_{s>0} s|\{x \in \mathbb{R}^d : |h(x)| > s\}|^{1/d} < \infty.$$

Clearly, $L^d \subset L^{d,w}$, but not vice versa, e.g. $h(x) = |x|^{-1}$ is in $L^{d,w}$ but not in L^d.

4. The De Giorgi mollifier $E_\varepsilon \equiv E_\varepsilon^d$ on $\mathbb{R}^d$:

$$E_\varepsilon f(x) := e^{\varepsilon\Delta} f(x), \quad x \in \mathbb{R}^d, \quad \varepsilon > 0,$$

where $f \in L^1_{\rm loc}$.

The Friedrichs mollifier $E_\varepsilon \equiv E_\varepsilon^d$ on $\mathbb{R}^d$:

$$E_\varepsilon f(x) := \eta_\varepsilon * f(x),$$

where $\eta_\varepsilon(y) := \frac{1}{\varepsilon^d}\eta\left(\frac{y}{\varepsilon}\right)$, $\varepsilon > 0$ and

$$\eta(y) := \begin{cases} c\exp\left(\frac{1}{|y|^2-1}\right) & \text{if } |y| < 1, \\ 0, & \text{if } |y| \geqslant 1, \end{cases}$$

with constant c adjusted to $\int_{\mathbb{R}^d} \eta(x)dx = 1$.

3. Form-bounded drifts. Semigroup in L^2

First, we discuss sufficient conditions for existence of an operator realization of $-\Delta + b \cdot \nabla$ generating a strongly continuous semigroup in L^2.

Assume that a Borel measurable vector field $b : \mathbb{R}^d \to \mathbb{R}^d$ with $|b| \in L^2_{\rm loc}$ satisfies inequality

$$\|b\varphi\|_2^2 \leq \delta \|\nabla \varphi\|_2^2 + c_\delta \|\varphi\|_2^2 \quad \text{for all } \varphi \in W^{1,2} \tag{3.1}$$

for finite constants $\delta > 0$ and $0 \leq c_\delta < \infty$.

Definition 3.1. Such vector fields b are called *form-bounded* (written as $b \in \mathbf{F}_\delta$).

Inequality (3.1) can be re-stated as an operator norm inequality

$$\|b(\lambda - \Delta)^{-\frac{1}{2}}\|_{2 \to 2} \leq \sqrt{\delta} \tag{3.2}$$

with $\lambda \equiv \lambda_\delta = c_\delta/\delta$.

Using the quadratic (or Cauchy-Schwarz) inequality

$$|\langle b \cdot \nabla \varphi, \varphi \rangle| \leq \varepsilon \|b\varphi\|_2^2 + \frac{1}{4\varepsilon}\|\nabla \varphi\|_2^2, \quad \varepsilon > 0, \tag{3.3}$$

one can see that the form-boundedness condition (3.1) *with* $\delta < 1$ is what is needed to verify conditions of the Lax-Milgram theorem for the bilinear form $\tau[\varphi, \psi] := \lambda\langle \varphi, \psi \rangle + \langle \nabla \varphi, \nabla \psi \rangle + \langle b \cdot \nabla \varphi, \psi \rangle$ defined on the real space $W^{1,2}$. That is, coercivity for all $\lambda \geq c_\delta/2\sqrt{\delta}$:

$$\begin{aligned} |\tau[\varphi, \varphi]| &\geq \lambda \|\varphi\|_2^2 + \|\nabla \varphi\|_2^2 - \varepsilon \|b\varphi\|_2^2 - \frac{1}{4\varepsilon}\|\nabla \varphi\|_2^2 \\ &\quad \text{(we apply (3.1) and select } \varepsilon = \frac{1}{2\sqrt{\delta}}) \\ &= (\lambda - \frac{c_\delta}{2\sqrt{\delta}})\|\varphi\|_2^2 + (1 - \sqrt{\delta})\|\nabla \varphi\|_2^2 \end{aligned} \tag{3.4}$$

and boundedness

$$|\tau[\varphi, \psi]| \leq C\|\varphi\|_{W^{1,2}}\|\psi\|_{W^{1,2}}. \tag{3.5}$$

So, by the Lax-Milgram theorem, there exists a unique weak solution to elliptic equation

$$(\mu - \Delta + b \cdot \nabla)u = f, \quad f \in L^2.$$

Furthermore, form-boundedness (3.1) with $\delta < 1$ ensures that the sesquilinear form of $\lambda - \Delta + b \cdot \nabla$ defined on the complex space $W^{1,2}$ is sectorial, and hence by the KLMN theorem[1] it determines a unique closed densely defined operator $\Lambda \equiv \Lambda_2(b)$,

$$\Lambda \supset (-\Delta + b \cdot \nabla) \upharpoonright C_c^\infty(\mathbb{R}^d), \tag{3.6}$$

generating a strongly continuous quasi contraction semigroup in L^2.

Examples. Let us mention some elementary sufficient conditions for form-boundedness.

1. If $|b| \in L^d$ (or $|b| \in L^d + L^\infty$, i.e. the sum of two functions, one in L^d and the other one in L^∞), then $b \in \mathbf{F}_\delta$ with δ that can be chosen arbitrarily small (at expense of increasing c_δ, see Appendix B).

[1] Kato-Lions-Lax-Milgam-Nelson theorem, see [Ka, Ch. VI], [O, Ch.1]

2. There are form-bounded vector fields that have stronger singularities than the ones covered by the class L^d. For instance, by Hardy's inequality

$$\left(\frac{d-2}{2}\right)^2 \||x|^{-1}\varphi\|_2^2 \leq \|\nabla\varphi\|_2^2, \quad \varphi \in W^{1,2},$$

the vector field

$$b(x) = \pm\sqrt{\delta}\frac{d-2}{2}\frac{x}{|x|^2}, \tag{3.7}$$

is form-bounded with $c_\delta = 0$. The constant in Hardy's inequality is sharp, and the last vector field is not in $\mathbf{F}_{\delta'}$ for any $\delta' < \delta$ regardless of the value of $c_{\delta'}$.

As we explain below, the value of constant δ determines weak solvability of SDEs (see (4.1) below), and thus is a key characteristics of the vector field b. However, the dependence of the solution theory of SDEs on δ is not visible if one deals only with $|b| \in L^d$. In this sense, the vector fields b with entries in L^d are sub-critical.

3. More generally, if vector field b belongs to the scaling-invariant Morrey class $M_{2+\varepsilon}$ for some $\varepsilon > 0$ arbitrarily small, i.e.

$$|b| \in L^{2+\varepsilon}_{\rm loc} \quad \text{and} \quad \|b\|_{M_{2+\varepsilon}} := \sup_{r>0, x\in\mathbb{R}^d} r\left(\frac{1}{|B_r|}\int_{B_r(x)} |b|^{2+\varepsilon}dx\right)^{\frac{1}{2+\varepsilon}} < \infty, \tag{3.8}$$

then b is form-bounded with $\delta = c\|b\|_{M_{2+\varepsilon}}$ for appropriate constant $c = c(d,\varepsilon)$. See [F], see also [CFr]. Note that

$$\|\cdot\|_{M_q} \leq \|\cdot\|_{M_{q_1}} \quad \text{if } q < q_1,$$

so Morrey class becomes larger as q becomes smaller (and so we are interested in the class M_q with q close to 2). This class contains all $|b| \in L^d$ and $|b| \in L^{d,w}$ (see definition in Section 2). It also contains, for every $\epsilon > 0$, vector fields b such that $|b| \notin L^{2+\epsilon}_{\rm loc}$.

On the other hand, it is easy to show, by considering translates of a bump function, that if $b \in \mathbf{F}_\delta$ (say, with $c_\delta = 0$), then $|b| \in M_2$.

4. If $|b|^2$ is in the Kato class of potentials $\mathbf{K}^d_\delta$, then vector field b is form-bounded. Recall that $V \in \mathbf{K}^d_\delta$ if

$$V \in L^1_{\rm loc} \quad \text{and} \quad \lim_{\lambda\to+\infty} \|(\lambda - \Delta)^{-1}|V|\|_\infty \leq \delta. \tag{3.9}$$

The Kato class condition first appeared in a 1961 article by M. Birman as an elementary sufficient condition for the form-boundedness of a potential V [Bi, Sect. 2].

Some other sufficient conditions for the form-boundedness of b, including those refining condition $b \in M_{2+\varepsilon}$, are given in Appendix B.

Let us also note that the sum $b_1 + b_2$ of form-bounded vector fields $b_1 \in \mathbf{F}_{\delta_1}$, $b_2 \in \mathbf{F}_{\delta_2}$ is form-bounded with form-bound $\delta = (\sqrt{\delta_1} + \sqrt{\delta_2})^2$ (cf. (3.2)). In particular, vector field

$$b(x) := \sum_{k=1}^\infty c_k \frac{x - a_k}{|x - a_k|^2}$$

with $\sum_{k=1}^\infty |c_k|^{\frac{1}{2}} < \infty$ and $\{a_k\}$ constituting e.g. a dense subset of $\mathbb{R}^d$ is form-bounded.

The form-boundedness condition is well known in the literature on spectral and scattering theory of Schrödinger operators, in particular, dealing with the questions of self-adjointness, estimates on the number of bound states, resolvent convergence, see e.g. [BS, MV, Si][2]. Regarding Kolmogorov operator (3.6), one can show, with little additional effort, that the semigroup $e^{-t\Lambda}$ is a Markov semigroup in L^2 (the proof can be found e.g. [KiS2]). In the probabilistic context, however, one of the basic problems is to construct a Markov process that inherits the essential properties of the Brownian motion. In particular, it is natural to expect that the constructed Markov semigroup would be strongly continuous on the space C_∞ of continuous functions vanishing at infinity endowed with the sup-norm, i.e. that it would be a Feller semigroup. However, to show this, one needs to verify the strong continuity in the norm of C_∞, which is, of course, a lot more rigid that the norm of L^2 where one defines the form-boundedness of the drift. In this regard, let us note that the strong continuity of this semigroup in L^p for any finite $p \geq 2$, on the contrary, presents no problem. Indeed, since $e^{-t\Lambda}$ is Markov, we have $\|e^{-t\Lambda}f\|_\infty \leq \|f\|_\infty$, $f \in L^2 \cap L^\infty$, so one can use an interpolation argument to define

$$e^{-t\Lambda_p} := \Big[e^{-t\Lambda} \upharpoonright L^2 \cap L^p\Big]^{\rm clos}_{L^p \to L^p}. \tag{3.10}$$

The left-hand side is a quasi contraction strongly continuous semigroup on L^p (see the proof e.g. in [LS]), its generator Λ_p is an appropriate operator realization of $-\Delta + b \cdot \nabla$ in L^p. So, the difficulty is in the strong continuity in C_∞. A major advancement came with the fundamental paper [KS] of Kovalenko-Semënov that, among other results, presented a construction of a Feller realization of $-\Delta + b \cdot \nabla$ in C_∞ for form-bounded b using an $L^2 \to L^\infty$ iteration procedure. To the best of the author's knowledge, surprisingly, the first reaction to this result was [Ki3], almost three decades later.

Apart from the results described in the present paper, we also refer to [CFKZ, FK] regarding form-boundedness appearing in probabilistic settings.

Form-boundedness and similar conditions, sometimes supplemented with other hypothesis on b, also appear in the regularity theory of elliptic and parabolic equations, which includes the Harnack inequality, Gaussian and non-Gaussian heat kernel bounds. See [AD, H, KiS6, KiV, LZ, Ph, S1] and references therein.

Regarding the necessity of the form-boundedness condition (3.1) with $\delta < 1$ for the existence of L^2 semigroup theory of $-\Delta + b \cdot \nabla$, let us mention the following consequence of the result in [MV]. Let b be a distributional vector field. The sesqulinear form $-\Delta + b \cdot \nabla$ is bounded, i.e.

$$|\langle \nabla \varphi, \nabla \psi \rangle + \langle b \cdot \nabla \varphi, \psi \rangle| \leq C \|\varphi\|_{W^{1,2}} \|\psi\|_{W^{1,2}} \tag{3.11}$$

for some constant C, for all $\varphi, \psi \in C_c^\infty$, if and only if b can be represented as $b = b^{(1)} + b^{(2)}$, where $b^{(1)} \in \mathbf{F}_\delta$ for some δ and $b^{(2)}$ is divergence-free and is in the class[3] BMO^{-1}. Thus, since in

[2] If one is willing to ignore different roles played by the positive and the negative parts of potential V in the theory of Schrödinger operator $-\Delta + V$, then V is form-bounded if $\| |V|^{\frac{1}{2}} \varphi\|_2^2 \leq \delta \|\nabla \varphi\|_2^2 + c_\delta \|\varphi\|_2^2$, $\varphi \in W^{1,2}$.

[3] i.e. the components of $b^{(2)}$ satisfy

$$b_k^{(2)} = \sum_{i=1}^d \nabla_i F_{ik}, \quad 1 \leq k \leq d,$$

for a matrix F with entries $F_{ik} = -F_{ki} \in \mathrm{BMO}$. Recall that a function $f \in \mathrm{BMO}$ if $f \in L^1_{\rm loc}$ and $\|f\|_{\rm BMO} := \sup_Q \frac{1}{|Q|} \int_Q |f - (f)_Q| dx < \infty$, where the supremum taken over all cubes $Q \subset \mathbb{R}^d$ with sides parallel to the axes, and $(f)_Q$ is the average of f over Q.

this paper we are interested in conditions on $|b|$, i.e. not assuming any additional structure of b such as zero divergence, our condition (3.1) is essentially necessary for (3.11) to hold. However, (3.11) is not synonymous with the existence of a realization of $-\Delta + b \cdot \nabla$ in L^2 generating a strongly continuous semigroup. Should we require or expect (3.11) to hold in order to have L^2 semigroup theory? As we explain below, for $-\Delta + b \cdot \nabla$ the answer is "no": in the next section we will abandon (3.11) and will go beyond the class $\mathbf{F}_\delta$. However, we will have (3.11) when will be dealing with SDE (1.3) having discontinuous diffusion coefficients. See Section 11.

Concerning the difference between a popular condition[4] $|b| \in L^d$ and more general condition $b \in \mathbf{F}_\delta$, let us note the following: if v is a weak solution of the elliptic equation

$$(\lambda - \Delta + b \cdot \nabla)v = f, \quad \lambda > 0, \quad f \in C_c^\infty$$

with $|b| \in L^d$ and $v \in W^{1,r}$ for r large (which is valid by a classical result), then, by Hölder's inequality,

$$\Delta v \in L^{\frac{rd}{d+r}}_{\rm loc}.$$

However, for $b \in \mathbf{F}_\delta$, one can only say that

$$\Delta v \in L^{\frac{2d}{d+2}}_{\rm loc}$$

(one can in fact show that $v \in W^{2,2}$). Thus, in case $b \in \mathbf{F}_\delta$, any $W^{2,p}$ estimate on the solution v of the elliptic equation for p large is out of question.

Remark 3.1. That being said, if $\delta < 1$, then one has

$$v \in \mathcal{W}^{1+\frac{2}{q},p}, \quad p \in [2, \frac{2}{\sqrt{\delta}}[, \quad q > p$$

(see Theorem 6.1 below). After applying the Sobolev embedding theorem, one obtains $|\nabla v| \in L^\gamma$ for $\gamma < \frac{dp}{d-2}$ arbitrarily close to $\frac{dp}{d-2}$ (depending on how close q is to p). Thus, although p can be as large as one wants provided that δ is chosen sufficiently small, one never arrives at $|\nabla v| \in L^\infty$. For a form-bounded drift b, the gradient of v is in general unbounded.

4. Sharp solvability

The constant c_δ in (3.1) controls the growth of the semigroup $e^{-t\Lambda}$ as $t \to +\infty$, see (6.13), and allows to include in the class $\mathbf{F}_\delta$ bounded vector fields. It is of secondary interest to us in this paper.

The constant δ in the assumption $b \in \mathbf{F}_\delta$, on the contrary, is very important to us since it determines weak solvability of SDE (1.1). Moreover, there is a quantitative dependence between δ and the regularity properties of solutions to the corresponding elliptic and parabolic equations, see Theorem 6.1 and other results below.

The following example, analysed in detail in [BFGM], shows that δ cannot be too large. Consider SDE

$$X_t = -\sqrt{\delta}\frac{d-2}{2}\int_0^t |X_r|^{-2} X_r dr + \sqrt{2} W_t, \tag{4.1}$$

[4]As is well known, on the Lebesgue scale, $|b| \in L^d$ is the best possible condition providing the solvability of (1.1).

which corresponds to the choice of attracting drift

$$b(x) = \sqrt{\delta}\frac{d-2}{2}\frac{x}{|x|^2} \in \mathbf{F}_\delta \tag{4.2}$$

and the initial point $x = 0$ in SDE (1.1). Then, if

$$\delta > 4\Big(\frac{d}{d-2}\Big)^2, \tag{4.3}$$

the SDE (4.1) does not have a weak solution. Indeed, suppose that a weak solution exists. Then $X(t) = (X_t^1, \dots, X_t^d)$ is a continuous semimartingale with cross-variation $[X^i, X^k]_t = 2\delta_{ik}t$. By Itô's formula,

$$|X_t|^2 = -2\int_0^t X_s b(X_s)ds + 2\sqrt{2}\int_0^t X_s dW_s + 2\int_0^t d[W,W]_s,$$

i.e.

$$|X_t|^2 = -2\sqrt{\delta}\frac{d-2}{2}\int_0^t \mathbf{1}_{X_s\neq 0}ds + 2\sqrt{2}\int_0^t X_s dW_s + 2td.$$

One has $\int_0^t \mathbf{1}_{X_s=0}ds = 0$ almost surely (see details in [BFGM]), so

$$|X_t|^2 = 2(d - \sqrt{\delta}\frac{d-2}{2})\int_0^t \mathbf{1}_{X_s\neq 0}ds + 2\sqrt{2}\int_0^t X_s dW_s \quad \text{a.s.}$$

Therefore, $X_t^2 \geq 0$ is a local supermartingale if $\sqrt{\delta}\frac{d-2}{2} > d$. Then a.s. $X_0 = 0 \Rightarrow X_t = 0$, which contradicts to $[X^1, X^1]_t = 2t$. This argument was used earlier in [CE] in the one-dimensional setting. [BFGM] furthermore showed if $\delta > 4$, then a trajectory started outside of the origin arrives at $x = 0$ in finite time with positive probability; in this regard, see also [W].

Although at the first sight this counterexample can be interpreted (and sometimes was) as a counterexample showing the optimality of the condition $|b| \in L^d$, we argued in [KiS1] that this is a counterexample to admissible values of constant δ. In fact, we have the following theorem.

Theorem 4.1. *Let $b \in \mathbf{F}_\delta$ with*

$$\delta < 4. \tag{C_1}$$

Then, for every $x \in \mathbb{R}^d$, the SDE

$$X_t = x - \int_0^b b(X_s)ds + \sqrt{2}W_t, \quad t \geq 0, \tag{4.4}$$

has a martingale solution.

Theorem 4.1 was proved in [KiS4]. (In fact, it was proved there for time-inhomogeneous form-bounded drifts b with $\delta < 4$, see Definition 9.1.)

Comparing "$\delta > 4(\frac{d}{d-2})^2$" in (4.3) and "$\delta < 4$", one sees that the result is essentially sharp in high dimensions.

Let us note that the well-posedness of SDEs and parabolic equations with δ reaching the critical value (up to the strict inequality) but for b having additional structure (e.g. $b = \nabla V$ for appropriate potential V such as $V(x) = c\log|x|$), was also addressed by Fournier-Jourdan [FJ] (in dimension 2), Bresch-Jabin-Wang [BJW], see also references therein. A crucial feature of Theorem 4.1 is that it attains the critical threshold $\delta = 4$, up to the strict inequality, for the entire class of form-bounded vector fields, i.e. without any assumptions on the structure of b.

Let us explain where does "$\delta<4$" come from (and also how one can handle $1\leq\delta<4$ given that the KLMN theorem requires $\delta<1$). Multiplying the corresponding to (4.4) parabolic equation $(\partial_t-\Delta+b\cdot\nabla)u=0$ by $u|u|^{p-2}$, integrating by parts and using the quadratic inequality and form-boundedness, one obtains that the admissible p that give e.g. an energy inequality turn out to be $p>2/(2-\sqrt{\delta})$. In fact, it was proved in [KS] that if $b=b(x)$ has form-bound $\delta<4$, then there exists a realization of $-\Delta+b\cdot\nabla$ in L^p, $p>2/(2-\sqrt{\delta})$ generating a quasi contraction strongly continuous semigroup there. The interval of contraction solvability can be extended to $[2/(2-\sqrt{\delta}),\infty[$ and is sharp, see [KiS2]. Now, as $\delta\uparrow 4$, this interval disappears, and with it disappears the theory of operator $-\Delta+b\cdot\nabla$ (see, however, the next section).

The proof of Theorem 4.1 is based on the following analytic result, which allows to use the standard tightness argument (see [RZh1]) to construct a martingale solution of (4.4). Namely, put

$$\rho(x)=(1+\kappa|x|^{-2})^{-\beta},\quad \beta>\frac{d}{4}, \tag{4.5}$$

with $\kappa>0$ sufficiently small. Let u be the classical solution to Cauchy problem

$$(\partial_t-\Delta+b\cdot\nabla)u=|\mathsf{h}|f,\quad u(0)=0, \tag{4.6}$$

where $f\in C_c$ and $b\in\mathbf{F}_\delta\cap C_c^\infty$, $\delta<4$ and $\mathsf{h}\in\mathbf{F}_\nu\cap C_c^\infty$, $\nu<\infty$. Fix $T>0$ and $1<\theta<\frac{d}{d-1}$. For all $p>\frac{2}{2-\sqrt{\delta}}$, $p\geq 2$, there exists a constant C independent of smoothness of b and h such that

$$\|u\|_{L^\infty([0,T]\times\mathbb{R}^d)}\leq C\sup_{z\in\mathbb{Z}^d}\biggl(\int_0^T\Bigl\langle\bigl(\mathbf{1}_{\{|\mathsf{h}|\geq 1\}}+\mathbf{1}_{\{|\mathsf{h}|<1\}}|\mathsf{h}|^p\bigr)^{\theta'}|f|^{p\theta'}\rho_z^2\Bigr\rangle\biggr)^{\frac{1}{p\theta'}} \tag{4.7}$$

where $\rho_z(x):=\rho(x-z)$.

Now, let b_n be smooth vector fields having compact support, e.g. defined by (6.2) below, approximating b in the sense of (6.3), (6.4). Fix $x\in\mathbb{R}^d$. By a classical result, there exist strong solutions X^n to SDEs

$$X_t^n=x-\int_0^t b_n(X_s^n)ds+\sqrt{2}W_t,\quad n=1,2,\dots,$$

where $\{W_t\}_{t\geq 0}$ is a Brownian motion in $\mathbb{R}^d$ on a fixed complete probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbf{P})$. Then (4.7) yields

$$\biggl|\mathbf{E}\int_{t_0}^{t_1}|\mathsf{h}(X_s^n)|f(X_s^n)ds\biggr|\leq C\sup_{z\in\mathbb{Z}^d}\biggl(\int_{t_0}^{t_1}\Bigl\langle\bigl(\mathbf{1}_{\{|\mathsf{h}|\geq 1\}}+\mathbf{1}_{\{|\mathsf{h}|<1\}}|\mathsf{h}|^p\bigr)^{\theta'}|f|^{p\theta'}\rho_z^2\Bigr\rangle\biggr)^{\frac{1}{p\theta'}}, \tag{4.8}$$

where $0\leq t_0<t_1\leq T$. We now apply (4.8) with $\mathsf{h}=b_n$ and $f\equiv 1$ (here $f\in C_c\Rightarrow f\equiv 1$ using Fatou's Lemma):

$$\begin{aligned}\mathbf{E}\int_{t_0}^{t_1}|b_n(X_s^n)|ds&\leq C\sup_{z\in\mathbb{Z}^d}\biggl(\int_{t_0}^{t_1}\Bigl\langle\bigl(\mathbf{1}_{\{|b_n|\geq 1\}}+\mathbf{1}_{\{|b_n|<1\}}|b_n|^p\bigr)^{\theta'}\rho_z^2\Bigr\rangle\biggr)^{\frac{1}{p\theta'}}\\&\leq C_0(t_1-t_0)^\mu\quad\text{for generic }\mu>0\text{ and }C_0.\end{aligned} \tag{4.9}$$

The latter allows to verify tightness of probability measures

$$\mathbb{P}_x^n:=(\mathbf{P}\circ X^n)^{-1}$$

on $(C([0,T],\mathbb{R}^d),\mathcal{B}_t)$, so for every $x \in \mathbb{R}^d$ there exists a subsequence $\{\mathbb{P}_x^{n_k}\}$ and a probability measure $\mathbb{P}_x$ such that

$$\mathbb{P}_x^{n_k} \to \mathbb{P}_x \text{ weakly.} \tag{4.10}$$

Another application of (4.8) but with $\mathsf{h} = b_{n_1} - b_{n_2}$ allows to conclude that $\mathbb{P}_x$ is a martingale solution of (4.4), see [KiS4] for details.

Estimate (4.7) is proved using De Giorgi's iterations in L^p for $p > 2/(2-\sqrt{\delta})$. Thus, $p > 2$ if $1 < \delta < 4$. In this regard, let us note that passing to L^p right away, using the fact that $u^{\frac{p}{2}}$ is a sub-solution, and then applying to $u^{\frac{p}{2}}$ the standard De Giorgi's iterations in L^2, does not allow to treat $1 \le \delta < 4$. We have to follow the iteration procedure from the very beginning and adjust it accordingly.

Earlier, De Giorgi's method in L^2 was used by Zhang-Zhao [ZZh], Zhao [Zh], Röckner-Zhao [RZh1] to prove, among other results, weak well-posedness of (4.4) with b having not too singular divergence and satisfying

$$|b| \in L^q_{\rm loc}(\mathbb{R}_+, L^r + L^\infty), \quad \frac{d}{r} + \frac{2}{q} < 2.$$

Remark 4.1. Looking at the counterexample (4.1)-(4.3) and Theorem 4.1, one can draw an analogy with the celebrated result of Baras-Goldstein for Schroödiner operator

$$-\Delta - V_0, \quad V_0(x) = \delta\frac{(d-2)^2}{4}|x|^{-2}$$

on $\mathbb{R}^d$, $d \ge 3$. This V_0 is a form-bounded potential, i.e. $\langle V_0\varphi,\varphi\rangle \le \delta\langle\nabla\varphi,\nabla\varphi\rangle$ for all $\varphi \in W^{1,2}$ (this is Hardy's inequality). If $0 < \delta < 1$, then the self-adjoint operator realization H of $-\Delta - V_0$ on L^2, defined e.g. via the KLMN theorem, satisfies

$$e^{-tH} = s\text{-}L^2\text{-}\lim_{\varepsilon\downarrow 0} e^{-tH(V_\varepsilon)},$$

where $V_\varepsilon(x) = \delta\frac{(d-2)^2}{4}|x|_\varepsilon^{-2}$, $|x|_\varepsilon^2 := |x|^2 + \varepsilon$, $\varepsilon > 0$. For $\delta > 1$, however, by the result in [BG] (see also [GZa]),

$$\lim_{\varepsilon\downarrow 0} e^{-tH(V_\varepsilon)} f(x) = \infty, \quad t > 0, \quad x \in \mathbb{R}^d, \quad f \ge 0,\ f \not\equiv 0,$$

i.e. all positive solutions of the corresponding parabolic equation explode instantly at any point. This phenomenon is not observable on any $V_0 \in L^{\frac{d}{2}}$ since any such potential has arbitrarily small form-bound (similarly to how any b with $|b| \in L^d$ has arbitrarily small form-bound δ).

4.1. **Critical magnitude of drift singularities.** It turns out that one still has a strong[5] solution (i.e. semigroup) theory of parabolic equation $(\partial_t - \Delta + b\cdot\nabla)u = 0$ with $b \in \mathbf{F}_\delta$ in the critical borderline case $\delta = 4$, although not in L^p, as in the previous section, but in an Orlicz space. Orlicz spaces are known to appear in various borderline situations in analysis (e.g. Trudiner's theorem).

We will work over the d-dimensional torus Π^d obtained as the quotient of $[-\frac{1}{2},\frac{1}{2}]^d$. This is not a technical assumption since the volume of the torus enters the estimates. Still, since we are interested in the local singularities of b, working on a torus is sufficient for our purposes. The definition of form-bounded vector fields does not change, except that now we integrate over

[5]Here "strong" refers to the differentiability of the solution in time.

the torus. Moreover, the examples of form-bounded vector fields mentioned earlier in the paper remain essentially intact, see [BO, G]. In this section, $\langle \cdot, \cdot \rangle$ denotes integration over torus Π^d.

Put

$$\Phi(y) := \cosh(y) - 1, \quad y \in \mathbb{R}$$

Expanding cosh in the Taylor series, one sees right away that $\Phi(y) = \Phi(|y|)$. This function is convex on $\mathbb{R}$, $\Phi(y) = 0$ if and only if $y = 0$, $\Phi(y)/y \to 0$ if $y \to 0$ and $\Phi(y)/y \to \infty$ if $y \to \infty$. Therefore, the space $\mathcal{L}_\Phi = \mathcal{L}_\Phi(\mathbb{R}^d)$ of real-valued measurable functions f on Π^d satisfying

$$\|f\|_\Phi := \inf \{c > 0 \mid \langle \Phi(\frac{f}{c}) \rangle \leq 1\} < \infty, \tag{4.11}$$

is a Banach space with respect to Orlicz norm $\| \cdot \|_\Phi$. See e.g. [AF, Ch. 8].

Let L_Φ denote the closure $\mathcal{L}_\Phi$ of the subspace of smooth functions on Π^d. This is our Orlicz space.

We note that

$$\| \cdot \|_\Phi \geq \frac{1}{(2p)!} \| \cdot \|_{2p}, \quad p = 1, 2, \dots, \tag{4.12}$$

i.e. we are dealing with an Orlicz norm that is stronger that any L^p norm.

Let u_n be the classical solution to Cauchy problem

$$\begin{cases} (\partial_t - \Delta + b_n \cdot \nabla) u_n = 0 \text{ on } [0, \infty[\times \Pi^d, \\ \qquad\qquad u_n(0, \cdot) = f(\cdot) \in C^\infty(\Pi^d), \end{cases}$$

where b_n are bounded smooth vector fields such that $b_n \in \mathbf{F}_\delta$ with the same c_δ and $b_n \to b$ in $L^2(\Pi^d)$ (e.g. obtained upon applying to b the De Giorgi mollifier on Π^d). Let $e^{-t\Lambda(b_n)}$, $\Lambda(b_n) := -\Delta + b_n \cdot \nabla$ denote the corresponding semigroup, i.e.

$$e^{-t\Lambda(b_n)} f := u_n(t).$$

On the smooth initial functions, $[0, \infty[\ni t \mapsto e^{-t\Lambda(b_n)} f$ is strongly continuous in the norm of L_Φ since it is strongly continuous in the norm of L^∞.

Theorem 4.2. *Let $b \in \mathbf{F}_\delta$, $0 < \delta \leq 4$. The following are true:*

(i) *For all $n \geq 1$, $f \in C^\infty(\Pi^d)$,*

$$\|e^{-t\Lambda(b_n)} f\|_\Phi \leq e^{2\frac{c_\delta}{\sqrt{\delta}} t} \|f\|_\Phi, \quad t \geq 0.$$

(ii) *There exists a strongly continuous quasi contraction semigroup $e^{-t\Lambda(b)}$ on L_Φ such that, for every $f \in C^\infty(\Pi^d)$,*

$$\|e^{-t\Lambda(b)} f - e^{-t\Lambda(b_n)} f\|_\Phi \to 0 \quad \text{ as } n \to \infty \text{ loc. uniformly in } t \geq 0.$$

It follows that $e^{-t\Lambda(b)}$ is a positivity preserving L^∞ contraction. Its generator $\Lambda(b)$ is the appropriate operator realization of the formal operator $-\Delta + b \cdot \nabla$ in L_Φ.

This semigroup is unique in the sense that it does not depend on the choice of smooth vector fields $\{b_n\}$, $b_n \to b$ in $L^2(\Pi^d)$, as long as they do not increase constants δ, c_δ.

(iii) *The following energy inequality holds for $u = e^{-t\Lambda(b_n)} f$:*

$$\frac{1}{2} \sup_{s \in [0,t]} \langle e^{u^p(s)} \rangle + 4\frac{(p-1)}{p} \langle (\nabla u^{\frac{p}{2}})^2 e^{u^p} \rangle \leq \langle e^{f^p} \rangle, \quad p = 2, 4, \dots$$

provided $\frac{c_\delta}{\sqrt{\delta}}t < \frac{1}{2}$*; the last constraint can be removed using the semigroup property.*

The last assertion is noteworthy, since, at the first sight, it seems like one can reach $\delta = 4$ only at the cost of killing off the dispersion term.

Theorem 4.2 was proved in [Ki6]. The following calculation illustrates the main observation behind this result. Below we are looking for integral bounds on u_n that can depend on δ and c_δ, but not on the smoothness or boundedness of b_n. Put $u = u_n$, $b = b_n$. Replacing v by $v = e^{-\lambda t}u$, $\lambda \geq 0$, we will deal with Cauchy problem

$$(\lambda + \partial_t - \Delta + b \cdot \nabla)v = 0, \quad v(0) = f.$$

We multiply the equation by e^v and integrate:

$$\lambda\langle v, e^v\rangle + \langle \partial_t(e^v - 1)\rangle + 4\langle(\nabla e^{\frac{v}{2}})^2\rangle + 2\langle be^{\frac{v}{2}}, \nabla e^{\frac{v}{2}}\rangle = 0.$$

By quadratic inequality,

$$\lambda\langle v, e^v\rangle + \langle \partial_t(e^v - 1)\rangle + 4\langle(\nabla e^{\frac{v}{2}})^2\rangle \leq \alpha\langle b^2 e^v\rangle + \frac{1}{\alpha}\langle(\nabla e^{\frac{v}{2}})^2\rangle.$$

Now, selecting $\alpha = \frac{1}{\sqrt{\delta}}$ and using $b \in \mathbf{F}_\delta$, we obtain

$$\lambda\langle v, e^v\rangle + \langle \partial_t(e^v - 1)\rangle + (4 - 2\sqrt{\delta})\langle(\nabla e^{\frac{v}{2}})^2\rangle \leq \frac{c_\delta}{\sqrt{\delta}}\langle e^v\rangle.$$

Using $\delta \leq 4$, we obtain after integrating in time from 0 to t,

$$\lambda\int_0^t \langle v, e^v\rangle ds + \langle e^{v(t)} - 1\rangle \leq \langle e^f - 1\rangle + \frac{c_\delta}{\sqrt{\delta}}\int_0^t \langle e^v\rangle ds.$$

Replacing in the last inequality u by $-u$ and adding up the resulting inequalities, we obtain

$$\lambda\int_0^t \langle v\sinh(v)\rangle ds + \langle\cosh(v(t)) - 1\rangle \leq \langle\cosh(f) - 1\rangle + \frac{c_\delta}{\sqrt{\delta}}\int_0^t \langle\cosh(v)\rangle ds.$$

Finally, applying $v\sinh(v) \geq \cosh(v) - 1$, we arrive at

$$(\lambda - \frac{c_\delta}{\sqrt{\delta}})\int_0^t \langle\cosh(v) - 1\rangle ds + \langle\cosh(v(t)) - 1\rangle \leq \langle\cosh(f) - 1\rangle + \frac{c_\delta}{\sqrt{\delta}}t, \tag{4.13}$$

where at the last step we have used the fact that volume $|\Pi^d| = 1$. Let $\lambda \geq \frac{c_\delta}{\sqrt{\delta}}$. Estimate (4.13) suggests that one should work in the topology determined by the "norm" $\langle\cosh(v) - 1\rangle$ or, better, in the corresponding Orlicz space L_Φ.

5. Three approaches to constructing Feller semigroup for $-\Delta + b \cdot \nabla$

We have at our disposal the following approaches to constructing Feller semigroup associated with the Kolmogorov operator $-\Delta + b \cdot \nabla$ with $b \in \mathbf{F}_\delta$. These approaches require $\delta < 1$ ($\ll 1$ in high dimensions).

(1) By using the iteration procedure of [KS] for solutions u_n to elliptic equations

$$(\mu - \Delta + b_n \cdot \nabla)u_n = f, \quad f \in C_c,$$

with bounded smooth drifts b_n approximating $b = b(x)$ (in the sense of (6.2), (6.3) below). It yields inequality

$$\|u_n - u_m\|_\infty \leq B\|u_n - u_m\|_2^\gamma \quad \text{for some } \gamma > 0 \text{ independent of } n,\, m.$$

The latter, in turn, transfers the verification of the Cauchy criterion in C_∞ to a much easier to deal with[6] space L^2. The convergence of the iteration procedure depends on the uniform in n gradient bound

$$\|\nabla u_n\|_{\frac{qd}{d-2}} \leq C\|f\|_q, \quad q > d-2,$$

which was also established in [KS], a pioneer work on the elliptic regularity theory of $-\Delta + b\cdot\nabla$ with form-bounded b, for $q > 2 \vee (d-2)$. This bound requires $\delta < 1 \wedge (\frac{2}{d-2})^2$.

We describe this approach, or rather its relatively recent counterpart for time-inhomogeneous form-bounded drift $b = b(t,x)$, in Section 9. In the time-inhomogeneous case, one obtains a Feller evolution family (see definition in Section 9), and the convergence of the iteration procedure depends on the uniform in n gradient bound for a $q > d$:

$$\sup_{t\in[s,T]} \|\nabla u_n(t)\|_q^q + c\int_s^T \|\nabla u_n\|^q_{\frac{qd}{d-2}} dt \leq C\|\nabla f\|_q^q, \tag{5.1}$$

where u_n is the solution to parabolic equation $(\partial_t - \Delta + b_n(t,x)\cdot\nabla u_n) = 0$, $u_n(0) = f \in C_c^1$, and positive constants c, C are independent of n. See (9.12). The time-inhomogeneous form-bounded vector fields are defined as: $|b| \in L^2_{\rm loc}(\mathbb{R}\times\mathbb{R}^d)$ and there exist a constant $\delta > 0$ and a function $0 \leq g \in L^1_{\rm loc}(\mathbb{R})$ such that for a.e. $t \in \mathbb{R}$

$$\|b(t,\cdot)\varphi\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + g(t)\|\varphi\|_2^2 \tag{5.2}$$

for all $\varphi \in W^{1,2}$ (Definition 9.1).

(2) By constructing the Feller resolvent (Theorem 6.1):

$$\begin{aligned}(\mu - \Delta + b\cdot\nabla)^{-1} f &:= (\mu-\Delta)^{-1} f \\ &\quad - (\mu-\Delta)^{-\frac{1}{2}-\frac{1}{q}} Q_p (1+T_p)^{-1} G_p (\mu-\Delta)^{-\frac{1}{2}+\frac{1}{r}} f,\end{aligned}$$

where $\mu > 0$, on functions $f \in C_\infty \cap L^p$ for $r < p < q$. The operators Q_p, T_p, G_p are bounded on L^p and, moreover, $\|T_p\|_{p\to p} < 1$ under appropriate assumptions on δ, so the geometric series converges. With p chosen larger than $2\vee(d-2)$, we select q sufficiently close to p so that, by the Sobolev embedding theorem, the free Bessel potential $(\mu-\Delta)^{-1/2-1/q}$ will take us from L^p to C_∞. The difficulty in this approach is ensuring boundedness of Q_p, T_p, G_p in L^p given that form-boundedness is an L^2 assumption on $|b|$.

We note that $\|T_p\|_{p\to p} < 1$ for $p > 2\vee(d-2)$ provided that $\delta < 1 \wedge (\frac{2}{d-2})^2$, i.e. we arrive at the same constraint on δ as in the previous approach.

This approach was developed later in [Ki1, Ki2]. It is arguably simpler than (1), due to the use of the linear structure of $-\Delta + b\cdot\nabla$. It also gives an explicit representation for the Feller resolvent of $-\Delta + b\cdot\nabla$. Throughout most of this paper, we pursue approach (2).

In the case of time-inhomogeneous form-bounded $b = b(t,x)$, one constructs solution to the inhomogeneous parabolic equation $(\mu + \partial_t - \Delta + b\cdot\nabla)u = f$ on $\mathbb{R}\times\mathbb{R}^d$ as

$$\begin{aligned}(\mu + \partial_t - \Delta + b\cdot\nabla)^{-1} f &:= (\mu+\partial_t-\Delta)^{-1} \\ &\quad - (\mu+\partial_t-\Delta)^{-\frac{1}{2}-\frac{1}{q}} Q_p (1+T_p)^{-1} G_p (\mu+\partial_t-\Delta)^{-\frac{1}{2}+\frac{1}{r}} f,\end{aligned}$$

[6]Because it has much weaker topology than C_∞ and because form-boundedness is an L^2 assumption on $|b|$.

for appropriately defined *parabolic* operators Q_p, T_p, G_p. Armed with this result, one can obtain the sought Feller evolution family for $\partial_t - \Delta + b \cdot \nabla$, also in an explicit form. This approach is developed in Section 16.

The two methods of constructing Feller semigroup, (1) and (2), are quite different, but they give results of more or less the same strength (see Remark 6.1). However, for the larger class of *weakly form-bounded vector fields*, described in Sections 14-15, only approach (2) is available. That being said, the iteration procedure in (1) has degrees of flexibility that have not been fully explored yet.

Let us now present the third approach to constructing Feller semigroup (Feller evolution family) associated with $-\Delta + b \cdot \nabla$:

(3) Using De Giorgi's method. Let us show that $\{u_n\}$, solutions to parabolic equations $(\partial_t - \Delta + b_n(t,x) \cdot \nabla u_n) = 0$, $u_n(0) = f \in C_c^1$, constitute a Cauchy sequence in $L^\infty([0,T] \times \mathbb{R}^d)$. For simplicity, let us assume that b has compact support (this is not necessary, see Remark 5.1 below).

Set $g := u_n - u_m$. We have

$$\partial_t g - \Delta g + b_n \cdot \nabla g = -(b_n - b_m) \cdot \nabla u_m, \quad g(0) = 0.$$

This is a Cauchy problem for an inhomogeneous parabolic equation of the same form as (4.6) (with $\mathsf{h} = b_n - b_m$, $f = \nabla u_m$; the fact that these are vector-valued functions is not an obstacle). Therefore, by (4.7),

$$\|g\|_{L^\infty([0,T]\times\mathbb{R}^d)} \leq C \sup_{z \in \mathbb{Z}^d} \biggl(\int_0^T \Bigl\langle \bigl(\mathbf{1}_{\{|b_n - b_m| \geq 1\}} + \mathbf{1}_{\{|b_n - b_m| < 1\}} |b_n - b_m|^p\bigr)^{\theta'} |\nabla u_m|^{p\theta'} \rho_z^2 \Bigr\rangle \biggr)^{\frac{1}{p\theta'}}, \tag{5.3}$$

where $\theta < \frac{d}{d-1}$ is chosen close to $\frac{d}{d-1}$, so that $\theta' > d$ is close to d. Using Hölder's inequality, we obtain that

$$\begin{aligned} &\|g\|_{L^\infty([0,T]\times\mathbb{R}^d)} \\ &\leq C \sup_{z \in \mathbb{Z}^d} \biggl(\int_0^T \Bigl\langle \bigl(\mathbf{1}_{\{|b_n - b_m| \geq 1\}} + \mathbf{1}_{\{|b_n - b_m| < 1\}} |b_n - b_m|^p\bigr)^{\frac{s'}{p}} \rho_z^2 \Bigr\rangle \biggr)^{\frac{1}{s'}} \biggl(\int_0^T \langle |\nabla u_m|^s \rho_z^2 \rangle \biggr)^{\frac{1}{s}} \end{aligned}$$

for some $s > p\theta' > pd$ is close to pd (upon selecting θ' close to d). Since b has compact support, by the Dominated convergence theorem the first multiple in the RHS tends to zero as $n, m \to \infty$. Therefore,

$$\|g\|_{L^\infty([0,T]\times\mathbb{R}^d)} = \|u_n - u_m\|_{L^\infty([0,T]\times\mathbb{R}^d)} \to 0 \quad \text{as } n, m \to \infty \tag{5.4}$$

if

$$\sup_m \|\nabla u_m\|_{L^s([0,T]\times\mathbb{R}^d)} < \infty \tag{5.5}$$

for $s > pd$ is close to pd. By the gradient estimate (5.1), after applying the interpolation inequality, we have

$$\sup_m \|\nabla u_m\|_{L^{\frac{qd}{d-2+\frac{4}{d+2}}}([0,T]\times\mathbb{R}^d)} \leq C \|\nabla f\|_q, \tag{5.6}$$

so (5.5) holds with $s=\frac{qd}{d-2+\frac{4}{d+2}}$, hence "$s>pd$" gives us constraint

$$\frac{qd}{d-2+\frac{4}{d+2}}>dp, \quad p>\frac{2}{2-\sqrt{\delta}}. \tag{5.7}$$

Additionally, the gradient estimate (5.1) imposes its own constraint on q, see Remark 9.2. These two constraints on q (which has to be as small as possible for admissible δ to be as large as possible) ensure (5.4) and hence the existence of the Feller semigroup (evolution family).

The proof of (4.7) in [KiS4], which gives us (5.3), also uses the assumption $p\geq 2$, so q in (5.7) is not as small as one hopes. Hence, we obtain more restrictive assumption on δ than we have in approaches (1) and (2) (see Sections 6 and 9). It is possible, however, to weaken "$p\geq 2$" by modifying (5.3), which we will address elsewhere.

Remark 5.1. To exclude the assumption that b has compact support we estimate

$$\begin{aligned}
&\|g\|_{L^\infty([0,T]\times\mathbb{R}^d)} \\
&\leq C \sup_{z\in\mathbb{Z}^d,|z|\leq R}\biggl(\int_0^T\Bigl\langle\bigl(\mathbf{1}_{\{|b_n-b_m|\geq 1\}}+\mathbf{1}_{\{|b_n-b_m|<1\}}|b_n-b_m|^p\bigr)^{\frac{s'}{p}}\rho_z^2\Bigr\rangle\biggr)^{\frac{1}{s'}}\biggl(\int_0^T\langle|\nabla u_m|^s\rho_z^2\rangle\biggr)^{\frac{1}{s}} \\
&+\sup_{z\in\mathbb{Z}^d,|z|>R}\biggl(\int_0^T\Bigl\langle\bigl(\mathbf{1}_{\{|b_n-b_m|\geq 1\}}+\mathbf{1}_{\{|b_n-b_m|<1\}}|b_n-b_m|^p\bigr)^{\frac{s'}{p}}\rho_z^2\Bigr\rangle\biggr)^{\frac{1}{s'}}\biggl(\int_0^T\langle|\nabla u_m|^s\rho_z^2\rangle\biggr)^{\frac{1}{s}},
\end{aligned}$$

where the second sum can be made as small as needed by selecting R sufficiently large. To show this, one needs to use instead of (5.6) the estimate

$$\sup_m\int_0^T\langle|\nabla u_m|^s\rho_z^2\rangle^{\frac{1}{s}}\leq C\langle|\nabla f|^q\rho_z^2\rangle^{\frac{1}{q}}, \quad s=\frac{qd}{d-2+\frac{4}{d+2}}$$

with C independent of m and z (this is a consequence of (10.3)). Here $\langle|\nabla f|^q\rho_z^2\rangle$ is small if $|z|>R$ for R large compared to the support of f. Now, for R fixed sufficiently large, we can treat the first sum in the same way as in the case of a compact support b.

It should be noted that the methods of constructing a Feller realization of $-\Delta+b\cdot\nabla$ actually yield other regularity results. This is discussed in the rest of this paper.

6. Basic result on weak well-posedness of SDEs with singular drift

1. The theory of SDE

$$dX_t=-b(X_t)dt+\sqrt{2}dW_t, \quad X_0=x\in\mathbb{R}^d, \tag{6.1}$$

with $b\in\mathbf{F}_\delta$, becomes more detailed as form-bound δ gets smaller. Namely, if $\delta<1\wedge(\frac{2}{d-2})^2$, then there is a realization of $-\Delta+b\cdot\nabla$ in C_∞ that generates a Feller semigroup. The latter, in turn, determines weak solutions of (1.1). For every $x\in\mathbb{R}^d$, the constructed weak solution is unique in some large classes (e.g. in the class of weak solutions satisfying Krylov-type estimates, or in the class of weak solutions that can be obtained via a "reasonable approximation procedure"). See Theorems 6.1 and 6.2 below.

Set $b^{\frac{2}{p}} := b|b|^{-1+\frac{2}{p}}$. For given $p \in [2,\infty[$ and $1 \le r < p < q < \infty$, define operators ($\mu > 0$)

$$G_p(r) := b^{\frac{2}{p}} \cdot \nabla(\mu-\Delta)^{-\frac{1}{2}-\frac{1}{r}},$$
$$Q_p(q) \upharpoonright \mathcal{E} := (\mu-\Delta)^{-\frac{1}{2}+\frac{1}{q}}|b|^{1-\frac{2}{p}},$$
$$T_p \upharpoonright \mathcal{E} := b^{\frac{2}{p}} \cdot \nabla(\mu-\Delta)^{-1}|b|^{1-\frac{2}{p}}.$$

We define the last two operators on $\mathcal{E} := \bigcup_{\varepsilon>0} e^{-\varepsilon|b|}L^p$, a dense subspace of L^p, to remove any question regarding the summability of $|b|^{1-\frac{2}{p}}f$, $f \in L^p$, on which we act with the Bessel potential.

Set

$$c_{\delta,p} := \Big(\frac{p}{2}\delta + \frac{p-2}{2}\sqrt{\delta}\Big)^{\frac{1}{p}}\Big(p-1-(p-1)\frac{p-2}{2}\sqrt{\delta} - \frac{p(p-2)}{4}\delta\Big)^{-\frac{1}{p}}.$$

Lemma 6.1 ([Ki2])**.** *Let $b \in \mathbf{F}_\delta$. Then for every $p \in [2,\infty[$, there exists $\mu_0 = \mu_0(d,p,q)$ such that the following is true for all $\mu \ge \mu_0$:*

(i) $T_p \upharpoonright \mathcal{E}$ admits extension by continuity to L^p, denoted by T_p. One has

$$\|T_p\|_{p\to p} \le c_{\delta,p}.$$

In particular, if $\delta < 1$, then

$$\|T_p\|_{p\to p} < 1 \quad \text{for every } p \in [2, \frac{2}{\sqrt{\delta}}].$$

(ii) $Q_p(q) \upharpoonright \mathcal{E}$ admits extension by continuity to L^p, denoted by $Q_p(q)$.

(iii) $G_p(r)$ is bounded on L^p.

Lemma 6.1 is a key result needed to prove Theorem 6.1. Its proof, which uses only elementary arguments, is included in Appendix A.

Let us fix the following approximation of b by smooth vector fields having compact supports:

$$b_n := c_n\eta_{\varepsilon_n} * (\mathbf{1}_n b), \tag{6.2}$$

where $\mathbf{1}_n$ is the indicator of $\{x \mid |x| \le n, |b(x)| \le n\}$ and η_{ε_n} is the Friedrichs mollifier (Section 2), and we choose $\varepsilon_n \downarrow 0$ sufficiently rapidly so that, for appropriate $c_n \uparrow 1$, one has

$$b_n \to b \quad \text{in } L^2_{\rm loc}(\mathbb{R}^d,\mathbb{R}^d) \tag{6.3}$$

and

$$b_n \in \mathbf{F}_\delta \quad \text{with some } c_\delta \text{ independent of } n = 1,2,\dots \tag{6.4}$$

see Appendix C.1. Actually, in the next theorem any bounded smooth b_n satisfying (6.3), (6.4) will do, not necessarily the ones defined by (6.2), but at the moment we save ourselves some efforts by considering (6.2) since these are, essentially, cutoffs of b. Later, however, we will consider any bounded smooth $\{b_n\}$ satisfying (6.3), (6.4). This is important because it will give us a uniqueness result on its own: the constructed semigroups or weak solutions to SDEs will not depend on the choice of approximating $\{b_n\}$, as long as they satisfy (6.3), (6.4). In this regard, it is worth introducing the following definition:

Definition 6.1. We say that vector fields b_n satisfying (6.4) are uniformly (in n) form-bounded.

Theorem 6.1. *Let $b \in \mathbf{F}_\delta$, $\delta < 1$. There exists $\mu_0 > 0$ such that the following is true:*

(*i*) *For every $p \in [2, \frac{2}{\sqrt{\delta}}[$, for all $\mu \geq \mu_0$ the function*[7]

$$u = (\mu - \Delta)^{-1} f - (\mu - \Delta)^{-\frac{1}{2} - \frac{1}{q}} Q_p(q)(1 + T_p)^{-1} G_p(r)(\mu - \Delta)^{-\frac{1}{2} + \frac{1}{r}} f, \quad f \in L^p \tag{6.5}$$

is a weak solution to equation

$$(\mu - \Delta + b \cdot \nabla) u = f \tag{6.6}$$

i.e.

$$\mu\langle u, \psi\rangle + \langle \nabla u, \nabla \psi\rangle + \langle b \cdot \nabla u, \psi\rangle = \langle f, \psi\rangle \quad \text{for all } \psi \in C_c^\infty.$$

Moreover, if $f \in L^p \cap L^2$, then u is the unique in $W^{1,2}$ weak solution to (6.6).

(*ii*) *It follows from* (6.5) *that*

$$u \in \mathcal{W}^{1 + \frac{2}{q}, p} \ \ (\textit{Bessel potential space}), \quad q > p. \tag{6.7}$$

In particular, if $\delta < 1 \wedge \big(\frac{2}{d-2}\big)^2$, then in the interval $p \in [2, \frac{2}{\sqrt{\delta}}]$ we can select $p > d - 2$, and then select q sufficiently close to p, so that by the Sobolev embedding theorem u is Hölder continuous (possibly after a modification on a measure zero set), with the Hölder continuity exponent less than but arbitrarily close to $1 - \frac{d-2}{p}$.

(*iii*) *The operator-valued function in* (6.5)

$$\Theta_p(\mu, b) := (\mu - \Delta)^{-1} - (\mu - \Delta)^{-\frac{1}{2} - \frac{1}{q}} Q_p(q)(1 + T_p)^{-1} G_p(r)(\mu - \Delta)^{-\frac{1}{2} + \frac{1}{r}},$$

defined on $\{\mu \geq \mu_0\}$, takes values in $\mathcal{B}(\mathcal{W}^{-1 + \frac{2}{r}, p}, \mathcal{W}^{1 + \frac{2}{q}, p})$.

(*iv*) *Let $\delta < 1 \wedge \big(\frac{2}{d-2}\big)^2$. Fix $p \in]d-2, \frac{2}{\sqrt{\delta}}[$ if $d \geq 4$, or $p \in [2, \frac{2}{\sqrt{\delta}}[$ if $d = 3$. Then*

$$(\mu + \Lambda_{C_\infty}(b))^{-1} := \big(\Theta_p(\mu, b) \upharpoonright L^p \cap C_\infty\big)^{\rm clos}_{C_\infty \to C_\infty}, \quad \mu \geq \mu_0, \tag{6.8}$$

determines the resolvent of a Feller generator on C_∞. This semigroup satisfies

$$e^{-t\Lambda_{C_\infty}(b)} = s\text{-}C_\infty\text{-}\lim_n e^{-t\Lambda_{C_\infty}(b_n)} \quad \textit{locally uniformly in } t \geq 0,$$

where bounded smooth b_n are defined by (6.2) *and the approximating operators $\Lambda_{C_\infty}(b_n) := -\Delta + b_n \cdot \nabla$ with domain $D(\Lambda_{C_\infty}(b_n)) := (1 - \Delta)^{-1} C_\infty$ are, by a classical result, Feller generators.*

(*v*) *Feller semigroup $e^{-t\Lambda_{C_\infty}(b)}$ is conservative, i.e. its integral kernel $e^{-t\Lambda_{C_\infty}}(x, \cdot)$ satisfies*

$$\langle e^{-t\Lambda_{C_\infty}(b)}(x, \cdot)\rangle = 1 \quad \text{for all } x \in \mathbb{R}^d, t > 0. \tag{6.9}$$

Excluding the cases where b is sufficiently regular, one has little information about the domain $D(\Lambda_{C_\infty}(b))$ of the Feller generator $\Lambda_{C_\infty}(b)$. But, it is easily seen, one can be certain that $C_c^2 \not\subset D(\Lambda_{C_\infty}(b))$ already if $|b| \in L^\infty - C_b$.

Remark 6.1. The construction of the Feller semigroup via the iteration procedure of [KS] requires gradient bounds on solutions u_n of $(\mu - \Delta + b_n \cdot \nabla) u_n = f$, $f \in C_c^\infty$. Namely, it is established in [KS, proof of Lemma 5] that if $\delta < 1 \wedge (\frac{2}{d-2})^2$, then for $p > 2 \vee (d-2)$ close to $2 \vee (d-2)$

$$\|\nabla |\nabla u_n|^{\frac{p}{2}}\|_2^2 \leq K \|f\|_p^p, \tag{6.10}$$

[7] Note that after expanding $(1 + T_p)^{-1}$ in (6.5) in the geometric series, we obtain the formal Neumann series representation for u.

for a constant K independent of n. (We discuss a parabolic analogue of (6.10) below.) Then, applying the Sobolev embedding theorem twice, one obtains that u_n has Hölder continuity exponent $1-\frac{d-2}{p}$ (independent of n). Taking into account that p satisfies a strict inequality, one thus obtains the same Hölder continuity result as in Theorem 6.1(*ii*). However, there is substantial difference between gradient bounds (6.7) and (6.10): (6.7) allows to control order >1 derivatives of u_n, while (6.10) does not need additional strict inequalities such as "$q>p$" in (6.7) (clearly, if in (6.7) one could take $q=p$, then it would make the regularity result stronger, but one cannot do this).

Above we mentioned that the approach of Theorem 6.2 is somewhat simpler than [KS]. One reason is that it uses to the full extent the linear structure of the equation. There is another reason: in the iteration procedure of [KS] one shows that the solutions of the approximating equations constitute a Cauchy sequence, while in the proof of Theorem 6.2 one already has a candidate for the limit of this sequence, cf. (6.8).

The free Bessel potential $(\mu-\Delta)^{-\frac{1}{2}-\frac{1}{q}}$ in (6.5) provides a "trampoline" from L^p to C_∞. One advantage of working in L^p with p large is that it simplifies the proof of approximation results as e.g. in assertion (*iv*) of Theorem 6.1, since it is easier to prove convergence in L^p than in C_∞. In fact, this "trampoline" was used in [Ko] who considered bounded drifts of α-stable process. There, however, working in L^p is a matter of convenience (one can also stay entirely in C_∞ while dealing with unbounded drifts, see [BC]). Our goal, however, is to transition from an L^2 assumptions on the drift (i.e. the form-boundedness), via L^p, to a semigroup in C_∞. It is the transition from L^2 to L^p with p large that is the most difficult one.

The operator-valued function $\mu\mapsto\Theta_p(\mu,b)\in\mathcal{B}(L^p)$ in Theorem 6.1 is itself the resolvent of the generator of a quasi contraction semigroup in L^p. In fact, this semigroup coincides with the semigroup $e^{-t\Lambda_p(b)}$ constructed via (3.10), and satisfies

$$e^{-t\Lambda_p(b)}=s\text{-}C_\infty\text{-}\lim_n e^{-t\Lambda_p(b_n)}\quad\text{locally uniformly in } t\geq 0, \tag{6.11}$$

where $\Lambda_p(b_n)=-\Delta+b_n\cdot\nabla$ having domain $\mathcal{W}^{2,p}$. See [Ki2] for details.

By the way, it is easy to write a similar to $\Theta_p(\mu,b)$ operator-valued function representation for the resolvent of $-\Delta-\nabla\cdot b$ with $b\in\mathbf{F}_\delta$, and to modify the proofs in [Ki2] to work for this operator.

Combining (6.11) with Theorem 6.1, one of course obtains

$$e^{-t\Lambda_{C_\infty}(b)}\upharpoonright L^2\cap C_\infty=e^{-t\Lambda_p(b)}\upharpoonright L^2\cap C_\infty=e^{-t\Lambda(b)}\upharpoonright L^2\cap C_\infty, \tag{6.12}$$

where the last semigroup is provided by the KLMN theorem.

The resolvent representations of the type $\Theta_p(\mu,b)$ were considered earlier for Schrödinger operators with form-bounded potentials, to obtain information about Sobolev regularity of their domains and hence regularity of their eigenfunctions, see [BS, LS]. Let us note, however, that a direct comparison between Feller theories of Kolmogorov and Schrödinger operators is not possible, see Remark 14.2.

The semigroup $e^{-t\Lambda_p(b)}$ can, in fact, be defined via the limit (6.11) for all $\delta<4$, $p>\frac{2}{2-\sqrt{\delta}}$ [KS], and satisfies

$$\|e^{-t\Lambda_p(b)}f\|_q\leq ce^{t\omega_p}t^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}\|f\|_p,\quad t>0,\quad \omega_p:=\frac{c_\delta}{2(p-1)}, \tag{6.13}$$

for all $f \in L^q \cap L^p$. In view of (6.12), the same estimate on $e^{-t\Lambda_{C_\infty}(b)}f$, $f \in L^p \cap C_\infty$ is valid (of course, under the assumptions on δ and p of Theorem 6.1(iv)). Regarding the properties of $e^{-t\Lambda_p(b)}$ (in particular, regarding extending the interval $p \in]\frac{2}{2-\sqrt{\delta}}, \infty[$ to a larger interval), see [KiS2].

Nevertheless, both the upper and the lower Gaussian bounds on the heat kernel of $-\Delta + b \cdot \nabla$ are easily destroyed by form-bounded drifts. (The heat kernel is defined, up to a modification on a measure zero set, as the integral kernel $e^{-t\Lambda(b)}(x,y)$ of $e^{-t\Lambda_p(b)}$. The heat kernel does not depend on p.) In fact, it suffices to consider drifts

$$b(x) = \pm\sqrt{\delta}\frac{d-2}{2}|x|^{-2}x, \tag{6.14}$$

which are form-bounded. The singularity of (6.14) is so strong that it introduces an extra factor $\varphi_t(y)$ in the Gaussian bounds,

$$c_1 t^{-\frac{d}{2}} e^{-\frac{|x-y|^2}{c_2 t}} \varphi_t(y) \leq e^{-t\Lambda(b)}(x,y) \leq c_3 t^{-\frac{d}{2}} e^{-\frac{|x-y|^2}{c_4 t}} \varphi_t(y),$$

where $\varphi_t(t)$ either explodes or vanishes at the origin depending on the sign in front of $\sqrt{\delta}$ in (6.14) (moreover, the rate of explosion or vanishing is an explicit function of δ) [MNS]. Of course, if one considers a sum (or a series) of such drifts with singularities at different points (which is still form-bounded), the situation at the level of heat kernel bounds becomes even more complicated. A detailed discussion on this subject can be found in [KiS6] where the authors prove Gaussian lower and/or upper bound on the heat kernel of $-\nabla \cdot a \cdot \nabla + b \cdot \nabla$ with measurable uniformly elliptic matrix a and drift b that is form-bounded or even more singular, under additional constraints on $\operatorname{div} b$.

Let us also note that the two-sided Gaussian bound on the heat kernel of $-\Delta + b \cdot \nabla$ or of $-\nabla \cdot a \cdot \nabla + b \cdot \nabla$ hold, without any assumptions on $\operatorname{div} b$, when b is in the Kato class of vector fields[8] [Za2], or in the Nash class[9] [S2, KiS7], respectively. Moreover, one can go beyond the Kato class and prove Guassian bounds for distributional drifts, see [PZ, ZZh2]. Both the Kato class and the Nash class neither contain the class $\mathbf{F}_\delta$ nor are contained in it (but the Kato class of vector fields is contained in the class of *weakly form-bounded vector fields*, considered in the next section).

Remark 6.2. It is clear from (6.5) that we do not have (and cannot have) information about L^p summability of the second derivatives of u for $p > 2$ large (in this regard, see discussion before Theorem 4.1). However, we have weighted estimates on the second derivatives. Assume for simplicity that b is bounded and smooth, so we are looking for estimates with constants that do not depend on smoothness of b or its boundedness. It follows from (6.5) that

$$(|b|+1)^{-1+\frac{2}{p}}(\mu - \Delta)u = (|b|+1)^{-1+\frac{2}{p}} f - \left(\frac{|b|}{|b|+1}\right)^{1-\frac{2}{p}} (1+T_p)^{-1} T_p' \, (|b|+1)^{-1+\frac{2}{p}} f$$

[8]See definition in Section 14.

[9]$b = b(x)$ is in the Nash class $\mathbf{N}_\delta$ if $|b| \in L^2_{\rm loc}$ and

$$\inf_{h>0} \sup_{x \in \mathbb{R}^d} \int_0^h \sqrt{e^{t\Delta}|b|^2(x)} \, \frac{dt}{\sqrt{t}} \leq \delta$$

. It contains b with $|b| \in L^p$, $p > d$, and it also contains vector fields with $|b| \not\in L^{2+\varepsilon}_{\rm loc}$, $\varepsilon > 0$.

where $T_p' = b^{\frac{2}{p}} \cdot \nabla(\mu-\Delta)^{-1}(|b|+1)^{1-\frac{2}{p}}$ is bounded on L^p, just like T_p, and so

$$\|(|b|+1)^{-1+\frac{2}{p}}(\mu-\Delta)u\|_p \leq (1+C)\|(|b|+1)^{-1+\frac{2}{p}}f\|_p, \quad C<\infty.$$

Thus, if $p>2$, then at the points where $|b|$ is infinite the factor $(|b|+1)^{-1+\frac{2}{p}}$ vanishes, and so around the singular set of b the information about L^p summability of $(\mu-\Delta)u$ disappears, but in a controlled way.

Finally, let us add that in Theorem 6.1 we could also write

$$u = (\mu-\Delta)^{-1}f - (\mu-\Delta)^{-\frac{1}{2}-\frac{1}{q}}\hat{Q}_p(q)(1+\hat{T}_p)^{-1}\hat{G}_p(\mu-\Delta)^{-\frac{1}{2}+\frac{1}{r}}f,$$

where operators $\hat{Q}_p$, $\hat{T}_p$, $\hat{R}_p(r)$ have "classical" form but are bounded in the weighted L^p space with weight $|b|_1^2 := (|b|+1)^2$:

$$\begin{aligned}
\hat{Q}_p(q) = (\mu-\Delta)^{-\frac{1}{2}+\frac{1}{q}}b\,\cdot \quad &\text{is in } \mathcal{B}\big(\big[L^p(\mathbb{R}^d,|b|_1^2dx)\big]^d, L^p(\mathbb{R}^d)\big),\\
\hat{T}_p = \nabla(\mu-\Delta)^{-1}b\,\cdot \quad &\text{is in } \mathcal{B}\big(\big[L^p(\mathbb{R}^d,|b|_1^2dx)\big]^d\big),\\
\hat{G}_p(r) = \nabla(\mu-\Delta)^{-\frac{1}{2}-\frac{1}{r}} \quad &\text{is in } \mathcal{B}\big(L^p(\mathbb{R}^d), \big[L^p(\mathbb{R}^d,|b|_1^2dx)\big]^d\big)
\end{aligned}$$

where $[\cdot]^d$ denotes vector with d components. Their boundedness follows from the boundedness of Q_p, T_p, G_p on $L^p = L^p(\mathbb{R}^d, dx)$.

2. We are now in position to prove the following result on weak well-posedness of SDE

$$X_t = x - \int_0^t b(X_r)dr + \sqrt{2}W_t, \quad 0 \leq t \leq T \tag{6.15}$$

with $x\in\mathbb{R}^d$ fixed.

Theorem 6.2. *Let $b \in \mathbf{F}_\delta$ with $\delta < 1\wedge\big(\frac{2}{d-2}\big)^2$. Let $e^{-t\Lambda_{C_\infty}(b)}$ be the Feller semigroup constructed in Theorem 6.1. Fix $T>0$. The following is true:*

(i) There exist probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ on the canonical space $(C[0,T],\mathcal{B}_t)$ such that

$$\mathbb{E}_{\mathbb{P}_x}[f(X_t)] = (e^{-t\Lambda_{C_\infty}(b)}f)(x), \quad f\in C_\infty, \quad x\in\mathbb{R}^d.$$

For every $x\in\mathbb{R}^d$ the measure $\mathbb{P}_x$ is a weak solution to SDE (6.15)*.*

(ii) If δ is sufficiently small and, additionally, $|b|\in L^{\frac{d}{2}+\varepsilon}$ for some $\varepsilon>0$, then the constructed in (i) weak solution $\mathbb{P}_x$ belongs to and is unique in the class of weak solutions satisfying the following Krylov-type estimate for $q>\frac{d}{2}$ sufficiently close to $\frac{d}{2}$ (depending on how small ε is):

$$\mathbb{E}_{\mathbb{P}_x}\int_0^T |h(t,\omega_t)|dt \leq c\|h\|_{L^q([0,T]\times\mathbb{R}^d)}. \tag{6.16}$$

for all $h\in C_c(\mathbb{R}^{d+1})$, for generic constant c.

Remark 6.3. Arguing as in [KiS1], one can also prove the following "approximation uniqueness" result. If $\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d}$ is another weak solution to (6.15) such that

$$\mathbb{Q}_x = w\text{-}\lim_n \mathbb{P}_x(\tilde{b}_n) \quad \text{for every } x\in\mathbb{R}^d,$$

for some $\{\tilde{b}_n\}\subset\mathbf{F}_{\delta_1}$ with $\delta<1\wedge\big(\frac{2}{d-2}\big)^2$ and c_δ independent of n, then $\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d} = \{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$. In other words, the constructed weak solutions $\mathbb{P}_x$ of (6.15) are unique among those that can be obtained via an approximation procedure. Note that we do not require here any convergence of $\tilde{b}_n$ to b.

In Section 15 we will discuss analogues of Theorems 6.1, 6.2 for drifts that can be more singular than the form-bounded drifts. See Theorems 15.1, 15.2. This, however, will come at the cost of imposing more restrictive assumption on δ, and losing the possibility to include discontinuous diffusion coefficients, as we do for the form-bounded drifts in end of this section. Also, while the proof of the analogue of Lemma 6.1 in Section 15 (i.e. Lemma 15.1 there) relies on some operator inequalities for fractional powers of the Laplacian, the proof of Lemma 6.1 uses only elementary arguments.

Remark 6.4. The proof of Theorem 6.2 and the construction of the Feller semigroup in Theorem 6.1(*iv*)-(*v*) rely on the elliptic regularity result of Theorem 6.1(*i*)-(*iii*). In the next sections we will be working directly in the parabolic setting, thus avoiding the use of the Trotter approximation theorem and, generally speaking, arriving at shorter proofs. However, by working with resolvents and using the semigroup theory (i.e. Trotter's theorem), we can construct a Feller semigroup departing from L^p for a smaller p, which leads to less restrictive assumptions on δ. For instance, one can compare Theorem 6.1(*iv*) where p is chosen to be strictly greater than $2 \vee (d-2)$, and we require $\delta < 1 \wedge (\frac{2}{d-2})^2$, and Theorem 9.1 where p has to be strictly greater than d, and $\delta < \frac{1}{d^2}$; or, better, δ satisfies (C_5) below, which is still more restrictive than $\delta < 1 \wedge (\frac{2}{d-2})^2$.

7. Proof of Theorem 6.1

Assertion (*iii*) follows right away from Lemma 6.1. The first part of assertion (*i*), and assertion (*ii*), follow from (*iii*). The proof of the second part of (*i*), i.e. the characterization of u as the unique weak solution to the elliptic equation, is standard and we will attend to it in the end.

(*iv*) For every $n = 1, 2, \dots$, the operator-valued function $\Theta_p(\mu, b_n)$ is a pseudo-resolvent, i.e. it satisfies

$$\Theta_p(\mu, b_n) - \Theta_p(\eta, b_n) = (\nu - \mu)\Theta_p(\mu, b_n)\Theta_p(\nu, b_n), \quad \mu, \nu \geq \mu_0, \tag{7.1}$$

where μ_0 is from Lemma 6.1. Identity (7.1) is verified via direct calculation. See [Ki1, proof of Prop. 2] for details[10].

By the classical theory, for every $n = 1, 2, \dots$, the resolvent of the approximating operator $(\mu + \Lambda_{C_\infty}(b_n))^{-1}$ is defined on $\{\mu \geq \mu_n\}$, where μ_n depends e.g. on $\|b_n\|_\infty$. Our first observation is that we can replace μ_n by a μ_0 independent of n by establishing a link between $(\mu + \Lambda_{C_\infty}(b_n))^{-1}$ and the operator-valued function $\Theta_p(\mu, b_n)$. That is,

$$(\mu + \Lambda_{C_\infty}(b_n))^{-1} \upharpoonright \mathcal{S} = \Theta_p(\mu, b_n) \upharpoonright \mathcal{S} \quad \text{for all } \mu \geq \mu_0, \tag{7.2}$$

for some μ_0 independent of n. Indeed, we have

$$\Theta_p(\mu_n, b_n) \upharpoonright \mathcal{S} = (\mu_n + \Lambda_{C_\infty}(b_n))^{-1} \upharpoonright \mathcal{S}$$

for all sufficiently large μ_n. By $\Theta_p(\mu, b_n)\mathcal{S} \subset \mathcal{S}$, the previous identity and the resolvent identity (7.1),

$$\Theta_p(\mu, b_n) \upharpoonright \mathcal{S} = (\mu_n + \Lambda_{C_\infty}(b_n))^{-1}\big(1 + (\mu_n - \mu)\Theta_p(\mu, b_n)\big) \upharpoonright \mathcal{S}, \quad \mu \geq \mu_0,$$

[10]Alternatively, one can verify, using the KLMN theorem, that $\Theta_2(\mu, b_n)$, $\mu \geq \mu_0$, for a $\mu_0 > 0$ independent of n, is the resolvent of $-\Delta + b_n \cdot \nabla$ in L^2, and so (7.1) holds on $L^2 \cap L^p$ and hence on L^p. But we do not need the L^2 theory of $-\Delta + b \cdot \nabla$ in the proof of Theorem 6.1

so $\Theta_p(\mu,b_n) \upharpoonright \mathcal{S}$ is the right inverse of $\mu+\Lambda_{C_\infty}(b_n) \upharpoonright \mathcal{S}$ on $\mu \geq \mu_0$. Similarly, it is seen that $\Theta_p(\mu,b_n)|_{\mathcal{S}}$ is the left inverse of $\mu+\Lambda_{C_\infty}(b_n) \upharpoonright \mathcal{S}$ on $\mu \geq \mu_0$. This gives (7.2).

Second, let us show that for every $\mu \geq \mu_0$

$$\Theta_p(\mu,b)\mathcal{S} \subset L^p \cap C_\infty \quad \text{(after a modification on a measure zero set)}, \tag{7.3}$$

and

$$\Theta_p(\mu,b_n) \overset{s}{\to} \Theta_p(\mu,b) \text{ in } L^p \cap C_\infty \quad (n \to \infty). \tag{7.4}$$

The inclusion into C_∞ in (7.3) is immediate due to the factor $(\mu-\Delta)^{-\frac{1}{2}-\frac{1}{q}}$ in the definition of $\Theta_p(\mu,b_n)$, upon applying the Sobolev embedding theorem (here we use the assumption $p > 2 \vee (d-2)$ and the choice of $q>p$ close to p). The second assertion (7.3) is proved using again the Sobolev embedding theorem and the convergence

$$Q_p(q,b_n) \overset{s}{\to} Q_p(q,b), \quad T_p(b_n) \overset{s}{\to} T_p(b), \quad G_p(r,b_n) \overset{s}{\to} G_p(r,b).$$

The latter, in turn, follow from the Dominated convergence theorem since b_n defined by (6.2) are, essentially, cutoffs of b (for details, if needed, see the proof of [Ki1, Prop. 7]).

Third, we have

$$\sup_n \|\mu(\mu+\Lambda_{C_\infty}(b_n))^{-1}\|_{\infty\to\infty} \leq 1 \quad \text{ for all } \mu \geq \mu_0. \tag{7.5}$$

Indeed, for every $n=1,2,\dots$, the semigroup $e^{-t\Lambda_{C_\infty}(b_n)}$ is an L^∞ contraction, so, integrating $\|e^{-\mu t}e^{-t\Lambda_{C_\infty}(b_n)}f\|_\infty \leq e^{-\mu t}\|f\|_\infty$ in t from 0 to ∞ we arrive at (7.5).

Fourth, we note that

$$\mu\Theta_p(\mu,b_n) \overset{s}{\to} 1 \text{ in } C_\infty \text{ as } \mu \uparrow \infty \text{ uniformly in } n. \tag{7.6}$$

Indeed, in view of (7.5), since $\mathcal{S}$ is dense in C_∞, it suffices to prove that $\mu\Theta_p(\mu,b_n)f \to f$ in C_∞ as $\mu \uparrow \infty$ for all $f \in \mathcal{S}$. In turn, since $\lim_{\mu\to\infty}\|\mu(\mu-\Delta)^{-1}f-f\|_\infty=0$, it suffices to show that $\sup_n\|\mu\Theta_p(\mu,b_n)f-\mu(\mu-\Delta)^{-1}f\|_\infty \to 0$ as $\mu \uparrow \infty$. We have

$$\begin{aligned}\Theta_p(\mu,b_n)f-(\mu-\Delta)^{-1}f &= \\ &-(\mu-\Delta)^{-\frac{1}{2}-\frac{1}{q}}Q_p(q)\big(1+T_p\big)^{-1}b_n^{\frac{2}{p}}\cdot\nabla(\lambda-\Delta)^{-1}(\mu-\Delta)^{-1}(\lambda-\Delta)f,\end{aligned}$$

with $q>p$ where λ is sufficiently large but fixed. Note, that

$$\|(\mu-\Delta)^{-\frac{1}{2}-\frac{1}{q}}\|_{p\to\infty} \leq c\mu^{-\frac{1}{2}+\frac{d}{2p}-\frac{1}{q}},$$

and

$$\|b_n^{\frac{2}{p}}\cdot\nabla(\lambda-\Delta)^{-1}\|_{p\to p} \leq c_1$$

with c_1 independent of n and μ (since $\|G_p(r)\|_{p\to p} \equiv \|b_n^{\frac{2}{p}}\cdot\nabla(\lambda-\Delta)^{-\frac{1}{2}-\frac{1}{r}}\|_{p\to p}$ is uniformly bounded in n). Thus

$$\|\Theta_p(\mu,b_n)f-(\mu-\Delta)^{-1}f\|_\infty \leq C\mu^{-\frac{1}{2}+\frac{d}{2p}-\frac{1}{q}}\mu^{-1}\|(\lambda-\Delta)f\|_p.$$

Since $p>2\vee(d-2)$, we select q sufficiently close to p so that $-\frac{1}{2}+\frac{d}{2p}-\frac{1}{q}-1<-1$, and hence $\sup_n\|\mu\Theta_p(\mu,b_n)f-\mu(\mu-\Delta)^{-1}f\|_\infty \to 0$ as $\mu\uparrow\infty$, which yields (7.6).

We now prove assertion (*iv*) of the theorem using the Trotter approximation theorem (Appendix D). Its conditions

$$\sup_n \|(\mu+\Lambda_{C_\infty}(b_n))^{-1}\|_{\infty\to\infty} \leq \frac{1}{\mu} \quad \text{ for all } \mu \geq \mu_0,$$

$$\text{there exists } s\text{-}C_\infty\text{-}\lim_n(\mu+\Lambda_{C_\infty}(b_n))^{-1} \quad \text{for some } \mu \geq \mu_0,$$

$$\mu(\mu+\Lambda_{C_\infty}(b_n))^{-1} \to 1 \quad \text{in } C_\infty \text{ as } \mu\uparrow\infty \text{ uniformly in } n$$

are verified in (7.2)-(7.6). So, Trotter's theorem yields (iv) including the strong convergence of semigroups in C_∞.

(v) The proof of (6.9) uses the localized estimate

$$\big\|\rho(\mu+\Lambda_{C_\infty}(b_n))^{-1}h\big\|_\infty \leq C_3\|\rho h\|_p, \tag{7.7}$$

where

$$\rho(x) := (1+\kappa|x|^2)^{-\nu}$$

with $\nu > \frac{d}{2p}$ fixed (then $\rho\in L^p$) and $\kappa>0$ to be chosen sufficiently small. We comment on the proof of (7.7) below. Estimate (7.7) yields: for every fixed $x\in\mathbb{R}^d$ there is C such that

$$|(\mu+\Lambda_{C_\infty}(b_n))^{-1}h(x)| \leq C\|\rho h\|_p.$$

By considering an increasing sequence $h\uparrow 1-\mathbf{1}_{B_R(0)}$ we obtain

$$\langle(\mu+\Lambda_{C_\infty}(b_n))^{-1}(x,\cdot)(1-\mathbf{1}_{B_R(0)}(\cdot))\rangle \leq C\|\rho(1-\mathbf{1}_{B_R(0)}\|_p,$$

where, as is evident from the definition of ρ, the right-hand side can be made smaller than any ε uniformly in n by selecting radius $R>0$ sufficiently large.

Since $\langle(\mu+\Lambda_{C_\infty}(b_n))^{-1}(x,\cdot)\rangle=\mu^{-1}$, $n=1,2,\dots$, we have

$$\mu^{-1}-\varepsilon \leq \langle(\mu+\Lambda_{C_\infty}(b_n))^{-1}(x,\cdot)\mathbf{1}_{B_R(0)}(\cdot)\rangle \leq \mu^{-1}.$$

By passing to the limit in n we obtain $\mu^{-1}-\varepsilon \leq \langle(\mu+\Lambda_{C_\infty}(b))^{-1}(x,\cdot)\mathbf{1}_{B_R(0)}(\cdot)\rangle \leq \mu$. Finally, sending $R\to\infty$, and then $\varepsilon\downarrow 0$, we arrive at

$$\langle(\mu+\Lambda_{C_\infty}(b))^{-1}(x,\cdot)\rangle = \mu^{-1}$$

which gives us (6.9).

Regarding the proof of (7.7), we can either commute ρ with the operators that constitute $\Theta_p(\mu,b_n)\equiv(\mu+\Lambda_{C_\infty}(b_n))^{-1}$ (see [KiS1] for details), or we cannote that the equation for u_n, i.e. $(\mu-\Delta+b_n\cdot\nabla)u_n=h$, can be rewritten as

$$\mu\rho u_n-\Delta(\rho u_n)+\tilde{b}_n\cdot\nabla(\rho u_n)=\rho h+K,$$

where

$$\tilde{b}_n := b_n+2\frac{\nabla\rho}{\rho}, \qquad K=2\frac{(\nabla\rho)^2}{\rho}u_n+(-\Delta\rho)u_n+b_nu_n\cdot\nabla\rho.$$

Now, we apply bounds

$$|\nabla\rho(x)|\leq 2\nu\sqrt{\kappa}\rho(x), \quad \frac{|\nabla\rho(x)|^2}{\rho(x)}\leq 4\nu^2\kappa\rho(x), \quad |\Delta\rho(x)|\leq(4\nu^2+(4+2d)\nu)\kappa\rho(x), \quad x\in\mathbb{R}^d.$$

By selecting κ sufficiently small, one can make the form-bound of $\tilde{b}$ as close to the form-bound δ of b as needed. Furthermore, the first two terms in K can be absorbed by $\mu\rho u_n$ (at the expense of replacing μ with $\mu-\mu_1$ for appropriate $\mu_1=\mu_1(\kappa)>0$). We have

$$\rho u_n=(\mu-\mu_1+\Lambda_{C_\infty}(\tilde{b}_n))^{-1}(\rho h+b_nu_n\cdot\nabla\rho)=\Theta_p(\mu-\mu_1,\tilde{b}_n)(\rho h+b_nu_n\cdot\nabla\rho), \quad \mu>\mu_1,$$

so we can apply Lemma 6.1 to obtain (7.7). See details in [Ki5].

Returning to the proof of (i), we note that u_n coincides with the classical solution to $(\mu - \Delta + b_n \cdot \nabla)u_n = f$. Now, using convergence

$$Q_p(q,b_n) \xrightarrow{s} Q_p(q,b), \quad T_p(b_n) \xrightarrow{s} T_p(b), \quad G_p(r,b_n) \xrightarrow{s} G_p(r,b),$$

(see the prof of (7.4)) it is easy to pass to the limit in n in

$$\mu\langle u_n,\psi\rangle + \langle \nabla u_n, \nabla\psi\rangle + \langle b_n \cdot \nabla u_n, \psi\rangle = \langle f,\psi\rangle \quad \text{for all } \psi \in C_c^\infty,$$

which shows that u is a weak solution. A standard argument (i.e. the Lax-Milgram theorem) yields that u is the unique weak solution. $\square$

8. Proof of Theorem 6.2

(i) The following estimate will be needed: for all $h \in C_c$ and all $\mu \geq \mu_0$

$$\big\|(\mu + \Lambda_{C_\infty}(b))^{-1}|b_m|h\big\|_\infty \leq C_1\||b_m|^{\frac{2}{p}}h\|_p, \tag{8.1}$$

$$\|(\mu + \Lambda_{C_\infty}(b))^{-1}|b_m - b_n|h\|_\infty \leq C_2\big\||b_m - b_n|^{\frac{2}{p}}h\big\|_p. \tag{8.2}$$

for appropriate constants $C_i = C_i(\delta,p)$, $i=1,2$, where $p > 2 \vee (d-2)$ is fixed. These estimates follow right away from the construction of $(\mu + \Lambda_{C_\infty}(b))^{-1}$ via $\Theta_p(\mu,b)$ in Theorem 6.1(iv). Namely, since $|b_m|h \in C_c$,

$$\begin{aligned}(\mu + \Lambda_{C_\infty}(b))^{-1}|b_m|h &= \Theta_p(\mu,b)|b_m|h \\ &= (\mu-\Delta)^{-\frac{1}{2}-\frac{1}{q}}G_p(q,b_m)|b_m|^{\frac{2}{p}}h \\ &\quad - (\mu-\Delta)^{-\frac{1}{2}-\frac{1}{q}}Q_p(q,b)(1+T_p(b))^{-1}T_p(b,b_m)|b_m|^{\frac{2}{p}}h. \end{aligned} \tag{8.3}$$

The operator $T_p(b,b_m) := b^{\frac{2}{p}} \cdot \nabla(\mu-\Delta)|b_m|^{1-\frac{2}{p}}$ is "almost $T_p(b)$". In fact, repeating the proof of Lemma 6.1(i), we obtain $\|T_p(b,b_m)\|_{p\to p} \leq c'_\delta$ with constant c'_δ independent of m. Now, applying Lemma 6.1 in (8.3) and using the Sobolev embedding theorem (recall $p > 2 \vee (d-2)$ and $q > p$ is close to p), we obtain (8.1). The same argument gives (8.2).

By a standard result (see e.g. [BGe, Sect. I.9]), given a conservative Feller semigroup $e^{-t\Lambda_{C_\infty}(b)}$, there exist probability measures $\mathbb{P}_x$ ($x \in \mathbb{R}^d$) on $(D(\mathbb{R}_+,\mathbb{R}^d), \mathcal{B}'_t = \sigma(\omega_r, 0 \leq r \leq t))$, where $D([0,T],\mathbb{R}^d)$ is the space of right-continuous functions having left limits, and ω_t is the coordinate process on $D(\mathbb{R}_+,\mathbb{R}^d)$, such that

$$\mathbb{E}_x[f(\omega_t)] = e^{-t\Lambda_{C_\infty}(b)}f(x), \quad f \in C_\infty, \quad t > 0.$$

Here and below, $\mathbb{E}_x := \mathbb{E}_{\mathbb{P}_x}$. We will show that $\mathbb{P}_x$ are actually concentrated on $(C(\mathbb{R}_+,\mathbb{R}^d),\mathcal{B}_t)$.

For every $n = 1,2,\dots$, let $X_t^n = X_{t,x}^n$ denote the strong solution to the approximating SDE

$$X_t^n = x - \int_0^t b_n(s,X_s^n)ds + \sqrt{2}W_t, \quad x \in \mathbb{R}^d,$$

on a complete probability space $(\Omega,\ \mathcal{F},\ \mathcal{F}_t,\ \mathbf{P})$. Put $\mathbb{P}_x^n := (\mathbf{P}X^n)^{-1}$, $n = 1,2,\dots$, and set $\mathbb{E}_x^n := \mathbb{E}_{\mathbb{P}_x^n}$.

Fix $\mu \geq \mu_0$. In what follows, $0 < t \leq T < \infty$. For every $g \in C_c^2$, the following is true:

(**a**) $\mathbb{E}_x \int_0^t |b \cdot \nabla g|(\omega_s)ds < \infty$.

Indeed, since $b_n \to b$ everywhere outside of a measure zero set, we have by Fatou's lemma[11]

$$\begin{aligned}&\mathbb{E}_x \int_0^t \big|b\cdot\nabla g\big|(\omega_s)ds \\ &\leq \liminf_n \mathbb{E}_x \int_0^t \big|b_n\cdot\nabla g\big|(\omega_s)ds = \liminf_n \int_0^t e^{-s\Lambda_{C_\infty}(b)}\big|b_n\cdot\nabla g\big|(x)ds \\ &\leq e^{\mu T}\liminf_n (\mu+\Lambda_{C_\infty}(b))^{-1}|b_n||\nabla g|(x)\end{aligned}$$

Now, applying (8.1) with $h = |\nabla g|$, we obtain

$$\begin{aligned}\mathbb{E}_x \int_0^t \big|b\cdot\nabla g\big|(\omega_s)ds &\leq C_1 e^{\mu T}\liminf_n \langle |b_n|^2|\nabla g|^p\rangle^{\frac{2}{p}} \\ &= C_1 e^{\mu T}\langle |b|^2|\nabla g|^p\rangle^{\frac{2}{p}} < \infty \quad (\text{by } |b| \in L^2_{\rm loc}).\end{aligned}$$

(**b**)

$$\mathbb{E}_x \int_0^t (b_n\cdot\nabla g)(\omega_s)ds - \mathbb{E}^n_x \int_0^t (b_n\cdot\nabla g)(\omega_s)ds \to 0 \quad \text{ as } n\to\infty.$$

Indeed, we have:

$$\begin{aligned}&\mathbb{E}_x \int_0^t (b_n\cdot\nabla g)(\omega_s)ds - \mathbb{E}^n_x \int_0^t (b_n\cdot\nabla g)(\omega_s)ds \\ &= \int_0^t \Big(e^{-s\Lambda_{C_\infty}(b)} - e^{-s\Lambda_{C_\infty}(b_n)}\Big)(b_n\cdot\nabla g)(x)ds \\ &= \int_0^t \Big(e^{-s\Lambda_{C_\infty}(b)} - e^{-s\Lambda_{C_\infty}(b_n)}\Big)((b_n-b_m)\cdot\nabla g)(x)ds \\ &+ \int_0^t \Big(e^{-s\Lambda_{C_\infty}(b)} - e^{-s\Lambda_{C_\infty}(b_n)}\Big)(b_m\cdot\nabla g)(x)ds \\ &=: S_1 + S_2,\end{aligned}$$

where m is to be chosen. Reducing the estimates on the expectations of time integrals to the estimates on resolvents as in the proof of (**a**), we obtain:

$$S_1(x) \leq e^{\mu T}(\mu+\Lambda_{C_\infty}(b))^{-1}|(b_n-b_m)\cdot\nabla g|(x) + e^{\mu T}(\mu+\Lambda_{C_\infty}(b_n))^{-1}|(b_n-b_m)\cdot\nabla g|(x).$$

Using (8.2) and the convergence $b_n - b_m \to 0$ in $L^2_{\rm loc}$ as $n,m\uparrow\infty$, we obtain $S_1 \to 0$ as $n,m\uparrow\infty$. Now, let us fix a sufficiently large m. Since $e^{-s\Lambda_{C_\infty}(b)} = s\text{-}C_\infty\text{-}\lim_n e^{-s\Lambda_{C_\infty}(b_n)}$ uniformly in $0\leq s\leq T$ (i.e. assertion (*iv*) of Theorem 6.1), we have $S_2 \to 0$ as $n\uparrow\infty$. The proof of (**b**) is completed.

(**c**)

$$\mathbb{E}^n_x[g(\omega_t)] \to \mathbb{E}_x[g(\omega_t)]$$

[11]When applying Fatou's lemma, we use

$$\mathbb{E}_x \int_0^t \Big|b\cdot\nabla g\Big|(\omega_s)ds = \mathbb{E}_x \int_0^t \liminf_n \Big|b_n\cdot\nabla g\Big|(\omega_s)ds.$$

Indeed, $\xi := \Big|b\cdot\nabla g\Big| - \liminf_n \Big|b_n\cdot\nabla g\Big| = 0$ a.e. on $\mathbb{R}^d$, but $|\mathbb{E}_x\int_0^t \xi(\omega_s)ds| = 0$ (as follows e.g. by representing $\{\xi\neq 0\} = \cap_k U_k$ for a decreasing sequence of open sets U_k such that $|U_k|\downarrow 0$, smoothing out $\mathbf{1}_{U_k}$ by replacing it with $e^{\varepsilon_k\Delta}\mathbf{1}_{U_k}$ with $\varepsilon_k\downarrow 0$ rapidly, and then applying $\mathbb{E}_x\int_0^t e^{\varepsilon_k\Delta}\mathbf{1}_{U_k}(\omega_s)ds \leq e^{\mu T}(\mu+\Lambda_{C_\infty}(b))^{-1}e^{\varepsilon_k\Delta}\mathbf{1}_{U_k}(x) \leq Ce^{\mu T}\|e^{\varepsilon_k\Delta}\mathbf{1}_{U_k}\|_p \downarrow 0$ as $k\to\infty$. The last inequality follows from the construction of $(\mu+\Lambda_{C_\infty}(b))^{-1}$ via the operator-valued function $\Theta_p(\mu,b)$.

and

$$\mathbb{E}_x \int_0^t (b \cdot \nabla g)(\omega_s) ds - \mathbb{E}_x^n \int_0^t (b_n \cdot \nabla g)(\omega_s) ds \to 0 \quad \text{ as } n \to \infty.$$

The first convergence is direct a consequence of $e^{-s\Lambda_{C_\infty}(b)} = s\text{-}C_\infty\text{-}\lim_n e^{-s\Lambda_{C_\infty}(b_n)}$ uniformly in $0 \leq s \leq T$. The second convergence is a consequence of (**b**) and $\mathbb{E}_x \int_0^t ((b_n - b) \cdot \nabla g)(\omega_s) ds \to 0$ as $n \to \infty$, as follows from (8.2) upon applying Fatou's lemma in m there.

Now, since

$$M_{r,m}^g := g(\omega_r) - g(x) + \int_0^r (-\Delta g + b_m \cdot \nabla g)(\omega_t) dt$$

is a $\mathcal{B}'_r$-martingale under $\mathbb{P}_x^m$,

$$x \mapsto \mathbb{E}_x^m[g(\omega_r)] - g(x) + \mathbb{E}_x^m \int_0^r (-\Delta g + b_m \cdot \nabla g)(\omega_t) dt \quad \text{ is identically zero on } \mathbb{R}^d,$$

and so by (**c**)

$$x \mapsto \mathbb{E}_x[g(\omega_r)] - g(x) + \mathbb{E}_x \int_0^r (-\Delta g + b \cdot \nabla g)(\omega_t) dt \quad \text{ is identically zero in } \mathbb{R}^d.$$

Since $\{\mathbb{P}_x\}_{x \in \mathbb{R}^d}$ are determined by a Feller semigroup, and thus constitute a Markov process, we can conclude (see e.g. the proof of [Kr1, Lemma 2.2]) that

$$M_r^g := g(\omega_r) - g(x) + \int_0^r (-\Delta g + b \cdot \nabla g)(\omega_t) dt,$$

is a $\mathcal{B}'_r$-martingale under $\mathbb{P}_x$.

Let us show now that $\{\mathbb{P}_x\}_{x \in \mathbb{R}^d}$ are concentrated on $(C([0,T], \mathbb{R}^d), \mathcal{B}_t)$. Since ω_t is a semi-martingale under $\mathbb{P}_x$, Itô's formula yields, for every $g \in C_c^\infty(\mathbb{R}^d)$, that

$$g(\omega_t) - g(x) = \sum_{s \leq t} (g(\omega_s) - g(\omega_{s-})) + S_t, \tag{8.4}$$

where S_t is defined in terms of some integrals and sums of $(\partial_{x_i} g)(\omega_{s-})$ and $(\partial_{x_i} \partial_{x_j} g)(\omega_{s-})$ in s, see [CKS, Sect. 2] for details. Now, let A, B be arbitrary compact sets in $\mathbb{R}^d$ such that $\operatorname{dist}(A, B) > 0$. Fix $g \in C_c^\infty(\mathbb{R}^d)$ that separates A, B, say, $g = 0$ on A, $g = 1$ on B. Set

$$K_t^g := \int_0^t \mathbf{1}_A(\omega_{s-}) dM_s.$$

In view of (8.4), when evaluating K_t^g one needs to integrate $\mathbf{1}_A(\omega_{s-})$ with respect to S_t, however, one obtains zero since $(\partial_{x_i} g)(\omega_{s-}) = (\partial_{x_i} \partial_{x_j} g)(\omega_{s-}) = 0$ if $\omega_{s-} \in A$. Thus,

$$\begin{aligned} K_t^g &= \sum_{s \leq t} \mathbf{1}_A(\omega_{s-}) \, g(\omega_s) + \int_0^t \mathbf{1}_A(\omega_{s-}) (-\Delta g + b \cdot \nabla g)(\omega_s) ds \\ &= \sum_{s \leq t} \mathbf{1}_A(\omega_{s-}) \, g(\omega_s). \end{aligned}$$

Since M_t^g is a martingale, so is K_t^g. Thus, $\mathbb{E}_x \big[\sum_{s \leq t} \mathbf{1}_A(\omega_{s-}) g(\omega_s)\big] = 0$. Using the Dominated convergence theorem, we further obtain $\mathbb{E}_x \big[\sum_{s \leq t} \mathbf{1}_A(\omega_{s-}) \mathbf{1}_B(\omega_s)\big] = 0$, which yields the required. By the way, this construction, in a more general form, was used to control the jumps of stable process perturbed by a drift, see [CKS].

We denote the restriction of $\mathbb{P}_x$ from $(D([0,T],\mathbb{R}^d),\mathcal{B}'_t)$ to $(C([0,T],\mathbb{R}^d),\mathcal{B}_t)$ again by $\mathbb{P}_x$, and thus obtain that for every $x\in\mathbb{R}^d$ and all $g\in C_c^2(\mathbb{R}^d)$

$$M_r^g = g(\omega_r) - g(x) + \int_0^r (-\Delta g + b\cdot\nabla g)(\omega_t)dt, \quad \omega\in C([0,T],\mathbb{R}^d),$$

is a $\mathcal{B}_r$-martingale under $\mathbb{P}_x$. Thus, $\mathbb{P}_x$ is a martingale solution to (6.15).

To show that $\mathbb{P}_x$ is a weak solution it suffices to show that M_r^g is also a martingale for $g(x)=x_i$ and $g(x)=x_ix_j$ (proving along the way that $\mathbb{E}_x\int_0^t |b|(X(s))ds<\infty$), which can be done by following closely [KiS1, proof of Lemma 6].

(ii) is obtained via a simple modification of the proof of the uniqueness result of Theorem 16.2(iv) below. $\square$

9. Time-inhomogeneous form-bounded drifts and Feller theory via iterations

1. The following is the time-inhomogeneous counterpart of Definition 3.1.

Definition 9.1. A Borel measurable vector field $b:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}^d$ is said to be form-bounded if

$$|b|\in L^2_{\rm loc}(\mathbb{R}_+\times\mathbb{R}^d)$$

and there exist a constant $\delta>0$ and a function $0\leq g\in L^1_{\rm loc}(\mathbb{R}_+)$ such that for a.e. $t\in\mathbb{R}_+$

$$\|b(t,\cdot)\varphi\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + g(t)\|\varphi\|_2^2 \tag{9.1}$$

for all $\varphi\in W^{1,2}$.

This will be written as $b\in L^\infty\mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$.

An equivalent form of the a.e. inequality (9.1) is: for every $0<T<\infty$,

$$\int_0^T \|b(t)\psi(t)\|_2^2dt \leq \delta\int_0^T \|\nabla\psi(t)\|_2^2dt + \int_0^T g(t)\|\psi(t)\|_2^2dt$$

for all $\psi\in L^\infty(\mathbb{R}_+,W^{1,2})$.

Examples. The class of time-inhomogeneous form-bounded vector fields includes e.g. the critical Ladyzhenskaya-Prodi-Serrin class

$$|b|\in L^q_{\rm loc}(\mathbb{R}_+,L^r+L^\infty), \quad \frac{d}{r}+\frac{2}{q}\leq 1, \quad 2\leq q\leq\infty,$$

as well as vector fields having stronger spatial singularities, see Appendix B.

We fix an approximation of $b\in L^\infty\mathbf{F}_\delta+L^2_{\rm loc}$ by smooth bounded vector fields b_m that preserve the form-bound δ and have functions g_n locally uniformly bounded in $L^1(\mathbb{R}_+)$, i.e.

$$b_n\to b \quad \text{in } L^2_{\rm loc}(\mathbb{R}_+\times\mathbb{R}^d,\mathbb{R}^d) \tag{9.2}$$

and for all $t\geq 0$

$$\|b_n(t)\varphi\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + g_n(t)\|\varphi\|_2^2 \tag{9.3}$$

with g_n such that

$$\sup_n\int_0^T g_n(s)ds<\infty \quad \text{for any } 0<T<\infty. \tag{9.4}$$

Examples. It is easy to show that the following b_n, with $\varepsilon_n \downarrow 0$ sufficiently rapidly and $c_n \uparrow 1$ sufficiently slow, satisfy (9.2)-(9.4).

$$b_n := c_n E^{1+d}_{\varepsilon_n}(\mathbf{1}_n b),$$

where $\mathbf{1}_n$ is the indicator of $\{(t,x) \mid |b(t,x)| \leq n, |x| \leq n, |t| \leq n\}$ (say, b is extended by 0 to $t<0$), E^{1+d}_ε is the De Giorgi or the Friedrichs mollifier on $\mathbb{R} \times \mathbb{R}^d$. See details in Appendix C.1. Note that, by selecting $\varepsilon_n \downarrow 0$ rapidly, one can treat b_n as basically a cutoff of b times constant c_n.

Moreover, with some additional effort, one can simplify this approximation to

$$b_\varepsilon := E^1_\varepsilon E^d_\varepsilon b, \quad \varepsilon \downarrow 0,$$

where E^1_ε in the Friedrichs mollifier on $\mathbb{R}$, and E^d_ε is the De Giorgi or the Friedrichs mollifier on $\mathbb{R}^d$. See Appendix C.3.

The last approximation is important if one needs e.g. to transfer the form-boundedness assumption on "potential" $\operatorname{div} b$ to the uniform form-boundedness of $\operatorname{div} b_\varepsilon$ since then one commute div and the mollifiers, although we are not concerned with this here.

2. Our first goal is to construct the corresponding to $(\partial_t - \Delta + b \cdot \nabla)u = 0$, $b \in L^\infty \mathbf{F}_\delta + L^2_{\rm loc}$, Feller evolution family on $D_T = \{(s,t) \mid 0 \leq s \leq t \leq T\}$ for $T>0$ fixed, i.e. a family of operators $\{U^{t,s}\}_{(s,t)\in D_T}$ that are bounded on C_∞, and

1) $U^{t,r}U^{r,s} = U^{t,s}$, $r \in [s,t]$, and $U^{s,s} = \mathrm{I}$,
2) $\|U^{t,s}f\|_\infty \leq \|f\|_\infty$, $U^{t,s}[C^+_\infty] \subset C^+_\infty$,
3)
$$U^{r,s} = s\text{-}C_\infty\text{-}\lim_{t \downarrow r} U^{t,s}, \quad r \geq s,$$

and which will, additionally, satisfy: $u(t) := U^{t,s}f$ is the unique weak solution of $(\partial_t - \Delta + b\cdot\nabla)u = 0$, $u(s) = f \in C_\infty \cap L^2$.

The sought Feller evolution family is produced as the limit $L^\infty(D_T, L^\infty)$ of

$$U^{t,s}_n f(\cdot) := u_n(t,\cdot), \quad (s,t) \in D_T$$

and u_n is the classical solution to initial problem

$$(\partial_t - \Delta + b_n \cdot \nabla)u_n = 0, \quad u_n(s,\cdot) = f(\cdot) \in C^\infty_c. \tag{9.5}$$

We will prove the uniform convergence of the functions $\{(t,s,x) \mapsto U^{t,s}_n f(x)\}$ on $D_T \times \mathbb{R}^d$ by showing that they constitute a Cauchy sequence in $L^\infty(D_T, L^\infty)$. To that end, we will employ a parabolic variant of the iteration procedure of [KS]. This parabolic variant first appeared in [Ki3] and was recently refined in [KiS9].

Namely, subtracting the equations for u_m, u_n and setting

$$h := u_m - u_n,$$

one obtains

$$\partial_t h - \Delta h + b_m \cdot \nabla h + (b_m - b_n) \cdot \nabla u_n = 0, \quad h(s,\cdot) = 0 \tag{9.6}$$

Multiplying the last equation by $h|h|^{r-2}$, $r > \frac{2}{2-\sqrt{\delta}}$, integrating over $[s,T] \times \mathbb{R}^d$ and applying the Sobolev embedding theorem, we arrive at the inequality

$$c_d r^k \|h\|^r_{L^\infty([s,T],L^r)} + \|h\|^r_{L^r([s,T],L^{\frac{rd}{d-2}})} \leq Cr^{2k}e^{C_T} \sup_{\tau\in[s,T]} \|\nabla u_n(\tau)\|^2_q \int_s^T \|h(\tau)\|^{r-2}_{\frac{q}{q-2}(r-2)} d\tau,$$

where for a fixed $q > d$, for constants C, C_T that are independent of m, n. Hence, applying the interpolation inequality in the left-hand side and setting $K = Ce^{C_T}$, one obtains

$$\|h\|_{\frac{r}{1-\beta},\frac{rd}{d-2+2\beta}} \leq K^{\frac{1}{r}}(r^{\frac{1}{r}})^{2k}\Big(\sup_{\tau\in[s,T]}\|\nabla u_n(\tau)\|_q\Big)^{\frac{2}{r}}\|h\|^{1-\frac{2}{r}}_{L^{r-2}([s,T],L^{\frac{q}{q-2}(r-2)})}.$$

(we only need $\delta < 4$ to prove this inequality). Now, with appropriate choice of β, one can iterate this inequality in essentially the same way as it was done [KS] *provided that one has uniform in n bound on* $\sup_{\tau\in[s,T]}\|\nabla u(\tau)\|_q$ (see below), arriving at

$$\|u_m - u_n\|_{L^\infty([s,T],L^\infty)} \leq C_1\|u_m - u_n\|^\gamma_{L^{r_0}([s,T],L^{r_0})}$$

for $r_0 > \frac{2}{2-\sqrt{\delta}}$ and some $\gamma > 0$ (this strict inequality is the main concern of the iteration procedure). Now, a standard argument yields

$$\|u_m - u_n\|^\gamma_{L^{r_0}([s,T],L^{r_0})} \to 0 \quad \text{as } m,n\to\infty,$$

see e.g. [Ki3],[KiS9], and so we have our Cauchy sequence:

$$\|u_m - u_n\|_{L^\infty([s,T],L^\infty)} \to 0 \quad \text{ as } m,n\to\infty,$$

moreover, the convergence is uniform in $s \in [0,T]$. So, we can define the sought Feller evolution family by

$$U^{t,s}f := s\text{-}C_\infty\text{-}\lim_n u_n(t), \quad (s,t)\in D_T,$$

for $f \in C_c^\infty$, as was assumed above, and then extend operators $U^{t,s}$ to all $f \in C_\infty$ by continuity using the fact that $U^{t,s}$ inherits the L^∞ contraction property from $U_n^{t,s}$. Let us emphasize that the a priori assumption $f \in C_c^\infty$ is needed for the uniform in n bound on $\sup_{\tau\in[s,T]}\|\nabla u(\tau)\|_p$, $p > d$.

Remark 9.1. Working in the elliptic setting (i.e. as in the proof of Theorem 6.1 or in [KS]), after showing that solutions to the approximating elliptic equation converge in C_∞, one needs to verify the other conditions of the Trotter approximation theorem. This is not needed when one is working directly with the parabolic equation, so we arrive at shorter proofs even if $b = b(x)$, however, at expense of requiring smaller δ. We discussed this effect in Remark 6.4. It is fair to say that there is a fundamental difference between time-homogeneous and time-inhomogeneous cases when one is dealing with singular drifts.

3. To make the iteration procedure converge, one needs gradient bound

$$\sup_n \sup_{\tau\in[s,T]}\|\nabla u_n(\tau)\|_p < \infty, \quad \text{ for some } p > d \tag{9.7}$$

for f in a dense subset of C_∞ (e.g. for $f \in C_c^\infty$). To obtain such a bound, one can differentiate the initial problem (9.5). Namely, writing for brevity

$$u := u_n, \quad b := b_n$$

and

$$w := \nabla u, \quad w_i := \nabla_i u,$$

we obtain

$$\partial_t w_j - \Delta w_j + b\cdot\nabla w_j + (\nabla_j b)\cdot w = 0, \quad w_j(0) = \nabla_j f, \quad 1 \leq j \leq d. \tag{9.8}$$

Now, one needs to "wrap up" this system and, additionally, get rid of the derivative $\nabla_j b$. For instance, one can consider all products $w_{i_1} \dots w_{i_m}$ for m fixed and then sum them up, as was done in [BFGM] for solutions of stochastic transport equation and, after them, in [KSS]. However, we are interested in arguments that give less restrictive assumptions on δ. Another argument was used in [Kr2], although in a different situation dealing with a more sophisticated system of parabolic equations. This argument still imposes more restrictive assumption on δ than the arguments employed [Ki3] and [KiS9] (compare (C_3) with (C_4), (C_5)). However, it is quite nice and simple (and, again, works for other equations), so we describe it here. Assume that that the form-bound δ of b satisfies

$$\sqrt{\delta} < \frac{d-1}{d(d+1)}. \tag{C_3}$$

Put

$$w_\eta := \eta \cdot w, \quad \eta = (\eta_j)_{j=1}^d \in \mathbb{R}^d.$$

We differentiate (9.5) in the direction η, i.e. multiply (9.8) by η_j and add the resulting identities in $j = 1, \dots, d$ to obtain

$$\partial_t w_\eta - \Delta_x w_\eta - b \cdot \nabla_x w_\eta - \sum_{i=1}^d (\eta \cdot \nabla b_i) \nabla_{\eta_i} w_\eta = 0,$$

where in the last term we have used $w_i = \nabla_{\eta_i} w_\eta$. Given a function $g(x, \eta)$, we denote by $\langle g \rangle_x$, $\langle g \rangle_x$ the integral of g over $\mathbb{R}^d$ in x and in η, respectively. Let $\langle g \rangle_{x,\eta}$ denote the corresponding repeated integral over $\mathbb{R}^d \times \mathbb{R}^d$. Set

$$h(\eta) := (1 + \kappa |\eta|^2)^{-\theta},$$

where $\kappa > 0$ is fixed arbitrarily, and $\theta > \frac{d+q}{2}$ so that $\langle |\eta|^q h \rangle_\eta < \infty$. Let $q \geq 2$ (in the iteration procedure we need $q > d$). Also, without loss of generality, q is rational with odd denominator, so we, if needed, we can raise negative numbers of power q. Multiply the previous identity by $h w_\eta^{q-1}$ and integrate in $(x, \eta) \in \mathbb{R}^{2d}$ to obtain

$$\begin{aligned} &\frac{1}{q} \partial_t \langle h w_\eta^q \rangle_{x,\eta} + \frac{4(q-1)}{q^2} \langle h |\nabla w_\eta^{\frac{q}{2}}|^2 \rangle_{x,\eta} \\ &\qquad - \frac{2}{q} \langle b \cdot \nabla w_\eta^{\frac{q}{2}}, h w_\eta^{\frac{q}{2}} \rangle_{x,\eta} - \sum_{i=1}^d \langle (\eta \cdot \nabla b_i) \partial_{\eta_i} w_\eta, h w_\eta^{q-1} \rangle_{x,\eta} = 0, \end{aligned} \tag{9.9}$$

where b_i are the components of b. The last term in the left-hand side is dealt with as follows:

$$\begin{aligned} - \sum_{i=1}^d \langle (\eta \cdot \nabla b_i) \partial_{\eta_i} w_\eta, h w_\eta^{q-1} \rangle_{x,\eta} &= \sum_{i=1}^d \langle \eta b_i \nabla \partial_{\eta_i} w_\eta, h w_\eta^{q-1} \rangle_{x,\eta} \\ &\quad + (q-1) \sum_{i=1}^d \langle \eta b_i \partial_{\eta_i} w_\eta, h w_\eta^{q-2} \nabla w_\eta \rangle_{x,\eta} \\ &\quad \text{(now we integrate by parts in } \eta_i \text{ in the first term)} \\ &= -\langle b \cdot \nabla w_\eta, h w_\eta^{q-1} \rangle_{x,\eta} - \sum_{i=1}^d \langle \eta b_i \nabla w_\eta, (\partial_{\eta_i} h) w_\eta^{q-1} \rangle_{x,\eta}. \end{aligned}$$

Hence, (9.9) becomes

$$\frac{1}{q}\partial_t\langle hw_\eta^q\rangle_{x,\eta}+\frac{4(q-1)}{q^2}\langle h|\nabla w_\eta^{\frac{q}{2}}|^2\rangle_{x,\eta} \\ -\frac{4}{q}\langle b\cdot\nabla w_\eta^{\frac{q}{2}},hw_\eta^{\frac{q}{2}}\rangle_{x,\eta}-\frac{2}{q}\sum_{i=1}^d\langle\eta b_i\nabla w_\eta^{\frac{q}{2}},(\partial_{\eta_i}h)w_\eta^{\frac{q}{2}}\rangle_{x,\eta}=0,$$

so

$$\partial_t\langle hw_\eta^q\rangle_{x,\eta}+\frac{4(q-1)}{q}\langle h|\nabla w_\eta^{\frac{q}{2}}|^2\rangle_{x,\eta} \\ -4\langle b\cdot\nabla w_\eta^{\frac{q}{2}},hw_\eta^{\frac{q}{2}}\rangle_{x,\eta}+4\theta\sum_{i=1}^d\langle b_i\nabla w_\eta^{\frac{q}{2}},\frac{\kappa\eta_i\eta}{1+\kappa|\eta|^2}hw_\eta^{\frac{q}{2}}\rangle_{x,\eta}=0.$$

Hence

$$\partial_t\langle hw_\eta^q\rangle_{x,\eta}+\frac{4(q-1)}{q}\langle h|\nabla w_\eta^{\frac{q}{2}}|^2\rangle_{x,\eta}-(4+4\theta)\langle|b||\nabla w_\eta|^{\frac{q}{2}},hw_\eta^{\frac{q}{2}}\rangle_{x,\eta}\leq 0. \tag{9.10}$$

We estimate

$$\begin{aligned}\langle|b||\nabla w_\eta|^{\frac{q}{2}},hw_\eta^{\frac{q}{2}}\rangle_{x,\eta}&\leq\varepsilon\langle|b|^2,hw_\eta^q\rangle_{x,\eta}+\frac{1}{4\varepsilon}\langle h|\nabla w_\eta|^q\rangle_{x,\eta}\\&\leq\varepsilon\Big(\delta\langle h|\nabla w_\eta^{\frac{q}{2}}|^2\rangle+g(t)\langle w_\eta^q\rangle\Big)+\frac{1}{4\varepsilon}\langle h|\nabla w_\eta|^q\rangle_{x,\eta}\qquad\varepsilon:=\frac{1}{2\sqrt{\delta}}.\end{aligned}$$

Thus, integrating (9.10) from s to t, one obtains

$$\langle hw_\eta^q(t)\rangle_{x,\eta}+4\Big(\frac{q-1}{q}-\sqrt{\delta}(1+\theta)\Big)\int_s^t\langle h|\nabla w_\eta^{\frac{q}{2}}(\tau)|^2\rangle_{x,\eta}d\tau\leq\frac{1}{2\sqrt{\delta}}\int_s^t g(\tau)\langle w_\eta^q(\tau)\rangle d\tau+\langle h(\nabla f\cdot\eta)^q\rangle,$$

which allows to conclude

$$\sup_{t\in[s,T]}\langle hw_\eta^q(t)\rangle_{x,\eta}+\int_s^t\langle h|\nabla w_\eta^{\frac{q}{2}}(\tau)|^2\rangle_{x,\eta}d\tau\leq C\|\nabla f\|_q^q \tag{9.11}$$

for some for some $q>d$, for constant C independent of n, provided that

$$\frac{q-1}{q}-\sqrt{\delta}(1+\theta)>0\text{ for some }\theta>\frac{d+q}{2}\quad\Leftrightarrow\quad\sqrt{\delta}<\frac{d-1}{d(d+1)}.$$

It remains to derive (9.7) from (9.11). Put

$$A_{t,x}:=\{\eta\in\mathbb{R}^d\mid|\eta-\frac{w(t,x)}{|w(t,x)|}|<\frac{1}{2}\},\quad x\in\mathbb{R}^d$$

(if $w(t,x)=0$, fix $\eta=(1,0,\dots,0)$). Thus, $A_{t,x}$ is a ball of radius $\frac{1}{2}$ with centre placed at distance 1 from the origin. The angle between $w(t,x)$ and $\eta\in A_{t,x}$ is bounded from above by a generic constant, hence $|\eta\cdot w(t,x)|\geq c|w(t,x)|$ for some $c>0$ independent of (t,x), for all $\eta\in A_{t,x}$. Therefore, for all $t\in[s,T]$

$$\begin{aligned}\langle hw_\eta^q\rangle_{x,\eta}&=\langle\langle h(\eta)|\eta\cdot w(x)|^q\rangle_\eta\rangle_x\\&\geq\langle\langle h(\eta)|\eta\cdot w(x)|^q\mathbf{1}_{A_{t,x}}(\eta)\rangle_\eta\rangle_x\\&\geq\langle c^q|w(x)|^q\langle h(\eta)\mathbf{1}_{A_{t,x}}(\eta)\rangle_\eta\rangle_x=C\langle|w|^q\rangle_x,\quad C>0.\end{aligned}$$

Thus, we obtain from (9.11)

$$\sup_{t\in[s,T]}\|w(t)\|_q^q + c\int_s^T \|\nabla|w|^{\frac{q}{2}}\|_2^2 dt \leq C_T\|\nabla f\|_q^q, \quad c>0. \tag{9.12}$$

Hence, one can run the iteration procedure under the assumption (C_3).

4. The proofs of (9.12) in [Ki3, KiS9] choose a specific direction of the differentiation (following [KS] which, by the way, appeared earlier than the other papers cited above):

$$\eta = \frac{w}{|w|},$$

which maximizes the directional derivative $w_\eta = w \cdot \eta$. Put differently, one multiplies the parabolic equation in (9.5) by the test function

$$\varphi = -\nabla\cdot\Big(\frac{w}{|w|}|w|^{q-1}\Big) \tag{9.13}$$

and then integrates by parts (the same test function is used in the proof of Lemma 6.1). This choice of the test function (or direction) leads to better assumptions on δ than (C_3). Indeed, the identity

$$\langle \partial_t u, \varphi\rangle + \langle -\Delta u, \varphi\rangle + \langle b_m\cdot w, \varphi\rangle = 0,$$

yields

$$\frac{1}{q}\partial_t\|w\|_q^q + I_q + (q-2)J_q = \langle b_m \cdot w, \nabla\cdot(w|w|^{q-2})\rangle \tag{9.14}$$

where

$$I_q = \sum_{i=1}^d \langle|\nabla w_i|^2, |w|^{q-2}\rangle, \quad J_q = \langle|\nabla|w||^2, |w|^{q-2}\rangle$$

are the "good" terms, i.e. the right-hand side of (9.14) will be estimated in terms of I_q, J_q multiplied by coefficients that, thus, cannot be too large, hence our assumptions on δ. Namely, we represent

$$\begin{aligned}\langle b_m\cdot w, \nabla\cdot(w|w|^{q-2})\rangle &= \langle b_m\cdot w, \Delta u|w|^{q-2}\rangle + (q-2)\langle b_m\cdot w, |w|^{q-3}w\cdot\nabla|w|\rangle \\ &:= S_1 + S_2.\end{aligned}$$

Put $B_q = \langle (b_m\cdot w)^2, |w|^{q-2}\rangle$, then

$$S_2 \leq (q-2)B_q^{\frac{1}{2}}J_q^{\frac{1}{2}}, \tag{9.15}$$

where B_q is estimated using (9.3), i.e. the form-boundedness of b_m:

$$B_q \leq \frac{q\sqrt{\delta}}{2}J_q + g_m\|w\|_q^q. \tag{9.16}$$

Thus, S_2 is estimated in terms of J_q, and one can apply the resulting bound on S_2 in (9.14).

In [Ki3], after estimating S_1 as

$$|S_1| \leq B_q^{\frac{1}{2}}\langle|\Delta u|^2, |w|^{q-2}\rangle^{\frac{1}{2}}, \tag{9.17}$$

the factor $\langle|\Delta u|^2, |w|^{q-2}\rangle$ was bounded by I_q and J_q without appealing to the equation, by representing $|\Delta u|^2 = (\nabla \cdot w)^2$ and integrating by parts twice:

$$\langle |w|^{q-2}|\Delta u|^2\rangle = -\langle w\cdot\nabla|w|^{q-2}, \Delta u\rangle + \sum_{r=1}^{d}\left\langle w\cdot\nabla w_r, \nabla_r|w|^{q-2}\right\rangle + I_q,$$

where

$$|\langle w\cdot\nabla|w|^{q-2}, \Delta u\rangle| \leqslant (q-2)\left(\frac{1}{4\varkappa}\langle|w|^{q-2}|\Delta u|^2\rangle + \varkappa J_q\right), \quad \kappa > 0,$$

and

$$|\sum_{r=1}^{d}\left\langle w\cdot\nabla w_r, \nabla_r|w|^{q-2}\right\rangle| \leqslant (q-2)\left(\frac{1}{2}I_q + \frac{1}{2}J_q\right).$$

Hence

$$\left(1-\frac{q-2}{4\varkappa}\right)\langle|w|^{q-2}|\Delta u|^2\rangle \leqslant I_q + (q-2)\left(\varkappa J_q + \frac{1}{2}I_q + \frac{1}{2}J_q\right), \quad \varkappa > \frac{q-2}{4}. \tag{9.18}$$

The resulting from (9.17), (9.18) bound on $|S_1|$, combined with (9.15), (9.16), led in [Ki3] to the gradient estimate (9.12) for a $q > d$ close to d provided that

$$\sqrt{\delta} < \frac{1}{d}. \tag{C_4}$$

One important advantage of working with the test function (9.13) is that one can "evaluate" it by representing $\Delta u = \partial_t u + b_m \cdot \nabla u$, thus using the equation one more time. This is what [KS] did. The same can be done in the parabolic setting, and it leads to better assumptions on δ than (C_4) [KiS9]. Specifically, in dimensions $3 \le d \le 6$, one abandons estimate (9.17) and represents $\Delta u = \partial_t u + b_m \cdot \nabla u$ to evaluate

$$S_1 = -\frac{1}{q}\frac{d}{dt}\|w\|_q^q + B_q - \langle|\partial_t u|^2, |w|^{q-2}\rangle - (q-2)\langle|w|^{q-3}w\cdot\nabla|w|, \partial_t u\rangle,$$

which, upon applying the quadratic inequality, gives

$$S_1 \le -\frac{1}{q}\partial_t\|w\|_q^q + B_q + \frac{(q-2)^2}{4}J_q.$$

This bound, (9.15) and (9.16) applied in (9.14) give the desired gradient bound (9.12) for $\sqrt{\delta} \le (\sqrt{q-1} - \frac{q-2}{2})\frac{2}{q}$. In dimensions $d = 3$ and $d = 4$ this gives significantly less restrictive assumption on δ than (C_4), see (C_5) below.

In dimensions $d \ge 5$, one starts with two representations for S_1:

$$\begin{aligned} S_1 &= \langle -\partial_t u + \Delta u, |w|^{q-2}\Delta u\rangle \\ &= \langle|\Delta u|^2, |w|^{q-2}\rangle - \mathrm{Re}\langle\partial_t u, |w|^{q-2}\Delta u\rangle, \end{aligned}$$

$$\begin{aligned} S_1 &= \langle b_n\cdot w, |w|^{q-2}(\partial_t u + b_n\cdot w)\rangle \\ &= B_q + \langle b_n\cdot w, |w|^{q-2}\partial_t u\rangle. \end{aligned}$$

Equating the right-hand sides, one obtains

$$\begin{aligned}\langle|\Delta u|^2,|w|^{q-2}\rangle &= B_q + \langle\partial_t u, |w|^{q-2}(b_n\cdot w+\Delta u)\rangle \\ &= B_q + \langle\partial_t u, |w|^{q-2}(-\partial_t u+2\Delta u)\rangle \\ &= B_q - \langle|\partial_t u|^2,|w|^{q-2}\rangle + 2\langle\partial_t u, |w|^{q-2}\Delta u\rangle \\ &= B_q - \langle|\partial_t u|^2,|w|^{q-2}\rangle - \frac{2}{q}\frac{d}{dt}\|w\|_q^q - 2(q-2)\langle\partial_t u, |w|^{q-3}w\cdot\nabla|w|\rangle \\ &\leq B_q - \langle|\partial_t u|^2,|w|^{q-2}\rangle - \frac{2}{q}\frac{d}{dt}\|w\|_q^q + (q-2)^2J_q + \langle|\partial_t u|^2,|w|^{q-2}\rangle \\ &= B_q - \frac{2}{q}\partial_t\|w\|_q^q + (q-2)^2J_q.\end{aligned}$$

This estimate on $\langle|\Delta u|^2,|w|^{q-2}\rangle$, which is more efficient than (9.18), when applied in (9.17) leads, together with (9.15), (9.16), to the following. *If form-bound δ satisfies*

$$\begin{aligned} &d\geq 5 \quad \sqrt{\delta} \text{ satisfies } \tfrac{d\sqrt{\delta}}{2}\Big(\sqrt{\tfrac{d^2\delta}{4}+(d-2)^2}+d-2\Big)<d-1,\\ &d=4 \quad \sqrt{\delta}<\tfrac{2(\sqrt{3}-1)}{d}\approx 0.36602,\\ &d=3 \quad \sqrt{\delta}<\tfrac{2\sqrt{2}-1}{d}\approx 0.60947,\end{aligned} \tag{C_5}$$

then gradient estimate (9.12) *holds for a $q>d$ close to d.* See [KiS9] for the proof.

The assumption (C_5) is less restrictive than (C_4). In fact, if one assumes $\sqrt{\delta}=\frac{1}{d}$, then (9.12) holds even for $q=d+1$.

Remark 9.2. The gradient bounds in [KS, Ki3, KiS9] are proved not only for q close to d, but for the entire range of (δ,q) satisfying some algebraic inequalities. In particular, in [KiS9],

$$q-1-\frac{q\sqrt{\delta}}{2}\Big(\sqrt{\frac{q^2\delta}{4}+(q-2)^2}+q-2\Big)>0 \quad \text{if } d\geq 5$$

and, as was mentioned above, for $d=3,4$, $\sqrt{\delta}\leq\big(\sqrt{q-1}-\frac{q-2}{2}\big)\frac{2}{q}$.

Thus, we have the following result.

Theorem 9.1 ([Ki3], [KiS9])**.** *Assume that $b\in L^\infty\mathbf{F}_\delta+L^2_{\rm loc}(\mathbb{R}_+)$ with δ that satisfying (C_4) or, better, (C_5). Then the following is true:*

(i) *The limit*

$$U^{t,s}f := s\text{-}C_\infty\text{-}\lim_n U_n^{t,s}f \quad \textit{uniformly in } (s,t)\in D_T$$

exists for all $f\in C_c^\infty$ and satisfies $\|U^{t,s}f\|_\infty\leq\|f\|_\infty$. Upon extending operators $U^{t,s}$ by continuity to all $f\in C_\infty$, one obtains a Feller evolution family.

(ii) *The Feller evolution family $\{U^{t,s}\}_{(s,t)\in D_T}$ is unique in the sense that it does not depend on the choice of the approximation vector fields $\{b_n\}$, as long as they satisfy* (9.2), (9.3), (9.4).

(iii) *If $f\in C_\infty\cap L^2$, then $u(t):=U^{t,s}f$ is the unique weak solution of $(\partial_t-\Delta+b\cdot\nabla)u=0$, $u(s)=f$, in L^2.*

10. SDEs with time-inhomogeneous form-bounded drifts

We return to the discussion of weak well-posedness of SDE

$$X_t = x - \int_0^b b(s, X_s)ds + \sqrt{2}W_t, \quad t \geq 0, \tag{10.1}$$

where $x \in \mathbb{R}^d$ is fixed and $b \in L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$, where, we assume for simplicity, (C_4) holds.

We need to supplement Theorem 9.1 with a localized analogue of (9.12) for inhomogeneous parabolic equations, proved in [KiM1]. Let $\mathsf{f} \in L^\infty \mathbf{F}_\nu + L^2_{\rm loc}(\mathbb{R}_+)$, $\nu < \infty$, define f_k similarly to b_m in (9.2)-(9.4). Let $h \in C([s,T], \mathcal{S})$, $g \in C_c^\infty(\mathbb{R}^d)$. Fix $T > s$. Let $u = u_{m,k}$ be the solution to Cauchy problem on $[s,T]$

$$(\partial_t - \Delta + b_m \cdot \nabla)u = |\mathsf{f}_k| h, \quad u(s, \cdot) = g. \tag{10.2}$$

Then, for every $q \in]d, \delta^{-\frac{1}{2}}[$, there exist constants C and κ such that, for all $0 \leq s \leq r \leq T$,

$$\begin{aligned} \|u\|^q_{L^\infty([s,r], L^q_\rho)} + \|\nabla u\|^q_{L^\infty([s,r], L^q_\rho)} + \|\nabla |\nabla u|^{\frac{q}{2}}\|^2_{L^2([s,r], L^2_\rho)} \\ \leq C\big(\|\mathsf{f}|h|^{\frac{q}{2}}\|^2_{L^2([s,r], L^2_\rho)} + \|\nabla g\|^q_{L^q_\rho} + \|g\|^q_{L^q_\rho}\big). \end{aligned} \tag{10.3}$$

Here $\rho(x) := (1 + \kappa |x|^2)^{-\theta}$ ($x \in \mathbb{R}^d$), where $\theta > \frac{d}{2}$ is fixed, and $L^2_\rho := L^2(\mathbb{R}^d, \rho dx)$.

Define backward Feller evolution family ($0 \leq t \leq r \leq T$)

$$P^{t,r}(b) = U^{T-t, T-r}(\tilde{b}), \quad \tilde{b}(t,x) = b(T-t, x),$$

where $U^{t,s}$ is the Feller evolution family from Theorem 9.1. Using (10.3) with $\mathsf{h} = 0$ and arguing essentially as in the proof of Theorem 6.1(v), one obtains that $\{P^{t,r}(b)\}_{0 \leq t \leq r \leq T}$ is conservative, i.e. for all $x \in \mathbb{R}^d$ $\langle P^{t,r}(x, \cdot)\rangle = 1$. Now, by a standard result (see e.g. [GC, Ch. 2]), given a conservative backward Feller evolution family, there exist probability measures $\mathbb{P}_x$ ($x \in \mathbb{R}^d$) on $(D([0,T], \mathbb{R}^d), \mathcal{B}'_t)$, such that

$$\mathbb{E}_x[f(\omega_r)] = P^{0,r} f(x), \quad 0 \leq r \leq T.$$

Here and below, $\mathbb{E}_x := \mathbb{E}_{\mathbb{P}_x}$.

Let X^m_t ($m = 1, 2, \dots$) be the strong solution of

$$X^m_t = x - \int_0^t b_m(r, X^m_r)dr + \sqrt{2}W_t, \quad x \in \mathbb{R}^d$$

defined on some complete probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$.

We will require the following estimate: there exists a constant $C > 0$ independent of m, k such that

$$\sup_m \sup_{x \in \mathbb{R}^d} \mathbf{E} \int_s^r |b_k(t, X^m_t)| dt \leq C F(r-s) \tag{10.4}$$

for $0 \leq s \leq r \leq T$, where $F(h) := h + \sup_{s \in [0, T-h]} \int_s^{s+h} g(t)dt$. Here we assume, without loss of generality, that $b_k \in L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$ with the same function g as b (if not, then we can increase g, cf. (9.4)).

Indeed, let $v = v_{m,k}$ be the solution to the terminal-value problem

$$(\partial_t + \Delta - b_m \cdot \nabla)v = -|b_k|, \quad v(r, \cdot) = 0, \quad t \leq r. \tag{10.5}$$

By Itô's formula,

$$v(r,X_r^m)=v(s,X_s^m)+\int_s^r(\partial_t v+\Delta v-b_m\cdot\nabla v)(t,X_t^m)dt+\sqrt{2}\int_s^r\nabla v(t,X_t^m)dW_t.$$

Taking expectation, we obtain

$$\mathbb{E}\int_s^r|b_k(t,X_t^m)|dt=\mathbb{E}v(s,X_s^m).$$

Now, (10.3) applied to equation (10.5) (so, we reverse the direction of time and take $|\mathsf{f}|:=|b_k|$, $h\equiv 1$ and $g=0$) yields, upon applying the Sobolev embedding theorem,

$$\|v\|_{L^\infty([s,r]\times B_1(0))}\leq C_1\|b_k\sqrt{\rho}\|^2_{L^2([s,r],L^2)},$$

where 0 is the "centre" of the weight ρ. Thus, considering its translates $\rho_z:=\rho(x-z)$, we obtain

$$\|v\|_{L^\infty([s,r]\times\mathbb{R}^d)}\leq C_2\sup_{z\in\mathbb{Z}^d}\|b_k\sqrt{\rho_z}\|_{L^2([s,r],L^2)}.$$

Since $\mathbb{E}v(s,X_s^m)\leq\|v(s,\cdot)\|_\infty$, we obtain

$$\mathbb{E}\int_s^r|b_k(t,X_t^m)|dt\leq C_2\sup_{z\in\mathbb{Z}^d}\|b_k\sqrt{\rho_z}\|_{L^2([s,r],L^2)}.$$

Since $b_k\in L^\infty\mathbf{F}_\delta+L^2_{\rm loc}$, we have

$$\begin{aligned}\|b_k\sqrt{\rho_z}\|^2_{L^2([s,r],L^2)}&\leq\frac{\delta}{4}\int_s^r\langle\frac{|\nabla\rho_z|^2}{\rho_z}\rangle dt+\int_s^r g(t)\langle\rho_z\rangle dt\\&\quad\text{(we are using }|\nabla\rho|\leq\theta\sqrt{\kappa}\rho\text{ and }\|\sqrt{\rho}\|_2<\infty)\\&\leq CF(r-s)\end{aligned}$$

for $0\leq s\leq r\leq T$, so (10.4) follows.

Now, define probability measures $\mathbb{P}_x^n:=(\mathbf{P}\circ X^n)^{-1}$ on $\big(C([0,T],\mathbb{R}^d),\mathcal{B}_t\big)$, so (10.4) takes form

$$\sup_m\sup_{x\in\mathbb{R}^d}\mathbb{E}_{\mathbb{P}_x^m}\int_s^r|b_k(t,\omega_t)|dt\leq CF(r-s),$$

where ω_t is the coordinate process. We apply in (10.4) the convergence result of Theorem 9.1 (in m) and then Fatou's lemma (in k) to obtain

$$\mathbb{E}\int_s^r|b(t,\omega_t)|dt\leq CF(r-s)<\infty$$

(which is one of the requirements in the definition of a martingale solution). Arguing similarly, we obtain, for every $f\in C_c^2$,

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_x^m}\bigg|\int_0^r((b_m-b_n)\cdot\nabla f)(t,\omega_t)dr\bigg|&\leq C\|(b_m-b_n)|\nabla f|^{\frac{q}{2}}\|_{L^2([0,r],L^2)}\\&\to 0\quad(m,n\to\infty)\end{aligned}$$

since $b_m\to b$ in $L^2_{\rm loc}(\mathbb{R}_+\times\mathbb{R}^d)$ and f has compact support. This, and the convergence result of Theorem 9.1, allow to pass to the limit in the martingale problem for b_m in essentially the same way as in the proof of Theorem 6.2 to show that $\mathbb{P}_x$ is a martingale solution of (10.1) but on $(D([0,T],\mathbb{R}^d),\mathcal{B}'_t)$. The latter allows to prove, arguing again as in the proof of Theorem 6.2, that $\mathbb{P}_x$ are actually concentrated on continuous trajectories. We arrive at the following result.

Theorem 10.1. *Under the assumptions of Theorem 9.1, let us also assume*[12] (C_4). *The following is true:*

(i) *For every $x \in \mathbb{R}^d$, the probability measure $\mathbb{P}_x$ is a weak solution to SDE* (10.1).

(ii) $\mathbb{P}_x$ *satisfies, for all $\mathsf{f} \in L^\infty \mathbf{F}_\nu + L^2_{\rm loc}(\mathbb{R}_+)$, $\nu < \infty$, $h \in C([0,T],\mathcal{S})$, for all $q \in]d, \delta^{-\frac{1}{2}}[$, the estimate*

$$\mathbb{E}_{\mathbb{P}_x} \int_0^T |\mathsf{f}(r,\omega_r)h(r,\omega_r)|dr \leq c\|\mathsf{f}|h|^{\frac{q}{2}}\|^{\frac{2}{q}}_{L^2([0,T]\times\mathbb{R}^d)}. \tag{10.6}$$

On the other hand, if, for some $x \in \mathbb{R}^d$, $\mathbb{P}'_x$ is a martingale solution of (10.1) *that satisfies* (10.6) *for some $q \in]d, \delta^{-\frac{1}{2}}[$ with $\mathsf{f} = b$, then it coincides with $\mathbb{P}_x$.*

(iii) $\mathbb{P}_x$ *satisfies, for a given $\nu > \frac{d+2}{2}$, for all $h \in C([0,T],\mathcal{S})$ the following Krylov-type bound:*

$$\mathbb{E}_{\mathbb{P}_x} \int_0^T |h(r,\omega_r)|dr \leq c\|h\|_{L^\nu([0,T]\times\mathbb{R}^d)}. \tag{10.7}$$

On the other hand, if additionally $|b| \in L^{\frac{d+2}{2}+\varepsilon}_{\rm loc}(\mathbb{R}_+ \times \mathbb{R}^d)$ for some $\varepsilon > 0$ and δ is sufficiently small, then any martingale solution $\mathbb{P}'_x$ of (10.1) *that satisfies* (10.7) *for some $\nu > \frac{d+2}{2}$ sufficiently close to $\frac{d+2}{2}$ (depending on how small ε is) coincides with $\mathbb{P}_x$.*

The first two assertions of Theorem 10.1 were proved in [KiM1], the last assertion will be proved in the next section, in fact, for a substantially larger than $L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$ class of drifts.

Remark 10.1. One advantage of the uniqueness class in (ii), i.e.

$$\mathbb{E}_{\mathbb{P}_x} \int_0^T |b(r,\omega_r)h(r,\omega_r)|dr \leq c\|b|h|^{\frac{q}{2}}\|^{\frac{2}{q}}_{L^2([0,T]\times\mathbb{R}^d)} \tag{10.8}$$

for some $q \in]d, \delta^{-\frac{1}{2}}[$ is that it senses the value of δ. Namely, as δ becomes smaller, one can take q larger, and so the verification of (10.8), in principle, becomes easier (e.g. $|b|$ is bounded, then δ can be arbitrarily small and hence q can be arbitrarily large).

Regarding the proof of Theorem 10.1(i), let us note that we can alternatively use the tightness argument, also employed in the proof of Theorem 4.1, and then apply the convergence result of Theorem 9.1. See details in [KiM1]. The uniqueness results in assertions (ii), (iii), however, require gradient bounds (10.3).

11. "Form-bounded" diffusion coefficients

1. The results of the previous two sections can be extended to Itô and Stratonovich SDEs

$$X_t = x - \int_0^t b(s,X_s)ds + \sqrt{2}\int_0^t \sigma(s,X_s)dW_s, \quad \in \mathbb{R}^d \tag{11.1}$$

$$X_t = x - \int_0^t b(s,X_s)ds + \sqrt{2}\int_0^t \sigma(s,X_s)\circ dW_s, \tag{11.2}$$

[12] We can assume (C_5), but then we need to adjust interval $q \in]d, \delta^{-\frac{1}{2}}[$. For simplicitiy, we will not do this here.

where the drift $b : \mathbb{R}^d \to \mathbb{R}^d$ is form-bounded and the diffusion coefficient $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are bounded, uniformly elliptic and can be discontinuous. For time-homogeneous b and σ such extension was carried out in [KiS3]. Namely, let $a := \sigma\sigma^\top$ satisfy,

$$\sigma I \leq a \leq \xi I \quad \text{a.e. on } \mathbb{R}^d \tag{$H_{\sigma,\xi}$}$$

for some $0 < \sigma \leq \xi < \infty$, and assume that the entries a_{ij} of a have form-bounded derivatives, that is,

$$(\nabla_r a_{ij})_{i=1}^d \in \mathbf{F}_{\delta_{rj}} \tag{11.3}$$

for some $\delta_{rj} > 0$. Equivalently, since the entries of σ are bounded, we can replace (11.3) with $(\nabla_r \sigma_{ij})_{i=1}^d \in \mathbf{F}_{\delta'_{rj}}$ for appropriate δ'_{rj}.

Examples. 1. If $a \in W^{1,d}(\mathbb{R}^d, \mathbb{R}^{d\times d})$, then (11.3) holds with δ_{rj} that can be chosen arbitrarily small (Appendix B). More generally, if the derivatives of a_{ij} are in the Morrey class $M_{2+\varepsilon}$, then (11.3) holds.

2. Here is a concrete example of matrix a satisfying (11.3) and having a critical discontinuity at the origin:

$$a(x) = I + c\frac{x\otimes x}{|x|^2}, \qquad \text{the constant } c > -1. \tag{11.4}$$

Indeed, $\nabla_r a_{ij} = c\mathbf{1}_{r=i}\frac{x_j}{|x|^2} + c\mathbf{1}_{r=j}\frac{x_i}{|x|^2} + cx_ix_j\frac{2x_r}{|x|^4}$, so

$$|(\nabla_r a_{ij})_{i=1}^d| \leq 2|c||x|^{-1} \quad \Rightarrow \quad (\nabla_r a_{ij})_{i=1}^d \in \mathbf{F}_{\delta_{rj}}, \quad \delta_{rj} = (4c)^2/(d-2)^2$$

by the Hardy inequality. Another example is

$$a(x) = I + c(\sin\log(|x|))^2 e\otimes e, \quad e \in \mathbb{R}^d, |e| = 1$$

(indeed, $\nabla_r a_{ij} = 2c(\sin\log|x|)(\cos\log|x|)|x|^{-2}x_r\, e_ie_j$, so using that the Hardy vector field (3.7) is form-bounded one obtains the required).

More generally, (11.3) holds for a that is an infinite sum of such matrices (properly normalized so that the series converges) with their points of discontinuity constituting e.g. a dense subset of $\mathbb{R}^d$.

Without loss of generality, in $(H_{\sigma,\xi})$ $\sigma = 1$. In [KiS3], assuming that $b \in \mathbf{F}_\delta$ and

$$a \text{ satisfies } (H_{\sigma,\xi}) \text{ and } (11.3), \quad \nabla a \in \mathbf{F}_{\delta_a}, \tag{11.5}$$

where $(\nabla a)_k := \sum_{i=1}^d (\nabla_i a_{ik})$, with δ, δ_a and δ_{rn} satisfying, for some $q > 2 \vee (d-2)$,

$$1 - \frac{q}{4}(\sqrt{\gamma} + \|a - I\|_\infty\sqrt{\delta+\delta_a}) > 0, \tag{11.6}$$

where $\gamma := \sum_{r,n=1}^d \delta_{rn}$, and

$$\begin{aligned}(q-1)(1 - \frac{q\sqrt{\gamma}}{2}) &- (\sqrt{\delta+\delta_a}\sqrt{\delta_a} + \delta + \delta_a)\frac{q^2}{4} \\ &- (q-2)\frac{q\sqrt{\delta+\delta_a}}{2} - \|a-I\|_\infty\frac{q\sqrt{\delta+\delta_a}}{2} > 0,\end{aligned} \tag{11.7}$$

the authors constructed a Feller semigroup and proved an analogue of Theorem 6.2(i), including the "approximation uniqueness" result in Remark 6.3, for the Itô SDE

$$X_t = x - \int_0^t b(X_s)ds + \sqrt{2}\int_0^t \sigma(X_s)dW_s. \tag{11.8}$$

The result for the Stratonovich SDE

$$X_t = x - \int_0^t b(X_s)ds + \sqrt{2}\int_0^t \sigma(X_s) \circ dW_s, \tag{11.9}$$

in [KiS3] is valid under assumption (11.6), (11.7) but with δ replaced by $\delta+\delta_a+\delta_c$, where δ_c is the form-bound of

$$c := (c^i)_{i=1}^d, \quad \text{where } c^i := \frac{1}{\sqrt{2}}\sum_{r,j=1}^d (\nabla_r\sigma_{ij})\sigma_{rj}.$$

The assumptions (11.6), (11.7) imply that δ, δ_a and δ_{rj} cannot be too large. It is also easily seen that if $a = I$, then these assumptions reduce to $\delta < 1 \wedge (\frac{d}{d-2})^2$, i.e. then there exists $q > 2 \vee (d-2)$ such that (11.6), (11.7) hold.

The assumptions on diffusion coefficients σ of the form (11.3) go back to Veretennikov [V] who proved strong well-posedness of (11.1) for bounded measurable b and $\nabla_r\sigma_{ij} \in L^{2d}_{\rm loc}$. There are many other papers that consider assumptions of this type, see e.g. [Z3] and [Kr2, Kr3] who considered $\nabla_r\sigma_{ij} \in L^p_{\rm loc}$ $(p > d)$ and $\nabla_r\sigma_{ij} \in L^d_{\rm loc}$.

The construction of the Feller semigroup in [KiS3] is based on an extension of the iteration procedure described in Section 9 (in the elliptic setting) to solutions u_n of divergence-form equations

$$(\mu + \Lambda(a_n, b_n))u_n = f, \quad f \in C_c^\infty, \quad \mu \geq \mu_0, \tag{11.10}$$

where $\Lambda(a_n, b_n) = -\nabla \cdot a_n \cdot \nabla + b_n \cdot \nabla$, and uses the gradient bound

$$\|\nabla u_n\|_{\frac{qd}{d-2}} \leq K\|f\|_q, \quad K \text{ is independent of } n. \tag{11.11}$$

Here a_n, b_n are bounded and smooth, $\{b_n\}$ is uniformly (in n) form-bounded, and $\{a_n\}$ satisfy the same assumptions as a above (thus, with constants independent on n). This iteration procedure and (11.11) yield Feller semigroup $e^{-t\Lambda_{C_\infty}(a,b)}$. Then $e^{-t\Lambda_{C_\infty}(a,\nabla a+b)}$ is the sought Feller semigroup that produces weak solution to Itô SDE (11.8), where we used the identity

$$-a_n \cdot \nabla^2 + b_n \cdot \nabla = -\nabla \cdot a_n \cdot \nabla + (\nabla a_n + b_n) \cdot \nabla. \tag{11.12}$$

For Stratonovich SDE (11.9) one needs Feller semigroup $e^{-t\Lambda_{C_\infty}(a,\nabla a+b-c)}$.

Remark 11.1. For instance, the approximating vector fields b_n can be defined via (6.2), so they are uniformly (in n) in $\mathbf{F}_\delta$. The approximating matrices a_n can be defined via

$$a_n = \eta_{\varepsilon_n} * a,$$

where η_{ε_n} is the Friedrichs mollifier, and $\varepsilon_n \downarrow 0$. To see that a_n are such that ∇a_n are indeed uniformly in $\mathbf{F}_{\delta_a}$ and satisfy (11.3) with the same δ_{rn}, also uniformly in n, one can apply the result of Appendix B.3. In [KiS3], there was an additional cutoff function under the mollifier, which is not necessary.

Let us add that already the task of proving the uniform in n gradient bound (11.11) for solutions u_n to the divergence-form equation (11.10) (which was the original interest of the authors of [KiS3]) leads to condition (11.3).

Theorem 9.1 and the existence part of Theorem 10.1 can be extended to SDEs (11.1) and (11.2) with time-inhomogeneous $b \in L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$ and time-inhomogeneous bounded σ such that $a = \sigma\sigma^\top$ is uniformly elliptic and satisfies

$$\nabla a \in L^\infty \mathbf{F}_{\delta_a} + L^2_{\rm loc}(\mathbb{R}_+), \quad (\nabla_r a_{ij})_{i=1}^d \in L^\infty \mathbf{F}_{\delta_{rj}} + L^2_{\rm loc}(\mathbb{R}_+)$$

for all $1 \leq r, j \leq d$, for appropriate δ_a and δ_{rj}. A direct extension of the weak uniqueness part of Theorem 10.1 to a_{ij} as above is problematic: one has to control second-order derivatives of u_n, but these are destroyed by form-bounded b. Thus, one needs extra assumptions both on b and a, such as Morrey class in [Kr5], see below.

2. Below we assume, for simplicity, that a, b are time-homogeneous, although most of the results cited below are valid for time-inhomogeneous coefficients.

The condition (11.3) puts a in the class VMO[13], see [Kr3]. Recall that matrix a is in VMO if

$$\sup_{B_r} \frac{1}{|B_r|} \int_{B_r} |a - (a)_{B_r}| dx \to 0 \quad \text{ as } \rho \downarrow 0,$$

where the supremum is taken over all balls of radius $\leq \rho$. Here $(a)_{B_r}$ denotes the average of a on B_r.

There is a very rich literature on well-posedness of parabolic equations and SDEs with VMO diffusion matrix and singular drift b satisfying more restrictive assumptions than the form-boundedness. The strongest result on weak well-posedness of SDE (11.1) with VMO diffusion coefficients is a very recent result of Krylov [Kr5] who proved that there exist positive constants θ (sufficiently small) and ρ_a such that if a in the BMO class with norm $\leq \theta$, i.e.

$$\sup_{B_r,\, r \leq \rho} \frac{1}{|B_r|} \int_{B_r} |a - (a)_{B_r}| dx \leq \theta \quad \text{ for all } \rho \leq \rho_a, \tag{11.13}$$

(e.g. if $a \in$ VMO) and

$$|b| \in M_{\frac{d}{2}+\varepsilon}, \quad \varepsilon > 0, \tag{11.14}$$

with sufficiently small norm, then (11.1) is weakly well-posed, i.e. the solution exists and is unique in a class similar to the one in Theorem 10.1.

In the case of constant diffusion coefficients σ one can prove weak well-posedness of (11.1) for substantially larger than $\mathbf{F}_\delta$ class of *weakly form-bounded drifts*, discussed in Sections 14-15. This class contains e.g.

$$|b| \in M_{1+\varepsilon} \tag{11.15}$$

with arbitrarily small $\varepsilon > 0$.

If we were to exploit the relationship between non-divergence and divergence form operators in the case the matrix a is sufficiently discontinuous, cf. (11.12), then the sesquilinear form of the divergence-form operator will have to satisfy

$$|\langle a \cdot \nabla \varphi, \nabla \psi \rangle + \langle b \cdot \nabla \varphi, \psi \rangle| \leq C \|\varphi\|_{W^{1,2}} \|\psi\|_{W^{1,2}}, \quad \varphi, \psi \in W^{1,2}, \tag{11.16}$$

which, by [MV], would make $\nabla a + b$ form-bounded (modulo a divergence-free component in BMO^{-1}, see the discussion around (3.11)). See also Remarks 14.1 and 15.6 below.

[13] In [KiS3] there is an incorrect statement that there are matrices a satisfying (11.3) and not contained in the VMO class.

12. Stochastic transport equation and strong solutions to SDEs

In [BFGM], Beck-Flandoli-Gubinelli-Maurelli presented, among many results, an approach to proving strong well-posedness of SDE

$$X_t = x - \int_0^t b(r, X_r)dr + \sigma W_t, \tag{12.1}$$

for a.e. $x \in \mathbb{R}^d$, where $\{W_t\}_{t\geq 0}$ is a d-dimensional Brownian motion in $\mathbb{R}^d$ defined on a complete filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbf{P})$, with b in the critical Ladyzhenskaya-Prodi-Serrin class. Below we will discuss the time-homogeneous case $b : \mathbb{R}^d \to \mathbb{R}^d$, so their assumption on b reads as $|b| \in L^d + L^\infty$. Their approach is based on a detailed regularity theory of the stochastic transport equation (STE)

$$\begin{gathered} du + b \cdot \nabla u dt + \sigma \nabla u \circ dW_t = 0 \quad \text{ on } (0,\infty) \times \mathbb{R}^d, \\ u|_{t=0} = f, \end{gathered} \tag{12.2}$$

where $u(t, x)$ is a scalar random field, $\sigma \neq 0$ is a constant, f is in L^p or $W^{1,p}$, $\circ$ is the Statonovich multiplication.

Speaking of the STE (12.2), let us mention that the Cauchy problem for the deterministic transport equation $\partial_t u + b \cdot \nabla u = 0$ is in general not well posed already for a bounded but discontinuous b. Moreover, in that case, even if the initial function f is regular, one cannot hope that the corresponding solution u will be regular immediately after $t = 0$. This, however, changes if one adds the noise term $\sigma \nabla u \circ dW_t$, $\sigma > 0$. For the stochastic STE (12.2), a unique weak solution exists and is regular for some discontinuous b. This effect of regularization and well-posedness by noise, demonstrated by the STE, attracted considerable interest in the past few years, as a part of the more general program of establishing well-posedness by noise for SPDEs whose deterministic counterparts arising in fluid dynamics are not well-posed, see Flandoli-Gubinelli-Priola [FGP], Gess-Maurielli [GM] for detailed discussions and references.

Let us make a few preliminary remarks regarding STE (12.2) in the case the drift is smooth.

1. Let $b \in C_c^\infty(\mathbb{R}^d, \mathbb{R}^d)$ and $f \in C_c^\infty$. Then there exists (see [Ku, Theorem 6.1.9]) a unique adapted strong solution to

$$u(t) - f + \int_0^t b \cdot \nabla u ds + \sigma \int_0^t \nabla u \circ dW_s = 0 \text{ a.s.}, \quad t \in [0, T], \tag{12.3}$$

given by

$$u(t) := f(\Psi_t^{-1}), \quad t \geqslant 0, \tag{12.4}$$

where $\Psi_t : \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is the stochastic flow for the SDE (12.1). The latter means that there exists $\Omega_0 \subset \Omega$, $\mathbb{P}(\Omega_0) = 1$, such that, for all $\omega \in \Omega_0$,

$$\Psi_t(\cdot, \omega)\Psi_s(\cdot, \omega) = \Psi_{t+s}(\cdot, \omega), \quad \Psi_0(x, \omega) = x,$$

for every $x \in \mathbb{R}^d$, the process $t \mapsto \Psi_t(x, \omega)$ is a strong solution to (12.1), and $\Psi_t(x, \omega)$ is continuous in (t, x), $\Psi_t(\cdot, \omega) : \mathbb{R}^d \to \mathbb{R}^d$ are homeomorphisms and $\Psi_t(\cdot, \omega)$, $\Psi_t^{-1}(\cdot, \omega) \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$.

2. Applying Itô's formula, one easily obtain that for every $\mu \geq 0$,

$$u(t) = e^{\mu t} f(\Psi_t^{-1}), \quad t \geqslant 0$$

solves

$$u(t) - f + \mu \int_0^t u ds + \int_0^t b \cdot \nabla u ds + \sigma \int_0^t \nabla u \circ dW_s = 0 \text{ a.s.}, \quad t \in [0,T]. \tag{12.5}$$

Thus, solutions of the Cauchy problems (12.3) and (12.5) differ by the factor $e^{-\mu t}$.

3. One can rewrite the equation in (12.5), using the identity relating Stratonovich and Itô integrals

$$\int_0^t \nabla u \circ dW_s = \int_0^t \nabla u dW_s - \frac{1}{2}\sum_{k=1}^d [\partial_{x_k} u, W^k]_t, \qquad W_t = (W_t^k)_{k=1}^d, \tag{12.6}$$

as

$$du + \mu u dt + b \cdot \nabla u dt + \sigma \nabla u dW_t - \frac{\sigma^2}{2}\Delta u dt = 0 \tag{12.7}$$

(the Itô form of the STE). Now, taking expectation, one obtains that $v := \mathbf{E}[u]$ solves Cauchy problem for the deterministic parabolic equation

$$\partial_t v + \mu v + b \cdot \nabla v - \frac{\sigma^2}{2}\Delta v = 0, \quad v|_{t=0} = f.$$

Let now b be discontinuous. The authors of [BFGM], in a sense, reversed (12.4), i.e. given a $|b| \in L^d$ (in the time-homogeneous case) they used their Sobolev regularity theory of (12.2) to prove strong well-posedness of SDE (12.1) for a.e. initial point $x \in \mathbb{R}^d$.

In [KSS], the authors extended the approach of [BFGM] to (time-homogeneous) form-bounded drifts b. We describe these results below.

Set

$$\rho(x) \equiv \rho_{\kappa,\theta}(x) := (1+\kappa|x|^2)^{-\theta}, \quad \kappa > 0, \quad \theta > \frac{d}{2}, \quad x \in \mathbb{R}^d.$$

Let $L^p_\rho \equiv L^p(\mathbb{R}^d, \rho dx)$. Denote by $\|\cdot\|_{p,\rho}$ the norm in L^p_ρ, and by $\langle \cdot,\cdot\rangle_\rho$ the inner product in L^2_ρ. Set

$$W^{1,2}_\rho := \{g \in W^{1,2}_{\rm loc} \mid \|g\|_{W^{1,2}_\rho} := \|g\|_{2,\rho} + \|\nabla g\|_{2,\rho} < \infty\}.$$

Fix $T > 0$ and put

$$\beta_{2q} := 1 + 4qd, \quad q = 1, 2, \dots$$

Theorem 12.1. *Let $b \in \mathbf{F}_\delta$ with $\sqrt{\delta} < \frac{\sigma^2}{2\beta_2}$. Let $p \geq 2$. Provided that κ is chosen sufficiently small, there are generic constants $\mu_1 \geq 0$, $C_1 > 0, C_2 > 0$ (i.e. they depend only on δ, c_δ, p and T) such that for any $\mu \geq \mu_1$, for every $f \in L^{2p}$ there exists a function $u \in L^\infty([0,T], L^2(\Omega, L^2_\rho))$ for which the following are true:*

(i) For a.e. $\omega \in \Omega$,

$$\nabla \int_0^T u(s,\cdot,\omega) ds \in L^2_{\rm loc}(\mathbb{R}^d, \mathbb{R}^d),$$

so

$$b \cdot \nabla \int_0^T u(s,\cdot,\omega) ds \in L^1_{\rm loc},$$

and for every test function $\varphi \in C_c^\infty$, we have a.s. for all $t \in [0,T]$,

$$
\begin{aligned}
&\langle u(t), \varphi\rangle - \langle f, \varphi\rangle \\
&+ \mu\langle \int_0^t uds, \varphi\rangle + \langle b\cdot\nabla \int_0^t uds, \varphi\rangle - \sigma\langle \int_0^t udW_s, \nabla\varphi\rangle + \frac{\sigma^2}{2}\langle \nabla \int_0^t uds, \nabla\varphi\rangle = 0. \qquad (12.8)
\end{aligned}
$$

(ii) For any sequence of smooth vector fields $b_m \in C_c^\infty(\mathbb{R}^d, \mathbb{R}^d)$, $m = 1, 2, \dots$, that are uniformly form-bounded in the sense that $b_m \in \mathbf{F}_\delta$ with c_δ independent of m, and are such that

$$b_m \to b \text{ in } L^2_{\rm loc}(\mathbb{R}^d, \mathbb{R}^d) \text{ as } m \to \infty,$$

we have for any initial function $f \in C_c^\infty$,

$$u_m(t) \to u(t) \quad \text{in } L^2(\Omega, L^2_\rho) \quad \text{uniformly in } t \in [0,T],$$

where u_m is the unique strong solution to (12.7) (with $b = b_m$) with initial condition $u_m|_{t=0} = f$.

The last property implies that u does not depend on the choice of the approximating sequence $\{b_m\}$ as long as it preserves the class of form-bounded vector fields. This can be viewed as a uniqueness result on its own.

The next theorem establishes Sobolev regularity of u up to the initial time $t = 0$.

Theorem 12.2. *Let $b \in \mathbf{F}_\delta$ with $\sqrt{\delta} < \frac{\sigma^2}{2\beta_2}$ and $f \in W^{1,4}$. Let κ be sufficiently small and μ_1 be the constant in Theorem 12.1 with $p = 2$. For $\mu \geq \mu_1$, let u be the process constructed in Theorem 12.1. There exists generic constant $\mu_2 \geq \mu_1$ such that for $\mu \geq \mu_2$, the following are true:*

(a) $\mathbf{E}u^2$, $\mathbf{E}|\nabla u|^2 \in L^\infty([0,T], L^2)$, *so* $u \in L^\infty([0,T], L^2(\Omega, W^{1,2}_\rho))$.

(b) *For any test function $\varphi \in C_c^\infty$, the process $t \mapsto \langle u(t), \varphi\rangle$ is $(\mathcal{F}_t)$-progressively measurable and has a continuous $(\mathcal{F}_t)$-semi-martingale modification that satisfies a.s. for every $t \in [0,T]$,*

$$
\begin{aligned}
&\langle u(t), \varphi\rangle - \langle f, \varphi\rangle \\
&+ \mu \int_0^t \langle u, \varphi\rangle ds + \int_0^t \langle b\cdot\nabla u, \varphi\rangle ds - \sigma \int_0^t \langle u, \nabla\varphi\rangle dW_s + \frac{\sigma^2}{2}\int_0^t \langle u, \Delta\varphi\rangle ds = 0. \qquad (12.9)
\end{aligned}
$$

Moreover, if $\sqrt{\delta} < \frac{\sigma^2}{2\beta_{2q}}$ for some $q = 1, 2, \dots$, then there exist generic constants $\mu_2(q) \geq \mu_1$ (with $\mu_2(1)$ equal to the μ_2 above) and $C_1 > 0$ such that when $\mu \geq \mu_2(q)$ and $f \in W^{1,4q}$, we have

$$\sup_{0\leq\alpha\leq 1} \big\|\mathbf{E}|\nabla u|^{2q}\big\|_{L^{\frac{2}{1-\alpha}}([0,T], L^{\frac{2d}{d-2+2\alpha}})} \leq C_1\|\nabla f\|_{4q}^{2q}. \qquad (12.10)$$

In particular, there exists generic $C_2 > 0$ such that

$$\sup_{t\in[0,T]} \mathbf{E}\langle \rho|\nabla u|^{2q}\rangle \leq C_2\|\nabla f\|_{4q}^{2q}.$$

If $2q > d$, then for a.e. $\omega \in \Omega$, $t \in [0,T]$, the function $x \mapsto u(t,x,\omega)$ is Hölder continuous, possibly after modification on a set of measure zero in $\mathbb{R}^d$ (in general, depending on ω).

The estimate (12.10) can be viewed as a counterpart of (9.12).

A function satisfying (a), (b) of Theorem 12.2 will be called a weak solution of Cauchy problem

$$
\begin{aligned}
&du + \mu\, udt + b\cdot\nabla udt + \sigma\nabla u \circ dW_t = 0 \quad \text{on } (0,\infty)\times\mathbb{R}^d, \\
&u|_{t=0} = f \in L^p, \quad p \geq 2.
\end{aligned} \qquad (12.11)
$$

This definition of weak solution is close to [BFGM, Definition 2.13].

Theorem 12.3. *Let $b \in \mathbf{F}_\delta$ with $\sqrt{\delta} < \frac{\sigma^2}{2\beta_2}$ and $f \in W^{1,4}$. Provided κ is sufficiently small, there exists generic $\mu_3 \geq 0$ such that for $\mu \geq \mu_3$, the Cauchy problem* (12.11) *has a unique weak solution in the class of functions satisfying* (a), (b) *of Theorem 12.2.*

Theorems 12.1-12.3 were proved in [KSS]. Theorem 12.2 extends a similar result in [BFGM] for (in the time-homogeneous case) $|b| \in L^d$. The proof of the uniqueness result in Theorem 12.3 adopts the method of [BFGM, Sect. 3].

It should be noted that the authors in [BFGM] prove their uniqueness result in a larger class of weak solutions (not requiring any differentiability, see [BFGM, Definition 3.3]) but under additional assumptions on b. Specialized to the time-dependent case, they assume that b satisfies

$$\operatorname{div} b \in L^d + L^\infty \tag{12.12}$$

in addition to $b \in L^d + L^\infty$. The latter is needed to establish (12.10) for solutions of the adjoint equation to the STE, i.e. the stochastic continuity equation (which allows to prove an even stronger result: the uniqueness of weak solution to the corresponding random transport equation), see [BFGM, Sect. 3].

Armed with Theorems 12.1, 12.2, one can repeat the argument in [BFGM, Sect. 4] to prove the following result. *Assuming that $b \in \mathbf{F}_\delta$ with δ sufficiently small, there exists a stochastic Lagrangian flow for SDE* (12.1)*, i.e. a measurable map $\Phi : [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ such that, for a.e. $x \in \mathbb{R}^d$, the process $t \mapsto \Phi_t(x,\omega)$ is a strong solution of the SDE* (12.1)*:*

$$\Phi_t(x,\omega) = x - \int_0^t b(s, \Phi_r(x,\omega))dr + \sigma W_t(\omega), \quad a.s., \quad t \in [0,T],$$

and $\Phi_t(x,\cdot)$ is $\mathcal{F}_t$-progressively measurable. If also $\sqrt{\delta} < \frac{\sigma^2}{2\beta_{2q}}$, $q = 1,2,\dots$, then $\Phi_t(\cdot,\omega) \in W^{1,2q}_{\mathrm{loc}}$ ($t \in [0,T]$) for a.e. $\omega \in \Omega$. Moreover, Φ_t is unique, i.e. any two such stochastic flows coincide a.s. for every $t > 0$ for a.e. x.

The restriction "for a.e. initial point" in the above strong existence result for SDE (12.1) can be removed using a different method discussed in the next section.

13. Strong well-posedness via Röckner-Zhao's approach

In [RZh2], Röckner and Zhao proved strong existence and uniqueness in a large class of strong solutions satisfying Krylov estimate (cf. (13.6) below) for SDE

$$X_t^x = x + \int_0^t b(s, X_s^x)ds + W_t, \quad 0 \leq t \leq T, \tag{13.1}$$

for every initial point $x \in \mathbb{R}^d$, provided that drift $b : \mathbb{R}^{d+1} \to \mathbb{R}^d$ satisfies the critical Ladyzhenskaya-Prodi-Serrin condition

$$b \in L^p(\mathbb{R}, L^q(\mathbb{R}^d)), \quad \frac{d}{q} + \frac{2}{p} \leq 1, \quad p > 2, \quad q \geq d. \tag{13.2}$$

Above $\{W_t\}_{0 \leq t \leq T}$ denotes, as before, a Brownian motion on a complete filtered probability space $(\Omega, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathcal{F}, \mathbf{P})$.

The method of Röckner-Zhao is different from the other methods used in the literature on strong well-posedness of (13.1) (cf. [BFGM] and [Kr1]-[Kr3]). Their proof of strong existence uses a relative compactness criterion for random fields on the Wiener-Sobolev space. Their proof of uniqueness uses their weak uniqueness result from [RZh1] and Cherny's theorem [C] (briefly, strong existence and weak uniqueness $\Rightarrow$ strong uniqueness). The method of [RZh2] is a far-reaching strengthening of the methods of Meyer-Brandis and Proske [MP], Mohammed-Nilsen-Proske [MNP] (for $b \in L^\infty(\mathbb{R} \times \mathbb{R}^d)$) and Rezakhanlou [R] (for b in the sub-critical Ladyzhenskaya-Prodi-Serrin class).

Let us give a brief outline of their method. For a given vector field $Y = (Y_i)_{i=1}^d : \mathbb{R}^k \to \mathbb{R}^m$, denote

$$\nabla Y = \nabla_x Y(x) := \begin{pmatrix} \nabla_1 Y_1 & \nabla_2 Y_1 & \dots & \nabla_k Y_1 \\ & & \dots & \\ \nabla_1 Y_m & \nabla_2 Y_m & \dots & \nabla_k Y_m \end{pmatrix}. \tag{13.3}$$

The proof of strong existence in [RZh2] is based on the following estimates. Let b be additionally bounded and smooth. For every $r \geq 1$, there exist constants K_1, K_2 independent of smoothness or boundedness of b such that

(*i*) $\|\nabla X_t^x - I\|_{L^{2r}(\mathbb{R}^d, L^r(\Omega))} \leq K_1 t^{\frac{1}{2r}}$ for all $0 \leq t \leq T$;

(*ii*) $\|D_s X_t^x - I\|_{L^{2r}(\mathbb{R}^d, L^r(\Omega))} \leq K_1 (t-s)^{\frac{1}{4r}}$ for a.e. $s \in [0,T]$ and $0 \leq s \leq t \leq T$, where $D_s X_t^x$ denotes the Malliavin derivative;

(*iii*) $\|D_s X_t^x - D_{s'} X_t^x\|_{L^{2r}(\mathbb{R}^d, L^r(\Omega))} \leq K_2 |s-s'|^{\frac{1}{4r}}$ for a.e. $s, s' \in [0,T]$ and $0 \leq s, s' \leq t \leq T$,

These estimates allow [RZh2] to apply the relative compactness criterion on the Wiener-Sobolev space in order to construct a strong solution to (13.1).

The first step in the proof of (*i*)-(*iii*) is to differentiate SDE (13.1), e.g. for (*i*)

$$\nabla X_t^x - I = \int_0^t \nabla b(s, X_s^x) \nabla X_s^x ds,$$

with the goal of iterating this identity and, in the end, obtaining an expression for the left-hand side that one can control. This goal is achieved using the following bound (for (*i*)): there exist positive generic constants C_0, K such that, for every $n \geq 1$,

$$\int_{\mathbb{R}^d} \left| \mathbf{E} \int_{\Delta_n(T_0,T_1)} \prod_{i=1}^n \nabla_{\alpha_i} f_i(t_i, X_{t_i}^x) dt_1 \dots dt_n \right|^2 dx \leq C_0 K^n (T_1 - T_0), \tag{13.4}$$

where $1 \leq \alpha_i \leq d$ ($i \geq 1$), $\nabla_i := \partial_{x_i}$, $\Delta_n(T_0, T_1) := \{(t_1, \dots, t_n) \mid T_0 \leq t_1 \leq \dots \leq t_n \leq T_1\}$, and functions f_i are taken to be the components of drift b. In [RZh2], (13.4) is proved using Sobolev regularity estimates for solutions of parabolic equations with distributional right-hand side. These estimates are quite strong and are interesting on their own.

In [KiM3], the authors noticed that (13.4) can be proved for time-inhomogeneous form-bounded drifts

$$b \in L^\infty \mathbf{F}_\delta + L^{2+\varepsilon}_{\rm loc}(\mathbb{R}), \quad \text{(i.e. with function } g_\delta \in L^{1+\varepsilon/2}_{\rm loc}(\mathbb{R})), \quad \varepsilon > 0,$$

that can have stronger spatial singularities than than the drifts in (13.2), using a different argument which applies repeatedly integration by parts, quadratic inequality and the form-boundedness of b. The rest of the proof of strong existence essentially repeats [RZh2]. This, combined with the weak uniqueness results from [KiM1] and [Ki5] via Cherny's theorem (see Sections 10 and 16), yields the following result.

Theorem 13.1 ([KiM3]). *Assume that $b \in L^\infty \mathbf{F}_\delta + L^{2+\varepsilon}_{\rm loc}(\mathbb{R})$ for some $\varepsilon > 0$. Also, assume that b has compact support. Then, provided that form-bound δ is sufficiently small, for every $x \in \mathbb{R}^d$, SDE* (13.1) *has a strong solution X_t^x. This strong solution satisfies the following Krylov-type bounds:*

1) *For a given $q \in]d, \delta^{-\frac{1}{2}}[$ and any vector field $\mathsf{g} \in L^\infty \mathbf{F}_{\delta_1} + L^{2+\varepsilon}_{\rm loc}(\mathbb{R})$, $\delta_1 < \infty$,*

$$\mathbf{E} \int_0^T |\mathsf{g}h|(\tau, X^x_{0,\tau}) d\tau \leq c \|\mathsf{g}|h|^{\frac{q}{2}}\|^{\frac{2}{q}}_{L^2([0,T]\times\mathbb{R}^d)} \quad \textit{for all } h \in C_c([0,T]\times\mathbb{R}^d). \tag{13.5}$$

2) *For a given $\mu > \frac{d+2}{2}$, there exists constant C such that*

$$\mathbf{E}\bigg[\int_0^T |h(\tau, X^x_{0,\tau})| d\tau\bigg] \leq C\|h\|_{L^\mu([0,T]\times\mathbb{R}^d)} \quad \textit{for all } h \in C_c([0,T]\times\mathbb{R}^d). \tag{13.6}$$

Solution X^x_t is unique among strong solutions to (13.1) *that satisfy* (13.5) *for some $q \in]d, \delta^{-\frac{1}{2}}[$ with $\mathsf{g} = 1$ and with $\mathsf{g} = b$. If, in addition to the hypothesis on b, one has $|b| \in L^{\frac{d+2}{2}+\epsilon}$ for some $\epsilon > 0$, then X^x_t is unique among strong solutions to* (13.1) *that satisfy* (13.6).

The assumption that b has compact support can be removed with an additional effort (using weight (4.5)).

Speaking of "$\varepsilon > 0$" in $b \in L^\infty \mathbf{F}_\delta + L^{2+\varepsilon}_{\rm loc}(\mathbb{R})$, this assumption does not allow us to include completely the critical Ladyzhenskaya-Prodi-Serrin class even with $p > 2$, see above, as is assumed in [RZh2]. It does include, however, the case that interests us the most: $p = \infty$, $q = d$. It also includes with case $p > 2$, $q = \infty$.

The weak well-posedness of SDE (13.1) is known to hold for larger classes of singular drifts than class $L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R})$ discussed in Theorem 13.1. This is the subject of the next three sections.

14. More singular than form-bounded. Semigroup in $\mathcal{W}^{\frac{1}{2},2}$

In Sections 14, 15 and 16 we expand the classes of singular drifts considered in Theorems 6.1, 6.2 and 9.1, 10.1, although at expense of requiring smaller δ.

Definition 14.1. A Borel measurable vector field $b : \mathbb{R}^d \to \mathbb{R}^d$ with $|b| \in L^1_{\rm loc}$ is said to be weakly form-bounded if there exists a constant $\delta > 0$ such that

$$\||b|^{\frac{1}{2}}(\lambda - \Delta)^{-\frac{1}{4}}\|_{2\to 2} \leq \sqrt{\delta}$$

for some $\lambda = \lambda_\delta \geq 0$. This is written as $b \in \mathbf{F}^{\frac{1}{2}}_\delta$.

There is an important difference between form-bounded vector fields and weakly form-bounded vector fields. Namely, when we dealt with $b \in \mathbf{F}_\delta$, we controlled the gradient term $b \cdot \nabla$ in the Kolmogorov operator $-\Delta + b \cdot \nabla$ using the quadratic (Cauchy-Schwarz) inequality

$$|\langle b \cdot \nabla\varphi, \varphi\rangle| \leq \varepsilon\|b\varphi\|_2^2 + \frac{1}{4\varepsilon}\|\nabla\varphi\|_2^2, \quad \varepsilon > 0 \tag{14.1}$$

(e.g. (3.3) in the verification of conditions of the Lax-Milgram theorem and the KLMN theorem in L^2, in the proof of Lemma 6.1, in the iteration procedure and gradient bound (9.7), etc). *We* *can no longer do this when dealing with* $b \in \mathbf{F}_\delta^{1/2}$ *if only because* $|b|$ *is in general no longer locally in* L^2*.*

The form-bounded vector fields are weakly form-bounded. To show this, let us recall that the condition $b \in \mathbf{F}_\delta$ can be stated as an operator-norm inequality $\||b|(\lambda-\Delta)^{-\frac{1}{2}}\|_{2\to 2} \leq \sqrt{\delta}$. The Heinz-Kato inequality [He] allows us to take square roots in the operators that constitute the left-hand side, so we arrive at $\||b|^{\frac{1}{2}}(\lambda-\Delta)^{-\frac{1}{4}}\|_{2\to 2} \leq \delta^{\frac{1}{4}}$, and hence

$$b \in \mathbf{F}_\delta \quad \Rightarrow \quad b \in \mathbf{F}^{\frac{1}{2}}_{\sqrt{\delta}}.$$

The opposite inclusion is invalid: there are weakly form-bounded vector fields that are not form-bounded:

Examples. 1. If $|b|$ belongs to scaling-invariant Morrey class $M_{1+\varepsilon}$, for arbitrarily fixed small $\varepsilon > 0$, i.e.

$$|b| \in L^{1+\varepsilon}_{\rm loc} \quad \text{ and } \quad \|b\|_{M_{1+\varepsilon}} := \sup_{r>0, x\in\mathbb{R}^d} r\biggl(\frac{1}{|B_r|}\int_{B_r(x)} |b|^{1+\varepsilon}dx\biggr)^{\frac{1}{1+\varepsilon}} < \infty, \tag{14.2}$$

then $b \in \mathbf{F}_\delta^{1/2}$. The proof of this inclusion follows right away from [A, Theorem 7.3].

Recalling that the class of form-bounded vector fields $\mathbf{F}_\delta$ satisfies $M_{2+\varepsilon} \subset \mathbf{F}_\delta \subset M_2$ (say, with $c_\delta = 0$), one can see that we gain quite a lot in admissible singularities of b by working with $\mathbf{F}_\delta^{1/2}$. In particular, we gain all b with $|b| \in M_{1+\varepsilon} - M_2$.

2. Recall that a vector field $b : \mathbb{R}^d \to \mathbb{R}^d$ is said to belong to the Kato class if $|b| \in L^1_{\rm loc}$ and

$$\|(\lambda-\Delta)^{-\frac{1}{2}}|b|\|_\infty \leq \sqrt{\delta} \tag{14.3}$$

for some $\delta > 0$ and $\lambda = \lambda_\delta \geq 0$. This is written as $b \in \mathbf{K}_\delta^{d+1}$. The Kato class vector fields are weakly form-bounded. Indeed, if $b \in \mathbf{K}_\delta^{d+1}$, then by duality one has

$$\||b|(\lambda-\Delta)^{-\frac{1}{2}}\|_{1\to 1} \leq \sqrt{\delta}. \tag{14.4}$$

Applying Stein's interpolation between (14.3) and (14.4), one has $\||b|^{\frac{1}{2}}(\lambda-\Delta)^{-\frac{1}{2}}|b|^{\frac{1}{2}}\|_{2\to 2} \leq \sqrt{\delta}$; in the left-hand side of the last inequality one has the norm of the product of $|b|^{\frac{1}{2}}(\lambda-\Delta)^{-\frac{1}{4}}$ and its adjoint. By a standard result this yields $\||b|^{\frac{1}{2}}(\lambda-\Delta)^{-\frac{1}{4}}\|_{2\to 2} \leq \delta^{\frac{1}{4}}$. Thus,

$$b \in \mathbf{K}_\delta^{d+1} \quad \Rightarrow \quad b \in \mathbf{F}^{1/2}_{\sqrt{\delta}}.$$

As was mentioned earlier, the Kato class $b \in \mathbf{K}_\delta^{d+1}$ with δ sufficiently small provides two-sided Gaussian bounds on the heat kernel of Kolmogorov operator $-\Delta + b\cdot\nabla$ [Za2]. The Kato class also provides uniqueness in law for SDE (15.1), see [BC]. There the authors required δ to be arbitrarily small. In fact, they show that under the Kato class assumption on b the gradient of solution of elliptic equation $(\mu - \Delta + b\cdot\nabla)v = f$ is bounded. The reader can compare this with Remark 3.1 concerning the gradient of v for a form-bounded b.

The Kato class $\mathbf{K}_\delta^{d+1}$ does not contain a popular class $|b| \in L^d$ (if only because there are vector fields b with $|b| \in L^d$ that destroy two-sided Gaussian bounds on $-\Delta + b\cdot\nabla$), and so it does not contain $\mathbf{F}_\delta$. On the other hand, the Kato class is not contained in $\mathbf{F}_\delta$. However, both form-bounded and Kato class vector fields are contained in $\mathbf{F}_\delta^{1/2}$.

Let us demonstrate one way to arrive at the condition $b \in \mathbf{F}_\delta^{1/2}$, $\delta < 1$. First, let $b \in \mathbf{F}_\delta$, $\delta < 1$, and, for brevity, assume that $c_\delta = 0$. Also, let b be bounded and smooth so that all manipulations with the equation are justified, but the constants in the estimates below will not depend on the smoothness or boundedness of b. One can prove the following two L^2 regularity results for Cauchy problem $(\partial_t - \Delta + b \cdot \nabla)u = 0$, $u(0) = f$ with such b:

– Multiplying $(\partial_t - \Delta + b \cdot \nabla)u = 0$ by u, integrating over $[0,t] \times \mathbb{R}^d$ and applying $b \in \mathbf{F}_\delta$ and quadratic inequality (14.1), we obtain energy inequality

$$\frac{1}{2}\|u(t)\|_2^2 + (1-\sqrt{\delta})\int_0^t \|\nabla u\|_2^2 dr \leq \frac{1}{2}\|f\|_2^2. \tag{14.5}$$

– Multiplying $(\partial_t - \Delta + b \cdot \nabla)u = 0$ by $-\Delta u$, integrating over $[0,t] \times \mathbb{R}^d$, we obtain

$$\frac{1}{2}\partial_t\|\nabla u\|_2^2 + \|\Delta u\|_2^2 + \langle b \cdot \nabla u, -\Delta u\rangle = 0,$$

where we further estimate, using $b \in \mathbf{F}_\delta$ with $c_\delta = 0$,

$$\begin{aligned}
|\langle b \cdot \nabla u, -\Delta u\rangle| &\leq \varepsilon \||b|\nabla u|\|_2 + \frac{1}{4\varepsilon}\|\Delta u\|_2^2 \\
&\leq \varepsilon\delta\|(-\Delta)^{\frac{1}{2}}|\nabla u|\|_2^2 + \frac{1}{4\varepsilon}\|\Delta u\|_2^2 \\
&\text{(use Beurling-Deny-type inequality } \|(-\Delta)^{\frac{1}{2}}|\nabla u|\|_2^2 \leq \|(-\Delta)^{\frac{1}{2}}\nabla u\|_2^2 \equiv \|\Delta u\|_2^2 \\
&\text{and select } \varepsilon = \frac{1}{2\sqrt{\delta}}) \\
&\leq \sqrt{\delta}\|\Delta u\|_2^2.
\end{aligned}$$

Thus, we obtain another "energy inequality":

$$\frac{1}{2}\|\nabla u(t)\|_2^2 + (1-\sqrt{\delta})\int_0^t \|\Delta u\|_2^2 dr \leq \frac{1}{2}\|\nabla f\|_2^2. \tag{14.6}$$

– Now, one can ask what happens if we multiply $(\partial_t - \Delta + b \cdot \nabla)u = 0$ by an intermediate test function $(-\Delta)^s u$, $0 < s < 1$. One obtains an intermediate result between (14.5) and (14.6), but for a larger class of vector fields b, which becomes maximal if one multiplies by $(-\Delta)^{\frac{1}{2}}u$:

$$\begin{aligned}
|\langle b \cdot \nabla u, (-\Delta)^{\frac{1}{2}}u\rangle| &= |\langle b^{\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}\nabla(-\Delta)^{\frac{1}{4}}u, |b|^{\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}(-\Delta)^{\frac{3}{4}}u\rangle| \qquad b^{\frac{1}{2}} := b|b|^{-\frac{1}{2}} \\
&\leq \||b|^{\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}\|_{2\to 2}\|\nabla(-\Delta)^{\frac{1}{4}}u\|_2 \,\||b|^{\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}\|_{2\to 2}\|(-\Delta)^{\frac{3}{4}}u\|_2 \\
&= \||b|^{\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}\|_{2\to 2}^2\|(-\Delta)^{\frac{3}{4}}u\|_2^2.
\end{aligned}$$

Thus, requiring

$$\||b|^{\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}\|_{2\to 2} \leq \sqrt{\delta} \quad (\text{i.e. } b \in \mathbf{F}_\delta^{1/2} \text{ with } \lambda = 0), \quad \delta < 1,$$

one obtains the following "energy inequality":

$$\frac{1}{2}\|(-\Delta)^{\frac{1}{4}}u(t)\|_2^2 + (1-\delta)\int_0^t \|(-\Delta)^{\frac{3}{4}}u\|_2^2 dr \leq \frac{1}{2}\|(-\Delta)^{\frac{1}{4}}f\|_2^2. \tag{14.7}$$

From the look of (14.7), it is seen that one needs to work with the chain of Bessel spaces

$$\mathcal{W}^{\frac{3}{2},2} \subset \mathcal{W}^{\frac{1}{2},2} \subset \mathcal{W}^{-\frac{1}{2},2} \tag{14.8}$$

rather than the standard Sobolev triple $W^{1,2} \subset L^2 \subset W^{-1,2}$ (above $\mathcal{W}^{-\frac{1}{2},2}$ is the dual of $\mathcal{W}^{\frac{3}{2},2}$ with respect to the inner product in $\mathcal{W}^{\frac{1}{2},2}$) . Of course, by doing that, one sacrifices (3.11) and loses the possibility to consider general operator $-\nabla \cdot a \cdot \nabla + b \cdot \nabla$ unless uniformly elliptic matrix a satisfies additional regularity assumptions that make $-\nabla \cdot a \cdot \nabla$ a bounded operator from $\mathcal{W}^{\frac{3}{2},2}$ to $\mathcal{W}^{-\frac{1}{2},2}$, see Remark 14.1. In fact, the following result is true:

Proposition 14.1. *Let $b \in \mathbf{F}_\delta^{\frac{1}{2}}$, $\delta < 1$. Then for every $f \in \mathcal{W}^{\frac{1}{2},2}$ there exists a unique weak solution to Cauchy problem*

$$(\partial_t + \lambda - \Delta + b \cdot \nabla)u = 0, \quad u(0+) = f \tag{14.9}$$

where λ is from the condition $b \in \mathbf{F}_\delta^{1/2}$, i.e. a unique in $L^\infty_{\rm loc}(]0,\infty[, \mathcal{W}^{\frac{1}{2},2}) \cap L^2_{\rm loc}(]0,\infty[, \mathcal{W}^{\frac{3}{2},2})$ function u satisfying

$$\int_0^\infty \langle (\lambda - \Delta)^{\frac{1}{4}} u, \partial_t (\lambda - \Delta)^{\frac{1}{4}} \varphi \rangle dt = \int_0^\infty \langle (\lambda - \Delta)^{\frac{3}{4}} u, (\lambda - \Delta)^{\frac{3}{4}} \varphi \rangle dt + \int_0^\infty \langle b \cdot \nabla u, (\lambda - \Delta)^{\frac{1}{2}} \varphi \rangle$$

for all $\varphi \in C_c^\infty(]0,\infty[, \mathcal{S})$ and w-$\mathcal{W}^{\frac{1}{2},2}$-$\lim_{t \downarrow 0} u(t) = f$. One has $u \in C(\mathbb{R}_+, \mathcal{W}^{\frac{1}{2},2})$, the following energy inequality holds:

$$\|u(t)\|^2_{\mathcal{W}^{\frac{1}{2},2}} + (1-\delta) \int_0^t \|u\|^2_{\mathcal{W}^{\frac{3}{2},2}} dr \leq \|f\|^2_{\mathcal{W}^{\frac{1}{2},2}}, \quad t \geq 0,$$

and $T^t f(\cdot) := u(t, \cdot)$ is a contraction strongly continuous in $\mathcal{W}^{\frac{1}{2},2}$ Markov semigroup. If $\{b_\varepsilon\}_{\varepsilon > 0}$ is a family of bounded smooth vector fields such that $b_\varepsilon \in \mathbf{F}_\delta^{1/2}$ with the same λ as b, $b_\varepsilon \to b$ in $[L^1_{\rm loc}]^d$ as $\varepsilon \to 0$, and if u_ε denotes the solution to Cauchy problem (14.9) with the vector field b_ε, then

$$u_\varepsilon \to u \quad \text{weakly in } L^2_{\rm loc}(\mathbb{R}_+, \mathcal{W}^{\frac{3}{2},2}) \text{ as } \varepsilon \to 0.$$

The proof uses the standard J.-L. Lions approach in the scale (14.8), and was carried out in [KiS8], in fact, in greater generality: for time-inhomogeneous $b \in L^\infty \mathbf{F}_\delta^{1/2}$, that is, satisfying for a.e. $t \in \mathbb{R}_+$ the operator inequality

$$\| |b(t)|^{\frac{1}{2}} (\lambda - \Delta)^{-\frac{1}{4}} \|_{2 \to 2} \leq \sqrt{\delta}$$

for some fixed $\lambda = \lambda_\delta$.

The above argument leading to the energy inequality (14.7) is not how the class $\mathbf{F}_\delta^{1/2}$ first appeared in the literature. The (L^p, L^q) estimate

$$\|e^{-t\Lambda_p(b)} f\|_q \leq c e^{t\omega_p} t^{-\frac{d}{2}(\frac{1}{p} - \frac{1}{q})} \|f\|_p, \quad t > 0, \quad \omega_p := \frac{c_\delta}{2(p-1)}, \tag{14.10}$$

can be proved separately for $b \in \mathbf{F}_\delta$, see (6.13), and for $b \in \mathbf{K}_\delta^{d+1}$ (as was mentioned above, for the Kato class one even has two-sided Gaussian bounds on the integral kernel $e^{-t\Lambda(b)}(x,y)$ of $e^{-t\Lambda_p(b)}$). It was noticed in [S1] that the validity of estimate (14.10) depends, in fact, only on the weaker condition $\| |b|^{\frac{1}{2}} (\lambda - \Delta)^{-\frac{1}{4}} \|_{2 \to 2} \leq \sqrt{\delta}$ with $\delta < 1$, which led to the introduction of the class $\mathbf{F}_\delta^{1/2}$. Also, [S1] proposed a way to construct a *quasi bounded* semigroup in L^2

associated with $-\Delta + b \cdot \nabla$ with weakly form-bounded $b \in \mathbf{F}_\delta^{1/2}$ by constructing its resolvent as the operator-valued function

$$\Phi(\zeta, b) := (\zeta - \Delta)^{-\frac{3}{4}}(1 + HS)^{-1}(\zeta - \Delta)^{-\frac{1}{4}}, \qquad \operatorname{Re}\zeta \geq \lambda_\delta,$$

where, by $b \in \mathbf{F}_\delta^{1/2}$, operators

$$H := (\zeta - \Delta)^{-\frac{1}{4}}|b|^{\frac{1}{2}}, \quad S := b^{\frac{1}{2}} \cdot \nabla(\zeta - \Delta)^{-\frac{3}{4}} \tag{14.11}$$

are bounded on L^2 with norm $\sqrt{\delta}$ each (for operator S, taking into account that $\nabla(\zeta - \Delta)^{-\frac{1}{2}}$ is bounded on L^2 with norm 1), and so $\Phi(\zeta, b)$ is bounded on L^2. Then, it is easily seen, $\Phi(\zeta, b)$ satisfies

$$\|\Phi(\zeta, b)\|_{2\to 2} \leq (1 - \delta)^{-1}|\zeta|^{-1}, \quad \text{ on } \{\operatorname{Re}\zeta \geq \lambda_\delta\}. \tag{14.12}$$

The proof that $\Phi(\zeta, b)$ is indeed the resolvent of the generator Λ of a quasi bounded strongly continuous semigroup[14] $e^{-t\Lambda}$ on L^2 uses the general Trotter approximation theorem. The latter, in practice, requires the uniform in n estimate (14.12) for $\Phi(\zeta, b_n)$, where b_n are approximating vector fields for b (cf. (15.6), (15.7)). In other words, it is essential for the construction that one is working with a holomorphic semigroup. We refer to [KiS2] for detailed discussion.

Note also that one no longer has (3.6), i.e.

$$\Lambda \not\supset -\Delta + b \cdot \nabla \upharpoonright C_c^\infty.$$

The reason is that for a weakly form-bounded b its norm $|b|$ is in general not in $L^2_{\rm loc}$.

If b is form-bounded or in the Kato class of vector fields – two standard assumptions – then one can construct a realization of $-\Delta + b \cdot \nabla$ as the generator of a strongly continuous semigroup *in some* L^p by invoking the KLMN theorem in L^2 or the Miyadera theorem in L^1, respectively (see e.g. [KiS2]). However these two theorems (and, generally speaking, the standard perturbation-theoretic tools) are inapplicable to $-\Delta + b \cdot \nabla$ in any L^p if b is in the class of weakly form-bounded drifts $\mathbf{F}_\delta^{1/2}$.

Remark 14.1. Let us comment on the assumptions on a measurable uniformly elliptic matrix a (i.e. $\sigma I \leq a \leq \xi I$ a.e. on $\mathbb{R}^d$ for $0 < \sigma \leq \xi < \infty$) that would allow to extend Proposition 14.1 to operator $-\nabla \cdot a \cdot \nabla + b \cdot \nabla$. If $b \in \mathbf{F}_\delta^{1/2}$, then it is easily seen that

$$b \cdot \nabla \in \mathcal{B}(\mathcal{W}^{\frac{3}{2},2}, \mathcal{W}^{-\frac{1}{2},2}).$$

The matrix a has to be such that

$$-\nabla \cdot a \cdot \nabla \in \mathcal{B}(\mathcal{W}^{\frac{3}{2},2}, \mathcal{W}^{-\frac{1}{2},2}) \tag{14.13}$$

(of course, if a is only measurable uniformly elliptic, then one only has $-\nabla \cdot a \cdot \nabla \in \mathcal{B}(W^{1,2}, W^{-1,2})$). Let us mention one elementary sufficient condition for (14.13). For simplicity we will stay at the a priori level (i.e. the matrix is smooth but the norm of the operator in (14.13) does not depend on smoothness of a). Also, assume that $a = I + a^0$ where a^0 has entries a^0_{ij} in $\mathcal{S}$. For given $\varphi \in \mathcal{W}^{\frac{3}{2},2}$, $\psi \in \mathcal{W}^{\frac{1}{2},2}$, we have

$$\begin{aligned}|\langle -\nabla \cdot a^0 \cdot \nabla\varphi, \psi\rangle| &= |\langle (1 - \Delta)^{\frac{1}{4}}a^0(1 - \Delta)^{-\frac{1}{4}} \cdot (1 - \Delta)^{\frac{1}{4}}\nabla\varphi, (1 - \Delta)^{-\frac{1}{4}}\nabla\psi\rangle| \\ &\leq \|(1 - \Delta)^{\frac{1}{4}}a^0(1 - \Delta)^{-\frac{1}{4}}\|_{2\to 2}\|\varphi\|_{\mathcal{W}^{\frac{3}{2},2}}\|\psi\|_{\mathcal{W}^{\frac{1}{2},2}},\end{aligned}$$

[14]But not a quasi contraction semigroup, as in Section 6.

where, in turn, by the Kato-Ponce inequality ($\equiv$ fractional Leibnitz rule) [GO], for all $1 \leq i,j \leq d$,

$$\|(1-\Delta)^{\frac{1}{4}}a^0_{ij}(1-\Delta)^{-\frac{1}{4}}f\|_2 \leq \|(1-\Delta)^{\frac{1}{4}}a^0_{ij}\|_{2d}\|(1-\Delta)^{-\frac{1}{4}}f\|_{\frac{2d}{d-1}} + \|a_{ij}\|_\infty\|f\|_2.$$

Thus, if

$$\|a^0_{ij}\|_{\mathcal{W}^{\frac{1}{2},2d}} \leq c < \infty$$

for all i,j, for a generic constant c (i.e. a constant that does not depend on the smoothness of a^0_{ij}), then, using $\|(1-\Delta)^{-\frac{1}{4}}f\|_{\frac{2d}{d-1}} \leq C\|f\|_2$, we obtain $\|(1-\Delta)^{\frac{1}{4}}a^0(1-\Delta)^{-\frac{1}{4}}\|_{2\to 2} \leq c'$ for a generic c', and hence (14.13) with the operator norm bounded by a generic constant.

Remark 14.2. At the level of Feller semigroups, one cannot really draw a parallel between the Kato class of potentials $\mathbf{K}^d_\delta$ (see (3.9)) and the Kato class of drifts $\mathbf{K}^{d+1}_\delta$. Indeed, in view of the results [OSSV], for the Schrödinger operator $-\Delta + V$ condition $V \in \mathbf{K}^d_\delta$ is, basically, necessary and sufficient for the Feller property ($\equiv$ strong continuity of the semigroup on C_∞) to hold. For the Kolmogorov operator $-\Delta + b\cdot\nabla$, condition $b \in \mathbf{K}^{d+1}_\delta$ is only sufficient for the Feller property; as Theorem 15.1 below shows, one can go much farther, to weakly form-bounded drifts.

15. Weakly form-bounded drifts and SDEs

In this section we construct the Feller semigroup for $-\Delta + b\cdot\nabla$ and prove weak well-posedness of SDE

$$X_t = x - \int_0^t b(X_r)dr + \sqrt{2}W_t, \quad t \geq 0, \tag{15.1}$$

for a fixed $x \in \mathbb{R}^d$, with $b : \mathbb{R}^d \to \mathbb{R}^d$ in the class of *weakly form-bounded vector fields* $\mathbf{F}^{1/2}_\delta$.

1. We construct the sought Feller generator by arguing essentially as in the proof of Theorem 6.1 (so, in particular, we do not use L^2 theory of $-\Delta + b\cdot\nabla$, $b \in \mathbf{F}^{1/2}_\delta$). First, we prove an analogue of Lemma 6.1 for weakly form-bounded drifts. Namely, for given $p \in]1,\infty[$, $1 \leq r < p < q < \infty$ and $\mu > 0$, define operators

$$\begin{aligned} G_p(r) &:= b^{\frac{1}{p}} \cdot \nabla(\mu - \Delta)^{-\frac{1}{2}-\frac{1}{2r}}, \\ Q_p(q) \upharpoonright \mathcal{E} &:= (\mu-\Delta)^{-\frac{1}{2}+\frac{1}{2q}}|b|^{1-\frac{1}{p}}, \\ T_p \upharpoonright \mathcal{E} &:= b^{\frac{1}{p}} \cdot \nabla(\mu-\Delta)^{-1}|b|^{1-\frac{1}{p}}. \end{aligned}$$

where, recall, $\mathcal{E} := \bigcup_{\varepsilon>0} e^{-\varepsilon|b|}L^p$ is a dense subspace of L^p. Notice that the power $\frac{2}{p}$ in the definition of operators $G_p(r)$, T_p, $Q_p(q)$ in Lemma 6.1 is now replaced with $\frac{1}{p}$. Set

$$m_d := \pi^{\frac{1}{2}}(2e)^{-\frac{1}{2}}d^{\frac{d}{2}}(d-1)^{\frac{1-d}{2}}, \qquad c_p := pp'/4.$$

Lemma 15.1. *Let $b \in \mathbf{F}^{1/2}_\delta$. For every $p \in]1,\infty[$, the following is true for all $\mu \geq \kappa_d\lambda_\delta$,*

(i) $T_p \upharpoonright \mathcal{E}$ admits extension by continuity to L^p, denoted by T_p. One has

$$\|T\|_{p\to p} \leq m_d c_p \delta.$$

In particular, if δ satisfies $m_d\delta < 1$, then for every $p \in I_\delta :=]\frac{2}{1+\sqrt{1-m_d\delta}}, \frac{2}{1-\sqrt{1-m_d\delta}}[$ one has $\|T_p\|_{p\to p} < 1$.

(ii) $Q_p(q) \upharpoonright \mathcal{E}$ *admits extension by continuity to* L^p*, denoted by* $Q_p(q)$*.*
(iii) $G_p(r)$ *is bounded on* L^p*.*

Lemma 15.1 was proved in [Ki1]. Let us demonstrate the proof of (i) to make it easier to compare Lemma 15.1 with Lemma 6.1 (proved in Appendix A). Define in L^2 operator $A = (\mu-\Delta)^{\frac{1}{2}}$, $D(A) = W^{1,2}$. This is a *symmetric* Markov generator. Therefore, we have for $p \in]1,\infty[$:

$$0 \leqslant u \in D(A_p) \quad \Rightarrow \quad u^{\frac{p}{2}} \in D(A^{\frac{1}{2}})$$

and the following inequality (sometimes called the Stroock-Varopoulos inequality) is valid:

$$c_p^{-1}\|A^{\frac{1}{2}}u^{\frac{p}{2}}\|_2^2 \leqslant \langle A_p u, u^{p-1}\rangle, \quad c_p := \frac{pp'}{4}, \quad p' = \frac{p}{p-1} \tag{15.2}$$

(see [LS, Theorem 2.1], see also [KiS5] for a useful vector-valued analogue of these inequalities). Here A_p is the generator of strongly continuous semigroup $e^{-tA_p} := [e^{-tA} \upharpoonright L^2 \cap L^p]^{\rm clos}_{L^p \to L^p}$, cf. discussion in the beginning of the previous section. Now, let u be the solution of equation[15] $A_p u = |b|^{1/p'}|f|$, $f \in \mathcal{E}$. The condition $b \in \mathbf{F}_\delta^{1/2}$ yields, provided $\mu \geq \lambda_\delta$,

$$\||b|^{\frac{1}{2}}u^{\frac{p}{2}}\|_2^2 \leq \delta\|A^{\frac{1}{2}}u^{\frac{p}{2}}\|_2^2.$$

Hence, by (15.2),

$$(c_p\delta)^{-1}\||b|^{\frac{1}{2}}u^{\frac{p}{2}}\|_2^2 \leqslant \langle A_p u, u^{p-1}\rangle.$$

Now, noting that $\||b|^{\frac{1}{2}}u^{\frac{p}{2}}\|_2^2 = \||b|^{\frac{1}{p}}u\|_p^p$ and using $A_p u = |b|^{\frac{1}{p'}}|f|$, we obtain

$$\||b|^{\frac{1}{p}}u\|_p^p \leqslant c_p\delta\|f\|_p\||b|^{\frac{1}{p}}u\|_p^{p-1}.$$

Thus, $\||b|^{\frac{1}{p}}u\|_p \leqslant c_p\delta\|f\|_p$, so we arrive at the estimate

$$\||b|^{\frac{1}{p}}A^{-1}|b|^{\frac{1}{p'}}|f|\|_p \leq c_p\delta\|f\|_p. \tag{15.3}$$

To end the proof of (i), it remains to apply in the definition of T_p the pointwise inequality (this is where the constant m_d comes from)

$$|\nabla_x(\mu-\Delta)^{-1}(x,y)| \leqslant m_d(\kappa_d^{-1}\mu-\Delta)^{-\frac{1}{2}}(x,y), \quad x,y \in \mathbb{R}^d,\ x \neq y, \tag{15.4}$$

where

$$\kappa_d := \frac{d}{d-1},$$

and then apply (15.3) to the result.

Remark 15.1. Similar estimates, without the gradient, were used earlier in [BS, LS] in the study of Schrödinger operators with form-bounded potentials.

Remark 15.2. By applying (15.4) in the definition of T_p we kill the gradient from the gradient term $b \cdot \nabla$. This allows us to apply to what is left the inequalities for symmetric Markov generators. By the way, this is why the interval I_δ of admissible p in Theorem 15.1 is symmetric, despite the fact that the operator $-\Delta + b \cdot \nabla$ is non-symmetric. In the proof of Lemma 6.1, when dealing with the condition $b \in \mathbf{F}_\delta$, we control the gradient in a more efficient way, which allows to impose less restrictive assumptions on δ than in Lemma 15.1. It is not clear at the

[15] We can also carry out the proof of Lemma 15.1 for bounded b_n as e.g. defined below, and then pass to the limit in n.

moment how to prove Lemma 15.1 without either resorting to (15.4) or restricting the class $\mathbf{F}_\delta^{1/2}$ to Morrey class $M_{1+\varepsilon}$, $\varepsilon>0$ (cf. the proof of Lemma 16.1 below).

The interval I_δ expands to $]1,\infty[$ as $\delta\downarrow 0$. In particular, if δ is sufficiently small, I_δ contains $p>d-1$, which is what will be needed to construct the resolvent of a Feller generator in terms of $Q_p(q)$, T_p, $G_p(r)$ and some "free" Bessel potentials using the Sobolev embedding theorem. This is what is done in Theorem 15.1 below.

Set

$$b_n := c_n\eta_{\varepsilon_n} * (\mathbf{1}_n b), \tag{15.5}$$

where $\mathbf{1}_n$ is the indicator of $\{x \mid |x|\leq n, |b(x)|\leq n\}$, η_{ε_n} is the Friedrichs mollifier, and we choose $\varepsilon_n\downarrow 0$ (sufficiently rapidly) so that, for appropriate $c_n\uparrow 1$ (sufficiently slow), one has

$$b_n\to b \quad \text{ in } L^1_{\rm loc}(\mathbb{R}^d,\mathbb{R}^d) \tag{15.6}$$

and

$$b_n\in \mathbf{F}_\delta^{1/2} \quad \text{ with some } \lambda_\delta \text{ independent of } n=1,2,\dots \tag{15.7}$$

see Appendix C.1. The following theorem was proved in [Ki3].

Theorem 15.1. *Let $b\in\mathbf{F}_\delta^{\frac12}$, $m_d\delta<1$. The following is true for all $\mu\geq\kappa_d\lambda_\delta$.*

(*i*) *For every $p\in I_\delta =]\frac{2}{1+\sqrt{1-m_d\delta}},\frac{2}{1-\sqrt{1-m_d\delta}}[$, the function*

$$u=(\mu-\Delta)^{-1}f-(\mu-\Delta)^{-\frac12-\frac{1}{2q}}Q_p(q)(1+T_p)^{-1}G_p(r)(\mu-\Delta)^{-\frac12+\frac{1}{2r}}f, \quad f\in L^p \tag{15.8}$$

is a weak solution to the elliptic equation

$$(\lambda-\Delta+b\cdot\nabla)u=f, \tag{15.9}$$

i.e. $b\cdot\nabla u\in L^1_{\rm loc}$

$$\mu\langle u,\psi\rangle+\langle\nabla u,\nabla\psi\rangle+\langle b\cdot\nabla u,\psi\rangle=\langle f,\psi\rangle \quad \textit{for all } \psi\in\mathcal{S}.$$

Moreover, if $f\in L^p\cap L^2$, then u is the unique in $\mathcal{W}^{\frac32,2}$ weak solution to (15.9).

(*ii*) *It follows from* (15.8) *that*

$$u\in\mathcal{W}^{1+\frac1q,p}, \quad q>p.$$

In particular, if $m_d\delta<\frac{4(d-2)}{(d-1)^2}$, then in the interval $p\in I_\delta$ we can select $p>d-1$, and then select q sufficiently close to p, so that by the Sobolev embedding theorem u is Hölder continuous.

(*iii*) *The operator-valued function in* (15.8)

$$\Theta_p(\mu,b):=(\mu-\Delta)^{-1}-(\mu-\Delta)^{-\frac12-\frac{1}{2q}}Q_p(q)(1+T_p)^{-1}G_p(r)(\mu-\Delta)^{-\frac12+\frac{1}{2r}}$$

on $\{\mu>\mu_0\}$ takes values in $\mathcal{B}(\mathcal{W}^{-1+\frac1r,p},\mathcal{W}^{1+\frac1q,p})$.

(*iv*) *Let δ satisfy $m_d\delta<\frac{4(d-2)}{(d-1)^2}$. Fix $p\in I_\delta$ such that $p>d-1$. Then*

$$(\mu+\Lambda_{C_\infty}(b))^{-1}:=\big(\Theta_p(\mu,b)\upharpoonright L^p\cap C_\infty\big)^{\rm clos}_{C_\infty\to C_\infty}, \quad \mu\geq\kappa_d\lambda,$$

determines the resolvent of a Feller generator on C_∞. This semigroup satisfies

$$e^{-t\Lambda_{C_\infty}(b)}=s\text{-}C_\infty\text{-}\lim_n e^{-t\Lambda_{C_\infty}(b_n)} \quad \textit{locally uniformly in } t\geq 0,$$

where b_n are defined by (15.5)*, and operators $\Lambda_{C_\infty}(b_n):=-\Delta+b_n\cdot\nabla$ with domain $D(\Lambda_{C_\infty}(b_n)):=(1-\Delta)^{-1}C_\infty$ are Feller generators.*

(*v*) *Feller semigroup* $e^{-t\Lambda_{C_\infty}(b)}$ *is conservative, i.e. its integral kernel* $e^{-t\Lambda_{C_\infty}}(x,\cdot)$ *satisfies*

$$\langle e^{-t\Lambda_{C_\infty}(b)}(x,\cdot)\rangle = 1 \quad \text{for all } x \in \mathbb{R}^d, t > 0.$$

For the proof, except for the part that concerns the weak solution to the elliptic equation, one can repeat the proof of Theorem 6.1 using Lemma 15.1 instead of Lemma 6.1.

Remark 15.3. The fact that u is a weak solution was proved in [Ki3]. Moreover, since $f \in L^2 \cap L^p$, we have

$$u = \Theta_2(\mu,b)f = \Phi_2(\mu,b)f,$$

so $u \in \mathcal{W}^{\frac{3}{2},2}$. (We could also prove that if in Theorem 15.1 $p=2$, then one can take $q=r=2$.) The proof of the uniqueness of u in $\mathcal{W}^{\frac{3}{2},2}$ goes as follows. Let $v \in \mathcal{W}^{\frac{3}{2}}$ be some weak solution of (15.9). Then, selecting $\varphi = (\mu-\Delta)^{-\frac{1}{2}}\eta$, $\eta \in C_c^\infty$, we have

$$\langle (\mu-\Delta)^{\frac{3}{4}}v, (\mu-\Delta)^{\frac{3}{4}}\eta\rangle + \langle S(\mu-\Delta)^{\frac{3}{4}}u, H^*(\mu-\Delta)^{\frac{3}{4}}\eta\rangle = \langle f, (\mu-\Delta)^{-\frac{1}{2}}\eta\rangle,$$

where $H = (\mu-\Delta)^{-\frac{1}{4}}|b|^{\frac{1}{2}}$, $S = b^{\frac{1}{2}}\cdot\nabla(\mu-\Delta)^{-\frac{3}{4}}$ are bounded on L^2 and $\|S\|_{2\to 2}$, $\|H^*\|_{2\to 2} \leq \sqrt{\delta}$ (for all $\mu \geq \lambda_\delta$ and thus for all $\mu \geq \kappa_d\lambda$), see (14.11) and the discussion after that formula. Thus, the quadratic form

$$\tau[v,\eta] := \langle (\mu-\Delta)^{\frac{3}{4}}v, (\mu-\Delta)^{\frac{3}{4}}\eta\rangle + \langle S(\mu-\Delta)^{\frac{3}{4}}v, H^*(\mu-\Delta)^{\frac{3}{4}}\eta\rangle$$

is bounded and coercive on $\mathcal{W}^{\frac{3}{2},2}$ (endowed with the norm $\|\eta\|_{\mathcal{W}^{\frac{3}{2},2}} := \|(\mu-\Delta)^{\frac{3}{2}}\eta\|_2$):

$$|\tau[v,\eta]| \leq (1+\delta)\|v\|_{\mathcal{W}^{\frac{3}{2},2}}\|\eta\|_{\mathcal{W}^{\frac{3}{2},2}}$$

and

$$|\tau[v,\eta]| \geq (1-\delta)\|v\|_{\mathcal{W}^{\frac{3}{2},2}}\|\eta\|_{\mathcal{W}^{\frac{3}{2},2}},$$

(the assumptions of Theorem 15.1 imply, of course, that $\delta < 1$). Extending τ to $\mathcal{W}^{\frac{3}{2},2}$, and returning to the discussion of the weak solution $u \in \mathcal{W}^{\frac{3}{2},2}$ constructed in (15.8), we obtain that $\tau[u-v,\eta] = 0$ for all $\eta \in \mathcal{W}^{\frac{3}{2},2}$, but on the other hand, by coercivity, $\tau[u-v,u-v] \geq (1-\delta)\|u-v\|^2_{\mathcal{W}^{\frac{3}{2},2}}$. Thus, $u=v$.

Remark 15.4. The operator-valued function $\mu \mapsto \Theta_p(\mu,b) \in \mathcal{B}(L^p)$ in Theorem 15.1 is the resolvent of the generator of a *quasi bounded* semigroup in L^p. This was proved in [Ki3].

Remark 15.5. There is a non-trivial difference between the resolvent representations Φ_2 and Θ_p. For instance, Θ_p is nonlinear in $|b|$ even if $p=2$, however Φ_2 is linear $|b|$. This circumstance was used in [Ki4] to extend Theorem 15.1 to measure-valued drifts of the form

$$b = \mathsf{f}dx + \mathsf{h}, \tag{15.10}$$

where f is a vector field in $\mathbf{F}_\delta^{1/2}$ and h is a $\mathbb{R}^d$-valued measure in the Kato class $\bar{\mathbf{K}}_\delta^{d+1}$, i.e.

$$\sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d}(\lambda-\Delta)^{-\frac{1}{2}}(x,y)|\mathsf{h}|_1(dy) \leq \delta \tag{15.11}$$

for some $\lambda = \lambda_\delta$. Here $|\mathsf{h}|_1$ denotes the sum of total variations of the components of h. This class contains, of course, the Kato class of vector fields $\mathbf{K}_\delta^{d+1}$ defined in the examples above. We could also define the class of weakly form-bounded measure-valued drifts. Indeed, Definition 14.1 can be stated as

$$\int_{\mathbb{R}^d} |b(x)|_1 \big[(\lambda-\Delta)^{-\frac{1}{4}} f(x)\big]^2 dx \leq \delta \|f\|_2^2, \quad f \in \mathcal{S},$$

where $|b|_1$ is the sum of absolute values of the components of b. Replacing absolute values by total variations, we arrive at a more general condition

$$\int_{\mathbb{R}^d} |b(dx)|_1 \big[(\lambda-\Delta)^{-\frac{1}{4}} f(x)\big]^2 \leq \delta \|f\|_2^2. \tag{15.12}$$

An example of a measure-valued b satisfying (15.12) is (15.10). In [Ki4], the additional constraint that b must be of the form (15.10) comes from the construction of a regularization of b that preserves weak form-bound δ. Although from purely analytic point of view (15.12) is an L^1 condition on $|b|$, from the operator-theoretic point of view (15.12) is still an L^2 condition, i.e. is an (L^2, L^2) operator norm inequality. Thus, (15.10) may be viewed as a way that (15.12) filters out the singular measure component h, because the latter satisfies, in the dual formulation of the Kato class, an (L^1, L^1) operator norm inequality (cf. (14.4)).

Let us add that, at the level of SDEs, Bass-Chen [BC] and Kim-Song [KSo] considered Kato class measure-valued drifts.

2. We now state our result on weak well-posedness of SDEs with weakly form-bounded drifts.

Theorem 15.2. *Let $b \in \mathbf{F}_\delta$ with $\delta < \frac{4(d-2)}{(d-1)^2}$. Let $e^{-t\Lambda_{C_\infty}(b)}$ be the Feller semigroup constructed in Theorem 15.1. Fix $T>0$. The following is true:*

(i) There exist probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ on the canonical space $(C[0,T], \mathcal{B}_t)$ such that

$$\mathbb{E}_{\mathbb{P}_x}[f(X_t)] = (e^{-t\Lambda_{C_\infty}(b)} f)(x), \quad f \in C_\infty, \quad x \in \mathbb{R}^d.$$

For every $x \in \mathbb{R}^d$ the measure $\mathbb{P}_x$ is a weak solution to SDE

$$X_t = x - \int_0^t b(X_r)dr + \sqrt{2} W_t, \quad 0 \leq t \leq T. \tag{15.13}$$

(ii) If $\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d}$ is another weak solution to (6.15) such that

$$\mathbb{Q}_x = w\text{-}\lim_n \mathbb{P}_x(\tilde{b}_n) \quad \textit{for every } x \in \mathbb{R}^d,$$

for some $\{\tilde{b}_n\} \subset \mathbf{F}_{\delta_1}$ with $\delta < \frac{4(d-2)}{(d-1)^2}$ and λ_δ independent of n, then $\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d} = \{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$.

This result was proved in [KiS1]. In fact, in the series of papers on well-posedness of SDEs with form-bounded and form-bounded-type drifts that are discussed in this work [KiS1] appeared first. In turn, [KiS1] was born out of the attempts to obtain a more detailed description of the corresponding Feller semigroup (i.e. the one constructed earlier in [Ki3]).

Remark 15.6. There are two reasons why one might want to assume form-boundedness of $b = b(x)$ and not its weak form-boundedness: the possibility to include discontinuous diffusion coefficients as in Section 11 and less restrictive assumptions on δ. (One can compare, using $\mathbf{F}_\delta \subset \mathbf{F}^{1/2}_{\sqrt{\delta}}$, the assumptions on δ in Theorem 6.2 and in Theorem 15.2)

The pointwise estimate (15.4) is also valid for the resolvent of $-\nabla \cdot a \cdot \nabla$ provided that the uniformly elliptic matrix a is Hölder continuous. If we were to extend Theorem 15.2 to non-constant diffusion coefficients in the spirit of Section 11, then we could require that a has Hölder continuous entries whose derivatives are weakly form-bounded.

For a form-bounded drift $b \in \mathbf{F}_\delta$, we had two types of gradient bounds on solution u to $(\mu - \Delta + b \cdot \nabla)u = f$, $f \in C_c^\infty$ (let us assume here, for simplicity, that b is bounded and smooth, so we discuss gradient bounds with constants that do not depend on boundedness or smoothness of b but depend only on δ and λ_δ). That is, we had

$$\|(\mu - \Delta)^{\frac{1}{2}+\frac{1}{q}} u\|_p^p \leq K_1 \|f\|_p^p, \quad q > p, \tag{15.14}$$

(Theorem 6.1(ii)) and

$$\|\nabla |\nabla u|^{\frac{p}{2}}\|_2^2 \leq K_2 \|f\|_p^p, \tag{15.15}$$

proved in [KS] using test function $\varphi = -\nabla \cdot (\nabla u |\nabla u|^{p-2})$. In both estimates $p \in [2, \frac{2}{\sqrt{\delta}}[$. These estimates were discussed in Remark 6.1.

Theorem 15.2(ii) provides an analogue of (15.14) for weakly form-bouned $b \in \mathbf{F}_\delta^{1/2}$:

$$\|(\mu - \Delta)^{\frac{1}{2}+\frac{1}{2q}} u\|_p^p \leq K_1 \|f\|_p^p, \quad q > p, \tag{15.16}$$

for $p \in I_\delta$. Does there exist an analogue of (15.15) for weakly form-bounded $b \in \mathbf{F}_\delta^{1/2}$? The answer is "yes". Let us note first that, by the solution representation (15.8),

$$(1 - c_p m_d \delta)\| |b|_1^{-\frac{1}{p'}} (\mu - \Delta) u\|_p \leq \| |b|_1^{-\frac{1}{p'}} f\|_p, \quad |b|_1 := |b| + 1 \tag{15.17}$$

(cf. Remark 6.2), where $c_p m_d \delta < 1$ since $p \in I_\delta$. Without loss of generality, p is rational with odd denominator, so that we can raise functions taking negative values to power p. Also, since all our assumptions on δ are strict inequalities, we may assume, without loss of generality, that $\|T_p(\mu, |b|_1)\|_{p \to p} \leq m_d c_p \delta$ for μ sufficiently large. (Here $T_p(\mu, |b|_1) = |b|_1^{1/p} (\mu - \Delta)^{-\frac{1}{2}} |b|_1^{1/p'}$). Now, we multiply equation $(\mu - \Delta + b \cdot \nabla)u = f$ by test function

$$\varphi := [(\mu - \Delta)^{\frac{1}{2}} u]^{p-1}$$

and integrate:

$$\langle (\mu - \Delta) u, [(\mu - \Delta)^{\frac{1}{2}} u]^{p-1} \rangle + \langle b \cdot \nabla, [(\mu - \Delta)^{\frac{1}{2}} u]^{p-1} \rangle = \langle f, [(\mu - \Delta)^{\frac{1}{2}} u]^{p-1} \rangle.$$

We treat each term separately:

1. We have

$$\langle b \cdot \nabla u, [(\mu - \Delta)^{\frac{1}{2}} u]^{p-1} \rangle = \langle b^{\frac{1}{p}} \cdot \nabla (\mu - \Delta)^{-1} |b|_1^{\frac{1}{p'}} |b|_1^{-\frac{1}{p'}} (\mu - \Delta) u, \big[|b|^{\frac{1}{p}} (\mu - \Delta)^{\frac{1}{2}} u\big]^{p-1} \rangle$$

$$= \langle b^{\frac{1}{p}} \cdot \nabla (\mu - \Delta)^{-1} |b|_1^{\frac{1}{p'}} |b|_1^{-\frac{1}{p'}} (\mu - \Delta) u, \big[|b|^{\frac{1}{p}} (\mu - \Delta)^{-\frac{1}{2}} |b|_1^{\frac{1}{p'}} |b|_1^{-\frac{1}{p'}} (\mu - \Delta) u\big]^{p-1} \rangle$$

(we use (15.4) and apply Hölder's inequality)

$$\leq m_d \| |b|^{\frac{1}{p}} (\kappa_d^{-1} \mu - \Delta)^{-\frac{1}{2}} |b|_1^{\frac{1}{p'}} \|_{p \to p} \| |b|^{\frac{1}{p}} (\mu - \Delta)^{-\frac{1}{2}} |b|_1^{\frac{1}{p'}} \|_{p \to p}^{p-1} \| |b|_1^{-\frac{1}{p'}} (\mu - \Delta) u\|_p^p$$

$$\leq m_d \|T_p(\kappa_d^{-1} \mu, |b|_1)\|_{p \to p} \|T_p(\mu, |b|_1)\|_{p \to p}^{p-1} \| |b|_1^{-\frac{1}{p'}} (\mu - \Delta) u\|_p^p$$

$$\leq m_d c_p^p \delta^p \| |b|_1^{-\frac{1}{p'}} (\mu - \Delta) u\|_p^p$$

(we apply (15.17))

$$\leq m_d c_p^p \delta^p (1 - c_p m_d \delta)^{-p} \|(|b| + 1)^{-\frac{1}{p'}} f\|_p^p \leq m_d c_p^p \delta^p (1 - c_p m_d \delta)^{-p} \|f\|_p^p.$$

2. Next,

$$\begin{aligned}\langle(\mu-\Delta)u,[(\mu-\Delta)^{\frac{1}{2}}u]^{p-1}\rangle &= \langle(\mu-\Delta)^{\frac{1}{2}}\,(\mu-\Delta)^{\frac{1}{2}}u,[(\mu-\Delta)^{\frac{1}{2}}u]^{p-1}\rangle \\ &\quad ((\mu-\Delta)^{\frac{1}{2}} \text{ is a symmetric Markov generator, so we apply (15.2)}) \\ &\geq \frac{4}{pp'}\|(\mu-\Delta)^{\frac{1}{4}}[(\mu-\Delta)^{\frac{1}{2}}u]^{\frac{p}{2}}\|_2^2 \\ &\geq \frac{2}{pp'}\|(\lambda-\Delta)^{\frac{1}{4}}[(\mu-\Delta)^{\frac{1}{2}}u]^{\frac{p}{2}}\|_2^2+\frac{C}{pp'}\|(\lambda-\Delta)^{\frac{1}{2}}u\|_p^p,\end{aligned}$$

where constant C is from the fractional Sobolev embedding theorem.

3. Also,

$$\langle f,[(\lambda-\Delta)^{\frac{1}{2}}u]^{p-1}\rangle \leq \|f\|_p\|(\lambda-\Delta)^{\frac{1}{2}}u\|_p^{p-1}.$$

Combining 1-3, we obtain the following result. Let $b\in\mathbf{F}_\delta^{1/2}$ with $m_d\delta<1$. Let $p\in I_\delta$. Then

$$\|(\lambda-\Delta)^{\frac{1}{4}}[(\lambda-\Delta)^{\frac{1}{2}}u]^{\frac{p}{2}}\|_2^2 \leq K\|f\|_p^p \tag{15.18}$$

for λ sufficiently large. The estimate (15.18) is the analogue of (15.15) for $b\in\mathbf{F}_\delta^{1/2}$. Note that it gives the same Hölder continuity of u as (15.16), cf. Remark 6.1.

The proof of Lemma 15.1, and hence the proof of Theorem 15.1, use inequalities for *symmetric* Markov generators, see Remark 15.2, and thus depend in an essential manner on the fact that we are working in the elliptic setting. Below we will treat time-inhomogeneous drifts at expense of restricting the class $\mathbf{F}_\delta^{1/2}$, but, from some points of view, not by much. At the same time, we will substantially strengthen all aspects of Theorems 9.1 and 10.1 except for their assumptions on δ.

16. Time-inhomogeneous drifts in Morrey class

In this section we consider drifts in the parabolic Morrey class E_q with integrability parameter $q>1$ that can be chosen arbitrarily close to 1. Define parabolic cylinder

$$C_r(t,x):=\{(s,y)\in\mathbb{R}^{d+1}\mid t\leq s\leq t+r^2,|x-y|\leq r\}$$

and, given a vector field $b:\mathbb{R}^{d+1}\to\mathbb{R}^d$ with components in $L^q_{\rm loc}(\mathbb{R}^{d+1})$, $q\in[1,d+2]$, set

$$\begin{aligned}\|b\|_{E_q} &:= \sup_{r>0,z\in\mathbb{R}^{d+1}} r\biggl(\frac{1}{|C_r|}\int_{C_r(z)}|b(t,x)|^q dtdx\biggr)^{\frac{1}{q}} \\ &= \sup_{r>0,z\in\mathbb{R}^{d+1}} r\biggl(\frac{1}{|C_r|}\int_{C_r(z)}|b(-t,x)|^q dtdx\biggr)^{\frac{1}{q}}.\end{aligned}$$

Definition. We say that a vector field b belongs to the parabolic Morrey class E_q if $\|b\|_{E_q}<\infty$.

One has

$$\|b\|_{E_q}\leq\|b\|_{E_{q_1}} \qquad \text{if } q<q_1.$$

So, the smaller is q the larger is Morrey class E_q.

If above $b=b(x)$, then one obtains the usual elliptic Morrey class M_q defined earlier.

Examples. 1. The critical Ladyzhenskaya-Prodi-Serrin class

$$|b| \in L^l(\mathbb{R}, L^p(\mathbb{R}^d)), \quad p \geq d, l \geq 2, \quad \frac{d}{p} + \frac{2}{l} \leq 1$$

is contained in E_q. To prove the inclusion it suffices to consider only the cases $l = 2$, $p = \infty$ and $l = \infty$, $p = d$ (see the argument in Appendix B(3)). In the former case the inclusion is trivial, in the latter case the inclusion follows using Hölder's inequality.

This example is strengthened in the next two examples.

2. Let $|b| \in L^{2,w}(\mathbb{R}, L^\infty(\mathbb{R}^d))$. Here and below, $L^{p,w}$ denotes weak Lebesgue spaces (Appendix B(4)). Then $b \in E_q$, $1 < q < 2$. Indeed, by a well known characterization of weak Lebesgue spaces, we have, setting $\tilde{b}(t) := \|b(t,\cdot)\|_{L^\infty(\mathbb{R}^d)}$,

$$r\biggl(\frac{1}{|C_r|}\int_{C_r} |b|^q dz\biggr)^{\frac{1}{q}} \leq Cr\biggl(\frac{1}{r^2}\int_t^{t+r^2} |\tilde{b}|^q ds\biggr)^{\frac{1}{q}} \leq C\|\tilde{b}\|_{L^{2,w}(\mathbb{R})}.$$

Hence, for example, a vector field b that satisfies

$$\|b(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} \leq \frac{C}{\sqrt{t}}, \quad t > 0 \tag{16.1}$$

(and defined to be zero for $t \leq 0$) is in E_q with $1 < q < 2$.

This example shows that the parabolic Morrey class E_q with $1 < q < 2$ contains vector fields that can have stronger singularities in the time variable than the vector fields in $L^\infty\mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R})$ considered in the previous section. Namely, if, for simplicity, b depends only on time, that it will be in E_q, $1 < q < 2$ if e.g. $|b(t)| \in L^{2,w}(\mathbb{R})$ (as (16.1) above). However, to be in $L^\infty\mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R})$ it would have to satisfy more restrictive condition $|b(t)| \in L^2_{\rm loc}(\mathbb{R})$.

3. By the well known inclusion of the weak Lebesgue space $L^{d,w}(\mathbb{R}^d)$ in M_q,

$$|b| \in L^\infty(\mathbb{R}, L^{d,w}(\mathbb{R}^d)) \quad \Rightarrow \quad b \in E_q \text{ with } 1 < q \leq d.$$

4. For every $\varepsilon > 0$, one can find $b \in E_q$ such that $|b|$ is not in $L^{q+\varepsilon}_{\rm loc}(\mathbb{R}^{d+1})$. So, selecting $q > 1$ close to 1, one obtains $b \in E_q$ that are not in $L^{1+\epsilon}_{\rm loc}(\mathbb{R}^{d+1})$, $\epsilon > 0$.

5. In view of the inclusion $\mathbf{F}_\delta$ (with $c_\delta = 0$) in M_2, we obtain that $\mathbf{F}_\delta \subset E_q$ with $1 < q \leq 2$. Furthermore, combining this with example 2, we obtain that

$$L^\infty\mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}) \subset E_q, \quad 1 < q \leq 2.$$

We now state our results for drifts b in E_q, $1 < q < 2$. Set for $0 < \alpha \leq 2$

$$(\lambda - \partial_t - \Delta)^{-\frac{\alpha}{2}} h(t,x) := \int_t^\infty \int_{\mathbb{R}^d} e^{-\lambda(s-t)} \frac{1}{(4\pi(s-t))^{\frac{d}{2}}} \frac{1}{(s-t)^{\frac{2-\alpha}{2}}} e^{-\frac{|x-y|^2}{4(s-t)}} h(s,y) ds dy, \tag{16.2}$$

$$(\lambda + \partial_t - \Delta)^{-\frac{\alpha}{2}} h(t,x) := \int_{-\infty}^t \int_{\mathbb{R}^d} e^{-\lambda(t-s)} \frac{1}{(4\pi(t-s))^{\frac{d}{2}}} \frac{1}{(t-s)^{\frac{2-\alpha}{2}}} e^{-\frac{|x-y|^2}{4(t-s)}} h(s,y) ds dy, \tag{16.3}$$

where $\lambda \geq 0$. By a standard result, if $\lambda > 0$, then these operators are bounded on $L^p(\mathbb{R}^{d+1})$, $1 \leq p \leq \infty$, with operator norm $\lambda^{-\frac{\alpha}{2}}$. If $\lambda > 0$, then $(\lambda \pm \partial_t - \Delta)^{-1}$ is the resolvent of a Markov generator on $L^p(\mathbb{R}^{d+1})$, $1 \leq p < \infty$, which we will denote by $\lambda \pm \partial_t - \Delta$, respectively. In particular, one has well defined fractional powers $(\lambda \pm \partial_t - \Delta)^{\frac{\alpha}{2}}$. We refer to [B, Go] for the properties of these operators.

Define for $p \in]1,\infty[$

$$G_p := b^{\frac{1}{p}}(\lambda + \partial_t - \Delta)^{-\frac{1}{2p}}, \qquad R_p := b^{\frac{1}{p}} \cdot \nabla(\lambda + \partial_t - \Delta)^{-\frac{1}{2}-\frac{1}{2p}},$$

$$Q_p \upharpoonright \mathcal{E} := (\lambda + \partial_t - \Delta)^{-\frac{1}{2p'}}|b|^{\frac{1}{p'}}$$

and

$$T_p \upharpoonright \mathcal{E} := R_p Q_p,$$

where $\mathcal{E} := \cup_{\varepsilon>0} e^{-\varepsilon|b|} L^p(\mathbb{R}^{d+1})$, a dense subspace of $L^p(\mathbb{R}^{d+1})$.

The following is an analogue of Lemmas 6.1 and 15.1.

Lemma 16.1. *Let $b = b_\mathfrak{s} + b_\mathfrak{b}$, satisfy (16.10). Then, for every $p \in]1,\infty[$, for all $\lambda > 0$, the operators G_p, R_p, Q_p admit extension to L^p by continuity, and thus so does T_p. Moreover,*

$$\|G_p\|_{L^p(\mathbb{R}^{d+1})\to L^p(\mathbb{R}^{d+1})}, \|R_p\|_{L^p(\mathbb{R}^{d+1})\to L^p(\mathbb{R}^{d+1})} \leq C_{d,p,q}\|b_\mathfrak{s}\|_{E_q}^{\frac{1}{p}} + c\lambda^{-\frac{1}{2p}}\|b_\mathfrak{b}\|_{L^\infty(\mathbb{R}^{d+1})}^{\frac{1}{p}} \tag{16.4}$$

$$\|Q_p\|_{L^p(\mathbb{R}^{d+1})\to L^p(\mathbb{R}^{d+1})} \leq C'_{p,q}\|b_\mathfrak{s}\|_{E_q}^{\frac{1}{p'}} + c'\lambda^{-\frac{1}{2p'}}\|b_\mathfrak{b}\|_{L^\infty(\mathbb{R}^{d+1})}^{\frac{1}{p'}}. \tag{16.5}$$

Remark 16.1. The operators G_p and R_p in Lemma 16.1 correspond operators $G_p(q)$, $R_p(r)$ in Lemmas 6.1 and 15.1 with "$q = r = p$", which is impossible in Lemmas 6.1 and 15.1 if one is dealing with form-bounded and weakly form-bounded drifts (they require $r < p < q$). As a consequence, Lemma 16.1 deals with the operator T_p very easily, since T_p is now a product of two bounded operators in L^p. Such decomposition of T_p is impossible for larger classes of form-bounded and weakly form-bounded drifts, see proofs of Lemmas 6.1 and 15.1.

The first estimate (16.4) follows from the boundedness of parabolic Riesz transforms (see [Go]) and the following result: let $|b| \in E_q$ for some $q > 1$ close to 1, then, for every $p \in]1,\infty[$, there exists a constant $c_{p,q}$ such that

$$\||b|^{\frac{1}{p}}(\pm\partial_t - \Delta)^{-\frac{1}{2p}}\|_{L^p(\mathbb{R}^{d+1})\to L^p(\mathbb{R}^{d+1})} \leq c_{p,q}\|b\|_{E_q}^{\frac{1}{p}}. \tag{16.6}$$

In the time homogeneous case $b = b(x)$, the estimate on $\||b|^{\frac{1}{p}}(\lambda - \Delta)^{-\frac{1}{2p}}\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)}$ in terms of the elliptic Morrey norm of $|b|$ is due to [A, Theorem 7.3]. Similar estimates in the parabolic case were obtained in [Kr3]. Lemma 16.1 is proved in [Ki5] by adapting the arguments from [Kr3, proof of Prop. 4.1].

The estimate (16.5) follows from (16.6) by duality.

Define

$$b_n := \mathbf{1}_n b, \tag{16.7}$$

where $\mathbf{1}_n$ is the indicator of the set $\{(t,x) \in \mathbb{R}^{d+1} \mid |(t,x)| \leq n, |b(t,x)| \leq n\}$. We can additionally mollify b_n to obtain a C^∞ smooth approximation of b such that the Morrey norm of the approximating vector field does not exceed $(1+\varepsilon)\|b\|_{E_q}$ for any fixed $\varepsilon > 0$. However, regularization (16.7) of b will suffice. In particular, we will be able to apply Itô's formula to solutions of parabolic equations with drift b_n.

Armed with Lemma 16.1, one obtains the following result [Ki5]. For every $p \in]1,\infty[$, there exist constants $c_{d,p,q}$ and $\lambda_{d,p,q}$ such that if

$$\|b_\mathfrak{s}\|_{E_q} < c_{d,p,q},$$

then, for every $\lambda \geq \lambda_{d,p,q}$, solutions $u_n \in L^p(\mathbb{R}^{d+1})$ to the approximating parabolic equations

$$(\lambda + \partial_t - \Delta + b_n \cdot \nabla)u_n = f, \quad f \in L^p(\mathbb{R}^{d+1})$$

converge in[16] $\mathbb{W}^{1+\frac{1}{p},p}(\mathbb{R}^{d+1})$ to

$$u := (\lambda + \partial_t - \Delta)^{-1} f - (\lambda + \partial_t - \Delta)^{-\frac{1}{2}-\frac{1}{2p}} Q_p(1+T_p)^{-1} R_p(\lambda + \partial_t - \Delta)^{-\frac{1}{2p'}} f. \tag{16.8}$$

Moreover, this u is a unique weak solution to $(\lambda + \partial_t - \Delta + b \cdot \nabla)u = f$, appropriately defined, see [Ki2]. If above $p > d+1$, then, by (16.8) and by the parabolic Sobolev embedding, the convergence is uniform on $\mathbb{R}^{d+1}$ and $u \in C_\infty(\mathbb{R}^{d+1})$.

Let us now construct a Feller evolution family. Let $\delta_{s=r}$ denote the delta-function in the time variable s. Put

$$(\lambda + \partial_t - \Delta)^{-1}\delta_{s=r} g(t,x) := \mathbf{1}_{t \geq r} e^{-\lambda(t-r)}(4\pi(t-r))^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{4(t-r)}} g(y)dy,$$

$$\nabla(\lambda + \partial_t - \Delta)^{-\frac{1}{2}-\frac{1}{2p'}}\delta_{s=r} g := \mathbf{1}_{t \geq r} e^{-\lambda(t-r)}(t-r)^{-\frac{1}{2}+\frac{1}{2p'}}(4\pi(t-r))^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla_x e^{-\frac{|x-y|^2}{4(t-r)}} g(y)dy.$$

Fix $T > 0$. For given $n = 1, 2, \dots$ and $0 \leq r < T$, let v_n denote the classical solution to Cauchy problem

$$\begin{cases} (\lambda + \partial_t - \Delta + b_n(t,x) \cdot \nabla)v_n = 0 \quad (t,x) \in]r,T] \times \mathbb{R}^d, \\ v_n(r, \cdot) = g(\cdot) \in C_\infty, \end{cases} \tag{16.9}$$

where b_n's are defined by (16.7). By a standard result, for every n, the operators

$$U_n^{t,r} g := v_n(t), \quad 0 \leq r \leq t \leq T$$

constitute a Feller evolution family on C_∞. Recall $D_T = \{0 \leq r \leq t \leq T\}$.

Theorem 16.1. *Let $b = b_{\mathfrak{s}} + b_{\mathfrak{b}}$, where*

$$|b_{\mathfrak{s}}| \in E_q \text{ for some } q > 1 \text{ close to } 1, \text{ and } |b_{\mathfrak{b}}| \in L^\infty(\mathbb{R}^{d+1}) \tag{16.10}$$

(indices $\mathfrak{s}$ and $\mathfrak{b}$ stand for "singular" and "bounded", respectively). Fix $p > d+1$. There exist constants $c_{d,p,q}$ and $\lambda_{d,p,q}$ such that if $\|b_{\mathfrak{s}}\|_{E_q} < c_{d,p,q}$, then the following are true:

(i) The limit

$$U^{t,r} := s\text{-}C_\infty(\mathbb{R}^d)\text{-}\lim_n U_n^{t,r} \quad \text{uniformly in } (r,t) \in D_T$$

exists and determines a Feller evolution family on $C_\infty(\mathbb{R}^d)$.

(ii) For every initial function $g \in C_\infty(\mathbb{R}^d) \cap W^{1,p}(\mathbb{R}^d)$, $v(t) := U^{t,r} g$, where $(r,t) \in D_T$, has representation

$$v = (\lambda + \partial_t - \Delta)^{-1}\delta_{s=r} g - (\lambda + \partial_t - \Delta)^{-\frac{1}{2}-\frac{1}{2p}} Q_p(1+T_p)^{-1} G_p S_p g, \tag{16.11}$$

where $S_p g := \nabla(\lambda + \partial_t - \Delta)^{-\frac{1}{2}-\frac{1}{2p'}}\delta_{s=r} g$ satisfies

$$\|S_p g\|_{L^p(\mathbb{R}^{d+1})} \leq C_{p,d} \|\nabla g\|_{L^p(\mathbb{R}^d)}.$$

(iii) As a consequence of (16.11) and the parabolic Sobolev embedding, we obtain

$$\sup_{(r,t) \in D_T, x \in \mathbb{R}^d} |v(t,x;r)| \leq C \|g\|_{W^{1,p}(\mathbb{R}^d)}.$$

[16]Here $\mathbb{W}^{\alpha,p}(\mathbb{R}^{d+1}) := (\lambda + \partial_t - \Delta)^{-\frac{\alpha}{2}} L^p(\mathbb{R}^{d+1})$ endowed with the norm $\|h\|_{\mathbb{W}^{\alpha,p}} := \|(\lambda + \partial_t - \Delta)^{-\frac{\alpha}{2}} h\|_p$.

Theorem 16.1 was proved in [Ki5].

Define backward Feller evolution family $(0 \leq t \leq r \leq T)$

$$P^{t,r}(b) = U^{T-t,T-r}(\tilde{b}), \quad \tilde{b}(t,x) = b(T-t,x),$$

where $U^{t,s}$ is the Feller evolution family from Theorem 16.1. As was explained in the previous section, weak well-posedness of SDE

$$X_t = x - \int_0^b b(s,X_s)ds + \sqrt{2}W_t, \quad t \geq 0, \tag{16.12}$$

follows from appropriate regularity results for the corresponding inhomogeneous parabolic equation (10.5). Indeed, the solution representations (16.8) and (16.11) can be combined and, furthermore, localized, which yields an analogue of gradient estimates (10.3) and thus allows to prove (see [Ki5]) the following result:

Theorem 16.2. *Under the assumptions of Theorem 16.1, the following are true:*

(*i*) *The backward Feller evolution family $\{P^{t,r}\}_{0 \leq t \leq r \leq T}$ is conservative, i.e. the density $P^{t,r}(x,\cdot)$ satisfies*

$$\langle P^{t,r}(x,\cdot)\rangle = 1 \quad \text{for all } x \in \mathbb{R}^d,$$

and determines probability measures $\mathbb{P}_x$, $x \in \mathbb{R}^d$ on $(C([0,T],\mathbb{R}^d),\mathcal{B}_t)$, such that

$$\mathbb{E}_x[f(\omega_r)] = P^{0,r}f(x), \quad 0 \leq r \leq T, \quad f \in C_\infty(\mathbb{R}^d).$$

(*ii*) *For every $x \in \mathbb{R}^d$, the probability measure $\mathbb{P}_x$ is a weak solution to* (16.12)*.*

(*iii*) *For every $x \in \mathbb{R}^d$ and f satisfying* (16.10)*, given a $p > d+1$ as in Theorem 16.1 (generally speaking, the larger p is the smaller $\|b_{\mathfrak{s}}\|_{E_q}$ has to be), there exists constant c such that for all $h \in C_c(\mathbb{R}^{d+1})$*

$$\mathbb{E}_x \int_0^T |\mathsf{f}(r,\omega_r)h(r,\omega_r)|dr \leq c\|\mathbf{1}_{[0,T]}|\mathsf{f}|^{\frac{1}{p}}h\|_{L^p(\mathbb{R}^{d+1})} \tag{16.13}$$

(in particular, one can take $\mathsf{f} = b$). On the other hand, any martingale solution $\mathbb{P}'_x$ to (16.12) *that satisfies for some $p > d+1$ as in Theorem 16.1 the estimate* (16.13) *for $\mathsf{h} := b$ coincides with $\mathbb{P}_x$.*

(*iv*) *For every $x \in \mathbb{R}^d$, given a $\nu > \frac{d+2}{2}$, there exists a constant c such that for all $h \in C_c(\mathbb{R}^{d+1})$ the following Krylov-type bound is true:*

$$\mathbb{E}_x \int_0^T |h(r,\omega_r)|dr \leq c\|\mathbf{1}_{[0,T]}h\|_{L^\nu(\mathbb{R}^{d+1})}. \tag{16.14}$$

On the other hand, if additionally $|b| \in L^{\frac{d+2}{2}+\varepsilon}_{\rm loc}(\mathbb{R}^{d+1})$ for some $\varepsilon > 0$ and $\|b_{\mathfrak{s}}\|_{E_q}$ is sufficiently small, then any martingale solution $\mathbb{P}'_x$ to (16.12) *that satisfies* (16.14) *for some $\nu > \frac{d+2}{2}$ sufficiently close to $\frac{d+2}{2}$ (depending on how small ε is) coincides with $\mathbb{P}_x$.*

We compared the uniqueness results of type (*iii*), (*iv*) in Remark 10.1.

Perhaps, the closest to our Theorems 16.1-16.2 results are contained in recent papers by Krylov [Kr5], [Kr6], [Kr7], which also allow to deal with discontinuous (VMO) diffusion coefficients, see literature review in the introduction.

The estimates (16.13), (16.14) follow from the same kind of solution representations as (16.8), (16.11) above, see [Ki5]. The proofs of the uniqueness results in (*iii*), (*iv*) are similar. Let us prove the uniqueness result in (*iv*). Suppose that we have two martingale solutions $\mathbb{P}_x^1$, $\mathbb{P}_x^2$ of (16.12) that satisfy, for $\nu > \frac{d+2}{2}$ close to $\frac{d+2}{2}$,

$$\mathbb{E}_x^i \int_0^T |h(t,\omega_t)|dt \leq c\|\mathbf{1}_{[0,T]}h\|_\nu, \quad h \in C_c(\mathbb{R}^{d+1}) \tag{16.15}$$

with constant c independent of h $(i=1,2)$. Here and below, $\mathbb{E}_x^1 := \mathbb{E}_{\mathbb{P}_x^1}$, $\mathbb{E}_x^2 := \mathbb{E}_{\mathbb{P}_x^2}$. Our goal is to show: for every $f \in C_c(\mathbb{R}^{d+1})$,

$$\mathbb{E}_x^1[\int_0^T f(t,\omega_t)dt] = \mathbb{E}_x^2[\int_0^T f(t,\omega_t)dt], \tag{16.16}$$

which implies $\mathbb{P}_x^1 = \mathbb{P}_x^2$.

So, let us prove (16.16). Let $u_n \in C([0,T], C_\infty(\mathbb{R}^d))$ be the solution to

$$(\partial_t + \Delta + b_n \cdot \nabla)u_n = f, \quad u_n(T,\cdot) = 0, \tag{16.17}$$

where, recall, $b_n = \mathbf{1}_n b$, and $\mathbf{1}_n$ is the indicator of $\{|b| \leq n\}$. Set $\tau_R := \inf\{t \geq 0 \mid |\omega_t| \geq R\}$, $R > 0$. By Itô's formula

$$\begin{aligned}\mathbb{E}_x^i u_n(T \wedge \tau_R, \omega_{T\wedge\tau_R}) = u_n(0,x) + \mathbb{E}_x^i \int_0^{T\wedge\tau_R} f(t,\omega_t)dt \\ + \mathbb{E}_x^i \int_0^{T\wedge\tau_R} \big[(b-b_n)\cdot\nabla u_n\big](t,\omega_t)dt\end{aligned} \tag{16.18}$$

$(i=1,2)$. We have

$$\begin{aligned}\left|\mathbb{E}_x^i \int_0^{T\wedge\tau_R} \big[(b-b_n)\cdot\nabla u_n\big](t,\omega_t)dt\right| &\leq \mathbb{E}_x^i \int_0^{T\wedge\tau_R} \big[|b|(1-\mathbf{1}_n)|\nabla u_n|\big](t,\omega_t)dt \\ &\text{(we are applying (16.15))} \\ &\leq c\|\mathbf{1}_{[0,T]\times B_R(0)}|b|(1-\mathbf{1}_n)|\nabla u_n|\|_\nu \\ &\leq c\|\mathbf{1}_{[0,T]\times B_R(0)}|b|(1-\mathbf{1}_n)\|_{s'}\|\nabla u_n\|_s, \qquad \frac{1}{s}+\frac{1}{s'}=\frac{1}{q}.\end{aligned}$$

At this point we note that $\tilde{u}_n(t) := e^{\lambda(T-t)}u_n(t)$ satisfies

$$(\lambda + \partial_t + \Delta + b_n\cdot\nabla)u_n = \mathbf{1}_{[0,T]}e^{\lambda(T-t)}f,$$

so a solution representation of type (16.8) (i.e. additionally taking into account the terminal value condition), see [Ki5], and the parabolic Sobolev embedding theorem, yield

$$\|\nabla u_n\|_s \leq C\|f\|_p \quad \text{for } s < \frac{d+2}{d+1}p \ \text{ close to } \frac{d+2}{d+1}p.$$

Assuming that the Morrey norm $\|b\|_{E_q}$ is sufficiently small, we can select p sufficiently large to make s' close to ν and hence close to $\frac{d+2}{2}$. To be more precise, we have by our assumption $|b| \in L^{\frac{d+2}{2}+\varepsilon}$ for some $\varepsilon > 0$, so we need $s' \geq \frac{d+2}{2}+\varepsilon$. Now, since $1-\mathbf{1}_n \to 0$ a.e. on $\mathbb{R}^{d+1}$ as $n \to \infty$, we have $\|\mathbf{1}_{[0,T]\times B_R(0)}|b|(1-\mathbf{1}_n)\|_{s'} \to 0$ as $n \to \infty$. Therefore,

$$\mathbb{E}_x^i \int_0^{T\wedge\tau_R} \big[(b-b_n)\cdot\nabla u_n\big](t,\omega_t)dt \to 0 \quad (n\to\infty).$$

We are left to note, using the convergence result in [Ki5, Cor. 2] (of the same type as Theorem 16.1(i)), that solutions u_n converge to a function $u \in C([0,T], C_\infty(\mathbb{R}^d))$. Therefore, we can pass to the limit in (16.18), first in n and then in $R \to \infty$, to obtain

$$0 = u(0,x) + \mathbb{E}_x^i \int_0^T f(t,\omega_t)dt \quad i = 1,2,$$

which gives (16.16). □

17. SDEs driven by α-stable process

In this section we deal with the SDE

$$X_t = x - \int_0^t b(X_s)ds + Z_t - Z_0, \quad t \geq 0, \quad x \in \mathbb{R}^d, \tag{17.1}$$

where Z_t be a rotationally symmetric α-stable process, $1 < \alpha < 2$, i.e. a Lévy process with characteristic function

$$\mathbb{E}[\exp(i\varkappa \cdot (Z_t - Z_0)] = \exp(-t|\varkappa|^\alpha) \quad \text{for every } \varkappa \in \mathbb{R}^d.$$

The drift $b : \mathbb{R}^d \to \mathbb{R}^d$ is in general locally unbounded.

Recall that a weak solution to (17.1) is a process X_t defined on some probability space having a.s. right continuous trajectories with left limits, such that $\int_0^t |b(X_s)|ds < \infty$ a.s. for every $t > 0$, and such that X_t satisfies (17.1) a.s. for a symmetric α-stable process Z_t.

A weak solution to (17.1), when it exists (e.g. if $|b| \in L^\infty$, see [Ko]), is called α-stable process with drift b. It plays a central role in the study of stochastic processes which, in contrast to the Brownian motion, can have long range interactions.

The operator behind SDE (17.1) is the non-local operator $(-\Delta)^{\frac{\alpha}{2}} + b \cdot \nabla$, i.e. one expects that the transition density of X_t solves the corresponding parabolic equation for $(-\Delta)^{\frac{\alpha}{2}} + b \cdot \nabla$.

We are interested in the same question as in the previous sections: what are the minimal assumptions on the local singularities of the vector field b, not assuming additional structure such as symmetry or existence of the divergence, such that, for an arbitrary initial point, there exists a unique (in appropriate sense) weak solution to (17.1)? This question has been extensively studied in the literature. By the results of Portenko [P2] and Podolynny-Portenko [PP], if

$$|b| \in L^p + L^\infty, \quad \text{for some } p > \frac{d}{\alpha - 1}, \tag{17.2}$$

then there exists a unique in law weak solution to (17.1). Although the exponent $\frac{d}{\alpha-1}$ is the best possible on the Lebesgue scale, the class (17.2) is far from being the maximal admissible: this result has been strengthened in [CKS, CW, KSo] where the authors consider b in the Kato class of vector fields $\mathbf{K}_\delta^{d,\alpha-1}$ (with δ arbitrarily small), i.e. $|b| \in L^1_{\rm loc}$ and

$$\big\|(\lambda + (-\Delta)^{\frac{\alpha}{2}})^{-\frac{\alpha-1}{\alpha}}|b|\big\|_\infty \leq \delta$$

for some $\lambda = \lambda_\delta \geq 0$. This class contains some vector fields b with $|b| \notin L^{1+\varepsilon}_{\rm loc}$, $\varepsilon > 0$, however, it does not contains the class $|b| \in L^{\frac{d}{\alpha-1}} + L^\infty$.

The Kato class $\mathbf{K}_\delta^{d,\alpha-1}$ with δ sufficiently small provides the standard bounds on the heat kernel of the fractional Kolmogorov operator

$$\Lambda(b) \supset (-\Delta)^{\frac{\alpha}{2}} + b \cdot \nabla,$$

i.e.

$$C^{-1}e^{-t(-\Delta)^{\frac{\alpha}{2}}}(x,y) \leq e^{-t\Lambda(b)}(x,y) \leq Ce^{-t(-\Delta)^{\frac{\alpha}{2}}}(x,y), \tag{17.3}$$

for all $x, y \in \mathbb{R}^d$ and $0 < t < T$ for a constant $C = C_T$ [BJ]. It was established in [CKS], among many other results, that the probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ determined by $e^{-t\Lambda(b)}$ solve the martingale problem for $(-\Delta)^{\frac{\alpha}{2}} + b\cdot\nabla$ with test functions in C_c^∞. The uniqueness in law of the weak solution to SDE (17.1) with $b \in \mathbf{K}_\delta^{d,\alpha-1}$ (with δ arbitrarily small) was established in [CW]. Let us also mention that [KSo] considered SDE (17.1) with Kato class measure-valued drift and established the corresponding heat kernel bounds.

We consider a larger class of *weakly form-bounded vector fields*:

Definition 17.1. A vector field $b : \mathbb{R}^d \to \mathbb{R}^d$ with entries in $L^1_{\rm loc}$ is said to be weakly form-bounded if there exist $\delta > 0$ such that

$$\||b|^{\frac{1}{2}}(\lambda + (-\Delta)^{\frac{\alpha}{2}})^{-\frac{\alpha-1}{2\alpha}}\|_{2\to 2} \leq \sqrt{\delta}$$

for some $\lambda = \lambda_\delta > 0$.

This will be written as $b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}$. This definition extends Definition 14.1 from the previous section corresponding to the case $\alpha = 2$.

Our assumptions concerning δ will involve only strict inequalities, so (using e.g. the Spectral theorem) we can re-state our hypothesis on the drift as

$$\||b|^{\frac{1}{2}}(\lambda - \Delta)^{-\frac{\alpha-1}{4}}\|_{2\to 2} \leq \sqrt{\delta}$$

for some $\lambda = \lambda_\delta > 0$.

Examples. 1. Using the fractional Sobolev inequality, it is not difficult to show that

$$|b| \in L^{\frac{d}{\alpha-1}} + L^\infty \quad \Rightarrow \quad b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}},$$

where $\delta > 0$ can be chosen arbitrarily small. More generally, vector fields with entries in $L^{\frac{d}{\alpha-1},w}$ (the weak $L^{\frac{d}{\alpha-1}}$ class) are weakly form-bounded:

$$|b| \in L^{\frac{d}{\alpha-1},\infty} + L^\infty \quad \Rightarrow \quad b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}},$$

with

$$\sqrt{\delta} = \Omega_d^{-\frac{\alpha-1}{2d}} \frac{2^{-\frac{\alpha-1}{2}}\Gamma(\frac{d-\alpha+1}{4})}{\Gamma(\frac{d+\alpha-1}{4})} \|b_1\|^{\frac{1}{2}}_{\frac{d}{\alpha-1},w},$$

where Ω_d is the volume of the unit ball $B(0,1) \subset \mathbb{R}^d$. The proof is obtained easily using [KPS, Corollary 2.9].

2. In particular, by the fractional Hardy inequality,

$$b(x) = \pm\sqrt{\delta}\kappa_{\alpha,d}|x|^{-\alpha}x, \tag{17.4}$$

where

$$\kappa_{\alpha,d} := 2^{\frac{\alpha-1}{2}-}\frac{\Gamma(\frac{d+\alpha-1}{4})}{\Gamma(\frac{d-\alpha+1}{4})},$$

is in $\mathbf{F}_\delta^{\frac{\alpha-1}{2}}$ (with $\lambda = 0$).

The drift (17.4) destroys the standard heat kernel bounds (17.3) (and so it is not in the Kato class). However, for such b sharp heat kernel bounds on $e^{-t\Lambda(b)}(x,y)$ exist but they depend explicitly on δ via an additional factor $\varphi_t(y)$,

$$C^{-1}e^{-t(-\Delta)^{\frac{\alpha}{2}}}(x,y)\varphi_t(y) \leq e^{-t\Lambda(b)}(x,y) \leq Ce^{-t(-\Delta)^{\frac{\alpha}{2}}}(x,y)\varphi_t(y), \quad x,y \in \mathbb{R}^d, \quad y \neq 0.$$

The factor $\varphi_t(y)$ either explodes at the origin or vanishes, depending on the sign of δ [KSSz], [KiS5].

3. The Kato class vector fields are weakly form-bounded:

$$b \in \mathbf{K}_\delta^{d,\alpha-1} \quad \Rightarrow \quad b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}.$$

Indeed, if $b \in \mathbf{K}_\delta^{d,\alpha-1}$, then by duality $\||b|(\lambda+(-\Delta)^{\frac{\alpha}{2}})^{-\frac{\alpha-1}{\alpha}}\|_{1\to1} \leq \delta$, and so by interpolation $\||b|^{\frac{1}{2}}(\lambda+(-\Delta)^{\frac{\alpha}{2}})^{-\frac{\alpha-1}{\alpha}}|b|^{\frac{1}{2}}\|_{2\to2} \leq \delta$, hence $b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}$.

4. The elliptic Morrey class:

$$|b|^{\frac{1}{\alpha-1}} \in M_{1+\varepsilon} \quad \Rightarrow \quad b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}$$

with δ depending on the Morrey norm of $|b|^{\frac{1}{\alpha-1}}$ (see definition (14.2)). Indeed, by [A, Theorem 7.3], one has $\||b|^{\frac{1}{2(\alpha-1)}}(\lambda-\Delta)^{-\frac{1}{4}}\|_{2\to2} \leq \delta^{\frac{1}{2(\alpha-1)}}$. Then, by the Heinz-Kato inequality (i.e. raising operators $|b|^{\frac{1}{2(\alpha-1)}}$ and $(\lambda-\Delta)^{-\frac{1}{4}}$ to power $\alpha-1<1$), we obtain $\||b|^{\frac{1}{2}}(\lambda-\Delta)^{-\frac{\alpha-1}{4}}\|_{2\to2} \leq \sqrt{\delta}$, i.e. $b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}$. This examples contains examples 1 and 2.

Remark 17.1. There is a rich literature on weak and strong well-posedness of SDE (17.1) (and its generalizations) in the case $0<\alpha\leq 1$, in which case $|b|$ is assumed to be (locally) Hölder continuous, say, with exponent β, and satisfy the balance condition $\alpha+\beta>1$ (sub-critical) or $\alpha+\beta=1$ (critical). See Zhao [Zh2], Song-Xie [SX] who considered the case $\alpha+\beta=1$. Regarding the corresponding heat kernel bounds, we refer to Xie-Zhang [XZ] and Menozzi-Zhang [MeZ] who proved the two-sided bound on the heat kernel of $(-\Delta)^{\frac{\alpha}{2}}+b\cdot\nabla$ in the case $\alpha+\beta>1$. Let us add that in the case $\alpha+\beta=1$ the behaviour of the heat kernel changes drastically, for instance, it can vanish, see [KMS] who considered the heat kernel of operator $\Lambda=(-\Delta)^{\frac{\alpha}{2}}-\kappa|x|^{-\alpha}x\cdot\nabla$, $\kappa>0$ (Hardy-type drift) and proved upper bound of the form

$$0 \leq e^{-t\Lambda}(x,y) \leq Ct^{-\frac{d}{\alpha}}[1\wedge t^{-\frac{\gamma}{\alpha}}|y|^{\gamma}], \quad t \in]0,1],$$

where the order of vanishing $\gamma \in]0,\alpha[$ is an explicit function of κ.

For a given vector field $b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}$, we fix a C^∞ smooth approximation

$$b_n := c_n\eta_{\varepsilon_n} * (\mathbf{1}_n b), \quad \varepsilon_n \downarrow 0, \quad n=1,2,\dots,$$

where $\mathbf{1}_n$ is the indicator of $\{x \in \mathbb{R}^d \mid |x| \leq n, |b(x)| \leq n\}$, η_ε is the Friedrichs mollifier. Selecting $\varepsilon_n \downarrow$ sufficiently rapidly and $c_n \uparrow 1$ sufficiently slow, one obtains that

$$b_n \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}, \quad n=1,2,\dots$$

with λ independent of n.

Fix constant $m_{d,\alpha}$ by the pointwise inequality

$$|\nabla_x(\mu+(-\Delta)^{\frac{\alpha}{2}})^{-1}(x,y)| \leq m_{d,\alpha}(\kappa^{-1}\mu+(-\Delta)^{\frac{\alpha}{2}})^{-\frac{\alpha-1}{\alpha}}(x,y) \tag{17.5}$$

for all $x, y \in \mathbb{R}^d$, $x \neq y$, and all $\mu > 0$, for some $\kappa = \kappa_{d,\alpha} > 0$. The following result was proved in [KiM2] (one can find there an elementary estimate on $m_{d,\alpha}$ from above).

Theorem 17.1. *Let $b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}$ with $\delta < m_{d,\alpha}^{-1} 4\big[\frac{d-\alpha}{(d-\alpha+1)^2} \wedge \frac{\alpha(d+\alpha)}{(d+2\alpha)^2}\big]$. The following is true.*

(i) The limit

$$s\text{-}C_\infty\text{-}\lim_n e^{-t\Lambda_{C_\infty}(b_n)} \quad (\text{loc. uniformly in } t \geq 0),$$

where

$$\Lambda_{C_\infty}(b_n) := (-\Delta)^{\frac{\alpha}{2}} + b_n \cdot \nabla \text{ with domain } D(\Lambda_{C_\infty}(b_n)) = \big(1 + (-\Delta)^{\frac{\alpha}{2}}\big)^{-1} C_\infty,$$

exists and determines a Feller semigroup $T^t =: e^{-t\Lambda_{C_\infty}(b)}$. Its generator Λ_{C_∞} is an operator realization of the formal operator $(-\Delta)^{\frac{\alpha}{2}} + b \cdot \nabla$ in C_∞.

(ii) There exists $\mu_0 \geq 0$ such that for all $\mu \geq \mu_0$, for every $p \in [2, p_+[$, $p_+ = \frac{2}{1-\sqrt{1-m_{d,\alpha}\delta}}$, and all $1 < r < p < q < \infty$,

$$(\mu + \Lambda_{C_\infty}(b))^{-1} \upharpoonright C_\infty \cap L^p \text{ extends by continuity to } \mathcal{B}(\mathcal{W}^{-\frac{\alpha-1}{r'},p}, \mathcal{W}^{1+\frac{\alpha-1}{q},p}).$$

In particular, if $p > d - \alpha + 1$, then

$$\big(\mu + \Lambda_{C_\infty}(b)\big)^{-1}[C_\infty \cap L^p] \subset C^{0,\gamma}, \quad \gamma < 1 - \frac{d-\alpha+1}{p}.$$

Also,

$$\big(\mu + \Lambda_{C_\infty}(b)\big)^{-1} \upharpoonright C_\infty \cap L^2 \text{ extends by continuity to } \mathcal{B}(\mathcal{W}^{-\frac{\alpha-1}{2},2}, \mathcal{W}^{\frac{\alpha+1}{2},2}). \tag{17.6}$$

(iii) $e^{-t\Lambda_{C_\infty}(b)}$ is conservative, i.e. $\int_{\mathbb{R}^d} e^{-t\Lambda_{C_\infty}(b)}(x,y)dy = 1 \ \ \forall\, x \in \mathbb{R}^d$.

Let $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ be the probability measures on the canonical space $(D[0,T], \mathcal{B}'_t)$ determined by $e^{-t\Lambda_{C_\infty}(b)}$, i.e.

$$\mathbb{E}_{\mathbb{P}_x}[f(X_t)] = (e^{-t\Lambda_{C_\infty}(b)}f)(x), \quad f \in C_\infty, \quad x \in \mathbb{R}^d.$$

(iv) For every $x \in \mathbb{R}^d$ and $t > 0$, $\mathbb{E}_{\mathbb{P}_x}\int_0^t |b(X_s)|ds < \infty$ and there exists a process Z_t with trajectories in $D(\mathbb{R}_+, \mathbb{R}^d)$, which is a symmetric α-stable process under each $\mathbb{P}_x$, such that

$$X_t = x - \int_0^t b(X_s)ds + Z_t - Z_0, \quad t \geq 0.$$

(v) The Feller property and property (17.6) determine $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ uniquely. That is, suppose that for every $x \in \mathbb{R}^d$ we are given a weak solution $\mathbb{Q}_x$ to SDE (17.1). Define for every $f \in C_c^\infty$

$$R_\mu^Q f(x) := \mathbb{E}_{\mathbb{Q}_x}\int_0^\infty e^{-\mu s} f(X_s)ds, \quad X_s \in D(\mathbb{R}_+, \mathbb{R}^d), \quad x \in \mathbb{R}^d, \quad \mu > \lambda_\delta.$$

If $R_\mu^Q C_c^\infty \subset C_b$ and $R_\mu^Q \upharpoonright C_c^\infty$ admits extension by continuity to $\mathcal{B}(\mathcal{W}^{-\frac{\alpha-1}{2},2}, L^2)$, then

$$\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d} = \{\mathbb{P}_x\}_{x\in\mathbb{R}^d}.$$

We stated assertions (i), (ii) in a form that is somewhat different from Theorems 6.1 or 15.1, but we could have stated it in the form of these theorems as well. The construction of the Feller semigroup (i) goes as in Theorems 6.1 (or rather as in Theorem 15.1 since we use pointwise bound (17.5)). Same for the embedding properties (ii), i.e. we can write an explicit operator-valued function representation for the resolvent as in the cited two theorems.

The proof of the probability conservation property in (*iii*) uses weighted estimates, similarly to the proof of Theorem 6.1 dealing with the local case $\alpha = 2$ (see Section 7, estimate (7.7) there). Set

$$\eta(x) := (1+|x|^2)^\nu, \quad 0<\nu<\frac{\alpha}{2}.$$

Denote $L^p_\eta := L^p(\mathbb{R}^d, \eta^2 dx)$, $\|\cdot\|^p_{p,\eta} := \langle |\cdot|^p \eta^2\rangle$.

Proposition 17.1. *Let $d \geq 3$, $b \in \mathbf{F}_\delta^{\frac{\alpha-1}{2}}$ with $\delta < m_{d,\alpha}^{-1} 4\big[\frac{d-\alpha}{(d-\alpha+1)^2} \wedge \frac{\alpha(d+\alpha)}{(d+2\alpha)^2}\big]$. There exist $0<\nu<\alpha/2$, $p > (d-\alpha+1) \vee (\frac{d}{2\nu}+2)$ and $\mu_0>0$ such that for every $h \in C_c$, $\mu \geq \mu_0$*

$$\|\eta^{-1}(\mu+\Lambda_{C_\infty}(b))^{-1}\eta h\|_\infty \leq K_1 \|h\|_{p,\eta}, \tag{E_1}$$

$$\|\eta^{-1}(\mu+\Lambda_{C_\infty}(b))^{-1}\eta |b_m| h\|_\infty \leq K_2 \||b_m|^{\frac{1}{p}} h\|_{p,\eta}, \tag{E_2}$$

$$\|\eta^{-1}|b_m|^{\frac{1}{p}}(\mu+\Lambda_{C_\infty}(b))^{-1}\eta |b_m| h\|_{p,\eta} \leq K_3 \||b_m|^{\frac{1}{p}} h\|_{p,\eta}, \tag{E_3}$$

where $K_i>0$, $i=1,2,3$, do not depend on $m=1,2,\dots$ The constant K_3 can be chosen arbitrarily small at expense of increasing μ_0.

The proof of these weighted estimates in [KiM2] is, however, quite different from the proof of Theorem 6.1 where one can control easily the commutator of the weight and the Laplacian. [KiM2] provides a different, rather interesting approach to the proof of Proposition 17.1, but we will not discuss it here. Note that if b has compact support then we can take $\eta \equiv 1$.

Let us describe the proof of (*iv*) in [KiM2], which uses the approach of [PP, P2] but in a weighted L^p space. Set

$$Z_t := X_t - X_0 - \int_0^t b(X_s)ds, \quad t \geq 0.$$

Our goal is to prove that under $\mathbb{P}_x$ the process Z_t is a symmetric α-stable process starting at 0. We use notation introduced in the beginning of the previous section. For brevity, write $e^{-t\Lambda(b)} = e^{-t\Lambda_{C_\infty}(b)}$.

1. *Define*

$$w(t,x,\varkappa) = \mathbb{E}_x\Big[e^{i\varkappa\cdot\left(X_t - \int_0^t b(X_s)ds\right)}\Big], \quad t\geq 0, \quad \varkappa \in \mathbb{R}^d.$$

Then w is a bounded solution to integral equation

$$w(t,x,\varkappa) = \int_{\mathbb{R}^d} e^{i\varkappa\cdot y} e^{-t\Lambda(b)}(x,y)dy - i\int_0^t \int_{\mathbb{R}^d} e^{-(t-s)\Lambda(b)}(x,z)(\varkappa\cdot b(z))w(s,z,\varkappa)dzds. \tag{17.7}$$

Indeed, in view of

$$e^{-i\cdot\varkappa\int_0^t b(X_\tau)d\tau} = 1 - i\int_0^t (\varkappa\cdot b(X_s))e^{-i\cdot\varkappa\int_s^t b(X_\tau)d\tau},$$

one has

$$\begin{aligned} w(t,x,\varkappa) &= \mathbb{E}_x\Big[e^{i\varkappa\cdot X_t}\Big] - i\int_0^t \mathbb{E}_x\Big[e^{i\varkappa\cdot X_t}(\varkappa\cdot b(X_s))e^{-i\cdot\varkappa\int_s^t b(X_\tau)d\tau}\Big]ds \\ &= \mathbb{E}_x\Big[e^{i\varkappa\cdot X_t}\Big] - i\int_0^t \mathbb{E}_x\Big[(\varkappa\cdot b(X_s))w(t-s,X_s,\varkappa)\Big]ds \\ &= \int_{\mathbb{R}^d} e^{i\varkappa\cdot y}e^{-t\Lambda(b)}(x,y)dy - i\int_0^t\int_{\mathbb{R}^d} e^{-s\Lambda(b)}(x,z)(\varkappa\cdot b(z))w(t-s,z,\varkappa)dzds. \end{aligned}$$

2. *Set* $\tilde{w}(t,x,\varkappa) := e^{i\varkappa\cdot x - t|\varkappa|^\alpha}$. *This is another bounded solution to* (17.7). Indeed, multiplying the Duhamel formula

$$e^{-t\Lambda}(x,y) = e^{-t(-\Delta)^{\frac{\alpha}{2}}}(x,y) + \int_0^t \langle e^{-(t-s)\Lambda}(x,\cdot)b(\cdot)\cdot\nabla.e^{-s(-\Delta)^{\frac{\alpha}{2}}}(\cdot,y)\rangle ds$$

(which is proved in [KiM2, Corollary 1(*iv*)]) by $e^{i\varkappa\cdot y}$ and then integrating in y, we obtain the required.

Next, let us show that a bounded solution to (17.7) is unique. We will need

3. *For every* $\varkappa \in \mathbb{R}^d$ *there exists* $T = T(\varkappa) > 0$ *such that the mapping*

$$(Hv)(t,x) := -i\int_0^t\int_{\mathbb{R}^d} e^{-(t-s)\Lambda(b)}(x,z)(\varkappa\cdot b(z))v(s,z)dsdz, \quad (t,x) \in [0,T]\times\mathbb{R}^d,$$

is a contraction on $L^p(\mathbb{R}^d, |b|\eta^{-p+2}dx; L^\infty[0,T])$ *(i.e. functions taking values in* $L^\infty[0,T]$*) for* p *as in Proposition 17.1.*

Indeed, we have

$$\begin{aligned} |Hv(t,x)| &\leq \left|\int_0^t \langle e^{-(t-s)\Lambda(b)}(x,\cdot)(\varkappa\cdot b(\cdot))v(s,\cdot)\rangle ds\right| \\ &\leq |\varkappa|\int_0^t \langle e^{-(t-s)\Lambda(b)}(x,\cdot)|b(\cdot)|^{\frac{1}{p'}}|b(\cdot)|^{\frac{1}{p}}|v(s,\cdot)|\rangle ds \\ &\leq |\varkappa|\int_0^t \langle e^{-(t-s)\Lambda(b)}(x,\cdot)|b(\cdot)|^{\frac{1}{p'}}|b(\cdot)|^{\frac{1}{p}}\sup_{\tau\in[0,T]}|v(\tau,\cdot)|\rangle ds \qquad (*)\end{aligned}$$

Let us note that, for every $x \in \mathbb{R}^d$,

$$|b(x)|^{\frac{1}{p}}\eta^{-1}(x)\sup_{t\in[0,T]}\int_0^t \langle e^{-(t-s)\Lambda(b)}(x,\cdot)|b(\cdot)|^{\frac{1}{p'}}\eta(\cdot)\rangle ds$$

(we are applying the Dominated Convergence Theorem)

$$|b(x)|^{\frac{1}{p}}\eta^{-1}(x)\sup_{t\in[0,T]}\lim_m\int_0^t \langle e^{-(t-s)\Lambda(b)}(x,\cdot)|b_m(\cdot)|^{\frac{1}{p'}}\eta(\cdot)\rangle ds,$$

where, in turn, the last term

$$\begin{aligned}&|b|^{\frac{1}{p}}\eta^{-1}\sup_{t\in[0,T]}\lim_m\int_0^t e^{-(t-s)\Lambda(b)}|b_m|^{\frac{1}{p'}}\eta ds \\ &\leq |b|^{\frac{1}{p}}\eta^{-1}e^{\mu T}\lim_m(\mu+\Lambda_{C_\infty}(b))^{-1}|b_m|^{\frac{1}{p'}}\eta \in \mathcal{B}(L^p_\eta) \quad \text{by Proposition 17.1}(E_3).\end{aligned}$$

Also by Proposition 17.1(E_3), selecting μ sufficiently large, and then selecting T sufficiently small, the $L^p_\eta \to L^p_\eta$ norm of the last operator can be made arbitrarily small. Applying this in ($*$), we obtain that H is indeed a contraction on $L^p(\mathbb{R}^d, |b|\eta^{-p+2}dx; L^\infty[0,T])$.

We have $L^\infty([0,T]\times\mathbb{R}^d) \subset L^p(\mathbb{R}^d, |b|\eta^{-p+2}dx; L^\infty[0,T])$ since (see [KiM2, Lemma 5.1]) $|b|\eta^{-p+2} \in L^1(\mathbb{R}^d)$. Combining the assertions of Steps 1-3, we obtain that for every $\varkappa \in \mathbb{R}^d$

$$w(t,x,\varkappa) = \tilde{w}(t,x,\varkappa) \quad \text{in } L^p(\mathbb{R}^d, |b|\eta^{-p+2}dx; L^\infty[0,T]),$$

and thus

$$w(t,x,\varkappa) = \tilde{w}(t,x,\varkappa) \quad \text{for a.e. } x \in \mathbb{R}^d$$

(although $t < T(\varkappa)$, one can get rid of this constraint using the reproduction property of $e^{-t\Lambda(b)}$, so without loss of generality $T \neq T(\varkappa)$). Now, applying the continuity of $\int_0^t e^{-s\Lambda_{C_\infty}(b)} b \cdot w ds$ on $\mathbb{R}_+ \times \mathbb{R}^d$ (this is [KiM2, Corollary 1(iii)]) in the RHS of (17.7), we obtain that for every $\varkappa \in \mathbb{R}^d$ $w(t,x,\varkappa)$ is continuous in t and x, and so $w = \tilde{w}$ everywhere. Thus, for all $t \leq T$, $x \in \mathbb{R}^d$

$$\mathbb{E}_x\Big[e^{i\varkappa \cdot \left(X_t - X_0 - \int_0^t b(X_s)ds\right)}\Big] = e^{-\varkappa \cdot x} w(t,x,\varkappa) = e^{-t|\varkappa|^\alpha}.$$

By a standard result, Z_t is a symmetric α-stable process. The proof of assertion (iv) is completed.

Appendix A. Proof of Lemma 6.1

The following proof was given in [Ki2].

Step 1. Let us show that

$$\|T_p\|_{p\to p} \leq c_{\delta,p}, \quad \mu \geq \mu_0,$$

which will give us assertion (i). We will also prove that operators

$$G_p = b^{\frac{2}{p}} \cdot \nabla(\mu - \Delta)^{-1}, \quad Q_p = (\mu - \Delta)^{-1}|b|^{1-\frac{2}{p}}$$

satisfy

$$\|G_p\|_{p\to p} \leq C_1 \mu^{-\frac{1}{2}+\frac{1}{p}}, \quad \|Q_p\|_{p\to p} \leq C_2 \mu^{-\frac{1}{2}-\frac{1}{p}}$$

The latter will be needed to prove assertions (ii) and (iii).

We will be using the operator-norm formulation of the form-boundedness condition:

$$\|b(\lambda - \Delta)^{-\frac{1}{2}}\|_{2\to 2} \leq \delta$$

for some $\lambda = \lambda_\delta$, see (3.2).

(**a**) Set $v := (\mu - \Delta)^{-1}|b|^{1-\frac{2}{p}} f$, $0 \leq f \in L^p$. Then

$$\begin{aligned}
\|T_p f\|_p^p &= \|b^{\frac{2}{p}} \nabla v\|_p^p = \langle |b|^2 |\nabla v|^p \rangle \\
&= \||b|(\lambda - \Delta)^{-\frac{1}{2}}(\lambda - \Delta)^{\frac{1}{2}} |\nabla v|^{\frac{p}{2}}\|_2^2 \qquad (\lambda = \lambda_\delta) \\
&\leq \||b|(\lambda - \Delta)^{-\frac{1}{2}}\|_{2\to 2}^2 \|(\lambda - \Delta)^{\frac{1}{2}} |\nabla v|^{\frac{p}{2}}\|_2^2 \\
&= \delta \|(\lambda - \Delta)^{\frac{1}{2}} |\nabla v|^{\frac{p}{2}}\|_2^2 = \delta\big(\lambda \|\nabla v\|_p^p + \|\nabla |\nabla v|^{\frac{p}{2}}\|_2^2\big).
\end{aligned}$$

It remains to prove the principal inequality

$$\delta\big(\lambda \|\nabla v\|_p^p + \|\nabla |\nabla v|^{\frac{p}{2}}\|_2^2\big) \leq c_{\delta,p}^p \|f\|_p^p, \tag{$*$}$$

and conclude that $\|T_p\|_{p\to p} \leq c_{\delta,p}$.

First, we prove an a priori variant of ($*$), i.e. for $v := (\mu - \Delta)^{-1}|b|^{1-\frac{2}{p}} f$ with $b = b_n$. Since our assumptions on δ involve only strict inequalities, we may assume, upon selecting appropriate $\varepsilon_n \downarrow 0$, that $b_n \in \mathbf{F}_\delta$ with the same $\lambda = \lambda_\delta$ for all n.

Set

$$w := \nabla v, \quad I_q := \sum_{r=1}^d \langle (\nabla_r w)^2 |w|^{p-2} \rangle, \quad J_q := \langle (\nabla |w|)^2 |w|^{p-2} \rangle.$$

We multiply $(\mu - \Delta) v = |b|^{1-\frac{2}{p}} f$ by $\phi := -\nabla \cdot (w|w|^{p-2})$ and integrate by parts to obtain

$$\mu \|w\|_p^p + I_p + (p-2) J_p = \langle |b|^{1-\frac{2}{p}} f, -\nabla \cdot (w|w|^{p-2}) \rangle, \tag{A.1}$$

where

$$\begin{aligned}&\langle |b|^{1-\frac{2}{p}}f, -\nabla\cdot(w|w|^{p-2})\rangle = \langle |b|^{1-\frac{2}{p}}f, (-\Delta v)|w|^{p-2} - (p-2)|w|^{p-3}w\cdot\nabla|w|\rangle \\ &(\text{use the equation } \; -\Delta v = -\mu v + |b|^{1-\frac{2}{p}}f) \\ &= \langle |b|^{1-\frac{2}{p}}f, (-\mu v + |b|^{1-\frac{2}{p}}f)|w|^{p-2}\rangle - (p-2)\langle |b|^{1-\frac{2}{p}}f, |w|^{p-3}w\cdot\nabla|w|\rangle.\end{aligned}$$

Remark A.1. Here we work with the same test function $\phi = -\nabla\cdot(w|w|^{p-2})$ as in [KS].

We have
1) $\langle |b|^{1-\frac{2}{p}}f, (-\mu v)|w|^{p-2}\rangle \leq 0$,
2) $|\langle |b|^{1-\frac{2}{p}}f, |w|^{p-3}w\cdot\nabla|w|\rangle| \leq \alpha J_p + \frac{1}{4\alpha}N_p$ $(\alpha>0)$, where $N_p := \langle |b|^{1-\frac{2}{p}}f, |b|^{1-\frac{2}{p}}f|w|^{p-2}\rangle$,
so, the RHS of (A.1) $\leq (p-2)\alpha J_p + (1+\frac{p-2}{4\alpha})N_p$, where, in turn,

$$\begin{aligned}N_p &\leq \langle |b|^2|w|^p\rangle^{\frac{p-2}{p}}\langle f^p\rangle^{\frac{2}{p}} \\ &\leq \frac{p-2}{p}\langle |b|^2|w|^p\rangle + \frac{2}{p}\|f\|_p^p \qquad (\text{use } b\in\mathbf{F}_\delta \Leftrightarrow \|b\varphi\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + \lambda\delta\|\varphi\|_2^2,\ \varphi\in W^{1,2}) \\ &\leq \frac{p-2}{p}\left(\frac{p^2}{4}\delta J_q + \lambda\delta\|w\|_p^p\right) + \frac{2}{p}\|f\|_p^p.\end{aligned}$$

Thus, applying $I_q \geq J_q$ in the LHS of (A.1), we obtain

$$(\mu-c_0)\|w\|_p^p+\left[p-1-(p-2)\left(\alpha+\frac{1}{4\alpha}\frac{p(p-2)}{4}\delta\right)-\frac{p(p-2)}{4}\delta\right]\frac{4}{p^2}\|\nabla|\nabla v|^{\frac{p}{2}}\|_2^2 \leq \left(1+\frac{p-2}{4\alpha}\right)\frac{2}{p}\|f\|_p^p,$$

where $c_0 = \frac{p-2}{p}\lambda\delta(1+\frac{p-2}{4\alpha})$. It is now clear that one can find a sufficiently large $\mu_0 = \mu_0(d,p,\delta) > 0$ so that, for all $\mu > \mu_0$, $(*)$ (with $b = b_n$) holds with

$$\begin{aligned}c_{\delta,p}^p &= \delta\frac{p^2}{4}\frac{\left(1+\frac{p-2}{4\alpha}\right)\frac{2}{p}}{p-1-(p-2)\left(\alpha+\frac{1}{4\alpha}\frac{p(p-2)}{4}\delta\right)-\frac{p(p-2)}{4}\delta} \qquad (\text{we select } \alpha = \frac{p}{4}\sqrt{\delta}) \\ &= \frac{\frac{p}{2}\delta + \frac{p-2}{2}\sqrt{\delta}}{p-1-(p-1)\frac{p-2}{2}\sqrt{\delta}-\frac{p(p-2)}{4}\delta},\end{aligned}$$

as claimed. Finally, we pass to the limit $n\to\infty$ using Fatou's Lemma. The proof of $(*)$ is completed.

Remark A.2. It is seen that $\sqrt{\delta} < \frac{2}{p} \Rightarrow c_{\delta,p} < 1$. We also note that the above choice of α is the best possible.

(**b**) Set $v = (\mu-\Delta)^{-1}f$, $0 \leq f \in L^p$. Then

$$\begin{aligned}&\|G_pf\|_p^p = \|b^{\frac{2}{p}}\cdot\nabla v\|_p^p \\ &(\text{we argue as in } (\mathbf{a})) \\ &\leq \delta(\lambda\|\nabla v\|_p^p + \|\nabla|\nabla v|^{\frac{p}{2}}\|_2^2),\end{aligned}$$

where, clearly, $\|\nabla v\|_p^p \leq \mu^{-\frac{p}{2}}\|f\|_p^p$. In turn, arguing as in (**a**), we arrive at $\mu\|w\|_p^p + I_p + (p-2)J_p = \langle f, -\nabla\cdot(w|w|^{p-2})$ $(w = \nabla v)$,

$$\mu\|w\|_p^p + (p-1)J_p \leq \langle f^2, |w|^{p-2}\rangle + (p-2)\langle f, |w|^{p-3}w\cdot\nabla|w|\rangle),$$

$$\mu\|w\|_p^p+(p-1)J_p\leq\langle f^2,|w|^{p-2}\rangle+(p-2)\big(\varepsilon J_p+\frac{1}{4\varepsilon}\langle f^2,|w|^{p-2}\rangle\big),\quad \varepsilon>0.$$

Selecting ε sufficiently small, we obtain

$$J_p\leq C_0\|w\|_p^{p-2}\|f\|_p^2.$$

Now, applying $\|w\|_p\leq\mu^{-\frac{1}{2}}\|f\|_p$, we arrive at $\|\nabla|\nabla v|^{\frac{p}{2}}\|_2^2\leq C\mu^{-\frac{p}{2}+1}\|f\|_p^p$. Hence, $\|G_pf\|_p\leq C_1\mu^{-\frac{1}{2}+\frac{1}{p}}\|f\|_p$ for all $\mu\geq\mu_0$.

(**c**) Set $v=(\mu-\Delta)^{-1}|b|^{1-\frac{2}{p}}f$ $(=Q_pf)$, $0\leq f\in L^p$. Then, multiplying $(\mu-\Delta)v=|b|^{1-\frac{2}{p}}f$ by v^{p-1}, we obtain

$$\mu\|v\|_p^p+\frac{4(p-1)}{p^2}\|\nabla v^{\frac{p}{2}}\|_2^2=\langle|b|^{1-\frac{2}{p}}f,v^{p-1}\rangle,$$

where we estimate the RHS using Young's inequality:

$$\langle|b|^{1-\frac{2}{p}}v^{\frac{p}{2}-1},fv^{\frac{p}{2}}\rangle\leq\varepsilon^{\frac{2p}{p-2}}\frac{p-2}{2p}\langle|b|^2v^p\rangle+\varepsilon^{-\frac{2p}{p+2}}\frac{p+2}{2p}\langle f^{\frac{2p}{p+2}}v^{\frac{p^2}{p+2}}\rangle\quad\varepsilon>0.$$

Using $b\in\mathbf{F}_\delta$ and selecting $\varepsilon>0$ sufficiently small, we obtain that for any $\mu_1>0$ there exists $C>0$ such that

$$(\mu-\mu_1)\|v\|_p^p\leq C\langle f^{\frac{2p}{p+2}}v^{\frac{p^2}{p+2}}\rangle,\qquad\mu>\mu_1.$$

Therefore, $(\mu-\mu_1)\|v\|_p^p\leq C\langle f^p\rangle^{\frac{2}{p+2}}\langle v^p\rangle^{\frac{p}{p+2}}$, so $\|v\|_p\leq C_2\mu^{-\frac{1}{2}-\frac{1}{p}}\|f\|_p$.

Step 2: We now use the results of Step 1 to prove assertions (*ii*) and (*iii*). Below we use the following formula: For every $0<\alpha<1$, $\mu>0$,

$$(\mu-\Delta)^{-\alpha}=\frac{\sin\pi\alpha}{\pi}\int_0^\infty t^{-\alpha}(t+\mu-\Delta)^{-1}dt.$$

We have

$$\begin{aligned}\|Q_p(q)f\|_p\leq&\ \|(\mu-\Delta)^{-\frac{1}{2}+\frac{1}{q}}|b|^{1-\frac{2}{p}}|f|\|_p\\ &\leq k_q\int_0^\infty t^{-\frac{1}{2}+\frac{1}{q}}\|(t+\mu-\Delta)^{-1}|b|^{1-\frac{2}{p}}|f|\|_pdt\\ &\text{(we use (}\mathbf{c}\text{))}\\ &\leq\ k_qC_2\int_0^\infty t^{-\frac{1}{2}+\frac{1}{q}}(t+\mu)^{-\frac{1}{2}-\frac{1}{p}}dt\ \|f\|_p=K_{2,q}\|f\|_p,\quad f\in\mathcal{E},\end{aligned}$$

where, clearly, $K_{2,q}<\infty$ due to $q>p$.

It suffices to consider the case $r>2$. We have

$$\begin{aligned}\|G_p(r)f\|_p\leq k_r&\int_0^\infty t^{-\frac{1}{2}-\frac{1}{r}}\|b^{\frac{2}{p}}\cdot\nabla(t+\mu-\Delta)^{-1}f\|_pdt\\ &\text{(we use (}\mathbf{b}\text{))}\\ &\leq k_rC_1\int_0^\infty t^{-\frac{1}{2}-\frac{1}{r}}(t+\mu)^{-\frac{1}{2}+\frac{1}{p}}dt\ \|f\|_p=K_{1,r}\|f\|_p,\quad f\in\mathcal{E},\end{aligned}$$

where, clearly, $K_{1,r}<\infty$ due to $r<p$. The proof of Lemma 6.1 is completed. □

Appendix B. Some examples of form-bounded vector fields

Below we list some sub-classes of the class of form-bounded vector fields, defined in elementary terms.

1. Let us prove that

$$b \in L^\infty(\mathbb{R}_+, L^d + L^\infty) \quad \Rightarrow \quad b \in L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$$

for appropriate δ and g (see Definition 9.1). Here we have, by definition, $b = b_1 + b_2$, where $b_1 \in L^\infty(\mathbb{R}_+, L^d)$, $b_2 \in L^\infty(\mathbb{R}_+, L^\infty)$. By Hölder's inequality, for a.e. $t \in \mathbb{R}_+$ and all $\varphi \in C_c^\infty$,

$$\begin{aligned}\|b(t)\varphi\|_2^2 &\leq (1+\varepsilon)\|b_1(t)\|_d^2\|\varphi\|^2_{\frac{2d}{d-2}} + (1+\varepsilon^{-1})\|b_2(t)\|_\infty^2\|\varphi\|_2^2 \qquad (\varepsilon > 0) \\ &\text{(apply the Sobolev embedding theorem)} \\ &\leq C_S(1+\varepsilon)\|b_1(t)\|_d^2\|\nabla\varphi\|_2^2 + (1+\varepsilon^{-1})\|b_2(t)\|_\infty^2\|\varphi\|_2^2.\end{aligned}$$

Thus, $b \in L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$ with

$$\delta := C_S(1+\varepsilon)\|b_1\|^2_{L^\infty(\mathbb{R}_+, L^d)}, \quad g(t) := (1+\varepsilon^{-1})\|b_2(t,\cdot)\|_\infty^2$$

(in this paper we mostly care about the value of δ, so ε should be chosen sufficiently small).

2. Next, let us show that

$$b \in C(\mathbb{R}_+, L^d + L^\infty) \quad \Rightarrow \quad b \in \mathbf{F}_\delta \quad \text{with } \delta \text{ that can be chosen arbitrarily small.}$$

Without loss of generality, let us carry out the proof for $b \in C(\mathbb{R}_+, L^d)$.

First, let $b = b(x)$. Since $|b| \in L^d$, one can represent for every $\varepsilon > 0$ $b = b_1 + b_2$, where $\|b_1\|_d < \varepsilon$ and $\|b_2\|_\infty < \infty$. (For instance, $b_2 = b\mathbf{1}_{|b|\leq m}$ and $b_1 = b - b_2$, so by the Dominated convergence theorem $\|b_1\|_d$ can be made arbitrarily small by selecting m sufficiently large.) Now the previous example applies and yields the required.

In the general case $b \in C(\mathbb{R}_+, L^d)$, the continuity of b in time allows us to represent $b(t,\cdot) = b_1(t,\cdot) + b_2(t,\cdot)$, where $\|b_1(t,\cdot)\|_d < \varepsilon$ for all $t \in [0,1]$ and b_2 is bounded on $[0,1]\times\mathbb{R}^d$. Repeating this on every interval $[n, n+1]$ ($n \geq 1$), one obtains $\|b_1\|_{L^\infty(\mathbb{R}_+, L^d)} < \varepsilon$ and $b_2 \in L^\infty_{\rm loc}(\mathbb{R}_+, L^\infty)$. In fact, the continuity in time is not necessary for the smallness of δ, e.g. consider $b(t,x) = c(t)b_0(x)$ where $c \in L^\infty(\mathbb{R}_+)$ is discontinuous and $|b_0| \in L^d$.

3. Any vector field

$$b \in L^p(\mathbb{R}_+, L^q), \quad \frac{d}{q} + \frac{2}{p} \leq 1, \quad p \geq 2, \quad q \geq d$$

is in $L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$ with appropriate δ. Indeed, e.g. in the more difficult case $\frac{d}{q} + \frac{2}{p} = 1$, we have by Young's inequality

$$|b(t,x)| = \frac{|b(t,x)|}{\langle |b(t,\cdot)|^q\rangle^{\frac{1}{q}}}\langle |b(t,\cdot)|^q\rangle^{\frac{1}{q}} \leq \frac{d}{q}\left(\frac{|b(t,x)|^q}{\langle |b(t,\cdot)|^q\rangle}\right)^{\frac{1}{d}} + \frac{2}{p}\left(\langle |b(t,\cdot)|^q\rangle^{\frac{1}{q}}\right)^{\frac{p}{2}},$$

where the first term is in $L^\infty(\mathbb{R}_+, L^d)$ (and so by the first example it is form-bounded) and the second term is in $L^2(\mathbb{R}_+, L^\infty)$ (the second term squared is to be absorbed by the function g). If $p < \infty$, $q > d$, then one can argue as in the previous example to show that δ can be chosen arbitrarily small.

4. The class $\mathbf{F}_\delta$ contains vector fields $b = b(x)$ with $|b|$ in $L^{d,w}$, i.e. the weak L^d class (Section 2). Indeed, by [KPS, Prop. 2.5, 2.6, Cor. 2.9], if $|b| \in L^{d,w}$, then $b \in \mathbf{F}_{\delta_1}$ with

$$\begin{aligned}\sqrt{\delta_1} &= \||b|(\lambda - \Delta)^{-\frac{1}{2}}\|_{2\to 2} \\ &\leq \|b\|_{d,w}\Omega_d^{-\frac{1}{d}}\||x|^{-1}(\lambda - \Delta)^{-\frac{1}{2}}\|_{2\to 2} \leq \|b\|_{d,w}\Omega_d^{-\frac{1}{d}}\frac{2}{d-2},\end{aligned}$$

where $\Omega_d = \pi^{\frac{d}{2}}\Gamma(\frac{d}{2}+1)$ is the volume of $B_1(0) \subset \mathbb{R}^d$.

5. The Chang-Wilson-Wolff class CW^2 consists of vector fields $b = b(x)$ such that

$$|b| \in L^2_{\rm loc}(\mathbb{R}^d)$$

and

$$\|b\|_{\mathrm{CW}^2} := \Big(\sup_Q \frac{1}{|Q|}\int_Q |b(x)|^2\, l(Q)^2 \xi(|b(x)|^2\, l(Q)^2)dx < \infty\Big)^{\frac{1}{2}} < \infty,$$

where $|Q|$ and $l(Q)$ are the volume and the side length of cube $Q \subset \mathbb{R}^d$, respectively, $\xi : \mathbb{R}_+ \to [1, \infty[$ is an increasing function such that

$$\int_1^\infty \frac{dx}{x\xi(x)} < \infty.$$

One has, for every $\varepsilon > 0$,

$$M_{2+\varepsilon} \quad \subsetneq \quad \mathrm{CW}^2 \quad \subsetneq \quad \mathbf{F}_\delta$$

with $\delta = \delta(\|b\|_{\mathrm{CW}^2})$, see [CWW].

5. More generally, vector fields in $L^\infty(\mathbb{R}_+, M_q)$, $q > 2$, or $L^\infty(\mathbb{R}_+, \mathrm{CW}^2)$ are form-bounded.

Appendix C. Smooth approximations of form-bounded (-type) vector fields

We construct two kinds of smooth approximations (or regularizations) of a vector field b: defined by mollifying a cutoff of b, or by mollifying b directly.

1. If $b = b(x)$ is either in the class $\mathbf{F}_\delta$ (form-bounded) or in the class $\mathbf{F}_\delta^{1/2}$ (weakly form-bounded), it is clear that multiplying b by the indicator $\mathbf{1}_m$ of $\{x \in \mathbb{R}^d \mid |b(x)| \leq m, |x| \leq m\}$ leaves us in the corresponding class with the same δ and λ_δ. This gives us a bounded, compact-support approximation $\{\mathbf{1}_m b\}_{m=1}^\infty$ of b (in $L^2_{\rm loc}$ or in $L^1_{\rm loc}$, respectively).

We can also go one step further and apply to $\mathbf{1}_m b$ a mollifier, which will still give us a uniformly form-bounded (or uniformly weakly form-bounded) approximation of b by vector fields whose components are in the Schwartz space $\mathcal{S}$. Below we provide the details of this simple construction for a time-inhomogeneous form-bounded vector field $b \in L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R}_+)$:

$$\|b(t,\cdot)\varphi\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + g(t)\|\varphi\|_2^2, \quad \varphi \in W^{1,2} \tag{$\bullet$}$$

for a.e. $t \in \mathbb{R}$, for a $0 \leq g \in L^1_{\rm loc}(\mathbb{R})$ (Definition 9.1).

Let us extend b to $\{t < 0\}$ by 0 and set

$$b_m := c_m E_\varepsilon(\mathbf{1}_m b), \tag{$\star$}$$

where $E_\varepsilon \equiv E_\varepsilon^{d+1}$ is the De Giorgi or Friedrichs mollifier on $\mathbb{R} \times \mathbb{R}^d$ (Section 2), $\mathbf{1}_m$ is the indicator of $\{(t,x) \in \mathbb{R}^{1+d} \mid |b(t,x)| \leq m, |x| \leq m, |t| \leq m\}$, and $\varepsilon_m \downarrow 0$ and $c_m \uparrow 1$ are to be chosen. Clearly,

$$b_m \in L^\infty \cap C^\infty(\mathbb{R}^{1+d}, \mathbb{R}^d), \tag{C.1}$$

also, provided that $\varepsilon_m \downarrow 0$ sufficiently rapidly,

$$b_m \to b \quad \text{in } L^2_{\rm loc}(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^d). \tag{C.2}$$

and, provided that $c_m \uparrow 1$ sufficiently slow,

$$\|b_m(t)\varphi\|_2^2 \leq \delta \|\nabla \varphi\|_2^2 + g(t)\|\varphi\|_2^2, \quad t \geq 0, \quad \varphi \in W^{1,2}, \tag{C.3}$$

i.e. $\{b_m\}$ are uniformly form-bounded.

Proof of (C.3). First, define $\tilde{b}_m = E_\varepsilon(\mathbf{1}_m b)$ and write

$$\tilde{b}_m = \mathbf{1}_m b + (\tilde{b}_m - \mathbf{1}_m b). \tag{$*$}$$

Clearly, the first term satisfies

$$\|\mathbf{1}_m b(t)\varphi\|_2^2 \leq \delta \|\nabla \varphi\|_2^2 + g(t)\|\varphi\|_2^2,$$

In turn, since $\mathbf{1}_m b$ has compact support and is in $L^\infty(\mathbb{R}_+, L^r)$ for any $r > d$, given any $\gamma_m \downarrow 0$ we can select $\varepsilon_m \downarrow 0$ so that $\|\tilde{b}_m - \mathbf{1}_m b\|^2_{L^r(\mathbb{R}_+ \times \mathbb{R}^d)} \leq \gamma_m$, so, in view of example 3 in Appendix B, the second term in $(*)$ satisfies

$$\|(\tilde{b}_m(t) - \mathbf{1}_m b(t))\varphi\|_2^2 \leq C\gamma_m \|\nabla \varphi\|_2^2.$$

Therefore,

$$\|\tilde{b}_m(t)\varphi\|_2^2 \leq \delta_m \|\nabla \varphi\|_2^2 + g(t)\|\varphi\|_2^2$$

with $\delta_m = (\sqrt{\delta} + \sqrt{C\gamma_m})^2$. Now, multiplying $\tilde{b}_m$ by $c_m = \frac{\delta}{\delta_m}$ (clearly, $c_m \uparrow 1$) and recalling that $b_m = c_m \tilde{b}_m$, we obtain (C.3). $\square$

2. In fact, do not need the cutoff function in $(\star)$ to construct a smooth approximation of a $b = b(x)$, $b \in \mathbf{F}_\delta$:

$$\|b\varphi\|_2^2 \leq \delta \|\nabla \varphi\|_2^2 + c_\delta \|\varphi\|_2^2, \quad \varphi \in W^{1,2}.$$

This observation is important e.g. if one needs to control the divergence of the approximating vector fields. In [KiS6] the authors defined

$$b_\varepsilon := E_\varepsilon b, \tag{$\star\star$}$$

where $E_\varepsilon \equiv E^d_\varepsilon := e^{\varepsilon \Delta}$ is the De Giorgi's mollifier on $\mathbb{R}^d$ and $\varepsilon \downarrow 0$ (at any rate). (We can also use Friedrichs' mollifier, see remark below.) We have

$$b_\varepsilon \in L^\infty \cap C^\infty(\mathbb{R}^d, \mathbb{R}^d), \tag{C.4}$$

$$\|b_\varepsilon \varphi\|_2^2 \leq \delta \|\nabla \varphi\|_2^2 + c_\delta \|\varphi\|_2^2, \quad \text{i.e. } b_\varepsilon \in \mathbf{F}_\delta \text{ with the same } c_\delta, \tag{C.5}$$

$$b_\varepsilon \to b \quad \text{in } L^2_{\rm loc}(\mathbb{R}^d, \mathbb{R}^d). \tag{C.6}$$

Proof of (C.4)-(C.6). We repeat the argument from [KiS6].

To prove (C.4), we represent $b_\varepsilon = E_{\varepsilon/2} E_{\varepsilon/2} b$, so it suffices to only prove that $|b_\varepsilon| \in L^\infty$. Indeed, we have, using Fatou's lemma,

$$\begin{aligned}|b_\varepsilon(x)| &\leq \liminf_n \langle e^{\varepsilon\Delta}(x,\cdot)\mathbf{1}_{B_n(0)}(\cdot)|b(\cdot)|\rangle \\ &\leq \liminf_n \langle e^{\varepsilon\Delta}(x,\cdot)\mathbf{1}_{B_n(0)}(\cdot)|b(\cdot)|^2\rangle^{\frac{1}{2}} \leq \big(\delta \langle |\nabla \sqrt{e^{\varepsilon\Delta}(x,\cdot)}|^2 \rangle + c_\delta\big)^{\frac{1}{2}},\end{aligned}$$

where $|\nabla_y \sqrt{e^{\varepsilon\Delta}(x,y)}| = (4\pi\varepsilon)^{-\frac{d}{4}} \frac{|x-y|}{4\varepsilon} e^{-\frac{|x-y|^2}{8\varepsilon}} \leq C\varepsilon^{-\frac{d}{4}-\frac{1}{2}} e^{-\frac{c|x-y|^2}{\varepsilon}}$, and so $|b_\varepsilon| \in L^\infty$ for each $\varepsilon > 0$.

Let us prove (C.5). Indeed, $|b_\varepsilon| \leq \sqrt{E_\varepsilon |b|^2}$, and so

$$\begin{aligned}\|b_\varepsilon \varphi\|_2^2 &\leq \langle E_\varepsilon |b|^2, \varphi^2\rangle = \|b\sqrt{E_\varepsilon \varphi^2}\|_2^2 \\ &\leq \delta \|\nabla \sqrt{E_\varepsilon \varphi^2}\|_2^2 + c_\delta \|\varphi\|_2^2, \quad \varphi \in W^{1,2},\end{aligned}$$

where

$$\begin{aligned}\|\nabla \sqrt{E_\varepsilon \varphi^2}\|_2 &= \|\frac{E_\varepsilon(|\varphi||\nabla|\varphi|)}{\sqrt{E_\varepsilon \varphi^2}}\|_2 \qquad (**) \\ &\leq \|\sqrt{E_\varepsilon |\nabla|\varphi||^2}\|_2 = \|E_\varepsilon |\nabla|\varphi||^2\|_1^{\frac{1}{2}} \\ &\leq \|\nabla|\varphi|\|_2 \leq \|\nabla\varphi\|_2,\end{aligned}$$

i.e. $b_\varepsilon \in \mathbf{F}_\delta$. (The fact that $\|b\sqrt{E_\varepsilon\varphi^2}\|_2 < \infty$ follows from $\mathbf{1}_{\{|b|\leq n\}} b \in \mathbf{F}_\delta$ and the inequality $\|\mathbf{1}_{\{|b|\leq n\}} b \sqrt{E_\varepsilon \varphi^2}\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + c_\delta\|\varphi\|_2^2$, using Fatou's lemma).

Regarding the proof of (C.6), let us only demonstrate that $b_\varepsilon - b \to 0$ in $L^2(B_1)$, $B_1 \equiv B_1(0)$. To that end, we fix some $R > 1$ and represent on B_1:

$$b_\varepsilon - b = I_1 + I_2, \qquad I_1 := E_\varepsilon(1 - \mathbf{1}_{B_R})b, \quad I_2 := E_\varepsilon(\mathbf{1}_{B_R} b) - \mathbf{1}_{B_R} b$$

(of course, on B_1, one has $b = \mathbf{1}_{B_R} b$ since $R > 1$). Then $I_2 \to 0$ in $L^2(B_1)$ since $\mathbf{1}_{B_R} b$ has compact support. In turn, $I_1 \to 0$ in $L^2(B_1)$ by the separation property of the Gaussian density, i.e. $e^{\varepsilon\Delta}(x,y) \to 0$ uniformly in $x \in B_1$ if $y \in \mathbb{R}^d - B_R$ (here we have used $R > 1$). ☐

Remark C.1. If we were to use the Friedrichs mollifier in ($\star\star$), then we would get smoothness of b_ε and convergence (C.6) from the usual properties of Friedrichs mollifiers. But we would have to be slightly more careful in ($**$) in order to avoid division by zero, e.g. replace $|f|^2$ by $|f|^2 + e^{-k|x|^2}$, carry out the estimates and then take $k \to \infty$.

3. The regularization ($\star\star$) can also be used to handle time-inhomogeneous form-bounded drifts $b \in L^\infty \mathbf{F}_\delta + L^2_{\rm loc}(\mathbb{R})$. That is, we can put

$$b_\varepsilon := E^1_\varepsilon E^d_\varepsilon b, \quad \varepsilon \downarrow 0,$$

where E^d_ε is the De Giorgi or Friedrichs mollifier on $\mathbb{R}^d$ (in the spatial variables) and E^1_ε is the Friedrichs mollifier on $\mathbb{R}$ (in the time variable; we use the Friedrichs mollifier here since, in the time variable, b is in only locally in $L^1(\mathbb{R})$, as is determined by our assumption on g, cf. ($\bullet$)).

This regularization $\{b_\varepsilon\}$ satisfies

$$|b_\varepsilon| \in L^\infty_{\rm loc}(\mathbb{R}, L^\infty(\mathbb{R}^d)), \quad b_\varepsilon \text{ are } C^\infty \text{ smooth}, \tag{C.7}$$

$$\|b_\varepsilon(t,\cdot)\varphi\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + g_\varepsilon(t)\|\varphi\|_2^2 \quad \text{ for all } t \in \mathbb{R}, \tag{C.8}$$

$$\sup_{\varepsilon>0} \int_{t_0}^{t_1} g_\varepsilon(s)ds < \infty \text{ for all finite } t_0, t_1, \text{ where } g_\varepsilon := E^1_\varepsilon g, \tag{C.9}$$

$$b_\varepsilon \to b \quad \text{ in } L^2_{\rm loc}(\mathbb{R}^{1+d}, \mathbb{R}^d). \tag{C.10}$$

Proof of (C.8). First, we regularize b only in the spatial variable: for every $t \in \mathbb{R}$, put

$$\tilde{b}_\varepsilon(t,\cdot) := E^d_\varepsilon b(t,\cdot).$$

Then, since for a.e. t $b(t,\cdot)$ is form-bounded on $\mathbb{R}^d$, we have by (C.5)

$$\|\tilde{b}_\varepsilon(t,\cdot)\varphi\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + g(t)\|\varphi\|_2^2 \tag{$***$}$$

for a.e. $t \in \mathbb{R}$. Next, recalling that $b_\varepsilon = E^1_\varepsilon \tilde{b}_\varepsilon$ and noting a pointwise inequality

$$|b_\varepsilon(t)|^2 \leq (E^1_\varepsilon |\tilde{b}_\varepsilon|^2)(t), \quad t \in \mathbb{R},$$

we estimate, for every $\varphi \in W^{1,2}(\mathbb{R}^d)$,

$$\begin{aligned}\langle |b_\varepsilon(t)|^2 \varphi^2 \rangle &\leq E^1_\varepsilon \langle |\tilde{b}_\varepsilon|^2 \varphi^2 \rangle(t) \\ &\quad \text{(we are applying } E^1_\varepsilon \text{ to both sides of the inequality } (***)) \\ &\leq \delta \|\nabla \varphi\|_2^2 + g_\varepsilon(t) \|\varphi\|_2^2.\end{aligned}$$

Thus, we arrive at (C.8). □

A similar construction was considered in [KiS9].

Appendix D. Trotter's approximation theorem

Consider a sequence $\{e^{-tA_k}\}_{k=1}^\infty$ of C_0 semigroups on a (complex) Banach space Y.

Theorem D.1 (H.F. Trotter [Ka, Ch. IX]). *Let* $\sup_k \|(\mu + A_k)^{-m}\|_{Y \to Y} \leq M(\mu - \omega)^{-m}$, *$m = 1, 2, \dots$, $\mu > \omega$, and*

$$s\text{-}\lim_{\mu \to \infty} \mu(\mu + A_k)^{-1} = 1 \quad \textit{uniformly in } k,$$

and let $s\text{-}\lim_k (\zeta + A_k)^{-1}$ exist for some ζ with $\operatorname{Re} \zeta > \omega$. Then there is a C_0 semigroup e^{-tA} such that

$$(z + A_k)^{-1} \xrightarrow{s} (z + A)^{-1} \quad \textit{for every } \operatorname{Re} z > \omega,$$

and

$$e^{-tA_k} \xrightarrow{s} e^{-tA}$$

uniformly in any finite interval of $t \geq 0$.

The first condition of the theorem is satisfied if e.g.

$$\sup_k \|(\mu + A_k)^{-1}\|_{Y \to Y} \leq (\mu - \omega)^{-1}, \quad \mu > \omega$$

(obviously) of if

$$\sup_k \|(z + A_k)^{-1}\|_{Y \to Y} \leq C|z - \omega|^{-1}, \quad \operatorname{Re} z > \omega,$$

see [Ka, IX.6.1]. The second condition is what can be verified in practice when one is dealing with quasi bounded semigroups.

References

[A] D. Adams, Weighted nonlinear potential theory, *Trans. Amer. Math. Soc.* **297** (1986), 73-94.

[AF] R. Adams and J. Fournier, "Sobolev Spaces", Second Edition, Elsevier, 2003.

[AD] D. Albrighton and H. Dong, Regularity properties of passive scalars with rough divergence-free drifts, arXiv:2107.12511.

[B] R. J. Bagby, Lebesgue spaces of parabolic potentials, *Illinois J. Math.* **15** (1971), 610-634.

[BC] R. Bass and Z.-Q. Chen, Brownian motion with singular drift. *Ann. Probab.*, 31 (2003), 791-817.

[BG] P. Baras and J. A. Goldstein, The heat equation with a singular potential. *Trans. Amer. Math. Soc.*, **284** (1984), 121-139.

[BFGM] L. Beck, F. Flandoli, M. Gubinelli and M. Maurelli, Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Electr. J. Probab.*, **24** (2019), Paper No. 136, 72 pp (arXiv:1401.1530).

[BS] A.G. Belyi and Yu.A. Semenov. On the L^p-theory of Schrödinger semigroups. II. *Sibirsk. Math. J.*, 31 (1990), p. 16-26; English transl. in *Siberian Math. J.*, 31 (1991), 540-549.

[BO] Á. Bényi and T. Oh, The Sobolev inequality on the torus revisited, *Publ. Math. Debrecen* **83**, no. 3 (2013), 359-374.

[Bi] M.S. Birman, "On the spectrum of singular boundary-value problems" (in Russian), *Mat. Sbornik* **55**(97) (1961), 125-174.

[BJ] K. Bogdan and T. Jakubowski, Estimates of heat kernel of fractional Laplacian perturbed by gradient operators, *Comm. Math. Phys.*, **271** (2007) 179-198.

[BGe] R.M. Blumenthal and R.K. Getoor, Markov Processes and Potential Theory. Pure and Applied Mathematics 29. Academic Press, New York.

[BJW] D. Bresch, P.-E. Jabin and Z. Wang, Mean field limit and quantitative estimates with singular attractive kernels, *Duke Math. J.* **172** (2023), no. 13, 2591-2641 (arXiv:2011.08022).

[CWW] S.Y.A. Chang, J.M. Wilson and T.H. Wolff, Some weighted norm inequalities concerning the Schrödinger operator, *Comment. Math. Helvetici*, **60** (1985), 217-246.

[CM] P.-É. Chaudru de Raynal and S. Menozzi, On multi-dimensional stable driven stochastic differential equations with Besov drift, *Electron. J. Probab.* **27** (2022), Paper No. 163, 52 pp. (arXiv:1907.12263).

[CJM] P.-É. Chaudru de Raynal, J.-F. Jabir and S. Menozzi, Multidimensional stable driven McKean-Vlasov SDEs with distributional interaction kernel: a regularization by noise perspective, arXiv:2205.11866.

[CKS] Z.-Q. Chen, P. Kim and R. Song, Dirichlet heat kernel estimates for fractional Laplacian with gradient perturbation, *Ann. Prob.*, **40** (2012), 2483-2538.

[CFKZ] Z.-Q. Chen, P. J. Fitzsimmons, K. Kuwae and T.-S. Zhang, Perturbation of symmetric Markov processes, *Probab. Theory Related Fields* **140** (2008), 239-275.

[CW] Z.-Q. Chen and L. Wang, Uniqueness of stable processes with drift, *Proc. Amer. Math. Soc*, **144** (2017), p. 2661-2675 (arXiv:1309.6414).

[C] A. S. Cherny, On the uniqueness in law and the pathwise uniqueness for stochastic differential equations, *Theory Probab. Appl.*, **46**(3) (2002), 406-419.

[CE] A. S. Cherny and H.-J. Engelbert. Singular Stochastic Differential Equations. *LNM 1858*. Springer-Verlag, 2005.

[CFr] F. Chiarenza and M. Frasca, A remark on a paper by C. Fefferman, *Proc. Amer. Math. Soc.*, **108** (1990), 407-409.

[D] H. Dong, Recent progress in the L^p theory for elliptic and parabolic equations with discontinuous coefficients, *Anal. Theory Appl.* 36 (2020), no. 2, 161-199 (arXiv:2006.03966).

[F] C. Fefferman, The uncertainty principle, *Bull. Amer. Math. Soc.* **9** (1983), 129-206.

[FK] P. J. Fitzsimmons and K. Kuwae, Non-symmetric perturbations of symmetric Dirichlet forms, *J. Funct. Anal.* **208** (2004), 140-162.

[FIR] F. Flandoli, E. Issoglio and F. Russo, Multidimensional stochastic differential equations with distributional drift, *Trans. Amerc. Math. Soc.*, **369** (2017), 1665-1688 (arXiv:1401.6010).

[FGP] F. Flandoli, M. Gubinelli and E. Priola, Well-posedness of the transport equation by stochastic perturbation, *Invent. Math.* **180** (2010), 1-53.

[FJ] N. Fournier and B. Jourdain, Stochastic particle approximation of the Keller-Segel and two-dimensional generalization of Bessel process, *Ann. Appl. Probab.* **27** (2017), 2807-2861.

[GM] B. Gess and M. Maurelli, Well-posedness by noise for scalar conservation laws, *Comm. Partial Differential Equations* **43** (2018), no. 12, 1702-1736.

[GZa] J. A. Goldstein and Qi. S. Zhang, Linear parabolic equation with strongly singular potentials, *Trans. Amer. Math. Soc.* **355** (2003), 197-211.

[Go] V.R. Gopala Rao, A characterization of parabolic function spaces, *Amer. J. Math.*, **99** (1977), 985-993.

[GO] L. Grafakos and S. Oh. The Kato-Ponce inequality. *Comm. Partial Diff. Equ.*, 39 (2014), 1128-1157.

[GC] A. Gulisashvili and J.A. van Casteren, Non-autonomous Kato Classes and Feynman-Kac Propagators, *World Scientific*, 2006.

[G] S. Gupta, Hardy and Rellich inequality on lattices, *Calc. Var. Partial Diff. Equations*, 62, article no. 81 (2023).

[H] T. Hara, A refined subsolution estimate of weak subsolutions to second order linear elliptic equations with a singular vector field, *Tokyo J. Math.*, **38**(1) (2015), 75-98.

[He] E. Heinz, *Beiträge zur Störungstheorie der Spektralzerlegung*, Math. Ann., **123** (1951) 415-438.

[J] P. Jin, Brownian motion with singular time-dependent drift. *J. Theoret. Probab.*, **30** (2017), 1499-1538 (arXiv:1710.05227).

[Ka] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, Berlin, Heidelberg, 1995.

[KSo] P. Kim and R. Song, Stable process with singular drift, *Stoc. Proc. Appl.* **124** (2014), 2479-2516.

[Ki1] D. Kinzebulatov, A new approach to the L^p-theory of $-\Delta + b \cdot \nabla$, and its applications to Feller processes with general drifts, *Ann. Sc. Norm. Sup. Pisa (5)*, **17** (2017), 685-711 (arXiv:1502.07286).

[Ki2] D. Kinzebulatov, Regularity theory of Kolmogorov operator revisited, *Canadian Bull. Math.* **64** (2021), 725-736 (arXiv:1807.07597).

[Ki3] D. Kinzebulatov, Feller evolution families and parabolic equations with form-bounded vector fields, *Osaka J. Math.*, **54** (2017), 499-516 (arXiv:1407.4861).

[Ki4] D. Kinzebulatov, Feller generators with measure-valued drifts, *Potential Anal.*, **48** (2018), 207-222.

[Ki5] D. Kinzebulatov, Parabolic equations and SDEs with time-inhomogeneous Morrey drift, arXiv:2301.13805.

[Ki6] D. Kinzebulatov, Laplacian with singular drift in a critical borderline case, arXiv:2309.04436.

[KiM1] D.Kinzebulatov and K.R.Madou, Stochastic equations with time-dependent singular drift, *J. Differential Equations*, **337** (2022), 255-293 (arXiv:2105.07312).

[KiM2] D. Kinzebulatov and K.R. Madou, On admissible singular drifts of symmetric α-stable process, *Math. Nachr.*, **295**(10) (2022), 2036-2064 (arXiv:2002.07001).

[KiM3] D. Kinzebulatov and K.R. Madou, Strong solutions of SDEs with singular (form-bounded) drift via Roeckner-Zhao approach, arXiv:2306.04825.

[KiS1] D. Kinzebulatov and Yu.A. Semënov, Brownian motion with general drift, *Stoch. Proc. Appl.*, **130** (2020), 2737-2750 (arXiv:1710.06729).

[KiS2] D. Kinzebulatov and Yu. A. Semënov, On the theory of the Kolmogorov operator in the spaces L^p and C_∞, *Ann. Sc. Norm. Sup. Pisa (5)* **21** (2020), 1573-1647 (arXiv:1709.08598).

[KiS3] D. Kinzebulatov and Yu. A. Semënov, Feller generators and stochastic differential equations with singular (form-bounded) drift, *Osaka J. Math.*, 58 (2021), 855-883 (arXiv:1904.01268).

[KiS4] D. Kinzebulatov and Yu. A. Semënov, Sharp solvability for singular SDEs, *Electr. J. Probab.*, **28** (2023), article no. 69, 1-15. (arXiv:2110.11232).

[KiS5] D. Kinzebulatov and Yu. A. Semënov, Fractional Kolmogorov operator and desingularizing weights, *Publ. Res. Inst. Math. Sci. Kyoto*, to appear (arXiv:2005.11199).

[KiS6] D. Kinzebulatov and Yu. A. Semënov, Heat kernel bounds for parabolic equations with singular (form-bounded) vector fields, *Math. Ann.*, **384** (2022), 1883-1929.

[KiS7] D. Kinzebulatov and Yu. A. Semënov, Kolmogorov operator with the vector field in Nash class, *Tohoku Math. J.*, **74**(4) (2022), 569-596 (arXiv:2012.02843).

[KiS8] D. Kinzebulatov and Yu. A. Semënov, Regularity for parabolic equations with singular non-zero divergence vector fields, arXiv:2205.05169.

[KiS9] D. Kinzebulatov and Yu.A. Semënov, Remarks on parabolic Kolmogorov operator, arXiv:2303.03993.

[KMS] D. Kinzebulatov, K.R. Madou and Yu.A. Semënov, On the supercritical fractional diffusion equation with Hardy-type drift, *J. d'Analyse Mathématique*, to appear (arXiv:2112.06329).

[KSS] D. Kinzebulatov, Yu. A. Semënov and R. Song, Stochastic transport equation with singular drift, *Ann. Inst. Henri Poincaré (B) Probab. Stat.*, to appear (arXiv:2102.10610).

[KSSz] D. Kinzebulatov, Yu. A. Semënov and K. Szczypkowski, Heat kernel of fractional Laplacian with Hardy drift via desingularizing weights, *J. London Math. Soc.*, 104 (2021), 1861-1900 (arXiv:1904.07368).

[KiV] D. Kinzebulatov and R. Vafadar, On divergence-free (form-bounded type) drifts, *Discrete Contin. Dyn. Syst. Ser. S.*, to appear (arXiv:2209.04537).

[Ko] T. Komatsu, On the martingale problem for generators of stable processes with perturbations, *Osaka J. Math.* **21** (1984), 113-132.

[KS] V. F. Kovalenko and Yu. A. Semënov, C_0-semigroups in $L^p(\mathbb{R}^d)$ and $C_\infty(\mathbb{R}^d)$ spaces generated by differential expression $\Delta + b \cdot \nabla$. (Russian) *Teor. Veroyatnost. i Primenen.*, **35** (1990), 449-458; translation in *Theory Probab. Appl.* **35** (1990), 443-453.

[KPS] V. F. Kovalenko, M. A. Perelmuter and Yu. A. Semënov, Schrödinger operators with $L_w^{\frac{1}{2}}(R^l)$-potentials, *J. Math. Phys.*, **22** (1981), 1033-1044.

[Kr1] N. V. Krylov, On diffusion processes with drift in L_d, *Probab. Theory Related Fields* **179** (2021), no. 1-2, 165-199 (arXiv:2001.04950).

[Kr2] N. V. Krylov, On strong solutions of Itô's equations with $A \in W^{1,d}$ and $B \in L^d$, *Ann. Probab.* **49** (2021), no. 6, 3142-3167 (arXiv:2007.06040).

[Kr3] N.V. Krylov, On strong solutions of Itô's equations with $D\sigma$ and b in Morrey classes containing L^d, *Ann. Probab.* **51** (2023), no. 5, 1729-1751 (arXiv:2111.13795).

[Kr4] N.V. Krylov, On parabolic Adams's, the Chiarenza-Frasca theorems, and some other results related to parabolic Morrey spaces, *Mathematics in Engineering*, **5**(2) (2022), 1-20 (arXiv:2110.09555).

[Kr5] N.V. Krylov, On weak solutions of time-inhomogeneous Itô's equations with VMO diffusion and Morrey drift, arXiv:2303.11238.

[Kr6] N.V. Krylov, Once again on weak solutions of time-inhomogeneous Itô's equations with VMO diffusion and Morrey drift, arXiv:2304.04634.

[Kr7] N.V. Krylov, On parabolic equations in Morrey spaces with VMO a and Morrey b, c, arXiv:2304.03736.

[KrR] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131 (2005), 154-196.

[Ku] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990.

[LS] V. A. Liskevich and Yu. A. Semënov, Some problems on Markov semigroups, *"Schrödinger Operators, Markov Semigroups, Wavelet Analysis, Operator Algebras" (M. Demuth et al.„ Eds.), Mathematical Topics: Advances in Partial Differential Equations, Vol. 11, Akademie Verlag, Berlin* (1996), 163-217.

[LZ] V. Liskevich and Q. S. Zhang, Extra regularity for parabolic equations with drift terms, *Manuscripta Math.* **113** (2004), 191-209.

[MK] A. J. Majda and P. R. Kramer, Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena, *Physics Reports* **314** (1999), 237-574.

[MV] V. G. Mazya and I. E. Verbitsky, Form boundedness of the general second-order differential operator, *Comm. Pure Appl. Math.* **59** (2006), 1286-1329.

[MeZ] S. Menozzi and X. Zhang, Heat kernel of supercritical non-local operators with unbounded drift, *J. Éc. polytech. Math.* **9** (2022), 537-579 (arXiv:2012.14475).

[MNS] G. Metafune, L. Negro and C. Spina, "Sharp kernel estimates for elliptic operators with second-order discontinuous coefficients", *J. Evol. Equ.* **18** (2018), 467-514.

[MP] T. Meyer-Brandis and F. Proske, Construction of strong solutions of SDE's via Malliavin calculus, *J. Funct. Anal.*, **258** (2010)(11), 3922-3953.

[MNP] A. Mohammed, T. Nilssen and F. Proske, Sobolev differentiable stochastic flows for SDEs with singular coefficients: applications to the transport equation, *Ann. Probab.*, **43**(3) (2015), 1535-1576.

[N] K. Nam, Stochastic differential equations with critical drifts. *Stoch. Proc. Appl.*, **130** (2020), 5366-5393 (arXiv:1802.00074).

[NU] A. I. Nazarov and N. N. Uraltseva, The Harnack inequality and related properties for solutions to elliptic and parabolic equations with divergence-free lower order coefficients, *Algebra i Analiz*, **23** (2011), 136-168.

[O] E.-M. Ouhabaz, Analysis of Heat Equations on Domains, Princeton Univ. Press, 2005.

[OSSV] E.-M. Ouhabaz, P. Stollmann, K.-Th. Sturm and J. Voigt, The Feller property for absorption semigroups, *J. Funct. Anal.* **138** (1996), 351-378.

[PZ] N. Perkowski and W. van Zuiljen, Quantitative heat kernel estimates for diffusions with distributional drift, *Potential Anal.*, https://doi.org/10.1007/s11118-021-09984-3 (2022)

[Ph] T. Phan, Local $W^{1,p}$-regularity estimates for weak solutions of parabolic equations with singular divergence-free drifts, *Electr. J. Differential Equations* (2017), Paper No. 75, 22 pp.

[P1] N. I. Portenko, Generalized Diffusion Processes. *AMS*, 1990.

[P2] N. I. Portenko, Some perturbations of drift-type for symmetric stable processes, *Random Oper. Stochastic Equations*, **2** (1994), 211-224.

[PP] S. I. Podolynny and N. I. Portenko, On multidimensional stable processes with locally unbounded drift, *Random Oper. Stochastic Equations*, **3** (1995), 113-124.

[Pr] E. Priola, Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math.* **49** (2012), 421-447.

[R] F. Rezakhanlou, Regular flows for diffusions with rough drifts, arXiv:1405.5856.

[RZh1] M. Röckner and G. Zhao, SDEs with critical time dependent drifts: weak solutions, *Bernoulli*, **29** (2023), 757-784 (arXiv:2012.04161).

[RZh2] M. Röckner and G. Zhao, SDEs with critical time dependent drifts: strong solutions, arXiv:2103.05803.

[S1] Yu. A. Semënov, Regularity theorems for parabolic equations, *J. Funct. Anal.*, **231** (2006), 375-417.

[S2] Yu. A. Semënov, On perturbation theory for linear elliptic and parabolic operators; the method of Nash, Proceedings of the Conference on Applied Analysis, April 19-21 (1996), Bâton-Rouge, Louisiana, *Contemp. Math.*, **221** (1999), 217-284.

[Si] B. Simon, Schrödinger semigroups, *Bull. Amer. Math. Soc.*, **7**(3) (1982), 447-526.

[SX] R. Song and L. Xie, Weak and strong well-posedness of critical and supercritical SDEs with singular coefficients, *J. Differential Equations* **362** (2023), 266-313 (arXiv:1806.09033).

[V] A. Yu. Veretennikov, Strong solutions and explicit formulas for solutions of stochastic integral equations, *Matematicheski Sbornik* (in Russian), **111**(3) (1980), 434-452, English translation in *Math. USSR-Sbornik*, **39**(3) (1981), 387-403.

[WLW] J. Wei, G. Lv and J.-L. Wu, On weak solutions of stochastic differential equations with sharp drift coefficients, arXiv:1711.05058.

[W] R. J. Williams, Brownian motion with polar drift, *Trans. Amer. Math. Soc.*, **292** (1985), 225-246.

[XXZZh] P. Xia, L. Xie, X. Zhang and G. Zhao, $L^q(L^p)$-theory of stochastic differential equations, *Stoch. Proc. Appl.* **130** (2020), 5188-5211 (arXiv:1908.01255).

[XZ] L. Xie and X. Zhang, Heat kernel estimates for critical fractional diffusion operators, *Studia Math.* **224** (2014), no. 3, 221-263.

[YZ] S. Yang and T. Zhang, Strong existence and uniqueness of solutions of SDEs with time dependent Kato class coefficients, arXiv:2010.11467.

[Za1] Q. S. Zhang, A strong regularity result for parabolic equations, *Comm. Math. Phys.* **244** (2004) 245-260.

[Za2] Q. S. Zhang, Gaussian bounds for the fundamental solutions of $\nabla(A\nabla u) + B\nabla u - u_t = 0$, *Manuscripta Math.* 93 (1997), 381-390.

[Z1] X. Zhang, Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients, *Electr. J. Prob.*, **16** (2011), 1096-1116.

[Z2] X. Zhang, Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients, *Stoch. Proc. Appl.*, **115**(11) (2005), 1805-1818.

[Z3] X. Zhang, Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients, *Electron. J. Probab.* **16** (2011), no. 38, 1096-1116.

[Z4] X. Zhang, Stochastic differential equations with Sobolev diffusion and singular drift and applications, *Ann. Appl. Probab.*, **26**(5) (2016), 2697-2732.

[Z5] X. Zhang, Stochastic differential equations with Sobolev drifts and driven by α-stable processes. *Ann. Inst. Henri Poincaré (B) Probab. Stat.* **49**(4) (2013), 1057-1079.

[ZZh] X. Zhang, G. Zhao, Stochastic Lagrangian path for Leray solutions of $3D$ Naiver-Stokes equations, *Comm. Math. Phys.*, **381**(2) (2021), 491-525.

[ZZh2] X. Zhang and G. Zhao, Heat kernel and ergodicity of SDEs with distributional drifts, arXiv:1710:10537.

[Zh] G. Zhao, Stochastic Lagrangian flows for SDEs with rough coefficients, arXiv:1911.05562.

[Zh2] G. Zhao, Weak uniqueness for SDEs driven by supercritical stable processes with Hölder drifts, *Proc. Amer. Math. Soc.* **147** (2019), 849-860.

[Zv] A. K. Zvonkin, A transformation of the phase space of a diffusion process that removes the drift, *Math. USSR Sbornik* **22** (1974), 129-149.