SDES WITH CRITICAL GENERAL DISTRIBUTIONAL DRIFTS: SHARP SOLVABILITY AND BLOW UPS

D. KINZEBULATOV AND R. VAFADAR

To the memory of Yu. A. Semënov

Abstract. We establish weak well-posedness for SDEs having discontinuous diffusion coefficients and general distributional drifts that may introduce blow up effects. Our drifts satisfy minimal assumptions, i.e. we assume only that the Cauchy problem for the Kolmogorov backward equation is well-posed in the standard Hilbert triple $W^{1,2} \hookrightarrow L^2 \hookrightarrow W^{-1,2}$. By a result of Mazya and Verbitsky, these assumptions are precisely those drifts that can be represented as the sum of a form-bounded component (encompassing, for example, Morrey or Chang-Wilson-Wolff drifts) and a divergence-free distributional component in the BMO^{-1} space of Koch and Tataru.

We apply our results to finite particle systems with strong attracting interactions immersed in a turbulent flow. This includes particle systems of Keller-Segel type. Crucially, in dimensions $d \geq 3$, we cover almost the entire admissible range of attraction strengths, reaching nearly to the blow-up threshold.

As a further application of our results for SDEs and of the theory of Bessel processes, we obtain an improved upper bound on the constant in the many-particle Hardy inequality. Consequently, the lower bound previously derived by Hoffmann-Ostenhof, Hoffmann-Ostenhof, Laptev, and Tidblom is shown to be close to optimal.

1. Introduction

The subject of this paper is the stochastic differential equation (SDE)

$$X_t = x - \int_0^t c(X_s)ds + \sqrt{2}\int_0^t \sigma(X_s)dB_s, \quad x \in \mathbb{R}^d, \tag{1.1}$$

Key words and phrases. Stochastic differential equations, singular and distributional drifts, divergence-free drifts, form-boundedness, BMO^{-1}, De Giorgi's method, particle systems, Keller-Segel model, blow ups.

The research of D.K. is supported by NSERC grant (RGPIN-2024-04236).

with critical distributional drift c and discontinuous diffusion coefficients σ (see Sections 5 and 6 for our precise setting). Here $\{B_t\}_{t\geq 0}$ denotes a d-dimensional Brownian motion. SDEs with singular drifts arise in various physical models, for instance, the passive-tracer model, where a time-dependent drift c is the velocity field obtained from the Navier–Stokes equations [MK], or finite-particle approximations of the Keller–Segel model of chemotaxis [CP, FJ]. In the latter case one observes blow ups, i.e. when even the weak existence for SDE (1.1) fails once one replaces c by $(1+\varepsilon)c$, $\varepsilon > 0$, in which case all particles collide a.s. in finite time and can stay "stuck" indefinitely (*sticky collision*). Both of these models are within the scope of the present work.

We postpone a survey of the recent literature on SDEs with singular drifts (much of it produced in the past decade) until after we have stated and discussed our assumptions on c.

Our main focus in this paper is on general drifts, i.e. not satisfying any special structural conditions such as control of the sign on $\operatorname{div} c$. So, on the one hand, our condition (A_1)-(A_4) (or, equivalently, (1.6)) on c includes the drifts in the scaling-invariant Morrey class

$$\sup_{r>0,x\in\mathbb{R}^d} r\left(\frac{1}{|B_r|}\int_{B_r(x)}|c|^{2+\varepsilon}dx\right)^{\frac{1}{2+\varepsilon}} < \infty, \tag{$M_{2+\varepsilon}$}$$

or, more generally, the drifts in the Chang-Wilson-Wolff class: for some $\alpha > 0$,

$$\sup_{r>0,x\in\mathbb{R}^d} \frac{1}{|B_r|}\int_{B_r(x)}|c|^2\, r^2\big(1+(\log^+|c|^2\, r^2)^{1+\alpha}\big) < \infty. \tag{1.2}$$

This, for example, provides us with fliexible means to construct interaction kernels in particle systems. Crucially, we reach the blow-up threshold for c, i.e. multiplying c by $1+\varepsilon$ takes us out of the weak existence regime. On the other hand, the same condition (A_1)-(A_4) $\Leftrightarrow$ (1.6) includes drifts arising in some of the physical models mentioned above, notably of the form

$$c = \nabla C, \quad C = -C^\top \text{ is an anti-symmetric matrix field with entries in } \mathrm{BMO}(\mathbb{R}^d). \tag{$\mathbf{BMO}^{-1}$}$$

The class $\mathbf{BMO}^{-1}$ thus consists of *divergence-free* vector fields. It was identified by Koch-Tataru [KT] as a large class of initial conditions for which one can prove the existence and uniqueness of mild solution to 3D Navier-Stokes equations, and which provides natural scale and translation invariant version of L^2 boundedness of this solution. This class contains divergence-free drifts with entries in Besov space of distributions $B^{-1+d/p}_{p,\infty}$, $p > d$, or Morrey class M_1 of Borel measurable drifts, i.e. $\langle |c|\mathbf{1}_{B_r(x)}\rangle \leq Cr^{d-1}$ with constant C independent of r or $x\in\mathbb{R}^d$. Thus, although we refer in this paper to $c\in\mathbf{BMO}^{-1}$ as distributional drifts, the class $\mathbf{BMO}^{-1}$ also contains quite singular Borel measurable vector fields such as

$$c(x) = \left(\frac{x_2}{x_1^2+x_2^2}, \frac{-x_1}{x_1^2+x_2^2}, 0, \dots, 0\right), \quad x\in\mathbb{R}^d, \tag{1.3}$$

which do not belong $[L^2_{\rm loc}]^d$.

The diffusion coefficients σ in Theorem 6.1 can have have critical discontinuities. This allows us to handle some systems of particles that interact via diffusion coefficients. To keep the introduction concise, we defer the discussion of non-constant diffusion coefficients to 6.

We test our results for general SDE (1.1) against some interacting particle systems:

Example 1.1 (Brownian particles in a turbulent flow)**.** We aim at describing the dynamics of N interacting particles in $\mathbb{R}^d$, $d\geq 3$, that experience strong mutual attraction while being advected by a turbulent flow whose divergence-free distributional velocity field belongs to $\mathbf{BMO}^{-1}(\mathbb{R}^d)$.

To this end, we work in $\mathbb{R}^{Nd}$ where N is large, and consider SDE

$$X_t = x_0 - \int_0^t c(X_s)ds + \sqrt{2}B_t, \quad x_0 = (x_0^1, \dots, x_0^N) \in \mathbb{R}^{Nd}, \tag{1.4}$$

where $X_t = (X_t^1, \dots, X_t^N)$, X_t^i is the position of the i-th particle at time t, $B_t = (B_t^1, \dots, B_t^N)$, $\{B_t^i\}_{t\geq 0}$, $(i = 1, \dots, N)$ are independent Brownian motions in $\mathbb{R}^d$. We further take

$$c = b + q$$

where:

– The first drift $b : \mathbb{R}^{Nd} \to \mathbb{R}^{Nd}$ is given component-wise by

$$b^i(x^1, \dots, x^N) := \frac{1}{N} \sum_{j=1, j\neq i}^{N} \sqrt{\kappa}\frac{d-2}{2} e_{i,j}(x) \frac{x^i - x^j}{|x^i - x^j|^2}, \quad x^i, x^j \in \mathbb{R}^d, \tag{1.5}$$

where $e_{ij} \in L^\infty(\mathbb{R}^{Nd})$, $\|e_{ij}\|_\infty \leq 1$. Setting $e_{ij} = 1$ introduces attraction between the particles arising, for example, in the Keller-Segel model of chemotaxis[1]. The parameter $\kappa > 0$ measures the strength of attraction.

– The second drift q is described by external divergence-free velocity field,

$$q(x^1, \dots, x^N) = \big(q_0(x^1), \dots, q_0(x^N)\big), \quad q_0 \in \mathbf{BMO}^{-1}(\mathbb{R}^d),$$

so, it is easily seen, $q \in \mathbf{BMO}^{-1}(\mathbb{R}^{Nd})$.

Either of our main results, Theorem 5.1 or Theorem 8.1 for $e_{ij} = 1$, applies to $c = b + q$ and provide, in particular, weak existence and approximation uniqueness for particle system (1.4). Both theorems impose dimension-independent conditions, and so the resulting constraint on κ does not degenerate as the number of particles N goes to infinity. See Examples 5.1, 8.1 where we detail the particle system (1.4). In fact, in Example 8.1 we show that when all $e_{ij} = 1$ Theorem 8.1 allows us to handle all $\kappa < 16$ for all N, which is close to the blow up threshold for (1.4), i.e. if κ is greater than a constant that is slightly larger than 16, then all particles collide in finite time and stay glued to each other.

Section 4 compares our results (in the case $q_0 = 0$ and $e_{ij} = 1$) with those of Cattiaux-Pédèches [CP], Fournier-Jourdain [FJ], Fournier-Tardy [FT] and Tardy [T] whose methods exploit the special structure of the drift (1.5) to obtain sharp, detailed results.

The search for the maximal admissible value of the strength of attraction κ before a blow up regime can be re-stated as the search for the minimal level of thermal excitation that prevents sticky collisions between the particles. The fact that the noise can turn local in time solutions into global ones can be viewed as another instance of regularization by noise (regarding the latter, see [Fl, FGP, FR]).

1.1. **Result #1.** In Theorem 5.1 and also in Theorem 6.1, we establish weak well-posedness of SDE (1.1) for the following class of drifts:

$$c = b + q, \tag{A_1}$$

where q is in general distribution-valued,

$$q \in \mathbf{BMO}^{-1} \qquad (\Rightarrow \quad \operatorname{div} q = 0), \tag{A_2}$$

[1]However, in this example we assume $d \geq 3$, but will be able to include the case $d = 2$ as well, albeit, at the moment, only partially, see Section 7.

and b is a Borel-measurable drift satisfying

$$b \in \mathbf{F}_\delta, \quad \text{i.e.} \quad |b| \in L^2_{\rm loc} \text{ and } \|b\varphi\|_2^2 \leq \delta \|\nabla \varphi\|_2^2 + c_\delta \|\varphi\|_2^2, \quad \forall \varphi \in W^{1,2}, \tag{A_3}$$

for some finite constants δ (important) and c_δ (only its finiteness is important). That is, b is *form-bounded*. The last condition covers several important cases. It includes the Morrey class $M_{2+\varepsilon}$ (with form-bound δ depending on the Morrey norm $\|b\|_{M_{2+\varepsilon}}$), the larger Chang-Wilson-Wolff class introduced in (1.2) and, when we work in $\mathbb{R}^{Nd}$, the many-particle drift (1.5) in Example 1.1 where one has $\delta = (N-1)^2 N^{-2}\kappa$. A fuller discussion and more examples appear in Section 3.

The constant δ measures the size or the "strength" of singularities. In Theorem 5.1 we impose a completely dimension-free condition on δ:

$$\delta < 4. \tag{A_4}$$

A key consequence is that the corresponding restriction on the attraction parameter $\kappa = N^2(N-1)^{-2}\delta$ in Example 1.1 does not degenerate as the number of particles N tends to infinity. (Theorem 8.1 relaxes this constraint on κ by taking into account the divergence of b.)

Our proofs make use of specific properties of the classes $\mathbf{F}_\delta$ and $\mathbf{BMO}^{-1}$, such as the compensated compactness estimates and results on BMO-multipliers.

As was mentioned, class $\mathbf{BMO}^{-1}$ also contains some quite singular Borel-measurable vector fields such as (1.3) that are not in $[L^2_{\rm loc}]^d$ and therefore are not form-bounded.

1.1.1. *Optimality of our conditions.* We claim that condition (A_1)-(A_4) is close to optimal, i.e. it cannot be substantially improved. This is justified by the following two observations.

– Mazya and Verbitsky [MV] proved that conditions (A_1)–(A_3) are *equivalent* to the estimate

$$|\langle c \cdot \nabla \varphi, \eta \rangle| \leq \alpha \|\nabla \varphi\|_2 \|\nabla \eta\|_2 \tag{1.6}$$

for all $\varphi, \eta \in C_c^\infty(\mathbb{R}^d)$, for some constant $\alpha > 0$ (up to replacing the homogeneous Sobolev spaces with their non-homogeneous counterparts, see Section 6). If $\alpha < 1$, then inequality (1.6) ensures that the KLMN theorem [Ka, Ch. 6,§2] applies and so the Cauchy problem for the Kolmogorov backward equation is weakly well-posed in the standard Hilbert tripe of Sobolev spaces $W^{1,2} \hookrightarrow L^2 \hookrightarrow W^{-1,2}$. To the extent that one can view the latter as a minimal theory of the Kolmogorov backward equation, the present paper bridges the Eulerian and Lagrangian descriptions of diffusion:

$$\text{Euler} \qquad \longleftrightarrow \qquad \text{Lagrange.}$$

Here:

- Eulerian viewpoint: one studies averaged quantities, such as temperature or concentration, governed by the heat equation or, more generally, by transport-diffusion PDEs.
- Lagrangian viewpoint: one follows individual molecules whose trajectories solve SDEs (Brownian motion in the simplest case).

Until quite recently, Lagrangian results required stronger regularity assumptions on the drift. One can argue that the Lagrangian picture offers finer detail.

We refer to the equivalence (A_1)-(A_3) $\Leftrightarrow$ (1.6) as the "generalized form-boundedness", although Mazya and Verbitsky [MV] call (1.6) simply the form-boundedness condition on c.

Let us comment on why one would expect the Kolmogorov equation to be well-posed in the standard Hilbert tripe. In fact, we already can substantially relax the conditions on drift b by requiring well-posedness of the Kolmogorov backward equation in the shifted triple of Bessel spaces $\mathcal{W}^{\frac{3}{2},2} \hookrightarrow \mathcal{W}^{\frac{1}{2},2} \hookrightarrow \mathcal{W}^{-\frac{1}{2},2}$, see Theorem 7.1. Doing so, however, forces us to drop the distributional part q, lose at least for now the almost optimal dimension-independent condition (A_4) and precludes discontinuous diffusion coefficients. By contrast, under (A_1)-(A_4) we can allow discontinuous diffusion coefficients σ for which the associated non-divergence operator can still be rewritten in divergence form (Theorem 6.1). Such diffusion coefficients keep us within the standard Hilbert tripe $W^{1,2} \hookrightarrow L^2 \hookrightarrow W^{-1,2}$, as discussed in Section 6.

– Near-optimality of the bound (A_4). Consider the SDE

$$X_t = x - \sqrt{\delta}\frac{d-2}{2}\int_0^t \frac{X_s}{|X_s|^2}ds + \sqrt{2}B_t, \tag{1.7}$$

where B_t is the d-dimensional Brownian motion, $d \geq 3$. The drift here, $b(x) = \sqrt{\delta}\frac{d-2}{2}\frac{x}{|x|^2}$, belongs to $\mathbf{F}_\delta$, see Section 3 for a detailed explanation. Theorems 5.1 and 8.1 show that if (A_4) is satisfied, then there exists a strong Markov family of weak solutions to SDE (1.7), and these weak solutions are unique in an appropriate sense. Moreover, in the critical case $\delta = 4$ there is still a sufficiently rich theory of the corresponding Kolmogorov backward PDE, see Remark 5.1. On the other hand, taking advantage of the anti-symmetry of the drift, one can show that $R_t = |X_t|^2$ is a squared Bessel process (see, e.g. [BFGM]) and therefore:

(a) If

$$\delta \geq 4\left(\frac{d}{d-2}\right)^2,$$

then for every initial point solution arrives at the origin in finite time with probability 1, and stays there indefinitely (that is, $R_t = 0$ for all $t \geq \tau$ for some finite stopping time τ). A simple argument shows that for such δ SDE (1.7) does not have a weak solution departing from $x = 0$, see e.g. [BFGM].

(b) If

$$4 < \delta < 4\left(\frac{d}{d-2}\right)^2,$$

then solution still visits the origin infinitely many times, but does not stay there, i.e. $\int_0^\infty R_t dt < \infty$ a.s.

In Section 4 we discuss similar counterexamples in the context of particle system inntroduced in Example 1.1.

Remark 1.1. The fact that (1.6) implies (A_1)–(A_3) is proved in [MV] by taking b and q from the Hodge-type decomposition of c:

$$b := -\nabla(1-\Delta)^{-1}\operatorname{div} c + (1-\Delta)^{-1}c, \tag{1.8}$$

$$q := -(1-\Delta)^{-1}\operatorname{curl} c. \tag{1.9}$$

Here, b caputres the attractive or the repulsive part of the drift in the dynamics of X_t.

Much more is already known when the entire drift belongs to the form-bounded class $\mathbf{F}_\delta$: one has strong well-posedness for the SDE, as well as well-posedness for the associated stochastic transport equation, see Theorem 5.2. How far these results extend when a non-zero distributional component q in $\mathbf{BMO}^{-1}$ is present is still unclear: if we insist on the decomposition (1.8), (1.9) above, Theorem 5.1 would treat only gradient-type drifts, which we want to avoid.

1.1.2. *Blow ups.* In Section 8 we will also discuss drifts having critical (form-bounded) divergence, which allows to relax the assumptions on the drift to some super-critical conditions, i.e. passing to the small scales actually increases the norm of the drift. The sub-critical/critical/super-critical classification of the spaces of vector fields, widely used in the literature, is recalled in Remark 8.2. It should be added, however, that this classification is not so relevant to the main body of the present paper (except for Section 8) because:

(a) Our main focus is on general drifts, for which one only has the dichotomy sub-critical/critical.

(b) This classification does not distinguish between critical spaces that do, and those that do not, reach blow-ups.

(Regarding (b), for example, both L^d and weak L^d spaces are critical, but only the weak L^d space contains drift $C\frac{x}{|x|^2}$ whose attracting singularity at the origin is strong enough to kill the weak well-posedness of the SDE if C is too large. In other words, in the critical case, a lot depends on the definition of the norm of the drift.)

If a critical class of drifts is broad enough to contain blow-up examples, a well-posedness theorem must include a smallness condition on the drift norm, such as e.g. (A_4).

Compared with blow-ups for the Navier–Stokes equations, the mechanism of blow-ups in particle systems of the kind treated in Example 1.1 is much better understood [CP, CPZ, FJ, FT, JL]. That said, as was noted in [FJ], at the time of writing of their article there was still a substantial gap between (i) the drift singularities that the general theory of SDEs with singular drifts could handle and (ii) the even stronger singularities they themselves had to treat. One of the purposes of the present work, together with [KS2, K6, KS6], is to close this gap:

$$\text{SDEs with general singular drift} \qquad \longrightarrow \qquad \text{particle systems,} \tag{1.10}$$

i.e. to bring the general theory of SDEs "up to the task" so that it can handle blow ups. See, in particular, Section 7 where we demonstrate how the weak well-posedness of the finite-particle Keller-Segel SDE in $\mathbb{R}^{2N}$ can, in principle, be reached from an earlier result in [KS1] on SDEs with general drifts. Interestingly, the path that leads to this passes through non-local operators (Theorem 7.1 and Corollary 7.1).

The same connection (1.10) was already pursued by Krylov and Röckner [KrR], but with a different interaction kernel: its attractive part can be extremely singular, but it is always dominated on average (not pointwise) by the repulsive part, so no blow-up occurs.

1.2. **Result #2.** As mentioned earlier, in Section 8 we will also discuss super-critical drifts (necessarily under additional assumptions on their divergence). Specifically, if the positive part of the divergence $\operatorname{div} b$ is a form-bounded potential, then this enables us to relax the form-boundedness condition on b to a super-critical form-boundedness condition:

$$|b|^{\frac{1+\nu}{2}} \in \mathbf{F}_\delta, \quad \nu \in]0,1[, \quad \delta < \infty.$$

In this case one still has weak existence for every initial point. This was alrady proved in [KS2]. Moreover, as it was outlined in [KS2], one can combine such a drift with an ordinary form-bounded component, again reaching the blow-up threshold, and may also include discontinuous diffusion coefficients of the type treated in Section in Section 6.

Generally speaking, in super-critical settings many standard regularity properties of the diffusion process are lost. Some of them can be saved, but to recover most of them one must impose additional critical conditions on the drift. In our framework, to establish e.g. the Markov property, Theorem 8.3 supplements the super-critical asumption with the multiplicative form-boundedness condition $b \in \mathbf{MF}_\delta$, i.e.

$$\langle |b|\varphi, \varphi\rangle \leq \delta \|\nabla \varphi\|_2 \|\varphi\|_2 + c_\delta \|\varphi\|_2^2, \quad \varphi \in C_c^\infty(\mathbb{R}^d), \quad \delta < \infty.$$

This condition is critical and is substantially more general than $b \in \mathbf{F}_\delta$ because somehow it implicitly presumes the existence of $\operatorname{div} b$. In contrast to the form-bounded class $\mathbf{F}_\delta$, the multiplicative form-bounded class $\mathbf{MF}_\delta$ can be completely characterized in elementary terms: $b \in \mathbf{MF}_\delta$ if and only if

$$\langle |b| \mathbf{1}_{B_r(x)} \rangle \leq C r^{d-1}$$

with constant C independent of r or $x \in \mathbb{R}^d$. The latter is the largest scaling-invariant Morrey[2] class M_1. In fact, we have

$$\begin{tikzcd} & {|b| \in M_1} \arrow[dl, leftrightarrow, "\text{form-bounded } (\operatorname{div} b)_+"'] \arrow[dr, "\operatorname{div} b=0"] & \\ b \in \mathbf{MF}_\delta \text{ for some finite } \delta \text{ and } c_\delta = 0 & & b \in \mathbf{BMO}^{-1} \end{tikzcd}$$

where the arrow $\to$ in this diagram means inclusion. (These inclusions were proved in [M, Theorem 1.4.7] and [MV2, Theorem V]. See discussion before Theorem 8.3 and the proof in Appendix D.) Choosing the inclusion into $\mathbf{MF}_\delta$ leads to an approach to studying the Kolmogorov backward equation based on "Caccioppoli's iterations" for establishing the classical Caccioppoli's inequality [KV], see Remark 8.1 for details; choosing the inclusion into $\mathbf{BMO}^{-1}$ amounts to absorbing the stream matrix Q (for $b = \nabla Q$) into the diffusion coefficients and obtaining a Caccioppoli-type inequality from there [H, SSŠZ]. We discuss this in Section 8.

1.3. **Result #3.** As an application of our results on SDEs (Theorem 8.1) and of the theory of Bessel processes, we obtain an upper bound on the best possible constant $C_{d,N}$ in the many-particle Hardy inequality

$$C_{d,N} \sum_{1 \leq i < j \leq N} \int_{\mathbb{R}^{Nd}} \frac{|\varphi(x)|^2}{|x^i - x^j|^2} dx \leq \int_{\mathbb{R}^{Nd}} |\nabla \varphi(x)|^2 dx \quad \forall \, \varphi \in C_c^\infty(\mathbb{R}^{Nd})$$

(Theorem 8.2). Our upper bound on $C_{d,N}$ improves the existing results by replacing factorial growth in the dimension with polynomial growth. It also shows that the lower bound on $C_{d,N}$ obtained earlier by Hoffmann-Onstenhof–Hoffmann-Onstenhof-Laptev-Tidblom [HHLT] is nearly optimal – at least in high dimensions.

To our knowledge, this is the first time a probabilistic argument is used in the analysis of the best possible constant in a Hardy inequality.

[2]For the theory of Morrey spaces and related estimates, see Adams-Xiao [AX] and Krylov [Kr5].

1.4. **Literature.** Let us now comment on the existing literature on SDEs with singular drifts, focusing mostly on general drifts.

1) The present paper continues [KS7, KS2, K5, K6, KS6]. This series of papers was initiated by Semënov and the first author in [KS1]. Building on the PDE results in [K1], that paper considered weak solutions to SDEs whose drifts that lie in a class even larger than $\mathbf{F}_\delta$, namely, the class of *weakly form-bounded* vector fields. This larger class contains, for example, the Morrey class $M_{1+\varepsilon}$, and therefore vector field (1.3); however, that result requires more restrictive condition $\delta < \frac{c}{d^2}$ (see Section 7). In all those earlier works one has $q = 0$ and $\sigma = I$ (excelpt [KS7]). Allowing a non-zero distributional $q \neq 0$ drift, as in (A_1)-(A_4), necessitated substantial modifications of the arguments in [KS2, K6, KS6], in particular the systematic use of compensated-compactness estimates.

2) If we restrict attention to Borel-measurable drifts, the most closely related results are the recent works of Krylov [Kr1, Kr2, Kr3, Kr4] and of Röckner and Zhao [RZ]. Their approaches differ from ours, both herein and in [KS7, KS2, K5, K6, KS6]. In particular, in [Kr1] Krylov treats a class of diffusion coefficients much larger than ours (his diffusion coefficients are in VMO class, or have small BMO norm), but a smaller class of drifts (namely, the Morrey class $M_{\frac{d}{2}+\varepsilon}$). The author's analysis requires estimates on second-order derivatives of solutions to the Kolmogorov backward equation, whereas we neither assume nor have such estimates under our assumptions on the drift. Krylov's drifts can also produce blow-up phenomena, but he assumes that the size of the singularity, measured by the Morrey norm, remains below a small, dimension-dependent constant. By contrast, our goal is to reach the maximal admissible strength of singularity (in Example 1.1, the interaction strength κ; or the supremum (1.2)). Other recent results of Krylov are described at various points in this work.

Let us add that, even if we restrict our attention to Borel measurable drifts, in our setting Zvonkin transform is not applicable and our drifts are not of Girsanov type.

3) Some techniques such as De Giorgi's method also connect our paper to the works of Zhang-Zhao [ZZ1] and Hao-Zhang [HZ], which focus on super-critical divergence-free drifts (in [HZ], distributional). Some of their results deal with non-zero divergence, but they do not reach blow ups.

4) There is rich literature on SDEs with general distributional drifts, also motivated by physical applications. In all the results we are aware of, the diffusion coefficients σ have to be at least Hölder continuous. Many of these works treat time-inhomogeneous drifts, however, in this brief discussion we will specify these results to time-homogeneous drifts. Flandoli-Issoglio-Russo [FIR] and Zhang-Zhao [ZZ2] study drifts that belong to Bessel potential spaces with negative index and employ the Zvonkin transform to obtain weak well-posedness for SDE (1.1). Chaudru de Raynal-Menozzi [CM] proved, among other results, weak well-posedness with drift b in Besov space $B^{\beta}_{p,q}$ with $-\frac{1}{2} < \beta \leq 0$ and $p > d$. They addressed, in particular, the problem of defining the product of two distributions $b \cdot \nabla v$, where v is a solution to the Kolmogorov backward equation. The latter, in turn, dictates their restriction $\beta > -\frac{1}{2}$. Earlier, Delarue-Diel [DD] and Cannizzaro-Chouk [CC] proved, in dimensions $d = 1$ and $d \geq 2$, respectively, well-posedness of the martingale problem for general distributional drifts in Hölder-Besov space $B^{\beta}_{\infty,\infty}$ for all $-\frac{2}{3} < \beta \leq 0$. In the more singular regime $-\frac{2}{3} < \beta \leq -\frac{1}{2}$ they assume that the drift can be enhanced to a rough distribution, in order to apply the theory of paracontrolled distributions. This allowed them to consider random drifts of the form $b = \nabla h$, where h is a solution of a KPZ-type equation, and thereby construct the polymer measure with white noise potential. As mentioned earlier, the class of divergence-free drifts $\mathbf{BMO}^{-1}$ contains drifts whose components lie in Besov space $B^{-1+d/p}_{p,\infty}$, $p > d$, i.e. the

exponent $\beta = -1 + d/p$ can go up to -1. This, however, cannot serve as a comparison with the previous cited results since by definition $\mathbf{BMO}^{-1}$ drifts are divergence-free, but this justifies to some extent why we do not address in this paper the problem of mutiplying distributions in the Kolmogorov equation, i.e. since our $\beta \not> -\frac{1}{2}$.

5) Let us also mention recent papers by Chaudru de Raynal-Jabir-Menozzi [CJM, CJM2] where the authors handle McKean-Vlasov SDEs with distributional Besov drifts. The regularization by convolution allows them to venture substantially farther in the assumptions on the drift, compared to the papers cited in 4). It is quite noteworthy, since they can start with a delta-function in the initial distribution.

The papers on general distributional drifts mentioned in 4) and 5) do not reach blow ups, in contrast to our Theorems 5.1, 6.1, 8.1. That said, the cited works cover some other highly irregular drifts that fall outside the scope of our hypotheses.

Finally, we mention Bresch-Jabin-Wang [BJW] who consider gradient-form singular attracting drifts that allow blow ups. Their focus is different, namely, it is quantiative propagation of chaos at the PDE level, but they are also interested in reaching the blow up thresholds. Their their drifts are not included in our setting.

6) Regarding super-critical drifts, we refer again to the earlier paper by Zhang-Zhao [ZZ1], as well as to recent papers by Hao-Zhang [HZ] and Gräfner-Perkowski [GP] where the authors consider super-critical distributional drifts. In [GP], the authors treat divergence-free drifts Besov drifts $b \not\in B^{-1}_{2d(d+2)^{-1},2}$, provided that the initial density is absolutely continuous with respect to the Lebesgue measure; they also consider quite irregular attracting/repulsing component of the drift, although it does not reach the blow ups.

Comprehensive surveys of the literature on SDEs with singular and distributional drifts can be found in [CM] and [HZ].

1.5. **Main instruments: De Giorgi's method, Trotter's theorem and compensated compactness estimates.** Since the drift c satisfying (1.6) is in general distribution-valued ($c = b + q$) we actually have to give a proper meaning to term $c(X_s)$ in SDE (5.1). We will do it in two ways:

– Theorem 5.1(iv): At the level of the martingale problem and with the Itô expansion, we have to define $(c \cdot \nabla v)(X_s)$ for suitable test functions. The usual test functions in C_c^∞ cannot be used, but it is still possible to find a sufficiently rich space of test functions (in particular, dense in C_∞; see notations in Section 2) that will give us a continuous martingale. In fact, this space of test functions will be the domain of the generator $\Lambda \supset -\Delta + c \cdot \nabla$ of a strongly continuous Feller semigroup in C_∞.

– Theorem 5.1(v): By constructing the "limiting drift" process (called "formal dynamics" in [CM]) for disperse initial data, assuming $\delta < 1$. This is a result of Bass-Chen type [BC], see also Zhang-Zhao [ZZ2], Chaudru de Raynal-Menozzi [CM] and Hao-Zhang [HZ] who have similar results. The proof will use convergence of the martingale solutions of the approximating SDEs (5.2) that will follow as a by-product of (a).

To construct the sought Feller semigroup, we will have to employ some deep results from the operator theory, the theory of PDEs and harmonic analysis:

- Trotter's approximation theorem,
- De Giorgi's method in L^p for $p > \frac{2}{2-\sqrt{\delta}}$, i.e. in the context of Example 1.1, p will be an explicit function of the strength of attraction between the particles κ.
- Compensated compactnes esimates and BMO-multipliers.

A crucial feature of Trotter's approximation theorem (Theorem 10.1) is that it requires no a priori knowledge of the limiting object, i.e. in our case, the limiting Feller semigroup. By contrast, in some other of our results on singular SDEs (such as Theorem 7.1) we first construct an explicit candidate for the Feller resolvent. This simplifies approximation arguments, but it also automatically provides strong gradient bounds that, in turn, introduce dimension-dependent restrictions on the form bound δ. As the dimension grows, the admissible range of δ shrinks to the empty set, making problematic possible applications to many-particle systems that exist in spaces of very large dimension. (Example 5.1 explains how a bound on δ translates into a bound on the attraction strength κ.) Using De Giorgi's method in the proof of Theorem 5.1 decouples gradient bounds from the bounds that provide the existence of the weak solution (i.e. tightness argument, construction of the Feller semigroup, etc). Consequently, we can impose dimension-free assumptions on δ, which is a crucial point for applications to many-particle systems.

De Giorgi's method for divergence-form operators with drift in $\mathbf{BMO}^{-1}$ has also been applied, in different contexts, by Hara [H] and Seregin-Silvestre-Šverak-Zlatoš [SSŠZ].

The problem of constructing Feller generator realization of $-\Delta + b\cdot\nabla$, $b \in \mathbf{F}_\delta$ was addressed earlier in [KoS, K1, K7, KS6] (under dimension-dependent conditions on δ except the last cited paper). For such b, one can also solve the classical martingale problem and, moreover, establish strong existence and conditional uniqueness, see Theorem 5.1; the Feller semigroup is one of several properties of the process. In the case of a distributional drift c, however, the Feller semigroup plays an even larger role since it also allows to address the difficulty with evaluating distributional drift along a trajectory.

1.6. **Structure of the paper.**

- Section 4 We begin with a preliminary discussion of blow-up thresholds for the particle system in 1.1. These thresholds both illustrate the sharpness of our SDE results and serve as input in Section 8, where they enter the proof of a sharper upper bound for the constant in the many-particle Hardy inequality (Theorem 8.2).
- Sections 5 Theorem 5.1 is our main result when the diffusion matrix is constant. Theorem 5.2 surveys what is currently known about SDEs with form-bounded drifts, meaning drifts with no distributional component in $\mathbf{BMO}^{-1}$, including strong solutions, conditional uniqueness, and regularity for the associated stochastic transport equation.
- Section 6 Theorem 6.1 extends the analysis to diffusion matrices that may have critical discontinuities.
- Section 7 Corollary 7.1 shows, in principle, how well-posedness of the finite-particle Keller–Segel system can be derived from an earlier result on SDEs with general singular drifts (Theorem 7.1). The proof of that theorem uses a more operator-theoretic approach, based on fractional resolvent representations, and covers the larger class of *weakly form-bounded drifts* that contains both the Morrey class $M_{1+\varepsilon}$ and the Kato class studied by Bass-Chen [BC]. We also comment in this section on general time-inhomogeneous drifts with critical singularities both in time and space.
- Section 8 Theorems 8.1 and 8.3 treat drifts whose divergence satisfies a critical (scaling-invariant) form-boundedness condition, with the distributional part set to zero. Under this assumption we can relax the constraint on $|b|$, allowing some super-critical form-boundedness conditions. As an application, Theorem 8.2 gives an improved upper bound on the constant in the many particle Hardy inequality.

Section 9 We make a few more comments regarding De Giorgi's method in L^p, the optimal choice of test function in the analysis of Kolmogorov backward equation, and time-inhomogeneous drifts.

Acknowledgements. We are grateful to Galia Dafni and Jean-François Jabir for very useful discussions.

2. Notations and auxiliary results

Throughout this work we use the following notations.

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

The space of d-dimensional vectors with entries in X is denoted by $[X]^d$. We reserve the upper index to denote the components $q^i \in X$ of $q \in [X]^d$.

The notation "$\overset{w}{\to}$ in X" stands for the weak covergence in X. Similarly for $[X]^d$.

We write

$$T = s\text{-}Y\text{-}\lim_n T_n$$

for T, $T_n \in \mathcal{B}(X,Y)$ if

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

By $T \upharpoonright X$ we denote the restriction of operator T to a subspace $X \subset D(T)$.

By $\big[T \upharpoonright X\big]^{\rm clos}_{Y\to Y}$ we denote the closure of the restriction $T \upharpoonright X$ (when it exists).

2. The space $L^p = L^p(\mathbb{R}^d, dx)$, $W^{1,p} = W^{1,p}(\mathbb{R}^d, dx)$ corresponds to the Lebesgue and to the Sobolev space, respectively. Let $L^p_\rho = L^p(\mathbb{R}^d, \rho(x)dx)$ denote the weighted Lebesgue space with weight ρ (defined by formula (2.1) below).

Set $\|\cdot\|_p := \|\cdot\|_{L^p}$ and denote operator norm $\|\cdot\|_{p\to q} := \|\cdot\|_{L^p\to L^q}$.

Given $1 < p < \infty$, we set $p' := \frac{p}{p-1}$.

Put

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

(all functions considered in this paper are real-valued). For vector fields b, $\mathsf{f} : \mathbb{R}^d \to \mathbb{R}^d$, we put

$$\langle b, \mathsf{f}\rangle := \langle b \cdot \mathsf{f}\rangle \qquad (\cdot \text{ is the scalar product in } \mathbb{R}^d).$$

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

C_b is the space of bounded continuous functions on $\mathbb{R}^d$ endowed with the sup-norm, and C_b^k is the subspace of bounded continuous functions with bounded continuous derivatives up to order k.

C_∞ is a closed subspace of C_b consisting of functions vanishing at infinity.

We denote by $\mathcal{S}$ the Schwartz space, and by $\mathcal{S}'$ the space of tempered distributions on $\mathbb{R}^d$.

Put

$$\Gamma_c(t,x) = \Gamma(ct,x) := (4c\pi t)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4ct}},$$

the Gaussian density.

Set

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

where c is adjusted to $\int_{\mathbb{R}^d}\gamma(x)dx = 1$, and put $\gamma_\varepsilon(x) := \frac{1}{\varepsilon^d}\gamma\left(\frac{x}{\varepsilon}\right)$, $\varepsilon > 0$, $x \in \mathbb{R}^d$. Define the De Giorgi mollifier of a function $h \in L^1_{\rm loc}$ (or a vector field with entries in $L^1_{\rm loc}$) by

$$E_\varepsilon h := e^{\varepsilon\Delta}h.$$

$B_r(x)$ denotes the open ball of radius r centered at $x \in \mathbb{R}^d$. If $x = 0$, we simply write B_r.

Given a function $f \in L^1_{\rm loc}$, we denote by $(f)_{B_r(x)}$ its average over the ball $B_r(x)$:

$$(f)_{B_r(x)} := \frac{1}{|B_r|}\int_{B_r(x)} f dx.$$

If $x = 0$, then we write $(f)_r \equiv (f)_{B_r}$.

We denote the positive and negative parts of function f by

$$(f)_+ := f \vee 0, \quad (f)_- := -(f \wedge 0).$$

Let $\nabla_i = \partial_{x_i}$, $1 \leq i \leq d$.

Define weight

$$\rho(y) \equiv \rho_{\epsilon_0}(y) = (1+\sigma|y|^2)^{-\frac{d+\epsilon_0}{2}}, \quad \varepsilon_0 > 0, \quad \sigma > 0. \tag{2.1}$$

This weight has property

$$|\nabla\rho| \leq \frac{d+\epsilon_0}{2}\sqrt{\sigma}\rho. \tag{2.2}$$

In the same way, $|\nabla\rho^{-1}|$ $(= \frac{|\nabla\rho|}{\rho^2})$ $\leq \frac{d+\epsilon_0}{2}\sqrt{\sigma}\rho^{-1}$. (The last two inequalities will allow us to replace all occurrences of $\nabla_i\rho$ or $\nabla_i\rho^{-1}$ resulting from the integration by parts in the analysis of PDEs in weighted spaces by the weight ρ itself or ρ^{-1}, respectively, times a constant that is proportional to $\sqrt{\sigma}$. This constant can be made arbitrarily small by fixing σ sufficiently small. We will use this to get rid of the terms containing $\nabla_i\rho$ or $\nabla_i\rho^{-1}$.) Put

$$\rho_x(y) := \rho(x-y).$$

3. Let $\mathbf{C}$ and $\mathbf{D}$ denote the canonical spaces of continuous and càdlàg trajectories from $[0,\infty[$ to $\mathbb{R}^d$ equipped with the uniform topology and the Skorohod topology, respectively, endowed with the natural filtration $\mathcal{B}_t = \sigma\{\omega_s \mid 0 \leq s \leq t\}$, where ω_t is the coordinate process.

Let $\mathcal{P}(\mathbf{C})$ and $\mathcal{P}(\mathbf{D})$ denote the space of probability measures on $\mathbf{C}$ and $\mathbf{D}$, respectively.

Recall that a probability measure $\mathbb{P}_{s,x}$ ($s \geq 0$, $x \in \mathbb{R}^d$) on $\mathbf{C}$ is said to be a classical martingale solution to SDE

$$X_t = x - \int_s^t b(r,X_r)dr + \sqrt{2}(B_t - B_s), \quad x \in \mathbb{R}^d, \quad t \geq s, \tag{2.3}$$

with a time-inhomogeneous drift $b \in L^1_{\rm loc}(\mathbb{R}^{1+d})^d$ if $\mathbb{P}_{s,x}[\omega_s = x] = 1$,

$$\mathbb{E}_{\mathbb{P}_{s,x}}\int_s^t |b(r,\omega_r)|dr < \infty$$

and for every $v \in C_c^2$

$$t \mapsto v(\omega_t) - x + \int_s^t (-\Delta + b\cdot\nabla)v(\omega_r)dr, \quad t \geq s,$$

is a continuous martingale with respect to $\mathbb{P}_{s,x}$.

If b does not depend on time, then we take $s=0$, and in the above martingale problem write $\mathbb{P}_x$.

4. *BMO functions.* (a) A function $g \in L^1_{\rm loc}$ is in class BMO $=$ BMO$(\mathbb{R}^d)$ if

$$\|g\|_{\rm BMO} := \sup_{x\in\mathbb{R}^d, R>0} \frac{1}{|B_R|}\int_{B_R(x)} |g-(g)_{B_R(x)}|dy < \infty.$$

One can also define the BMO semi-norm as

$$\|g\|_{\rm BMO} = \biggl(\sup_{x\in\mathbb{R}^d, R>0} \frac{1}{|B_R|}\int_{B_R(x)}\int_0^{R^2} |\nabla e^{t\Delta}g|^2 dtdy\biggr)^{\frac{1}{2}} \tag{2.4}$$

(this is Carleson's characterization of BMO functions).

(b) We will need the compensated compactness estimate of Coifman-Lions-Meyer-Semmes:

Proposition 2.1 ([CLMS])**.** *There exists a constant C_d such that for every anti-symmetric matrix field Q with entries in* BMO *one has*

$$|\langle Q\cdot\nabla f, \nabla g\rangle| \leq C_d\|Q\|_{\rm BMO}\|\nabla f\|_2\|\nabla g\|_2, \quad \forall\, f, g \in W^{1,2}.$$

(c) Qian-Xi [QX] employed in their work the following modification of the previous proposition. We will need it as well.

Proposition 2.2. *There exists a constant C_d such that, for each function $h \in$ BMO, $i=1,\dots,d$,*

$$|\langle h, g\nabla_i g\rangle| \leq C_d\|h\|_{\rm BMO}\|\nabla g\|_2\|g\|_2, \quad \forall\, g \in W^{1,2}.$$

(d) The following result on the multipliers in space BMO on $\mathbb{R}^d$ follows from the analysis of Nakai-Yabuta in [NY] (see, in particular, Sect. 5 in their paper).

Lemma 2.1. *Set $\xi(y) := \frac{y_i}{1+\sigma|y|^2}$ $(\sigma>0)$. Then we have*

$$\|\xi h\|_{\rm BMO} \leq C_{d,\sigma}\|h\|'_{\rm BMO} \quad \forall\, h \in {\rm BMO},$$

where $\|h\|'_{\rm BMO} := \langle \mathbf{1}_{B_1(0)}h\rangle + \|h\|_{\rm BMO}$ is the BMO*-norm.*

We will apply Lemma 2.1 in the proof of Theorem 5.1, to $\xi(y) = \frac{\nabla_i\rho}{\rho} = \frac{2\sigma y_i}{1+\sigma|y|^2}$, where ρ is the weight introduced above.

3. Singular drifts

3.1. **Definitions and examples.** The following classes of singular or distributional drifts are covered by our Theorems 5.1, 6.1, 8.1.

3.1.1. *Form-bounded vector fields.*

Definition 3.1. A vector field $b \in [L^2_{\rm loc}]^d$ is said to be form-bounded with *form-bound* $\delta>0$ (abbreviated as $b\in\mathbf{F}_\delta$) if

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

for some constant $c_\delta < \infty$.

The constant c_δ does not affect the well-posedness of SDE (5.1) (one needs $c_\delta > 0$ to include L^∞ drifts). By contrast, δ is crucial: if δ exceeds the critical threshold $\delta = 4$, then in general SDE (5.1) ceases to have global in time weak solution (a blow up occurs). On the other hand, as Theorems 5.1 and 5.2 show that wnenever $\delta < 4$, weak well-posedness holds. Moreover, as $\delta \downarrow 0$, this theory becomes more detailed, e.g. one has conditional weak uniqueness, strong well-posedness, etc. The form-bound δ appears explicitly in the examples below.

Examples 3.1. 1. If $b \in [L^d]^d + [L^\infty]^d$ (= sums of vector fields in $[L^d]^d$ and in $[L^\infty]^d$), then $b \in \mathbf{F}_\delta$ with form-bound δ that can be chosen arbitrarily small at the expense of increasing the constant c_δ. Indeed, for every $\epsilon > 0$ one can represent $b = b_1 + b_2$ with $\|b_1\|_d < \varepsilon$ and $\|b_2\|_\infty < \infty$. By the Sobolev embedding theorem,

$$\begin{aligned}\|b\varphi\|_2^2 &\leq 2\|b_1\|_d^2\|\varphi\|_{\frac{2d}{d-2}}^2 + 2\|b_2\|_\infty^2\|\varphi\|_2^2 \\ &\leq C_S 2\|b_1\|_d^2\|\nabla\varphi\|_2^2 + 2\|b_2\|_\infty^2\|\varphi\|_2^2,\end{aligned}$$

so $b \in \mathbf{F}_\delta$ with $\delta = C_S 2\epsilon$ and $c_\delta = 2\|b_2\|_\infty^2$.

2. The class of form-bounded vector fields contains the weak L^d class:

$$\|b\|_{d,\infty} := \sup_{s>0} s|\{y \in \mathbb{R}^d \mid |b(y)| > s\}|^{1/d} < \infty$$

$$\Rightarrow \quad b \in \mathbf{F}_\delta \quad \text{ with } \delta = \|b\|_{d,\infty}|B_1(0)|^{-\frac{1}{d}}\frac{2}{d-2}$$

with $c_\delta = 0$. This was proved in [KPS, Lemma 2.7].

3. The weak L^d class includes itself the Hardy drift:

$$b(x) = \pm\sqrt{\delta}\frac{d-2}{2}|x|^{-2}x \in \mathbf{F}_\delta \text{ with } c_\delta = 0 \quad (\text{but } b \notin \mathbf{F}_{\delta'} \text{ with any } \delta' < \delta, c_{\delta'} < \infty). \tag{3.2}$$

In fact, inclusion (3.2) is a re-statement of the usual Hardy inequality

$$\||x|^{-1}\varphi\|_2^2 \leq \frac{4}{(d-2)^2}\|\nabla\varphi\|_2^2.$$

The plus sign in (3.2) corresponds in SDE (5.1) to the attraction towards the origin, the minus corresponds to the repulsion.

4. The previous example can be refined using the weighted Hardy inequality of Hoffmann-Ostenhof–Laptev [HL]. Fix

$$0 \leq \Phi \in L^s(S^{d-1}) \quad \text{ for some } s \geq \frac{2(d-2)^2}{2(d-1)} + 1,$$

where S^{d-1} is the unit sphere in $\mathbb{R}^d$. If

$$|b(x)|^2 \leq \delta\frac{(d-2)^2}{4}c\frac{\Phi(x/|x|)}{|x|^2}, \qquad \text{where } c := \frac{|S^{d-1}|^{\frac{1}{q}}}{\|\Phi\|_{L^s(S^{d-1})}},$$

then $b \in \mathbf{F}_\delta$ with $c_\delta = 0$. Using this example, one can e.g. cut holes in the drift (3.2) while still controlling the value of δ.

5. One can also refine example (3.2) using a Hardy-type inequality of Felli-Marchini-Terracini [FMT, Lemma 3.5]. Assume that

$$|b(x)|^2 \leq \delta\frac{(d-2)^2}{4}\sum_{i=1}^\infty \frac{\mathbf{1}_{B_r(a_i)}(x)}{|x-a_i|^2}, \quad x \in \mathbb{R}^d,$$

where the loci of singularities $\{a_i\}_{i=1}^{\infty}$ are sufficiently "spread out":

$$\sum_{i=1}^{n} |a_i|^{-d+2} < \infty, \quad \sum_{k=1}^{\infty} |a_{i+k} - a_i|^{-d+2} \text{ is bounded uniformly in } i,$$

and $|a_i - a_m| \geq 1$ for all $i \neq m$. Then there exists r sufficiently small (so, the singularities are strictly local) so that $b \in \mathbf{F}_\delta$ with $c_\delta = 0$.

6. The following simple lemma applies, in particular, to the multi-particle Hardy drift $b : \mathbb{R}^{Nd} \to \mathbb{R}^{Nd}$ defined by (1.5) in Example 1.1.

Lemma 3.1 ([K6, Lemma 1]). *If $K \in \mathbf{F}_\kappa(\mathbb{R}^d)$, then the drift $b = (b_1, \dots, b_N) : \mathbb{R}^{Nd} \to \mathbb{R}^{Nd}$ with components defined by*

$$b_i(x^1, \dots, x^N) := \frac{1}{N} \sum_{j=1, j\neq i}^{N} K(x^i - x^j), \quad x^i, x^j \in \mathbb{R}^d$$

is in $\mathbf{F}_\delta$ with

$$\delta = \frac{(N-1)^2}{N^2}\kappa,$$

i.e. there is almost equality between δ and κ. (In the context of Example 1.1 κ is the strength of attraction between the particles.)

7. Critical Morrey class. Every vector field $b \in M_{2+\varepsilon}$ for some $\varepsilon > 0$ small, i.e.

$$\|b\|_{M_{2+\varepsilon}} := \sup_{r>0, x\in\mathbb{R}^d} r\left(\frac{1}{|B_r|}\int_{B_r(x)} |b(y)|^{2+\varepsilon} dy\right)^{\frac{1}{2+\varepsilon}} < \infty,$$

is in $\mathbf{F}_\delta$ with $\delta = C(d,\varepsilon)\|b\|_{M_{2+\varepsilon}}$ and $c_\delta = 0$. The constant $C = C(d,\varepsilon)$ depends on the constants in some fundamental inequalities of harmonic analysis [Fe]. There exist far-reaching and deep extensions of this inclusion due to Adams [A1] (Appendix C) and Chiarenza-Frasca [CF].

The Morrey class $M_{2+\varepsilon}$ is substantially larger than the weak L^d class, e.g. it includes vector fields having strong hypersurface singularities. It is easily seen that the class $M_{2+\varepsilon}$ gets larger as ε gets smaller. Note, however, that by passing through the Morrey class one to a large extent loses the control over the form-bound δ.

On the other hand, the form-bounded class $\mathbf{F}_\delta$ (with $c_\delta = 0$) is contained in the Morrey class M_2; this is not difficult to see by selecting cutoff functions as test function φ in the definition of $\mathbf{F}_\delta$. Thus, to summarize,

$$\cup_{\varepsilon>0} M_{2+\varepsilon} \quad \subsetneq \quad \cup_{\delta>0}\mathbf{F}_\delta \text{ (with } c_\delta = 0) \quad \subsetneq \quad M_2.$$

8. A larger class than $\cup_{\varepsilon>0} M_{2+\varepsilon}$ sub-class of $\mathbf{F}_\delta$ was found by Chang-Wilson-Wolff [CWW], that is, $|b| \in L^2_{\rm loc}(\mathbb{R}^d)$ and

$$\sup_{r>0, x\in\mathbb{R}^d} \frac{1}{|B_r|}\int_{B_r(x)} |b(y)|^2 r^2 \xi\big(|b(y)|^2 r^2\big) dy < \infty, \tag{3.3}$$

where $\xi : \mathbb{R}_+ \to [1,\infty[$ is a fixed increasing function such that

$$\int_1^\infty \frac{ds}{s\xi(s)} < \infty.$$

For instance, one can take $\xi(s) = 1 + (\log^+ s)^{1+\epsilon}$ or $\xi(s) = 1 + \log^+ s(\log\log^+ s)^{1+\epsilon}$ for some $\epsilon > 0$ (but not $\xi(s) = 1 + \log^+ s$).

The Lebesgue, weak Lebesgue, Morrey, and Chang–Wilson–Wolff classes are all elementary sub-classes of $\mathbf{F}_\delta$. We regard estimating an integral over a ball as an "elementary" calculation. This is admittedly subjective, but in practice one can carry it out quite easily for a concrete vector field. By contrast, computing an operator norm for a given vector field (or equivalently checking a quadratic-form inequality) often requires more advanced tools, such as various forms of Hardy's inequality.

9. There are also deep necessary and sufficient conditions for ensuring that $b \in \mathbf{F}_\delta$, such as the crtierion of Kerman-Sawyer [KSa] or the criterion of Mazya [M2]. The latter is

$$\langle \mathbf{1}_E |b|^2\rangle \leq K \mathrm{cap}(E) \quad \forall \, \text{compact } E \subset \mathbb{E}^d, \tag{3.4}$$

where, recall,

$$\mathrm{cap}(E) = \inf\{\|\nabla v\|_2 \mid v \in C_c^\infty, v \geq 1 \text{ on } E\}.$$

So, criterion (3.4) requires computing the capacity of an arbitrary compact set (one cannot restrict to dyadic cubes only). Kerman-Sawyer's inequality that needs to be verified is more complex than (3.4), but the calculations are confined to dyadic cubes, so it is in some sense more practical. That said, precisely because these are necessary and sufficient conditions, their verifications can be non-trivial (for example, verifying that the Chang-Wilson-Wolff class (3.3) satisfies these conditions, see [CWW]).

Efforts to characterize inequalities of type (3.1), known as trace inequalities, remains an active research area. The interest was originally motivated by the problems related to estimating the spectrum of Schrödinger operators, cf. [Fe, MV].

Finally, we note that we can combine the previous examples:

$$b_1 \in \mathbf{F}_{\delta_1}, \quad b_2 \in \mathbf{F}_{\delta_2} \quad \Rightarrow \quad b_1 + b_2 \in \mathbf{F}_\delta, \quad \delta = (\sqrt{\delta_1} + \sqrt{\delta_2})^2.$$

This, of course, extends to series of form-bounded drifts.

3.1.2. *Bounded mean oscillation*$^{-1}$.

Definition 3.2. A divergence-free vector field $q = (q_i)_{i=1}^d \in [\mathcal{S}']^d$ is said to be in class $\mathbf{BMO}^{-1}$ if there exist functions $Q^{ij} = -Q^{ji} \in \mathrm{BMO}(\mathbb{R}^d)$ such that

$$q^i = \sum_{j=1}^d \nabla_j Q^{ij}, \quad 1 \leq i \leq d,$$

i.e. $q = \nabla Q$, where ∇ is the row-divergence operator.

Examples of BMO functions include

$$g \in L^\infty \quad \text{ or } \quad g(x) = \log|p(x)| \text{ for a polynomial } p.$$

In the last example the BMO semi-norm does not depend on the coefficients of p.

Examples 3.2. 1. A divergence-free vector field q belongs to $\mathbf{BMO}^{-1}$ if and only if the caloric extensions of its components q^i, $i = 1, \dots, d$, satisfy

$$\sup_{x \in \mathbb{R}^d, R>0} \frac{1}{|B_R|} \int_{[0,R^2]\times B_R(x)} |e^{t\Delta} q^i|^2 dt dy < \infty,$$

where $e^{t\Delta}$ is the heat semigroup, see [KT].

Equivalently, q^i can be characterized as elements of homogeneous Triebel-Lizorkin space $\dot{F}^{-1}_{2,\infty}$.

2. The divergence-free vector fields with entries in Morrey space M_1, i.e. such that

$$\langle |q| \mathbf{1}_{B_r(x)} \rangle \leq C r^{d-1}$$

with constant C independent of r or $x \in \mathbb{R}^d$, are in $\mathbf{BMO}^{-1}$ [M, Sect. 3.4.5].

3. The divergence-free vector fields with entries in Besov space $B^{-1+d/p}_{p,\infty}$, $p > d$, are in $\mathbf{BMO}^{-1}$. This follows by recalling that $B^{-1+d/p}_{p,\infty}$ consists of tempered distribution h such that

$$\|e^{t\Delta} h\|_p \leq C t^{-\frac{1-\frac{d}{p}}{2}}, \quad 0 < t \leq 1,$$

see [KT] for details.

In [KT], Koch and Tataru established, among other results, the existence and the uniqueness of global in time mild solution to Cauchy problem for the 3D Navier-Stokes equations in the critical space Z of functions $v : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^3$ satisfying

$$\|v\|_Z := \sup_{t \geq 0} t^{\frac{1}{2}} \|v(t)\|_\infty + \bigg(\sup_{x \in \mathbb{R}^d, R>0} \frac{1}{|B_R|} \int_{[0,R^2] \times B_R(x)} |v(t,y)|^2 dt dy \bigg)^{\frac{1}{2}} < \infty, \tag{3.5}$$

provided that the initial data $v(0) \in \mathbf{BMO}^{-1}(\mathbb{R}^3)$ have sufficiently small norm. (An even larger $\mathbf{BMO}^{-1}$ type space is considered in the recent paper [CE].)

3.1.3. *Multiplicatively form-bounded vector fields.* This class of drifts was mentioned in the introduction (class $\mathbf{MF}_\delta$). We postpone its discussion until Section 8.

3.2. **Physical approximations.** We now introduce the classes of bounded smooth approximations of vector fields in $\mathbf{F}_\delta$ and $\mathbf{BMO}^{-1}$ that preserve the "structure constants" of the latter.

Definition 3.3. Given $b \in \mathbf{F}_\delta$, we denote by $[b]$ the set of sequences of vector fields $\{b_n\} \subset [C_b \cap C^\infty]^d$ such that

$$\left\{ \begin{array}{l} b_n \in \mathbf{F}_\delta \text{ with the same } c_\delta \text{ as } b, \\ b_n \to b \text{ in } [L^2_{\rm loc}]^d \end{array} \right. \tag{3.6}$$

(note that we can always increase c_δ, if needed; what matters is that c_δ does not depend on n).

For instance, the sequence $\{b_n\}$ defined by $b_n := E_{\varepsilon_n} b$ (De Giorgi mollifier E_ε is defined in Section 2) for any $\varepsilon_n \downarrow 0$ is in $[b]$. See [KS4] or [K6, Sect. 6] for the proof.

In the quantum-mechanical context, the form-boundedness condition on potential $|b|^2$ expresses smallness of the potential energy with respect to the kinetic energy in the system described by the Hamiltonian $-\Delta - |b|^2$. Our conditions on regularizations $\{b_n\}$ is thus that they, essentially, do not increase the potential energy.

Given a potential $0 \leq V \in L^1_{\rm loc}$, we write, with some abuse of notation, $V^{\frac{1}{2}} \in \mathbf{F}_{\delta_+}$ if

$$\langle V, \varphi^2 \rangle \leq \delta_+ \langle |\nabla \varphi|^2 \rangle + c_{\delta_+} \langle \varphi^2 \rangle \quad \forall\, \varphi \in C_c^\infty.$$

Definition 3.4. Let

$$b \in \mathbf{F}_\delta, \quad ({\rm div}\, b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+}, \quad ({\rm div}\, b)_- \in L^1 + L^\infty,$$

we denote by $[b]'$ the set of sequences of vector fields $\{b_n\} \subset [C^1_b \cap C^\infty]^d$ that satisfy the previous inclusions with the same constants δ, c_δ, δ_+, c_{δ_+} (so, independent of n) and such that

$$b_n \to b \text{ in } [L^2_{\rm loc}]^d, \quad {\rm div}\, b_n \to {\rm div}\, b \text{ in } L^1_{\rm loc}.$$

Once again, we can take e.g. $b_n := E_{\varepsilon_n} b$, see [K6, Sect. 6] for the proof.

We denote by H_ξ ($\xi > 0$) the set of bounded symmetric uniformly elliptic Borel measurable matrix fields $a : \mathbb{R}^d \to \mathbb{R}^{d\times d}$:

$$a = a^\top, \quad \xi I \leq a(x) \leq \xi^{-1} I \text{ for a.e. } x \in \mathbb{R}^d,$$

for I the $d \times d$ identity matrix.

Definition 3.5. Given a matrix field $a \in H_\xi$ and a vector field b such that $\nabla a + b \in \mathbf{F}_\delta$, we denote by $[a, b]$ the set of sequences $\{a_n\} \in H_\xi \cap [C_b \cap C^\infty]^{d\times d}$, $\{b_n\} \in [C_b \cap C^\infty]^d$ such that

$$\nabla a_n + b_n \in \mathbf{F}_\delta \text{ with } c_\delta \text{ independent of } n,$$

and

$$\nabla a_n + b_n \to \nabla a + b \text{ in } [L^2_{\rm loc}]^d, \quad a_n \to a \quad \text{ a.e. on } \mathbb{R}^d$$

as $n \to \infty$.

For example, we can take $a_n = E_{\varepsilon_n} a$, see [KS3, Sect. 4.4] for the proof.

Definition 3.6. Given $q = \nabla Q \in \mathbf{BMO}^{-1}$, where Q is the corresponding anti-symmetric matrix field with entries in BMO (so, $\operatorname{div} q = 0$), we denote by $[q]$ the set of sequences

$$\{q_m = \nabla Q_m \text{ for anti-symmetric } Q_m \in [C^\infty \cap W^{1,\infty}]^{d\times d}\}$$

such that

$$\begin{cases} \|Q_m\|_{\rm BMO} \leq C\|Q\|_{\rm BMO} \text{ for a constant } C \text{ independent of } m, \\ Q_m \to Q \text{ in } [L^s_{\rm loc}]^{d\times d} \text{ for any } 1 \leq s < \infty. \end{cases}$$

For example, one possible choice is

$$Q_m := E_{\varepsilon_m}(Q \wedge U_{\varepsilon_m} \vee V_{\varepsilon_m}),$$

where the maximum and the minimum are taken componentwise,

$$U_\varepsilon := (-c\log|x| + \varepsilon^{-1}) \wedge \varepsilon^{-1} \vee 0, \quad V_\varepsilon := (c\log|x| - \varepsilon^{-1}) \wedge 0 \vee (-\varepsilon^{-1}), \quad \varepsilon_m \downarrow 0$$

with c chosen so that $\|c\log|x|\|_{\rm BMO} \leq \|Q\|_{\rm BMO}$. The last two functions are compactly supported and are in BMO. Since BMO is a lattice, the components of the matrix fields Q_m (still anti-symmetric) are in BMO. This regularization of Q was employed earlier in [QX].

4. Preliminary discussion: blow up thresholds for Brownian particles

We continue discussing the particle system in Example 1.1, i.e.

$$X^i_t = x^i_0 - \sqrt{\kappa}\frac{d-2}{2}\frac{1}{N}\sum_{j=1, j\neq i}^N \int_0^t \frac{X^i_s - X^j_s}{|X^i_s - X^j_s|^2}ds + \int_0^t q_0(X^i_s)ds + \sqrt{2}B^i_t, \quad i = 1, \dots, N, \tag{4.1}$$

where $q_0 \in \mathbf{BMO}^{-1}(\mathbb{R}^d)$, $d \geq 3$.

1. First, we present positive well-posedness results for (4.1) that follow from Theorems 5.1 and 8.1.

2. Next, we exhibit counterexamples to well-posedness of particle system (4.1) when the attraction strength κ is too large. On one hand, these counterexamples test the sharpness of Theorems 5.1 and 8.1; on the other, they play a crucial role in the proof of Theorem 8.2, which provides an improved upper bound on the constant in the many-particle Hardy inequality.

3. Finally, we discuss the two- and the one-dimensional cases.

4.1. **Positive results.** Theorem 8.1 covers a large portion of the admissible range for the attraction parameter κ in (4.1):

$$\kappa \in \left[0, 16\left(1 \vee \frac{N}{1+\sqrt{1+\frac{3(d-2)^2}{(d-1)^2}(N-1)(N-2)}}\right)^2\right[. \tag{4.2}$$

Here, the right endpoint arises from Hoffmann-Ostenhof, Hoffmann-Ostenhof, Laptev and Tidblom's lower bound on the constant in the many-particle Hardy inequality [HHLT]. Since the optimal constant in that inequality is not yet known, one expects the true admissible interval for κ to be strictly larger. Note that in dimensions $d \geq 7$ the interval (4.2) reduces to $\kappa \in [0, 16[$.

By contrast, Theorem 5.1 handles only $\kappa \in [0, 4\frac{N^2}{(N-1)^2}[$, but it offers greater flexibility in modifying the interaction kernel in (4.1). For example, one can multiply the attracting interaction kernel by any function of L^∞ norm at most one, without altering the assumption on κ. In particular, one can "cut holes" in the interaction kernel so that the particles do not interact along certain directions (cf. Examples 3.1.4).

One can also describe how the particles behave as they approach collision. Let $p(t,x,y)$ be the transition density (heat kernel) of (4.1). For simplicity, set $q_0 = 0$. Then one obtains two-sided heat kernel bounds in $\mathbb{R}^{Nd}$

$$c_1 t^{-\frac{Nd}{2}} e^{-\frac{|x-y|^2}{c_2 t}} \varphi_t(y) \leq p(t,x,y) \leq c_3 t^{-\frac{Nd}{2}} e^{-\frac{|x-y|^2}{c_4 t}} \varphi_t(y), \tag{4.3}$$

where $\varphi_t \geq 1$ is the singular weight

$$\varphi_t(y) := \prod_{1 \leq i < j \leq N} \eta(t^{-\frac{1}{2}}|y^i - y^j|)$$

for a fixed function $1 \leq \eta \in C^2(]0,\infty[)$ such that

$$\eta(r) = \begin{cases} r^{-\sqrt{\kappa}\frac{d-2}{2}\frac{1}{N}} & 0 < r < 1, \\ 2, & r > 2. \end{cases}$$

The details will appear in [BK]. A weaker version of the upper bound in (4.3) omitting the exponential factor was proved in [K6].

4.2. **Counterexamples.** These were the positive results for (4.1) that we now balance with counterexamples, i.e. analogues of (a), (b) in Section 1.1.1. For simplicity, assume there is no divergence-free distributional component of the drift (i.e. $q_0 = 0$). The right endpoint of the interval (4.2) lies just below the first blow up threshold for (4.1), or the "non-sticky collisions threshold". More precisely, following [F], set

$$R_t := \frac{1}{4N}\sum_{i,j=1}^N |X_t^i - X_t^j|^2.$$

It is not difficult to see that R_t is a local squared Bessel process, i.e.

$$R_t = R_0 + 2\int_0^t \sqrt{|R_t|}dW_t + \mu t, \tag{4.4}$$

where W_t is a one-dimensional Brownian motion, and

$$\mu = (N-1)\left(d - \sqrt{\kappa}\frac{d-2}{4}\right)$$

is its “dimension”; the value of μ controls how often R_t hits zero. In what follows, by a collision of particles in (4.1) we mean a collision of all N particles at the same time, i.e. when $R_t = 0$. So, by standard theory of Bessel processes (see [RY, Ch. XI, §1]):

(a') If $\kappa \geq 16(\frac{d}{d-2})^2$, then there are a.s. sticky collisions in (4.1), i.e. the particles collide in finite time and stay clumped up. Furthermore, one can show that for $\kappa > 16(\frac{d}{d-2})^2$ the particle system ceases to have a weak solution, cf. (a) in Section 1.1.1.

(b') If $16\big(\frac{d}{d-2}\big)^2\big(1-\frac{2}{d(N-1)}\big) < \kappa < 16(\frac{d}{d-2})^2$, then there are a.s. non-sticky collisions in (4.1), i.e. particles collide infinitely many times, but $\int_0^\infty R_t dt < \infty$ a.s.

Comparing (a'), (b') with the admissible range (4.2) in Theorem 8.1, we see that the latter provides a result that is close to optimal.

4.3. **Two- and one-dimensional cases.** For the two-dimensional counterpart of (4.1), namely, the finite-particle approximation of the Keller-Segel model,

$$X_t^i = x^i - \sqrt{\varkappa}\int_0^t \frac{X_s^i - X_s^j}{|X_s^i - X_s^j|^2} ds + \sqrt{2}B_t^i, \quad x = (x^1, \dots, x^N) \in \mathbb{R}^{2N},$$

the attracting interaction kernel is no longer locally square intergrable and therefore is not form-bounded, so one cannot apply our Theorems 5.1, 8.1 (see, however, Section 7 where we address $d = 2$, but only to some extent). Cattiaux-Pédèches [CP] and Fournier-Jourdain [FJ], Fournier-Tardy [FT] exploited the special structure of the drift and constructed process $(X_t^1, \dots, X_t^N)$. They work either via a suitable Dirichlet form or by exhibiting a weak solution to the SDE, respectively. In particular, they cover the full critical range $\varkappa \in [0, 16[$, where 16 is the sticky collisons threshold in dimension two. In this context, note:

1) The arguments of [FJ, FT] work in dimensions $d \geq 3$ as well, and allow to handle $16 \leq \kappa < 16(\frac{d}{d-2})^2$ in (4.1), which includes non-sticky collisions, see (b') above.

2) Ohashi-Russo-Texeira [ORT] consider squared Bessel processes in the low-dimensional regime $0 < \nu < 1$ (i.e. non-sticky collisions). They characterize the process as the unique solution of SDE

$$dX_t = \frac{1-\nu}{2}\frac{dt}{X_t} + dB_t, \quad X_0 = x_0 > 0, \quad 0 < \nu < 1,$$

whose drift $b(x) = \frac{1}{x}$ is not locally in $L^1_{\rm loc}$, and thus has to be considered as a distributional drift.

It is not yet clear whether Theorems 5.1, 8.1 can be extended in some form to the non-sticky collisions part of the interval of admissible values of κ. However, taking into account the positive results in 1) and in 2), such extension is conceivable.

We refer to Cattiaux [C] and Fournier [F] for recent surveys on the Keller-Segel model and its finite particle approximations.

5. General distributional drifts

First, we consider SDE

$$X_t = x - \int_0^t \big(b(X_s) + q(X_s)\big) ds + \sqrt{2}B_t, \quad x \in \mathbb{R}^d, \quad t \geq 0. \tag{5.1}$$

with

$$b \in \mathbf{F}_\delta, \quad q \in \mathbf{BMO}^{-1} \ \ (\Rightarrow \operatorname{div} q = 0),$$

with the form-bound δ of b going all the way up to (but staying strictly less) the critical threshold $\delta = 4$. As mentioned in the introduction, by a result of Mazya and Verbitsky, this assumption on b and q is equivalent to having generalized form-boundedness (1.6) of $c = b + q$.

We fix bounded smooth approximations $\{b_n\} \in [b]$, $\{q_m\} \in [q]$ (as in Section 2). Consider the approximating SDEs

$$X_t^{n,m} = x - \int_0^t \big(b_n(X_s^{n,m}) + q_m(X_s^{n,m})\big)ds + \sqrt{2}B_t, \tag{5.2}$$

on a complete probability space $\mathfrak{F} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$, with B_t being a $\mathcal{F}_t$-Brownian motion. By classical theory, for every $x \in \mathbb{R}^d$ and every $n = 1, 2, \dots,$ there exists a pathwise unique strong solution $\{X_t^{n,m}\}_{t\geq 0}$ to (5.2). The corresponding Kolmogorov operators

$$\Lambda(b_n, q_m) := -\Delta + (b_n + q_m)\cdot\nabla, \quad D\big(\Lambda(b_n, q_m)\big) = (1-\Delta)^{-1}C_\infty.$$

generate strongly continuous Feller semigroups on C_∞ such that

$$e^{-t\Lambda(b_n,q_m)}f(x) = \mathbf{E}[f(X_t^{n,m})].$$

Set $\mathbb{P}_x^{n,m} := \mathbf{P}(X_t^{n,m})^{-1}$.

Theorem 5.1. *Let $d \geq 3$. Assume that b and q are, respectively, Borel measurable and distribution-valued vector fields $\mathbb{R}^d \to \mathbb{R}^d$ that satisfy*

$$\begin{cases} b \in \mathbf{F}_\delta \ \text{with } \delta < 4, \\ q \in \mathbf{BMO}^{-1}. \end{cases} \tag{5.3}$$

Let $\{b_n\} \in [b]$, $\{q_m\} \in [q]$ as in Definitions 3.3 and 3.6. The following statements are true:

(i) (Feller semigroup) *The limit*

$$s\text{-}C_\infty\text{-}\lim_n \lim_m e^{-t\Lambda(b_n,q_m)} \ \text{(loc. uniformly in } t \geq 0)$$

exists and determines a strongly continuous Feller semigroup, say, $e^{-t\Lambda} = e^{-t\Lambda(b,q)}$. (The order in which we take the limits is important.) The generator Λ of $e^{-t\Lambda}$ is an operator realization of the formal differential expression $-\Delta + (b+q)\cdot\nabla$ in C_∞.

(ii) (Approximation uniqueness) *The limit in (i) does not depend on the choice of $\{b_n\} \in [b]$ and $\{q_m\} \in [q]$.*

(iii) (Relaxed approximation uniqueness) *Let $\delta < 1$. If $\{b_n\} \in [b] \cap [L^2]^d$ does not necessarily converge strongly in $[L^2]^d$, but only weakly:*

$$b_n \xrightarrow{w} b \quad \text{in } [L^2]^d,$$

then we still have convergence of the approximating Feller semigroups to the same limit from (i):

$$e^{-t\Lambda(b_n,q)} \xrightarrow{s} e^{-t\Lambda(b,q)} \ \text{in } C_\infty \ \text{(loc. uniformly in } t \geq 0).$$

(Thus, when $\delta < 1$, we can extend quite substantially the class of admissible approximations of b.)

(iv) (Generalized martingale solution) *There exists a strong Markov family of probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ on the canonical space $\mathbf{C}$ of continuous trajectories such that*

$$e^{-t\Lambda(b,q)}f(x) = \mathbb{E}_{\mathbb{P}_x}[f(\omega_t)], \quad f \in C_\infty, \quad x \in \mathbb{R}^d,\ t \geq 0,$$

$$\mathbb{P}_x = w\text{-}\mathcal{P}(\mathbf{C})\text{-}\lim_n \lim_m \mathbb{P}_x^{n,m},$$

and for every test function v in the domain $D\big(\Lambda(b,q)\big)$, a dense subspace of C_∞, the process

$$t \mapsto v(\omega_t) - x + \int_0^t \Lambda(b,q)v(\omega_s)ds$$

is a continuous martingale under $\mathbb{P}_x$.

(Selecting test functions v from the domain of the Feller generator allows to address the problem of defining the term $\int_0^t q(X_s)\cdot\nabla v(X_s)ds$ in the martingale problem.)

(v) (Weak solution for the SDE with disperse initial data) *Let $\delta<1$ and let us also assume that b, q, b_n, q_m have supports in a ball of fixed radius. Given an initial (smooth) probability density ν_0 satisfying $\langle \nu_0^{2r}\rangle < \infty$ for some $1<r<\frac{1}{\sqrt{\delta}}$, there exist a probability space $\mathfrak{F}'=(\Omega',\mathcal{F}',\{\mathcal{F}'_t\}_{t\geq 0},\mathbf{P}')$ and a continuous process X_t on this space such that the limit*

$$A_t := L^2(\Omega')\text{-}\lim_n \lim_m \int_0^t \big(b_n(X_s)+q_m(X_s)\big)ds$$

exists, and we have a.s.

$$X_t = X_0 - A_t + \sqrt{2}B_t, \quad t>0,$$

for a $\mathcal{F}'_t$-Brownian motion B_t, for $\mathbf{P}'X_0^{-1}$ having density ν_0. (The compact support assumption can be removed with a few additional efforts at expense of requiring $\langle \nu_0^{2r}\rho^{-\alpha}\rangle < \infty$, where the weight ρ is defined by (2.1), for appropriate $\alpha>0$, see Remark 10.4.)

(vi) (Dispersion estimates and uniqueness of weak solution to Kolmogorov PDE) *We can descend from C_∞ to L^p and show that for every $p>\frac{2}{2-\sqrt{\delta}}$ the operators*

$$e^{-t\Lambda_p(b,q)} := \Big[e^{-t\Lambda(b,q)} \upharpoonright C^\infty \cap L^p\Big]^{\mathrm{clos}}_{L^p\to L^p}$$

are bounded on L^p and constitute a strongly continuous semigroup. Moreover, for all $\frac{2}{2-\sqrt{\delta}}<p\leq r<\infty$,

$$\|e^{-t\Lambda_p(b,q)}f\|_r \leq C_{\delta,d}e^{\omega_p t}t^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{r})}\|f\|_p, \quad f\in L^p, \quad \omega_p=\frac{c_\delta}{2(p-1)}.$$

The latter and the Dunford-Pettis theorem yields that $e^{-t\Lambda(b,q)}$, $t\geq 0$, are integral operators.

If $\delta<1$, then $v(t):=e^{-t\Lambda_2(b,q)}f$, $f\in L^2$, is the unique weak solution to Cauchy problem

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

in the standard Hilbert triple $W^{1,2}\hookrightarrow L^2\hookrightarrow W^{-1,2}$.

(vii) (Strong Feller property of the resolvents) *For every $f\in L^{p\theta}\cap L^{p\theta'}$, for every $x\in\mathbb{R}^d$, $u:=(\mu+\Lambda(b,q))^{-1}f$ satisfies*

$$\sup_{B_{\frac{1}{2}}(x)}|u| \leq K\bigg((\mu-\mu_0)^{-\frac{1}{p\theta}}\langle |f|^{p\theta}\rho_x\rangle^{\frac{1}{p\theta}} + (\mu-\mu_0)^{-\frac{\beta}{p}}\langle |f|^{p\theta'}\rho_x\rangle^{\frac{1}{p\theta'}}\bigg), \quad \mu>\mu_0>0,$$

for fixed $1<\theta<\frac{d}{d-2}$ and $p\geq 2$ such that $p>\frac{2}{2-\sqrt{\delta}}$. The constants K and μ_0 do not depend on f or x.

Additionally, Krylov-type bound (5.15) holds, see Remark 5.3.

In Theorem 5.1, we do not establish conditional weak uniqueness or strong solvability as in the next theorem. Nevertheless, Theorem 5.1 still covers a common scenario in physical models: one mollifies a singular drift and studies the corresponding SDE as an approximation of the true dynamics (for instance, when investigating long-term behavior such as sub- or super-diffusivity, see [ABK]). Although the mollified drift is smooth, one must ensure that:

– the resulting stochastic dynamics does not depend on the choice of the regularization of the drift (in particular, on the choice of the mollifier);

– in the limit one still gets a non-pathological diffusion process.

Assertions (*ii*), (*iii*) of Theorem 5.1 address the first point, and assertions (*i*), (*iv*)-(*vii*) the second.

In the next theorem $q = 0$, i.e. there is no distributional component of the drift. This theorem recollects results obtained previously by the first author and co-authors, and is included in this paper for the sake of completeness:

Theorem 5.2 (Case $q = 0$)**.** *Let $d \geq 3$. Assume that $b \in \mathbf{F}_\delta$ with $\delta < 4$. Let $\{b_n\} \in [b]$. Then, in addition to the assertions of Theorem 5.1, the following are true:*

(i) (Classical weak solution [KS2]) *For every $x \in \mathbb{R}^d$,*

$$\mathbb{E}_{\mathbb{P}_x} \int_0^1 |b(\omega_s)| ds < \infty$$

and, for every test function $v \in C_c^2$, the process

$$t \mapsto v(\omega_t) - x + \int_0^t (-\Delta + b \cdot \nabla) v(\omega_s) ds$$

is a continuous martingale with respect to $\mathbb{P}_x$. Thus, in the terminology used e.g. in [RY], *C_c^2 belongs to the domain of the extended generator of the Feller semigroup $e^{-t\Lambda(b)}$. Moreover,*

$$B_t(\omega) := \frac{1}{\sqrt{2}} \biggl(\omega_t - x + \int_0^t b(\omega_s) ds \biggr), \quad t \geq 0,$$

is a Brownian motion, so we have a weak solution to SDE (5.1).

If, additionally, $\delta < \frac{C}{d^2}$, then even more can be said:

(ii) (Strong solvability [KM1]) *Assume that b and $\{b_n\} \in [b]$ have supports in a fixed ball and let $c_\delta = 0$. Then the strong solutions X_t^n of the approximating SDE*

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

considered on a fixed probability space $\mathfrak{F} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbf{P})$, converge $\mathbf{P}$-a.s. to a strong solution X_t to

$$X_t = x - \int_0^t b(X_s) ds + \sqrt{2} B_t. \tag{5.4}$$

(The compact support assumption and the assumption $c_\delta = 0$ are of technical character and can be removed [KM4]*.) The proof is based on the Röckner-Zhao method* [RZ2].

(iii) (Krylov bound [K5]) *There exists constant $C = C(T,d,\delta,c_{\delta_1},\varepsilon)$ $(\varepsilon > 0)$ such that*

$$\mathbf{E}\bigg[\int_0^T |h(s,X_s)|ds\bigg] \leq C\|h\|_{L^{\frac{d}{2}+\varepsilon}([0,T]\times\mathbb{R}^d)} \tag{5.5}$$

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

(iv) (Krylov-type bound [KM2]) *Let $r \in]d, \delta^{-\frac{1}{2}}[$. Let $V \in \mathbf{F}_{\delta_1}$ be a form-bounded potential, that is, $V \in L^2_{\rm loc}$ and*

$$\langle V^2, \varphi^2\rangle \leq \delta_1 \langle |\nabla \varphi|^2\rangle + c_{\delta_1}\langle \varphi^2\rangle, \quad \varphi \in C_c^\infty, \tag{5.6}$$

for some $\delta_1 < \infty$ and $c_{\delta_1} < \infty$. Then there exists constant $C = C(T,d,r,\delta,c_\delta,\delta_1,c_{\delta_1})$ such that

$$\mathbf{E}\int_0^T |V(X_s)h(s,X_s)|ds \leq C\|V|h|^{\frac{r}{2}}\|^{\frac{2}{r}}_{L^2([0,T]\times\mathbb{R}^d)} \tag{5.7}$$

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

(v) (Conditional strong uniqueness [KM1]) *The strong solution X_t to (5.4) constructed in (ii) is unique among strong solutions that satisfy Krylov-type bound (5.7) for some $r \in]d, \delta^{-\frac{1}{2}}[$ both for $V = 1$ and $V = |b|$.*

If a form-bounded drift b additionally satisfies

$$|b| \in L^{\frac{d}{2}+\varepsilon} \quad \text{for some } \varepsilon > 0,$$

then X_t is unique among strong solutions to (5.4) that satisfy Krylov bound (5.5).

(vi) (Krylov bounds and conditional weak uniqueness [KM2, K5]) *Krylov bounds (5.5), (5.7) and the conditional uniqueness results in (v) hold for the weak solutions in (i).*

(vii) (Another kind of approximation uniqueness [KS7]) *Let $\delta < \frac{4}{(d-2)^2} \wedge 1$. Let $\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d}$ be a family of solutions to the martingale problem in (vii) that are constructed via approximation, i.e. are such that*

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

where $\tilde{b}_n \in \mathbf{F}_\delta \cap [C_b \cap C^\infty]^d$ with c_δ independent of n. Then

$$\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d} = \{\mathbb{P}_x\}_{x\in\mathbb{R}^d},$$

where $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ were constructed in (i). Here we do not require any convergence of $\tilde{b}_n$ to b.

(viii) (Gradient bounds [KoS, KS7, K7, KS8]) *Let $\delta < \frac{4}{(d-2)^2} \wedge 1$. Then the unique weak solution u to the elliptic equation*

$$(\mu - \Delta + b\cdot\nabla)u = f$$

satisfies, for every $r \in [2, \frac{2}{\sqrt{\delta}}[$,

$$\|\nabla u\|_r \leq K_1(\mu-\mu_0)^{-\frac{1}{2}}\|f\|_r, \quad \|\nabla|\nabla u|^{\frac{r}{2}}\|_2 \leq K_2(\mu-\mu_0)^{-\frac{1}{2}+\frac{1}{r}}\|f\|_r, \tag{5.8}$$

$$\|(\mu-\Delta)^{\frac{1}{2}+\frac{1}{s}}u\|_r \leq K\|(\mu-\Delta)^{-\frac{1}{2}+\frac{1}{\ell}}f\|_r, \quad \text{for all } 2 \leq \ell < r < s \tag{5.9}$$

for all μ greater than some generic constant μ_0.

The corresponding parabolic gradient bounds impose more restrictive conditions on δ. Namely, assume that form-bound δ satisfies, for some $r = d+\varepsilon$ (with this choice of r the Sobolev embedding theorem will give Hölder continuity of solution)

$$\sqrt{\delta} < \begin{cases} \left(\sqrt{r-1} - \frac{r-2}{2}\right)\frac{2}{r} & \text{in dimensions } d = 3,4, \\ (1-\mu)\frac{r-1}{r-2}\frac{1}{r} & \text{in dimensions } d \geq 5, \end{cases}$$

where $0 < \mu < 1$, $16\mu > (1-\mu)^4\frac{(r-1)^2}{(r-2)^4}$. (For instance, these assumptions on δ are satisfied if $\delta < \frac{1}{d^2}$, $d \geq 3$.) Then the unique weak solution v to Cauchy problem

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

satisfies

$$\sup_{0\leq s\leq t}\|\nabla v(s)\|_r^r + C_1\int_0^t \||\nabla v|^{\frac{r-2}{2}}\partial_s v\|_2^2 ds + C_2\int_0^t \langle|\nabla|\nabla v|^{\frac{r}{2}}|^2\rangle ds \leq e^{C_3 t}\|\nabla f\|_r^r \tag{5.10}$$

for constants $C_i > 0$ $(i=1,2,3)$ that depend only on d, δ and c_δ.

(ix) (Stochastic transport equation [KSS]) *If $b \in \mathbf{F}_\delta$ with $\delta < (1+4rd)^{-2}$, $r = 1,2,\dots$ Then, for every $v_0 \in W^{1,4r}$, there exists a unique weak solution to Cauchy problem for the stochastic transport equation*

$$dv + b\cdot\nabla v dt + \sqrt{2}\nabla v \circ dB_t = 0, \quad v|_{t=0} = f, \tag{5.11}$$

with $\circ$ denoting the Stratonovich multiplication. It satisfies

$$\sup_{0\leq\alpha\leq 1}\left\|\mathbb{E}|\nabla v|^{2r}\right\|_{L^{\frac{2}{1-\alpha}}([0,t],L^{\frac{2d}{d-2+2\alpha}})} \leq C_1 e^{C_2 t}\|\nabla f\|_{4r}^{2r}.$$

In particular, if $2r > d$, then by the Sobolev embedding theorem for a.e. $\omega \in \Omega$ the function $x \mapsto v(t,x,\omega)$ is Hölder continuous, possibly after modification on a set of measure zero in $\mathbb{R}^d$ (in general, depending on ω).

Example 5.1. Let us return to the problem of describing the dynamics $X_t = (X_t^1,\dots,X_t^N)$ of N interacting particles immersed in a velocity field in $\mathbf{BMO}^{-1}$, $d \geq 3$. That is, we are in the setting of Example 1.1, where, recall,

$$X_t = x_0 - \int_0^t \big(b(X_s) + q(X_s)\big)ds + \sqrt{2}B_t, \quad x_0 = (x_0^1,\dots,x_0^N) \in \mathbb{R}^{Nd}, \quad B_t = (B_t^1,\dots,B_t^N), \tag{5.12}$$

and $b(x) = (b^1(x),\dots,b^N(x))$, $q(x) = (q_0(x^1),\dots,q_0(x^N))$ $(x = (x^1,\dots,x^N) \in \mathbb{R}^{Nd})$,

$$b^i(x^1,\dots,x^N) := \frac{1}{N}\sum_{j=1,j\neq i}^N \sqrt{\kappa}\frac{d-2}{2}\frac{x^i - x^j}{|x^i - x^j|^2}, \quad q_0 \in \mathbf{BMO}^{-1}(\mathbb{R}^d).$$

Then, by Lemma 3.1, $b \in \mathbf{F}_\delta(\mathbb{R}^{Nd})$ with $\delta = \frac{(N-1)^2}{N^2}\kappa$. Also, as mentioned in the introduction, $q \in \mathbf{BMO}^{-1}(\mathbb{R}^{Nd})$. Therefore, if the strength of attraction between the particles κ satisfies

$$\kappa < 4\frac{N^2}{(N-1)^2} \tag{$\kappa_{\rm hyp}$}$$

(so that $\delta < 4$), then Theorem 5.1 applies and ensures the existence and the approximation uniqueness for this particle system. Importantly, the assumption on κ basically does not depend on the number of particles N (assumed to be large).

In the case $d=2$, which is of interest e.g. in the Keller-Segel model, the drift $b:\mathbb{R}^{2N}\to\mathbb{R}^{2N}$ defined above is not form-bounded. However, it is weakly form-bounded, which still allows us to say something about the corresponding SDE (5.12), see Section 7.

Remark 5.1 (Critical threshold $\delta=4$)**.** Let $b\in\mathbf{F}_4$. This case is not covered by Theorem 5.1. Assume additionally that b and the approximating vector fields $\{b_n\}\in[b]$ have supports in a fixed ball (or, more generally, decay sufficiently rapidly at infinity uniformly in n). Let $q\in\mathbf{BMO}^{-1}$. By following closely the proof of [K3, Theorem 1] (see also [KS6, Theorem 2]) i.e. using test function e^v-e^{-v} in the analysis of Cauchy problem for the Kolmogorov PDE $(\partial_t-\Delta+b\cdot\nabla)v=0$, one can show that the limit

$$s\text{-}L_{\cosh-1}\text{-}\lim_n\lim_m e^{-t\Lambda(b_n,q_m)}\ \text{(loc. uniformly in } t\geq 0)$$

exists and determines a strongly continuous Markov semigroup $e^{-t\Lambda(b)}$ on the Orlicz space

$$L_{\cosh-1}:=\text{the closure of the Schwartz space } \mathcal{S} \text{ with respect to norm}$$

$$\|f\|_{\cosh-1}=\inf\left\{c>0\mid\langle\cosh\frac{f}{c}-1\rangle\leq 1\right\}.$$

On the torus, in the case $q=0$ (i.e. no distributional component of the drift), [K3] established the following energy inequality for $v(t)=e^{-t\Lambda(b_n)}v_0$:

$$\frac{1}{2}\sup_{s\in[0,t]}\langle e^{v^p(s)}\rangle+4\frac{(p-1)}{p}\int_0^t\langle(\nabla v^{\frac{p}{2}})^2e^{v^p}\rangle ds\leq\langle e^{v_0^p}\rangle,\quad p=2,4,\dots,$$

provided $\frac{c_\delta}{\sqrt{\delta}}t<\frac{1}{2}$. The small time restriction can be removed using the semigroup property. One can compare this to the usual L^p energy inequality when $\delta<4$ (Remark 9.2). At first sight, letting $\delta\uparrow 4$ seems to eliminate the dispersion term; however, it turns out that one retains an energy inequality once appropriate exponential factors are included.

One also obtains uniqueness of weak solution to Cauchy problem for Kolmogorov PDE at least for sufficiently regular initial functions [K3].

In some sense, Orlicz space $L_{\cosh-1}$ can be viewed as the limit of L^p spaces, i.e. as $p>\frac{2}{2-\sqrt{\delta}}$ tends to ∞ as $\delta\uparrow 4$. See the end of Remark 9.2 for further discussion.

Remark 5.2 (On strong solutions and stochastic transport equation)**.** 1. Assertion (ii) on strong solvability was proved in [KM1] using the method of Röckner-Zhao [RZ2]. Their condition reads, for time-homogeneous drifts, as $|b|\in L^d+L^\infty$. For these drifts, the form-bound δ can be chosen arbitrarily small, so multiplying such drift by an arbitrarily large constant still leaves it admissible. This is a natural property for the applications to Navier-Stokes equations, i.e. the focus of Röckner and Zhao in [RZ2]. Their approach is based on the compactness criterion on the Wiener-Sobolev space. [KM1] extended (in fact, simplified) their argument to include form-bounded drifts. Also, following [RZ2], the proof of the conditional strong uniqueness in [KM1] combines Cherny's theorem [Ch] (which is, in a sense, the dual to the Yamada-Watanabe principle) and the conditional weak uniqueness from [KM2, K5], cf. Theorem 5.2(v) and (vi).

Let us also mention a very recent result of Krylov [Kr4] where he proved, using his method based on the Itô-Duhamel series (that can also be viewed as the Duhamel series for the stochastic transport equation) and his earlier results with Veretennikov, strong solvability for drifts in the Morrey class $M_{2+\varepsilon}$ and diffusion coefficients with derivatives in $M_{2+\varepsilon}$. His result thus extends a significant portion of [KM1] to some discontinuous diffusion coefficients. We discuss conditions of this type in the upcoming Section 6. Earlier, Krylov [Kr2] obtained strong well-posedness of the

SDE assuming additionally that the drift belongs to $M_{(2\vee\frac{d}{2})+\varepsilon}$, which is a subclass of $M_{2+\varepsilon}$ that, in higher dimensions, excludes some interesting drifts having singularities along hypersurfaces.

2. There is a well known link between the stochastic transport equation (5.11) and the SDE

$$X_t = x - \int_0^t b(X_r)dr + \sqrt{2}B_t. \tag{5.13}$$

Namely, when b is bounded and smooth, the solution v to the STE (5.11) can be represented as

$$v(t) = f(\Psi_t^{-1}), \quad t \geqslant 0, \tag{5.14}$$

where $\Psi_t : \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is the stochastic flow for the SDE (5.13) i.e. there exists $\Omega_0 \subset \Omega$, $\mathbb{P}(\Omega_0) = 1$, such that, for all $\omega \in \Omega_0$, $\Psi_t(\cdot,\omega)\Psi_s(\cdot,\omega) = \Psi_{t+s}(\cdot,\omega)$, $\Psi_0(x,\omega) = x$, and

1) for every $x \in \mathbb{R}^d$, the process $t \mapsto \Psi_t(x,\omega)$ is a strong solution to (5.13),

2) $\Psi_t(x,\omega)$ is continuous in (t,x), $\Psi_t(\cdot,\omega) : \mathbb{R}^d \to \mathbb{R}^d$ are homeomorphisms and $\Psi_t(\cdot,\omega)$, $\Psi_t^{-1}(\cdot,\omega) \in C^\infty(\mathbb{R}^d,\mathbb{R}^d)$.

Beck-Flandoli-Gubinelli-Maurelli [BFGM] reversed this connection when b is singular (for time-homogeneous drifts their condition reads as $|b| \in L^d + L^\infty$). They used the stochastic transport equation (5.11) to construct, for a.e. initial point $x \in \mathbb{R}^d$, a strong solution to SDE (5.13). Hence, having Theorem 5.2(ix), one can extend the argument of [BFGM] to $b \in \mathbf{F}_\delta$, see [KSS, Remark 1]. However, since this approach excludes a measure zero set of initial points, it does not imply Theorem 5.2(ii).

Remark 5.3 (Krylov-type bound in the distributional case)**.** Consider the assumptions of Theorem 5.1. Then the following a priori Krylov-type bound holds. Let $V \in \mathbf{F}_{\delta_1}$, $\delta_1 < \infty$, be a form-bounded potential (cf. (5.6)). Let $W = \operatorname{div} w$ for some vector field w whose components lie in BMO. Fix some smooth approximations $\{V_n\} \in [V]$, $\{W_m\} \in [W]$, defined just in Section 2. Fix $1 < \theta < \frac{d}{d-2}$ and $p \geq 2$ such that $p > \frac{2}{2-\sqrt{\delta}}$. Then, for all $f \in \mathcal{S}$,

$$\sup_{x\in\mathbb{R}^d}\left|\mathbb{E}_{\mathbb{P}_x}\int_0^1 (V_n + W_m)(\omega_s)f(\omega_s)ds\right| \leq K\|A_m\|_{p\theta} \vee \|A_m\|_{p\theta'}, \tag{5.15}$$

where

$$A_m = |w_m^i||\nabla f| + (1+|w_m^i|)|f|,$$

and constant K does not depend on n, m or f. Informally, this bound shows that a solution of SDE (5.1) cannot spend too much time near the singularities of V and W. The proof is essentially given in Proposition 10.1. (The latter is an elliptic estimate, so one needs to use identity $1 = e^{\mu s}e^{-\mu s}$ to arrive at (5.15).) There we take as V and W the components of vector fields b and q, but the proof extends to V and W right away since it does not exploit any interaction between the coefficients in the right-hand side and the drift term.

By running parabolic De Giorgi's iterations, one refines (5.15) to

$$\left|\mathbb{E}_{\mathbb{P}_x^{n,m}}\int_0^\varepsilon (V_n + W_m)(\omega_s)f(\omega_s)ds\right| \leq H(\varepsilon), \tag{5.16}$$

where $H(\varepsilon) \downarrow 0$ ($\varepsilon \downarrow 0$) is independent of n, m and f (and $x \in \mathbb{R}^d$). If we could place the absolute value under the integral, then, after taking $V = b^i$ and $W = q^i$, a standard argument would allow to conclude tightness of $\{\mathbb{P}_x^n\}$. However, since W is a distribution, we cannot do this. Still, a finer argument of Hao-Zhang [HZ] shows that one can conclude tightness of $\{\mathbb{P}_x^n\}$ from, basically, (5.16), by applying Itô's formula to $\sqrt{\sigma + |x-x_0|^2}$ with $\sigma > 0$ small, which allows to control the smallness of the incremenents of solutions of the approximating SDEs after taking $\sigma \downarrow 0$. That said, [HZ] need a tightness argument since they are dealing with divergence-free super-critical drifts, while we are dealing with general critical drifts and obtain stronger convergence results for $\{\mathbb{P}^n_x\}$ provided by the theory of Feller semigroups.

6. Diffusion coefficients with form-bounded ∇a

By [MV, Theorem 6.1], for a distributional vector field c on $\mathbb{R}^d$, one has

$$-\Delta + c \cdot \nabla \in \mathcal{B}(W^{1,2}, W^{-1,2}), \tag{6.1}$$

if and only if c admits a decomposition

$$c = b + q \quad \text{ for some } b \in \mathbf{F}_\delta \text{ and } q \in \mathbf{BMO}_\sharp^{-1}. \tag{6.2}$$

Here $\mathbf{BMO}_\sharp^{-1}$ consists of divergence-free vector fields $q = \nabla Q$ whose $n \times n$ anti-symmetric primitive $Q = (Q^{ij})_{i,j=1}^d$ satisfies

$$\|Q^{ij}\|_{\mathrm{BMO}_\sharp} = \sup_{x \in \mathbb{R}^d, 0<R\leq 1} \frac{1}{|B_R|} \int_{B_R(x)} |Q^{ij} - (Q^{ij})_{B_R(x)}| dy < \infty.$$

It follows that Theorem 5.1 covers exactly the same class of drifts that guarantee the embedding (6.1), up to replacing $\mathbf{BMO}_\sharp^{-1}$ with $\mathbf{BMO}^{-1}$, i.e. imposing rather mild assumptions on the growth of Q at infinity. In fact, [MV, Theorem 1] yields a similar necessary and sufficient condition for the embedding (6.1) for the homogeneous Sobolev spaces, in which case $\mathbf{BMO}_\sharp^{-1}$ in (6.2) gets replaced with $\mathbf{BMO}^{-1}$, and one has $c_\delta = 0$ in the form-boundedness condition for b.

As noted above, in light of the result of Mazya and Verbitsky [MV, Theorem 6.1], the drift conditions in Theorem 5.1 are essentially optimal if one expects the Kolmogorov operator $-\Delta + c\cdot\nabla$ to be $W^{1,2} \to W^{-1,2}$ bounded. In Theorem 7.1, however, we will be dealing with a larger than $\mathbf{F}_\delta$ class of Borel measurable drifts such that the Kolmogorov operator is $\mathcal{W}^{\frac{3}{2},2} \to \mathcal{W}^{-\frac{1}{2},2}$ (Bessel spaces) bounded. This apparent ambiguity and the question why one might expect Kolmogorov operator to be $W^{1,2} \to W^{-1,2}$ bounded is clarified by noting that both Theorem 5.1 and [MV, Theorem 6.1] are valid in a greater generality. Let a be a bounded uniformly elliptic symmetric matrix field on $\mathbb{R}^d$, i.e. $a \in H_\xi$ for some $\xi > 0$. The cited result of Mazya and Verbitsky states that one has

$$-\nabla \cdot a \cdot \nabla + c \cdot \nabla \in \mathcal{B}(W^{1,2}, W^{-1,2}) \quad \Leftrightarrow \quad c = b + q \text{ as in } (6.2), \tag{6.3}$$

where $-\nabla \cdot a \cdot \nabla$ does not let us deviate from the embedding $W^{1,2} \to W^{-1,2}$. In Theorem 6.1, we treat non-divergence form operators

$$-a \cdot \nabla^2 + c \cdot \nabla,$$

where $\nabla a \in [L^2_{\mathrm{loc}}]^{d \times d}$ is in $\mathbf{F}_\delta$. Since

$$-a \cdot \nabla^2 + c \cdot \nabla = -\nabla \cdot a \cdot \nabla + (\nabla a + c) \cdot \nabla,$$

[MV, Theorem 6.1] again applies. For such diffusion coefficients, which are considered in the next theorem, our condition on the drift is, arguably, optimal.

For our $a \in H_\xi$, define the row divergence operator $a \mapsto \nabla a$ via

$$(\nabla a)_j := \sum_{i=1}^d \nabla_i a_{ij}, \quad 1 \leq j \leq d.$$

Put $\sigma=\sqrt{a}$. In Theorem 6.1 below we consider SDE

$$X_t = x - \int_0^t \big(b(X_s)+q(X_s)\big)ds + \sqrt{2}\int_0^t \sigma(X_s)dB_s, \tag{6.4}$$

with $\nabla a + b \in \mathbf{F}_\delta$ and $q \in \mathbf{BMO}^{-1}$.

Example 6.1. For example, let

$$a(x) = I + c\frac{x\otimes x}{|x|^2}, \quad x\in\mathbb{R}^d, \quad c>-1.$$

Then $\nabla a(x) = c(d-1)\frac{x}{|x|^2}$, so, in view of Example 3.2,

$$\nabla a \in \mathbf{F}_{\delta_1}, \quad \delta_1 = \frac{4c^2(d-1)^2}{(d-2)^2}.$$

This matrix field produces diffusion coefficients σ with discontinuity at the origin that is strong enough to make the weak solution to SDE (6.4) (with $b=q=0$) arrive at the origin with positive probability, see e.g. [B, Ch. V, Sect. 3]. This example can be extended to a matrix field a in $\mathbb{R}^{Nd}$ similar to the many-particle drift (1.5), i.e. corresponding to N particles in $\mathbb{R}^d$ interacting via diffusion coefficients.

Let $\{a_n\},\{b_n\} \in [a,b]$ (Section 2), $\{q_m\}\in[q]$ and $\sigma_n = \sqrt{a_n}$. Define the approximating Kolmogorov operators

$$\Lambda(a_n,b_n,q_m) := -a_n\cdot\nabla^2 + (b_n+q_m)\cdot\nabla, \quad D\big(\Lambda(a_n,b_n,q_m)\big) = (1-\Delta)^{-1}C_\infty.$$

By classical theory, these are generators of strongly continuous Feller semigroups on C_∞. One has

$$e^{-t\Lambda(a_n,b_n,q_m)}f(x) = \mathbf{E}[f(X_t^{n,m})],$$

where $X_t^{n,m}$ is the unique strong solution to SDE

$$X_t^{n,m} = x - \int_0^t \big(b_n(X_s^{n,m}) + q_m(X_s^{n,m})\big)ds + \sqrt{2}\int_0^t \sigma_n(X_s^{n,m})dB_s, \tag{6.5}$$

considered on a fixed complete probability space $\mathfrak{F} = (\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbf{P})$, B_t is a d-dimensional Brownian motion on this space. Set $\mathbb{P}_x^{n,m} := \mathbf{P}(X_t^{n,m})^{-1}$.

Theorem 6.1. *Let $d\geq 3$. Let $a\in H_\xi$ and let b and q be, respectively, Borel measurable and distribution-valued vector fields $\mathbb{R}^d\to\mathbb{R}^d$ that satisfy*

$$\begin{cases} \nabla a + b \in \mathbf{F}_\delta \text{ with } \delta<\xi^2, \\ q\in\mathbf{BMO}^{-1}. \end{cases}$$

Let $\{a_n\},\{b_n\}\in[a,b]$, $\{q_m\}\in[q]$. The following are true:

(i) (Feller semigroup) *The limit*

$$s\text{-}C_\infty\text{-}\lim_n \lim_m e^{-t\Lambda(a_n,b_n,q_m)} \quad \textit{(loc. uniformly in } t\geq 0)$$

exists and determines a strongly continuous Feller semigroup on C_∞, say, $e^{-t\Lambda} = e^{-t\Lambda(a,b,q)}$, where thus $\Lambda \supset -a\cdot\nabla^2 + (b+q)\cdot\nabla$ in C_∞.

(ii) (Approximation uniqueness) *The limit in (i) does not depend on the choice of $\{a_n\}$, $\{b_n\}$ and $\{q_m\}$.*

(iii) (Relaxed approximation uniqueness) *In fact, we can replace the strong convergence of b_n to b in $[L^2]^d$ by the weak convergence*

$$b_n \overset{w}{\rightarrow} b \quad \text{in } [L^2]^d$$

to have convergence of the Feller semigroups $e^{-t\Lambda(a_n,b_n,q)} \overset{s}{\rightarrow} e^{-t\Lambda(a,b,q)}$ in C_∞ (loc. uniformly in $t \geq 0$).

(iv) (Generalized martingale solution) *There exists a strong Markov family of probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ on the canonical space $\mathbf{C}$ of continuous trajectories ω such that*

$$e^{-t\Lambda(a,b,q)}f(x) = \mathbb{E}_{\mathbb{P}_x}[f(\omega_t)], \quad f \in C_\infty, \quad x \in \mathbb{R}^d,\ t \geq 0,$$

$$\mathbb{P}_x = w\text{-}\mathcal{P}(\mathbf{C})\text{-}\lim_n \lim_m \mathbb{P}_x^{n,m},$$

and for every v in the domain $D\big(\Lambda(a,b,q)\big) \subset C_\infty$ of operator $\Lambda(a,b,q) \supset -a\cdot\nabla^2 + (b+q)\cdot\nabla$, a dense subspace of C_∞, the process

$$t \mapsto v(\omega_t) - x + \int_0^t \Lambda(a,b,q)v(\omega_s)ds$$

is a continuous martingale with respect to $\mathbb{P}_x$.

(v) (Weak solution with disperse initial data) *Assume additionally that b, q and b_n, q_m have supports in a fixed ball of finite radius. Given an initial (smooth) probability density ν_0 satisfying $\langle \nu_0^{2r}\rangle < \infty$ for some $1 < r < \delta^{-\frac{1}{2}}$, there exist a probability space $\mathfrak{F}' = (\Omega', \mathcal{F}', \{\mathcal{F}'_t\}_{t\geq 0}, \mathbf{P}')$ and a continuous process X_t on this space such that the limit*

$$A_t := L^2(\Omega')\text{-}\lim_n \lim_m \int_0^t \big(b_n(X_s) + q_m(X_s)\big)ds$$

exists, we have a.s.

$$X_t = X_0 - A_t + \sqrt{2}\int_0^t \sigma(X_s)dB_s, \quad t > 0,$$

for a $\mathcal{F}'_t$-Brownian motion B_t, and $\mathbf{P}'X_0^{-1}$ has density ν_0.

(vi) (Dispersion estimates and uniqueness of weak solution to Kolmogorov PDE) *For every $p > \frac{2}{2-\xi^{-1}\sqrt{\delta}}$, the operators*

$$e^{-t\Lambda_p(a,b,q)} := \Big[e^{-t\Lambda(a,b,q)} \upharpoonright C^\infty \cap L^p\Big]^{\text{clos}}_{L^p\to L^p}$$

are bounded on L^p, constitute a strongly continuous semigroup, and for all $\frac{2}{2-\xi^{-1}\sqrt{\delta}} < p \leq r < \infty$,

$$\|e^{-t\Lambda_p(a,b,q)}f\|_r \leq C_{\delta,d}e^{\omega_p t}t^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{r})}\|f\|_p, \quad f \in L^p,$$

and so, by the Dunford-Pettis theorem, $e^{-t\Lambda(a,b,q)}$, $t \geq 0$, are integral operators.

Furthermore, $v(t) := e^{-t\Lambda_2(a,b,q)}f$, $f \in L^2$, is the unique weak solution to Cauchy problem

$$(\partial_t - a\cdot\nabla^2 + (b+q)\cdot\nabla)v = 0, \quad v|_{t=0} = f,$$

in the standard Hilbert triple $W^{1,2} \hookrightarrow L^2 \hookrightarrow W^{-1,2}$.

(vii) (Strong Feller property of the resolvents) *For every $f \in L^{p\theta} \cap L^{p\theta'} \cap C_\infty$, for every $x \in \mathbb{R}^d$, solution $u = (\mu + \Lambda(a,b,q))^{-1} f$ to the elliptic Kolmogorov equation $(\mu - a \cdot \nabla^2 + (b+q) \cdot \nabla)u = f$ satisfies*

$$\sup_{B_{\frac{1}{2}}(x)} |u| \leq K\bigg((\mu-\mu_0)^{-\frac{1}{p\theta}} \langle |f|^{p\theta} \rho_x \rangle^{\frac{1}{p\theta}} + (\mu-\mu_0)^{-\frac{\beta}{p}} \langle |f|^{p\theta'} \rho_x \rangle^{\frac{1}{p\theta'}} \bigg),$$

for fixed $1 < \theta < \frac{d}{d-2}$ and $p > \frac{2}{2-\xi^{-1}\sqrt{\delta}}$, for all μ strictly greater than certain μ_0. The constant K does not depend on f or x.

Also, a priori Krylov-type bound analogous to (5.15) holds.

Theorem 6.1, when applied to $a = I$ (so $\xi = 1$), imposes a more restrictive condition on the form-bound δ than Theorem 5.1, i.e. $\delta < 1$ instead of $\delta < 4$. In fact, in the assumptions of Theorem 6.1 but with $\delta < 4\xi^2$, after passing to some subsequence $\{n_k\}$ we still get Feller semigroup

$$e^{-t\Lambda(a,b,q)} = s\text{-}C_\infty\text{-}\lim_k \lim_m e^{-t\Lambda(a_{n_k}, b_{n_k}, q_m)} \text{ (loc. uniformly in } t \geq 0).$$

That is, Theorem 6.1, with the exception of the approximation uniqueness, remains valid. Also, assuming that $\delta < \xi^2$ and $q = 0$, many assertions of Theorem 5.2 remain valid:

$(viii)$ (Classical weak solution) The proof is discussed in details in [KS2].

(ix) (Strong solvability) The method of Röckner-Zhao [RZ2] yields, after a few modifications in the spirit of [KM1], strong existence for SDE with diffusion coefficients satisfying

$$\nabla a_{ij} \in \mathbf{F}_{\nu_{ij}}, \quad \text{form-bounds } \nu_{ij} < c_d \text{ for some constant } c_d \downarrow 0 \text{ as } d \uparrow \infty, \tag{6.6}$$

the details will appear in [KM4]; this gives another proof of a recent result of Krylov in [Kr4].

(x), (xi) (Krylov bound) and (Krylov-type bound) The proof follows by combining the arguments in [K2] and [KS7].

(xiv), (xv) (Another kind of approximation uniqueness and gradient bounds (5.8)) These were proved in [KS7].

Earlier, Veretennikov [V] and Zhang [Z], Zhang-Zhao [ZZ2] established strong well-posedness for diffusions coefficients having derivatives in L^p with p strictly larger than $2d$ or d, respectively. These assumptions, however, make diffusion coefficients Hölder continuous, so they exclude e.g. Example 6.1.

In the case $q = 0$, [KS7] proved weak existence, gradient bounds of type (5.8) and an analogue of Theorem 5.2(vii) ("another kind of approximation uniqueness") for SDE (6.4) under dimension-dependent conditions on the form-bounds of b and ∇a_{ij}, i.e. (6.6). Theorem 6.1 shows that the assumptions on both the diffusion coefficients and drift can be weakened and made dimension-independent.

Similar to (6.6) conditions on diffusion coefficients on the scale of Morrey spaces are considered by Krylov, see [Kr1, Kr2].

7. Weakly form-bounded drifts and Keller-Segel finite particles

In the previous two sections we tested out results for SDEs with form-bounded drifts against the interacting particle system in Example 1.1. This, however, was limited to dimensions $d \geq 3$. In dimension $d = 2$, which is of interest e.g. in the study of the Keller-Segel model of chemotaxis, the particle system in Example 1.1 is more difficult to handle since its drift (1.5) is not in $L^2_{\rm loc}(\mathbb{R}^2)$

and, thus, is not form-bounded. We can address this issue, at least to some extent, by pursuing a different approach to proving weak well-posedness of SDEs. It works for substantially larger class of weakly form-bounded drifts.

DEFINITION 7.1. A vector field $b \in [L^1_{\rm loc}]^d$ is said to be weakly form-bounded if there exists constant $\delta > 0$ such that

$$\langle |b|\varphi, \varphi\rangle \leq \delta \|(\lambda - \Delta)^{\frac{1}{4}}\varphi\|_2^2, \quad \forall\, \varphi \in C_c^\infty,$$

for some $\lambda = \lambda_\delta \geq 0$. This will be abbreviated as $b \in \mathbf{F}_\delta^{\frac{1}{2}}$.

Example 7.1. 1. Morrey class $M_{1+\varepsilon}$ is a large subclass of $\mathbf{F}_\delta^{\frac{1}{2}}$ defined in elementary terms:

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

The inclusion follows by D. R. Adams' theorem [A1, Theorem 7.3]. The value of δ will be proportional to the Morrey norm, with a coefficient that depends on the constants in some fundamental inequalities of Harmonic Analysis.

2. The class of form-bounded drifts $\mathbf{F}_{\delta^2}$ considered in the previous section, i.e.

$$\langle |b|^2\varphi, \varphi\rangle \leq \delta^2\|(-\Delta)^{\frac{1}{2}}\varphi\|_2^2 + c_{\delta^2}\|\varphi\|_2^2 \quad \left(= \delta^2\|(\lambda-\Delta)^{\frac{1}{2}}\varphi\|_2^2, \;\; \lambda = \frac{c_{\delta^2}}{\delta^2}\right),$$

is a proper subclass of $\mathbf{F}_\delta^{1/2}$. This is seen easily by appyling Heinz' inequality. Alternatively, one can invoke the inclusion $\mathbf{F}_\delta$ (with $c_\delta = 0$) $\subset M_2$, see Examples 3.1 in Section 1, and, next, apply $M_2 \subset M_{1+\varepsilon}$ if $\varepsilon < 1$, and then use the previous example. That said, if one follows this path, one to a large extent loses the control over the value of the form-bound δ, which is in our focus in this paper.

3. It is instructive to compare how $\mathbf{F}_\delta$ and $\mathbf{F}_\delta^{1/2}$ handle the weak L^d class. Namely, for $|b| \in L^{d,\infty}$, we verify, using [KPS, Prop. 2.5, 2.6, Cor. 2.9],

$$\begin{aligned} d \geq 2, \quad b \in \mathbf{F}_\delta^{\frac{1}{2}} \text{ with } \sqrt{\delta} &= \||b|^{\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}\|_{2\to 2} \leqslant \|(|b|^*)^{\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}\|_{2\to 2} \\ &\leqslant \left(\|b\|_{d,\infty}\Omega_d^{-\frac{1}{d}}\right)^{\frac{1}{2}}\||x|^{-\frac{1}{2}}(-\Delta)^{-\frac{1}{4}}\|_{2\to 2} = \left(\|b\|_{d,\infty}\Omega_d^{-\frac{1}{d}}\right)^{\frac{1}{2}} 2^{-\frac{1}{2}}\frac{\Gamma\left(\frac{d-1}{4}\right)}{\Gamma\left(\frac{d+1}{4}\right)}, \end{aligned}$$

where $\Omega_d = \pi^{\frac{d}{2}}\Gamma(\frac{d}{2}+1)$, and $|b|^*$ is the symmetric decreasing rearrangement of $|b|$. Similarly,

$$\begin{aligned} d \geq 3, \quad b \in \mathbf{F}_{\delta_1} \text{ with } \sqrt{\delta_1} &= \||b|(-\Delta)^{-\frac{1}{2}}\|_{2\to 2} \\ &\leqslant \|b\|_{d,\infty}\Omega_d^{-\frac{1}{d}}\||x|^{-1}(-\Delta)^{-\frac{1}{2}}\|_{2\to 2} \\ &\leqslant \|b\|_{d,\infty}\Omega_d^{-\frac{1}{d}}2^{-1}\frac{\Gamma\left(\frac{d-2}{4}\right)}{\Gamma\left(\frac{d+2}{4}\right)} = \|b\|_{d,\infty}\Omega_d^{-\frac{1}{d}}\frac{2}{d-2}. \end{aligned}$$

In particular, using [KPS, Cor. 2.9], we have

$$d \geq 2, \quad x|x|^{-2} \in \mathbf{F}_\delta^{\frac{1}{2}}, \quad \sqrt{\delta} = 2^{-\frac{1}{2}}\frac{\Gamma\left(\frac{d-1}{4}\right)}{\Gamma\left(\frac{d+1}{4}\right)}, \tag{7.1}$$

$$d \geq 3, \quad x|x|^{-2} \in \mathbf{F}_{\delta_1}, \quad \sqrt{\delta_1} = \frac{2}{d-2}. \tag{7.2}$$

In fact, (7.2) coincides with the classical Hardy inequality.

4. An important proper subclass of $\mathbf{F}_\delta^{1/2}$ that is not contained in the Morrey class $M_{1+\varepsilon}$, regardless of how small $\varepsilon > 0$ is, is the Kato class. The Kato class consists of vector fields $b \in [L^1_{\rm loc}]^d$ such that

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

for some $\delta > 0$ and $\lambda = \lambda_\delta \geq 0$ (The inclusion Kato class $\subset \mathbf{F}_\delta^{1/2}$ follows e.g. by duality and interpolation.) SDEs with Kato class drifts were treated by Bass-Chen [BC], who studied Brownian motion on fractals such as the Sierpinski gasket. To be more precise, they considered measure-valued b with the total variation $|b|$ satisfying the Kato class condition. Moreover, when δ is sufficiently small, one obtains two-sided Gaussian bounds on the heat kernel of $-\Delta + b\cdot\nabla$ [Za]. Note that the Kato class does not contain $[L^d]^d$, but, for every fixed $\varepsilon > 0$, it contains some vector fields that are not in $[L^{1+\varepsilon}_{\rm loc}]^d$.

The proof of the next theorem is based on the resolvent representation (7.6) where we, crucially, work with the fractional powers $|b|^{\frac{1}{r}}$ for $r > d-1$ (so, we *take advantage of the fact that b is not distributional or measure-valued*).

Theorem 7.1 ([K1, KS1]). *Let $d \geq 2$. Assume that $b \in \mathbf{F}_\delta^{1/2}$ with weak form-bound δ satisfying*

$$\delta < m_d^{-1}\begin{cases} \frac{4(d-2)}{(d-1)^2} & \text{if } d \geq 4, \\ 1 & \text{if } d = 2,3, \end{cases}$$

where $m_d := \pi^{\frac{1}{2}}(2e)^{-\frac{1}{2}}d^{\frac{d}{2}}(d-1)^{\frac{1-d}{2}}$. The following are true:

(i) (Weak solution to SDE) *There exist a strong Markov family of probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ on the canonical space of continuous trajectories $\mathbf{C}$ that deliver, for every $x \in \mathbb{R}^d$, a weak solution to SDE*

$$X_t = x - \int_0^t b(X_r)dr + \sqrt{2}B_t. \tag{7.3}$$

(ii) (Feller semigroup)

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

is a strongly continuous Feller semigroup on C_∞.

(iii) (Uniqueness of weak solution to Kolmogorov backward PDE [KS5]) *$v(t) := e^{-t\Lambda(b)}f$, $f \in C_\infty \cap L^2$, is the unique weak solution to Cauchy problem*

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

in the "shifted" triple of Bessel potential spaces $\mathcal{W}^{\frac{3}{2},2} \hookrightarrow \mathcal{W}^{\frac{1}{2},2} \hookrightarrow \mathcal{W}^{-\frac{1}{2},2}$. This result yields approximation uniqueness for $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$.

(iv) (Another kind of approximation uniqueness) *If $\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d}$ is another weak solution to* (7.3) *such that*

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

for some $\{\tilde{b}_n\} \subset \mathbf{F}_{\delta_1}^{1/2} \cap [C_b \cap C^\infty]^d$ with $\delta < \frac{4(d-2)}{(d-1)^2}$ if $d \geq 4$ or $m_d\delta < 1$ if $d = 2,3$, and λ_δ independent of n, then $\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d} = \{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$.

(v) (Elliptic gradient bounds) $u := (\mu+\Lambda(b))^{-1}f$, $f \in C_\infty \cap L^r$, $r \in]d-1, \frac{2}{1-\sqrt{1-m_d\delta}}[$, *satisfies*

$$\|(\mu-\Delta)^{\frac{1}{2}+\frac{1}{2s}}u\|_r \leq K\|(\mu-\Delta)^{-\frac{1}{2}+\frac{1}{2\ell}}f\|_r, \quad \text{for all } 1 \leq \ell < r < s,$$

for all μ greater than a generic μ_0 (cf. (5.9)). In particular, since $r > d-1$, we can select s sufficiently close to r so that by the Sobolev embedding theorem u is Hölder continuous.

Corollary 7.1 (Finite particle approximation of the elliptic-parabolic Keller-Segel model)**.** *In $\mathbb{R}^{2N}$, consider SDE*

$$X_t = x_0 - \int_0^t b(X_s)ds + \sqrt{2}B_t, \quad x_0 = (x_0^1, \dots, x_0^N) \in \mathbb{R}^{2N}, \tag{7.4}$$

where $B_t = (B_t^1, \dots, B_t^N)$ is a Brownian motion in $\mathbb{R}^{2N}$, and

$$b_i(x^1, \dots, x^N) := \frac{\sqrt{\kappa}}{N} \sum_{j=1, j\neq i}^{N} \frac{x^i - x^j}{|x^i - x^j|^2}. \tag{7.5}$$

Then, provided that $\kappa < \frac{C}{N^3}$, the assertions of Theorem 7.1 are valid for this particle system.

Proof. Thus defined drift $b : \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ is in $\mathbf{F}_\delta^{1/2}$. In fact, it is in the Morrey class $M_{1+\varepsilon}$, a subclass of $\mathbf{F}_\delta^{1/2}$. To see this, it suffices to prove this inclusion for a single term

$$\mathbb{R}^{2N} \ni (x^1, \dots, x^N) \mapsto \frac{x^1 - x^2}{|x^1 - x^2|^2}$$

which, after a change of variable, reduces to proving that the scalar function $(x^1, \dots, x^N) \mapsto |x^1|^{-1}$ is in the Morrey class $M_{1+\varepsilon}$. Put $C_r(x) = D_r(x^1) \times \dots \times D_r(x^N)$ (the direct product of N discs centered at x^i). We have

$$\begin{aligned}
\||x^1|^{-1}\|_{M_{1+\varepsilon}} &\leq c \sup_{r>0} r \left(\frac{1}{r^{2N}} \langle \mathbf{1}_{C_r(0)} |x^1|^{-(1+\varepsilon)} \rangle \right)^{\frac{1}{1+\varepsilon}} \\
&= c \sup_{r>0} r \left(\frac{1}{r^2} \int_{D_r(0)} |x^1|^{-(1+\varepsilon)} dx^1 \right)^{\frac{1}{1+\varepsilon}} \\
&= c \sup_{r>0} r \left(\frac{1}{r^2} \int_0^r t^{-(1+\varepsilon)+1} dt \right)^{\frac{1}{1+\varepsilon}} = c \sup_{r>0} r \left(\frac{1}{r^2} \frac{r^{-\varepsilon+1}}{-\varepsilon+1} \right)^{\frac{1}{1+\varepsilon}} < \infty.
\end{aligned}$$

□

The main, quite unacceptable drawback of Corollary 7.1 is that the condition on κ degenerates as $N \to \infty$. Fournier-Jourdain [FJ], Fournier-Tardy [FT] and Tardy [T] exploit the special form of the interaction kernel in (7.5) and establish for (7.4), among other results, weak existence and the existence of mean field limit as $N \to \infty$ for all $\kappa \in [0, 16[$, where 16 is the sticky collisions threshold for (7.4), (7.5). One can also apply the Dirichlet forms approach, see Cattiaux-Pédèches [CP]. Already the weak existence results of [FJ] are thus much stronger than Corollary 7.1. Our point here, however, is different. Corollary 7.1 shows that one, in fact, can reach the Keller-Segel finite particle system (7.4), (7.5) by applying results on general singular SDEs. (Fournier and Jourdain noted that, at the time of writing, the strongest known SDEs results for general singular drift did not apply (7.5). This was indeed true, but only until the preprint [KS1] appeared a few months later; unfortunately, at the time of writing [KS1] we were not aware of papers [CP, FJ].)

One advantage of Theorem 7.1, compared to [CP, FJ, FT, T], is that we can easily modify the drift in Corollary 7.1. For example, multiplying each interaction kernel by a function with L^∞ norm less or equal to one does not affect the conclusion.

Remarks. 1. Replacing condition $b \in \mathbf{F}_\delta$ by more general condition $b \in \mathbf{F}_\delta^{1/2}$ comes at a cost. Although one can still include some diffusion coefficients, these may no longer be discontinuous (cf. [K4, Sect. 14]). Also, distributional drifts $q \in \mathbf{BMO}^{-1}$ are out of reach. It is, however, possible to consider drifts $b+q$, where $b \in \mathbf{F}_\delta^{1/2}$ and q is measure-valued with total variation in the Kato class [K8].

2. In Theorem 7.1, we construct the candidate for resolvent of the Feller generator *a priori.* It is the following formal Neumann series for

$$\mu + \Lambda \supset \mu - \Delta + b \cdot \nabla,$$

possibly after a modification on a measure set [K1]:

$$(\mu+\Lambda)^{-1}f := (\mu-\Delta)^{-1}f - (\mu-\Delta)^{-\frac{1}{2}-\frac{1}{2s}}Q_r(1+T_r)^{-1}G_r(\mu-\Delta)^{-\frac{1}{2}+\frac{1}{2\ell}}f, \tag{7.6}$$

where $f \in L^r \cap C_\infty$, and

$$Q_r := (\mu-\Delta)^{-\frac{1}{2}+\frac{1}{2s}}|b|^{\frac{1}{r'}}, \quad G_r := b^{\frac{1}{r}} \cdot \nabla(\mu-\Delta)^{-\frac{1}{2}-\frac{1}{2\ell}} \quad \text{are bounded on } L^r,$$

$$T_r := b^{\frac{1}{r}} \cdot \nabla(\mu-\Delta)^{-1}|b|^{\frac{1}{r'}} \quad \text{bounded on } L^r,$$

where

$$b^{\frac{1}{r}} := |b|^{-1+\frac{1}{r}}b, \quad \ell, s \text{ satisfy } 1 \leq \ell < r < s,$$

and, of course, one gets stronger regularity result by choosing ℓ, s close to r. The proof of the boundedness of Q_r, G_r and T_r is based on the Stroock-Varopoulos inequalities for symmetric Markov generators, i.e. this is an elliptic argument. The smallness condition on δ in Theorem 7.1 ensures $\|T_r\|_{r\to r} < 1$, so that $(1+T_r)^{-1}$ converges as geometric series in L^r. Earlier, similar estimates were employed in [BS, LS] to refine the L^2 theory of Schrödinger operators with the usual form-bounded potentials to an L^r theory, which allowed the authors, for instance, to obtain additional information about the Sobolev regularity of the eigenfunctions of Schrödinger operators.

3. A priori, it is not clear why (7.6) should determine the resolvent of a strongly continuous semigroup in L^r. The latter is, in fact, true: the proof uses Hille's theory of pseudoresolvents [K1]. When $r=2$, one can give a different proof using Lions' variational approach, but it requires working in a quintuple of Hilbert spaces (instead of the usual triple) [KS3].

4. Having an explicit candidate for the limiting object, i.e. the Feller resolvent, greatly simplifies the approximation arguments. In Theorem 5.1, no such representation is available, so one must rely on Trotter's approximation theorem, whose key feature is that it does not require any a priori representation of the limiting operator.

5. Although these broad assumptions on b destroy the usual L^r estimates for second-order derivatives of solution u to $(\mu-\Delta+b\cdot\nabla)u=f$, one can still use (7.6) to obtain *some* L^r bounds on $\nabla^2 u$. However, either one needs restriction $r<d$ or these estimates are valid only in weighted space $L^r(\mathbb{R}^d, (1+|b(x)|)-r+1dx)$. (The latter follows by applying $(1+|b|)^{-\frac{1}{r'}}(\mu-\Delta)$ to (7.6); note that with the information about the second derivatives of u disappears at the points where $|b|$ is infinite, but in a controlled way, see [K4] for more detailed discussion.)

The elliptic-parabolic Keller-Segel model of chemotaxis is the following nonlinear Kolmogorov forward equation in $\mathbb{R}^2$:

$$\partial_t\eta - \Delta + \mathrm{div}\,\big(\eta K \star \eta\big) = 0, \quad \eta|_{t=0} = \eta_0 \geq 0, \quad K(x) = \sqrt{\kappa}\frac{x}{|x|^2}, \tag{7.7}$$

where η_t is the concentration density of the chemoattractant, $\langle \eta_t \rangle = \langle \eta_0 \rangle = 1$, and $(K \star \eta)(s,x) = \langle K(x-\cdot)\eta(s,\cdot)\rangle$. The corresponding diffusion process X_t, i.e. $\eta_t = \mathrm{Law}\, X_t$, satisfies the non-linear SDE

$$X_t = X_0 - \int_0^t (K \star \eta)(s, X_s)ds + \sqrt{2}B_t, \quad \eta_0 = \mathrm{Law}\, X_0, \tag{7.8}$$

which arises as the mean field limit of the finite particle systems in Corollar 7.1 (see [FJ]). A crucial question in studying the Keller-Segel model is whether one can cover the entire range of admissible values of the strength of attraction κ, beyond which a blow up occurs in finite time, i.e. starting with X_0 having density that is absolutely continuous with respect to the Lebesgue measure, a delta-function develops in finite time [CPZ, JL]. Once again, we refer to [CP, FJ, FT, T] for very detailed results in this direction.

Theorem 7.1 covers only time-homogeneous drifts and therefore does not apply directly to SDE (7.8). In [K5], a time-inhomogeneous analogue of Theorem 7.1 addresses this issue. It provides, in particular, conditional weak uniqueness for the SDE and strong a priori gradient bounds for a large class of McKean-Vlasov PDEs. Compared to Theorem 7.1, the result in [K5] restricts somewhat the class of admissible drifts: one requires

$$b_1 + b_2, \quad b_2 \in L^\infty(\mathbb{R}^{1+d}), \quad b_1 \in E_{1+\varepsilon}$$

for some $\varepsilon > 0$, where

$$\|b_1\|_{E_{1+\varepsilon}} := \sup_{r>0, z \in \mathbb{R}^{1+d}} r\bigg(\frac{1}{|C_r|}\int_{C_r(z)} |b_1(t,x)|^{1+\varepsilon}dtdx\bigg)^{\frac{1}{1+\varepsilon}}$$

is the parabolic Morrey norm. Here

$$C_r(t,x) := \{(s,y) \in \mathbb{R}^{1+d} \mid t \leq s \leq t + r^2, |x-y| \leq r\}$$

is the parabolic cylinder. The reason for this restriction is that, at least at the moment, we have to replace the elliptic argument based on the Stroock-Varopoulos inequalities by an argument based on a parabolic variant of Adams' estimate in Lemma C.1. Nonetheless, we have partial results suggesting that [K1] and [KS1] can be extended to time-inhomogeneous, weakly form-bounded drifts, thus eliminating the need for Morrey-class assumptions.

For integer powers of $|b|$, the parabolic Adams' estimate was obtained by Krylov [Kr6] who needed it for somewhat different purposes, i.e. to handle discontinuous diffusion coefficients. However, the proof in [Kr6] can be extended to fractional powers $b^{\frac{1}{r}}$, $|b|^{\frac{1}{r'}}$, as is needed in [K5] to run analogous to (7.6) representations for the Duhamel series for $\partial_t - \Delta + b \cdot \nabla$.

Drifts in the class $E_{1+\varepsilon}$ can have strong singularities in time, for instance,

$$|b_1(t,x)| \leq |t - t_0|^{-1/2}, \quad (t,x) \in \mathbb{R}^{1+d}$$

or, more generally, in $L^{2,\infty}(\mathbb{R}, L^\infty(\mathbb{R}^d))$, i.e. in the weak L^2 class in time. Let us also add that the $|t|^{-\frac{1}{2}}$ blow up rate in time is essentially embedded in the Koch-Tataru class (3.5).

The method of the fractional resolvent representations of Neumann (Duhamel) series was also applied in [K7] and [K5]. That said, the proof of $\|T_r\|_{r\to r} < 1$ in [K7] is quite different and does not use Stroock-Varopoulos or Adams' estimates, thereby achieving the least restrictive conditions on δ.

Finally, if one assumes extra spatial regularity of the drift, such as Hölder continuity with exponent $\frac{2}{s}-1$ with $1<s<2$, then it suffices to require Lorentz $L^{s,1}$ regularity in time, see [HWY].

The extension of Theorem 7.1 to SDEs driven by isotropic α-stable process was obtained in [KM3]; see also [K9] regarding time-inhomogeneous drifts in this non-local setting.

8. Critical divergence and the best constant in many-particle Hardy inequality

1. We can substantially relax condition ($\kappa_{\rm hyp}$) on the strength of attraction between the particles in Example 5.1 by employing the many-particle Hardy inequality of [HHLT] and the following variant of Theorem 5.1.

Let $b \in [L^1_{\rm loc}]^d$ be a vector field with divergence $\operatorname{div} b \in L^1_{\rm loc}$. Let $(\operatorname{div} b)_+$ denote the positive part of $\operatorname{div} b$.

Definition 8.1. We say that "potential" $(\operatorname{div} b)_+$ is form-bounded, and write $(\operatorname{div} b)_+^{1/2} \in \mathbf{F}_{\delta_+}$, if

$$\langle (\operatorname{div} b)_+, \varphi^2\rangle \leq \delta_+ \langle |\nabla \varphi|^2\rangle + c_{\delta_+}\langle \varphi^2\rangle \quad \forall\, \varphi \in C_c^\infty$$

for some constants δ_+ and c_{δ_+}.

Theorem 8.1. *Let*

$$\begin{cases} b \in \mathbf{F}_\delta \text{ with } \delta<\infty, \quad (\operatorname{div} b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+} \text{ with } \delta_+<4, \quad (\operatorname{div} b)_- \in L^1+L^\infty, \\ q \in \mathbf{BMO}^{-1}. \end{cases}$$

Let $\{b_n\} \in [b]'$ (see Definition 3.4), $\{q_m\} \in [q]$. Then Theorem 5.1 and Theorem 5.1(i) remain valid.

Example 8.1. Let us establish weak well-posedness of particle system (5.12) using Theorem 8.1 rather than Theorem 5.1. The difference between the two theorems is in the assumptions on the drift $b(x)=(b^1(x),\dots,b^N(x))$,

$$b^i(x):=\frac{1}{N}\sum_{j=1, j\neq i}^N \sqrt{\kappa}\frac{d-2}{2}\frac{x^i-x^j}{|x^i-x^j|^2}, \quad x=(x^1,\dots,x^N).$$

We already know that this drift is form-bounded, but what matters in Theorem 8.1 is the form-bound of potential

$$(\operatorname{div} b(x))_+ = \operatorname{div} b(x) = \sqrt{\kappa}\frac{(d-2)^2}{N}\sum_{1\leq i<j\leq N}\frac{1}{|x^i-x^j|^2}.$$

To verify the form-boundedness of $(\operatorname{div} b)_+$, we invoke the many-particle Hardy inequality: for $d \geq 3$, all $N \geq 2$,

$$C_{d,N}\sum_{1\leq i<j\leq N}\int_{\mathbb{R}^{Nd}}\frac{|\varphi(x)|^2}{|x^i-x^j|^2}dx \leq \int_{\mathbb{R}^{Nd}}|\nabla \varphi(x)|^2 dx \tag{8.1}$$

for all $\varphi \in W^{1,2}(\mathbb{R}^{Nd})$, where, from now on, $C_{d,N}$ denotes the best possible constant in (8.1). Hence

$$(\operatorname{div} b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+}, \quad \delta_+ = \sqrt{\kappa}\frac{(d-2)^2}{N}C_{d,N}^{-1}.$$

To the best of our knowledge, the problem of finding the exact value of $C_{d,N}$ is still open. It is not difficult to obtain a crude lower bound on $C_{d,N}$ by summing up the ordinary Hardy inequalities for the inverse square potential $x^i \mapsto |x^i - x^j|^{-2}$, each in its own copy of $\mathbb{R}^d$. However, as is pointed out by Hoffmann-Ostenhof, Hoffmann-Ostenhof, Laptev and Tidblom in [HHLT], this lower bound on $C_{d,N}$ is quite suboptimal. They provided a finer argument that gives a much better lower bound

$$C_{d,N} \geq (d-2)^2 \max\left\{\frac{1}{N}, \frac{1}{1+\sqrt{1+\frac{3(d-2)^2}{2(d-1)^2}(N-1)(N-2)}}\right\}. \tag{8.2}$$

Therefore, it suffices for us to require

$$\kappa < 16 \tag{$\kappa_{\rm hyp2}$}$$

which guarantees $\delta_+ < 4$ and allows us to apply Theorem 8.1. Comparing ($\kappa_{\rm hyp2}$) with ($\kappa_{\rm hyp}$) in Example 5.1, we see that Theorem 8.1 improves the admissible range for κ by nearly a factor of four.

This gain comes at the cost: we restrict possible modifications of the interaction kernel. Indeed, if we use Theorem 5.1 then we can multiply each interaction kernel by a discontinuous function with the L_∞ norm less or equal to one, and the many-particle drift b remain form-bounded with the same form-bound δ. By contrast, if we use if we use Theorem 8.1, such modification of the interaction kernel is problematic since it will affect the divergence of the drift.

2. We argue that the relationship between the many-particle Hardy inequality and the particle system

$$X_t^i = x_0^i - \sqrt{\kappa}\frac{d-2}{2}\frac{1}{N}\sum_{j=1, j\neq i}^N \int_0^t \frac{X_s^i - X_s^j}{|X_s^i - X_s^j|^2} ds + \sqrt{2}B_t^i \tag{8.3}$$

goes both ways. Namely, we can use the counterexample in (a') of Section 4 to the weak well-posedness of (8.3), i.e. when the strength of attraction κ is too large, to obtain an *upper bound* on the best possible constant $C_{d,N}$ in (8.1).

Theorem 8.2 (An upper bound on the constant in the many particle Hardy inequality (8.1))**.**

$$C_{d,N} \leq \frac{d(d-2)}{N}.$$

Proof. By Theorem 8.1 and the calculation in the previous example, (8.3) has a weak solution for every initial configuration of the particles provided that

$$\sqrt{\kappa}\frac{(d-2)^2}{N}C_{d,N}^{-1} < 4.$$

On the other hand, by the counterexample in (a') of Section 4, if $\kappa > 16(\frac{d}{d-2})^2$, then (8.3) does not have a weak solution, so we must have

$$4\frac{d}{d-2}\frac{(d-2)^2}{N}C_{d,N}^{-1} \geq 4,$$

otherwise a weak solution would exist. This gives the sought upper bound on $C_{d,N}$. □

In [HHLT], the authors also provided, among other results, the following upper bound:

$$C_{d,N} \leq \frac{2d}{2(N-1)}\pi^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right). \tag{8.4}$$

Their argument uses particular test functions in (8.1). Corollary 8.2 improves (8.4) for all $d \geq 3, N \geq 3$. Natually, it gives the same asymptotics in N, but it improves substantially the dependence on the dimension d, i.e. we now have polynomial growth versus factorial growth in d. Of course, the simplicity of the proof of Theorem 8.2 is only seeming: we apply Theorem 8.1 whose proof uses De Giorgi's method.

Theorem 8.2 shows that the lower bound (8.2) of [HHLT] is close to optimal, at least in high dimensions.

3. We can relax the assumptions on b in Theorem 8.1 as follows.

Definition 8.2 (Multiplicative form-boundedness). A vector field $b \in [L^1_{\rm loc}(\mathbb{R}]^d$ is said to be multiplicatively form-bounded if

$$\langle |b|\varphi, \varphi\rangle \leq \delta \|\nabla \varphi\|_2 \|\varphi\|_2 + c_\delta \|\varphi\|_2^2 \qquad \forall\, \varphi \in C_c^\infty.$$

This will be abbreviated as $b \in \mathbf{MF}_\delta$.

Once again, the constant c_δ plays a secondary role when it comes to handling local singularities of b (e.g. $c_\delta > 0$ allows to include L^∞ drifts).

Mazya [M, Sect. 1.4.7] proved that

$$\langle |b|\varphi, \varphi\rangle \leq \delta \|\nabla \varphi\|_2 \|\varphi\|_2 \;\; \forall \varphi \in C_c^\infty \text{ for some } \delta < \infty \quad \Leftrightarrow \quad \sup_{r>0, x \in \mathbb{R}^d} \langle |b| \mathbf{1}_{B_r(x)}\rangle \leq C r^{d-1}$$

for some $C < \infty$, i.e. there is a complete characterization of $\mathbf{MF}_\delta$ in terms of Morrey spaces:

$$\cup_{\delta > 0} \mathbf{MF}_\delta \text{ (with } c_\delta = 0) \quad = \quad M_1 \tag{8.5}$$

(see Appendix D for the proof). Let us emphasize that for the class of form-bounded vector fields one only has inclusions

$$M_{2+\varepsilon} \quad \subset \quad \cup_{\delta>0} \mathbf{F}_\delta \text{ (with } c_\delta = 0) \quad \subset \quad M_2, \tag{8.6}$$

where $\varepsilon > 0$ is fixed arbitrarily small, i.e. there is no complete characterization of $\mathbf{F}_\delta$ in terms of Morrey spaces. See discussion in Section 3.

Comparing (8.5) and (8.6), one sees that one gains quite a lot in admissible singularities of b by passing from form-bounded drifts to multiplicatively form-bounded drifts. Of course, this comes at expense of imposing conditions on $\operatorname{div} b$.

Theorem 8.3. *The following are true:*

(i) (Classical martingale solutions [KS2]) *If*

$$|b|^{\frac{1+\nu}{2}} \in \mathbf{F}_\delta \quad \nu \in]0,1], \quad \delta < \infty,$$

and

$$(\operatorname{div} b)_+^{1/2} \in \mathbf{F}_{\delta_+}, \quad \delta_+ < 4, \qquad (\operatorname{div} b)_- \in L^1 + L^\infty, \tag{8.7}$$

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

$$X_t = x - \int_0^t b(X_s) ds + \sqrt{2} B_t$$

has a martingale solution $\mathbb{P}_x$.

(ii) (Approximation uniqueness and Markov property [K6]) *If, in addition to the assumptions of (i), $b \in \mathbf{MF}_\delta$ for some $\delta < \infty$, then there exists $0 < \gamma < 1$ such that, regardless of the choice of $\{b_n\} \in [b]$ (defined in the same way as in Section 3.2, i.e. to preserve the structure constants of b), provided that $\{b_n\}$ additionally satisfies*

$$b_n \to b \quad \text{in } [L^{1+\gamma}]^d,$$

we have convergence

$$\mathbb{P}_x^n \to \mathbb{P}_x \quad \text{weakly in } \mathcal{P}(\mathbf{C}),$$

of the martingale solutions $\{\mathbb{P}_x^n\}$ to the approximating SDEs

$$X_t^n = x - \int_0^t b_n(X_r^n)dr + \sqrt{2}B_t. \tag{8.8}$$

Furthermore, $\{\mathbb{P}_x\}_{x \in \mathbb{R}^d}$ is a Markov family.

(iii) (Feller semigroup) *In the assumptions of (ii),*

$$T_t f(x) := \mathbb{E}_{\mathbb{P}_x}[f(\omega_t)], \quad f \in C_\infty$$

is a strongly continuous Feller semigroup on C_∞, say, $T_t =: e^{-t\Lambda}$, where the generator Λ is thus appropriate operator realization of the formal operator $-\Delta + b \cdot \nabla$ in C_∞.

The novelty is in assertion (iii). Its proof uses the Trotter approximation theorem in the same way as the proof of Theorem 5.1.

Remark 8.1 (L^2 vs $L^{1+\gamma}$, $\gamma < 1$)**.** If, in the setting of Theorem 8.3(i), we additionally require $b \in [L^2_{\rm loc}(\mathbb{R}^d)]^d$, then it is also possible to prove a.e. approximation uniqueness. The last condition is actually satisfied if time-inhomogeneous b is a Leray-Hopf solution of the 3D Navier-Stokes equations, i.e. then one has $b \in L^\infty([0,1],[L^2_{\rm loc}(\mathbb{R}^d)]^d)$. This was explored by a number of authors, see the next Remark 8.2. There are, however, other classes of solutions to 3D N-S equations that are not uniformly in t square integrable, such as the critical class (3.5) of Koch and Tataru. So, we are interested in finding different additional conditions on b that do not require square integrability, but still allow us to prove, among other results, the approximation uniqueness. This is the condition $b \in \mathbf{MF}_\delta$ in Theorem 8.3(ii).

The proof of the approximation uniqueness in Theorem 8.3(ii) uses an $L^{\frac{1+\gamma}{\gamma}}(\mathbb{R}^d)$ gradient bound on solutions of the corresponding elliptic Kolmogorov equation (Lemma 10.3). It is proved by means of the Gehring-Giaquinta-Modica's lemma (Lemma 20.1), so γ can be estimated explicitly, see Remark 20.1.

To use Gehring-Giaquinta-Modica's lemma, we need Caccioppoli's inequality. The proof of Caccioppoli's inequality for multiplicatively form-bounded drifts employs an extra iteration procedure ("Caccioppoli's iterations") which was introduced in our previous paper [KV] to study regularity of solutions of Dirichlet problem for the drift-diffusion equation. Namely, for $v = (u-k)_+$ and cutoff function η one has

$$\begin{aligned}\langle b \cdot \nabla u, \eta v\rangle &= \frac{1}{2}\langle b \cdot \nabla v^2, \eta\rangle \\ &= -\frac{1}{2}\langle b \cdot \nabla \eta, v^2\rangle \\ &\leq \frac{1}{2}\langle |b|, \psi^2\rangle, \quad \text{where } \psi := \sqrt{|\nabla \eta|}v.\end{aligned}$$

By $b \in \mathbf{MF}_\delta$ (for simplicity, take $c_\delta = 0$),

$$\begin{aligned}\langle |b|, \psi^2 \rangle &\leq \delta \|\nabla (v\sqrt{|\nabla \eta|})\|_2 \|v\sqrt{|\nabla \eta|}\|_2 \\ &\leq \delta \bigg(\|(\nabla v)\sqrt{|\nabla \eta|}\|_2 + \|v \nabla \sqrt{|\nabla \eta|}\|_2 \bigg) \|v\sqrt{|\nabla \eta|}\|_2,\end{aligned}$$

so

$$\langle |b|, \psi^2 \rangle \leq \frac{C_1}{r_2 - r_1} \|(\nabla v)\mathbf{1}_{B_{r_2}}\|_2 \|v\mathbf{1}_{B_{r_2}}\|_2 + \frac{C_1}{(r_2 - r_1)^2} \|v\mathbf{1}_{B_{r_2}}\|_2^2,$$

provided that η is equal to 1 on B_{r_1}, is zero outside of B_{r_2}, and its derivatives satisfy appropriate estimates. The first term in the RHS contains both ∇v and the indicator function of the ball of larger radius, so we cannot simply apply Cauchy-Schwarz' inequality to obtain the Caccioppoli inequality. But it is possible to arrive at the Caccioppoli inequality by iterating over a sequence of intermediate balls with radii between r_1 and r_2.

Remark 8.2 (Critical divergence, super-critical drift)**.** Some results for the operator $-\Delta + b \cdot \nabla$ depend only on $\operatorname{div} b$. For instance, if $(\operatorname{div} b)_+^{1/2} \in \mathbf{F}_{\delta_+}$, $\delta_+ < 4$, then the solution v to the Kolmogorov backward equation

$$(\partial_t - \Delta + b \cdot \nabla)v = 0, \quad v|_{t=0} = v_0,$$

satisfies, for all $\frac{2}{2-\sqrt{\delta_+}} < q \leq p \leq \infty$, the dispersion estimate

$$\|v(t)\|_p \leq Ce^{\omega t} t^{-\frac{d}{2}(\frac{1}{p} - \frac{1}{q})} \|v_0\|_q, \quad t > 0.$$

The proof is due to J. Nash, see [KS3] for details. While one needs smoothness and boundeness of b and v_0 to carry out integration by parts, the constants C and ω depend only on d and δ_+. They and do not depend on any integral characteristics of b.

However, to obtain more detailed information about the diffusion process with drift b, one must impose some conditions on b. For instance, in Theorem 8.1 we also required $b \in \mathbf{F}_\delta$.

Between these two types of assumptions, there are intermediate conditions, such as in assertion (i) of Theorem 8.3. In this assertion, selecting ν close to zero, one can treat b that can be essentially twice more singular than the vector fields in $\mathbf{F}_\delta$. This is a super-critical condition on the drift in the sense of scaling. Let us recall the sub-critical/critical/super-critical classification of the spaces of vector fields. Given a vector field b, put $b_\lambda(x) := \lambda b(\lambda x)$. Let Y be a translation-invariant Banach space of distribution-valued vector fields b such that

$$\|b_\lambda\|_Y = \lambda^a \|b\|_Y.$$

Now,

$a > 0$, then Y is sub-critical, i.e. passing to the small scales decreases the norm,

$a = 0$, then Y is critical,

$a < 0$, then Y is super-critical.

In the last two cases "zooming in" does not change the norm or makes the norm larger. For example, L^p is sub-critical, critical or super-critical according to whether $p > d$, $p = d$ or $p < d$. This classification is widely used in the study of Navier-Stokes equations. There one applies it to spaces of solutions or initial data.

The super-critical condition on b appearing in assertion (i) was introduced in the work of Q.S. Zhang [Za2]. He considered the time-inhomogeneous counterpart of $|b|^{\frac{1+\nu}{2}} \in \mathbf{F}_\delta$, namely, $b \in [L^{1+\nu}_{\rm loc}(\mathbb{R}^{1+d})]^d$ and for a.e. $t \in \mathbb{R}$

$$\langle |b(t,\cdot)|^{1+\nu}\varphi,\varphi\rangle \leq \delta\|\nabla\varphi\|_2^2 + g_\delta(t)\|\varphi\|_2^2 \qquad \forall\, \varphi \in C_c^\infty(\mathbb{R}^d) \tag{8.9}$$

where $0 \leq g_\delta \in L^1_{\rm loc}(\mathbb{R})$ describes how irregular b can be in time. He established, among other results, local boundedness of any weak solution to the parabolic equation

$$(\partial_t - \Delta + b\cdot\nabla)v = 0 \text{ on } \mathbb{R}^{1+d},$$

provided that $\operatorname{div} b \leq 0$ and

$$b \in [L^2_{\rm loc}(\mathbb{R}^{1+d})]^d. \tag{8.10}$$

The last condition is satisfied if b is taken to be a Leray-Hopf solution of 3D Navier-Stokes equations, which motivated [Za2].

The proof of Theorem 8.3(i) uses a tightness estimate for solutions of the approximating SDEs with bounded smooth drifts (cf. (8.16)). The proof of that estimate, in turn, uses the idea from [Za2] for handling cutoff functions in presence of b satisfying (8.9).

The first result on SDEs with suprcritical divergence-free drifts belongs to X. Zhang and G. Zhao [ZZ1]. They considered

$$X_t = x - \int_s^t b(r,X_r)dr + \sqrt{2}(B_t - B_s), \quad t \geq s, \tag{8.11}$$

with divergence-free drift b additionally satisfying the square integrability condition (8.10), and included in the super-critical Ladyzhenskaya-Prodi-Serrin condition

$$|b| \in L^q([0,T],L^p(\mathbb{R}^d)), \quad p,q \geq 2, \quad \frac{d}{p}+\frac{2}{q} < 2. \tag{SLPS}$$

They proved that for every initial data $(s,x) \in \mathbb{R}^{1+d}$ the SDE (8.11) has a weak solution satisfying a Krylov type estimate. Moreover, using hypothesis (8.10), they proved that outside a measure zero set of (s,x) one has approximation uniqueness and a.s. Markov property for these weak solutions. Consequently, the weak well-posedness result of [ZZ1] justifies the passive tracer model in the Leray-Hopf setting under the a priori assumption (SLPS). They also allow the positive part $(\operatorname{div} b)_+$ of the divergence of b to be singular, provided it satisfies condition (SLPS) (possibly with different exponents p, q). In recent paper [HZ], Z. Hao and X. Zhang extended the results in [ZZ1] to divergence-free super-critical distributional drifts.

Let us add that condition (8.10) is quite powerful. For instance, if one assumes only $\operatorname{div} b \leq 0$ and (8.10), then it is already sufficient to prove weak uniqueness results for the backward Kolmogorov equation in L^1 [GS].

It is easy to show, using Hölder's inequality, that (SLPS) is a subclass of (8.9). It is a proper subclass. Indeed, (8.9) contains some vector fields having strong hypersurface singularities that are not covered by (SLPS).

The main focus of Theorem 8.3(i) was reaching the blow up threshold for δ_+ for $(\operatorname{div} b)_+$. As a by-product, it also closes the gap (at the level of weak existence for SDE (8.11)) between the hypotheses on the drift in [ZZ1] and in [Za2].

There remains nontrivial (as it seems to us) work left to establish weak existence for SDEs whose drifts lie in an even larger class of super-critical divergence-free drifts, namely, those considered by Q.S. Zhang in [Za3]:

$$\big\langle |b(t)| \log(1+|b(t)|)^2\varphi,\varphi\big\rangle \leq \delta\|\varphi\|_2^2 + g_\delta(t)\|\varphi\|_2^2 \qquad \forall\, \varphi \in C_c^\infty(\mathbb{R}^d) \quad \text{for a.e. } t \in \mathbb{R}. \tag{8.12}$$

Remark 8.3 (Heat kernel bounds)**.** Although in the results mentioned so far it is the singular positive part of $\operatorname{div} b$ that presents an obstacle to the well-posedness of the SDE, something nice that can be said about the case of positive divergence. Namely, assume that $a \in H_\xi$, i.e. we have a bounded symmetric uniformly elliptic matrix field. Let

$$b \in \mathbf{F}_\delta, \quad \delta < 4\xi^2,$$

and

$$\operatorname{div} b \geq 0.$$

Then the heat kernel $p(t,x,y)$ of $-\nabla\cdot a\cdot\nabla + b\cdot\nabla$, defined as the integral kernel of the corresponding C_0 semigroup in L^p, $p > \frac{2}{2-\nu^{-1}\sqrt{\delta}}$, constructed via a suitable regularization of a and b, satisfies, possibly after a modification on a measure zero set, the Gaussian lower bound

$$c_1\Gamma_{c_2}(t, x-y)e^{-c_3 t} \leq p(t,x,y), \tag{8.13}$$

where $\Gamma_c(t,x) := (4\pi ct)^{-\frac{d}{2}} e^{-\frac{|x|^2}{ct}}$ and $c_1, c_2 > 0$, $c_3 \geq 0$, see [KS4].

Under the above assumptions on b there is no Gaussian upper bound on $p(t,x,y)$. In fact, the counterexample is given by the Brownian particles considered in Example 1.1, see the end of Section 4. Consequently, the proof of (8.13) does not reply on a Gaussian upper bound which, to the best of our knowledge, is the first result of this type.

Remark 8.4 (More on the case $\operatorname{div} b = 0$)**.** 1. Let $b = b(x)$. In the case $\operatorname{div} b = 0$, there is an alternative approach to the proof of the approximation uniqueness for $b \in \mathbf{MF}_\delta$. Namely, Mazya-Verbitsky [MV, Theorem 5] proved equivalence

$$|\langle b\varphi, \varphi\rangle| \leq \delta\|\nabla\varphi\|_2\|\varphi\|_2 \quad \forall\, \varphi \in C_c^\infty \quad \Leftrightarrow \quad b = \nabla Q \text{ for some } \quad Q \in [\mathrm{BMO}]^{d\times d}, \tag{8.14}$$

where the LHS is, clearly, more general that $b \in \mathbf{MF}_\delta$. So, one can use this representation for divergence-free b, put the anti-symmetric matrix Q in the diffusion coefficients, and then prove uniqueness of the weak solution to Cauchy problem for the Kolmogorov parabolic equation by working in the standard Hilbert triple $W^{1,2} \hookrightarrow L^2 \hookrightarrow W^{-1,2}$, see [QX].

2. Assuming that $b \in \mathbf{MF}_\delta$, $\operatorname{div} b = 0$, Semënov [S] proved two-sided Gaussian bounds

$$C_1\Gamma_{c_2}(t-s, x-y) \leq p(t,s,x,y) \leq C_3\Gamma_{c_4}(t-s, x-y), \tag{8.15}$$

where $p(t,s,x,y)$ is the heat kernel of the Kolmogorov operator $-\nabla\cdot a\cdot\nabla + b\cdot\nabla$ with measurable symmetric uniformly elliptic a. He used Moser's method to prove the upper bound. His proof of the lower bound is based on a substantial modification of Nash's method. Next, Qian-Xi [QX] established two-sided Gaussian bounds for all $q = \nabla Q \in \mathbf{BMO}^{-1}$. The approach of [S] can be extended to the drifts having singular divergence, and allows to obtain lower and upper Gaussian bounds that hold either independently or simulatneously, cf. Remark 8.3.

Having two-sided Gaussian bounds greatly simplifies the analysis of the corresponding diffusion process, e.g. one obtains right away the Feller propagator, the continuity of trajectories follows easily from the Kolmogorov continuity criterion. Let us show how the tightness of the martingale solutions $\mathbb{P}_x^n$ of the approximating SDEs (8.8) follows from the upper Gaussian bound (our underlying goal here is to demonstrate how natural condition $b \in \mathbf{MF}_\delta$ is). The tightness argument requires, as its point of departure, the estimate

$$\mathbb{E}_{\mathbb{P}^n_x} \int_{t_0}^{t_0+\varepsilon} |b_n(\omega_r)| dr \leq H(\varepsilon), \quad 0 \leq t_0 \leq 1, 0 < \varepsilon < 1, \tag{8.16}$$

for a continuous function H such that $H(\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$. By Itô's formula, $\mathbb{E}_{\mathbb{P}^n_x} \int_{t_0}^{t_1} |b_n(\omega_r)| dr = u_n(t_0, X^n_{t_0})$, where u_n solves $\partial_t u_n + \Delta u_n - b_n \cdot \nabla u_n + |b_n| = 0$, $u_n(t_1) = 0$. After reversing the direction of time, we can deal instead with the initial-value problem

$$(\partial_t - \Delta + b_n \cdot \nabla) v_n = |b_n|, \quad v_n|_{t=0} = 0. \tag{8.17}$$

(with some abuse of notation, we continue denoting the drift by b_n). By the Duhamel formula,

$$v(t,x) = \int_0^t \langle p(t,s,x,\cdot) |b_n(\cdot)| \rangle ds,$$

so, applying the upper Gaussian bound on the heat kernel $p = p_n$, we obtain (put $\Gamma_{t,x} := \Gamma_{c_4}(t-s, x-\cdot)$):

$$\begin{aligned}
|v(t,x)| &\leq C_3 \int_0^t \langle |b| \sqrt{\Gamma_{t,x}}, \sqrt{\Gamma_{t,x}} \rangle ds \\
&\quad \text{(we use } b \in \mathbf{MF}_\delta) \\
&\leq C_3 \int_0^t \Big(\delta \|\nabla \sqrt{\Gamma_{t,x}}\|_2 \|\sqrt{\Gamma_{t,x}}\|_2 + c_\delta \|\sqrt{\Gamma_{t,x}}\|_2^2 \Big) ds \\
&= C_3 \int_0^t \Big(\delta \|\nabla \sqrt{\Gamma_{t,x}}\|_2 + c_\delta \Big) ds \\
&= C_3 \int_0^t \Big(\delta \sqrt{\frac{d}{8c_4} \frac{1}{t-s}} + c_\delta \Big) ds,
\end{aligned} \tag{8.18}$$

which yields (8.16) for $H(t) = C\sqrt{t} + c_\delta t$.

The above argument works for time-inhomogeneous multiplicatively form-bounded drifts: $|b| \in L^1_{\rm loc}(\mathbb{R}^{1+d})$ and for a.e. $t \in \mathbb{R}$

$$\langle |b(t,\cdot)| \varphi, \varphi \rangle \leq \delta \|\nabla \varphi\|_2 \|\varphi\|_2 + g_\delta(t) \|\varphi\|_2^2 \qquad \forall\, \varphi \in C_c^\infty, \tag{8.19}$$

where $g_\delta \geq 0$ (measuring the singularity of b in time) is assumed to be in the weak L^2 class on $\mathbb{R}$, that is, $\int_{t_0}^{t_1} g(s) ds \leq c\sqrt{t_1 - t_0}$ with constant c independent of t_0, $t_1 \in \mathbb{R}$.

Let us note that the upper Gaussian bound that we assumed above is valid e.g. if $b \in \mathbf{MF}_\delta$ and ${\rm div}\, b \leq 0$ or, more generally, $({\rm div}\, b)_+$ is in the Kato class, see [KS4] regarding the time-homogeneous case and [KS5] regarding the time-inhomogeneous case (8.19).

We conclude by returning to super-critical drifts. Q.S. Zhang [Za2] and Qian-Xi [QX2] also obtained non-Gaussian upper bounds on the heat kernel of $-\Delta + b \cdot \nabla$ with supercritical b. It would be interesting to "test" the optimality of these bounds and try to dedeuce the tightness estimate of [ZZ1] from them, in the same way as it was done above for $b \in \mathbf{MF}_\delta$ and the Gaussian upper bound. Of course, the calculations become more complicated due to the lack of a scaling invariance in these bounds.

9. Further remarks

Remark 9.1 (De Giorgi's method in L^p)**.** This remark concerns the proofs of Theorems 5.1, 5.2, 6.1 and 8.1. We run De Giorgi's method in L^p for $p > \frac{2}{2-\sqrt{\delta}}$, and so we need $\delta < 4$. The condition

$$p > \frac{2}{2-\sqrt{\delta}}$$

comes from the following elementary calculation for the Kolmogorov backward equation. Let $b \in \mathbf{F}_\delta$ and $q \in \mathbf{BMO}^{-1}$ be additionally bounded and smooth, and let us assume for simplicity that $c_\delta = 0$ (otherwise we need to add a constant term in the Kolmogorov equation to absorb $c_\delta > 0$). Consider the Cauchy problem $(\partial_t - \Delta + (b+q)\cdot\nabla)v = 0$, $v|_{t=0} = v_0 \in C_c^\infty$. Without loss of generality, $v_0 \geq 0$, and so $v \geq 0$. Multiply equation by v^{p-1} and integrate by parts. Since $\operatorname{div} q = 0$, one finds

$$\frac{1}{p}\langle \partial_t v^p\rangle + \frac{4(p-1)}{p^2}\langle |\nabla v^{\frac{p}{2}}|^2\rangle + \frac{2}{p}\langle b\cdot\nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\rangle = 0,$$

or, equivalently,

$$\partial_t \langle v^p\rangle + \frac{4(p-1)}{p}\langle |\nabla v^{\frac{p}{2}}|^2\rangle = -2\langle b\cdot\nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\rangle.$$

Applying the Cauchy-Schwarz inequality in the last term gives

$$\langle \partial_t v^p\rangle + \frac{4(p-1)}{p}\langle |\nabla v^{\frac{p}{2}}|^2\rangle \leq 2\biggl(\alpha\langle |b|^2, v^p\rangle + \frac{1}{4\alpha}\langle |\nabla v^{\frac{p}{2}}|^2\rangle\biggr).$$

Now, applying $b \in \mathbf{F}_\delta$ and selecting $\alpha = \frac{1}{2\sqrt{\delta}}$, we obtain

$$\langle \partial_t v^p\rangle + \left[\frac{4(p-1)}{p} - 2\sqrt{\delta}\right]\langle |\nabla v^{\frac{p}{2}}|^2\rangle \leq 0.$$

To keep the dispersion term positive, one needs $\frac{4(p-1)}{p} - 2\sqrt{\delta} > 0$, i.e. $p > \frac{2}{2-\sqrt{\delta}}$. Hence we need $\delta < 4$. This calculation reappears (in slightly different form, e.g. for sub-solutions of the Kolmogorov equation) in the proof of Theorem 5.1. Counterexamples discussed in the introduction show that $\delta < 4$ is sharp at least in high dimensions.

The observation that one should work in L^p, $p > \frac{2}{2-\sqrt{\delta}}$, was made already in [KoS] in the context of the L^p semigroup theory of $-\Delta + b\cdot\nabla$, $b \in \mathbf{F}_\delta$. See also [CPZ, JL] where the authors use energy methods in L^p to study regularity of solutions of the Keller-Segel equation, although the optimal choice of p is not really discussed in these papers.

The point of departure for De Giorgi's method is the Caccioppoli inequality. For illustration purposes, assume that $b \in \mathbf{F}_\delta$ has form-bound $\delta < 1$, and so we can take $p = 2$. Also, let $q = 0$. Then Proposition 17.1 gives: for every $k \in \mathbb{R}$, the positive part $w = (u-k)_+$ of $u - k$, where $(\mu - \Delta + b\cdot\nabla)u = f$, satisfies

$$\langle |\nabla w|^2, \eta^2\rangle \leq K_1\langle w^2, |\nabla\eta|^2\rangle + K_2\langle (f-\mu w)^2, \eta^2\rangle, \tag{9.1}$$

where η is a smooth cutoff function, constant K depends only on d and δ, not on k. De Giorgi's method turns this Caccioppoli's inequality into the Hölder continuity of u. (One can draw a loose parallel between Caccioppoli's inequality and the form-boundedness of the gradient ∇w, in which case De Giorgi's method becomes an embedding theorem from f to u into the space of Hölder continuous functions. This is admittedly speculative, but we can to some extent justify this by referring to a recent result of Krylov in [Kr3] on the regularity theory of inhomogeneous parabolic equations in $\mathbb{R}^{d+1}$ with drift, solution and its derivatives in appropriate Morrey classes, so to establish Hölder continuity of solutions he appeals to the Campanato embedding theorem rather than the Sobolev embedding theorem.)

Remark 9.2 (On the choice of test function)**.** It is interesting to abstract away the calculation in the beginning of the previous remark to see if there are other test functions that allow for a similar energy analysis of Cauchy problem $(\partial_t - \Delta + b \cdot \nabla)v = 0$, $v|_{t=0} = v_0$, where $0 \leq v_0 \in C_c^\infty$. We seek test functions $\varphi(v) = \varphi(v(t))$ such that $\varphi \geq 0$, $\varphi' \geq 0$ and there exists a "conjugate" function ψ satisfying $\psi(0) = 0$ and

$$\psi' = \sqrt{\varphi'}, \quad \varphi = a\psi\psi' \tag{9.2}$$

for some constant $a > 0$ to be chosen (hence $\psi \geq 0$). Multiplying the parabolic equation by $\varphi(v)$ and integrating by parts, we obtain

$$\begin{aligned}\langle \partial_t v, \varphi(v)\rangle &= a\langle \partial_t v, \psi(v)\psi'(v)\rangle = \frac{a}{2}\langle \partial_t(\psi(v))^2\rangle, \\ \langle -\Delta v, \varphi(v)\rangle &= \langle \nabla v, \varphi'(v)\nabla v\rangle = \langle |\nabla \psi(v)|^2\rangle,\end{aligned}$$

and

$$\begin{aligned}\langle b\cdot\nabla v, \varphi(v)\rangle &= \langle b \cdot \nabla\psi(v), a\psi(v)\rangle \\ &\leq a\biggl(\alpha\langle |b|^2, (\psi(v))^2\rangle + \frac{1}{4\alpha}\langle |\nabla\psi(v)|^2\rangle\biggr) \\ &\text{(take } \alpha = 1/2\sqrt{\delta} \text{ and apply } b \in \mathbf{F}_\delta \text{ (with } c_\delta = 0)) \\ &\leq a\sqrt{\delta}\langle |\nabla\psi(v)|^2\rangle.\end{aligned}$$

Thus, we obtain an energy inequality of the form

$$\frac{a}{2}\partial_t\langle (\psi(v))^2\rangle + \bigl(1 - a\sqrt{\delta}\bigr)\langle |\nabla\psi(v)|^2\rangle \leq 0.$$

In order for the disperson term to remain non-negative, we need to take in (9.2) $a = \frac{1}{\sqrt{\delta}}$ (or smaller, but the equality is least restrictive on φ). So far, no constraint on δ has appeared, but it will appear once we solve the resulting (from (9.2)) system

$$\begin{cases} \psi' = \sqrt{\varphi'}, \\ \varphi = \frac{\psi\psi'}{\sqrt{\delta}} \end{cases} \qquad \Rightarrow \qquad \varphi' = \frac{1}{\sqrt{\delta}}(\psi\psi'' + (\psi')^2),$$

so that ψ must satisfy

$$\psi\psi'' = (\sqrt{\delta} - 1)(\psi')^2.$$

Assuming $\psi > 0$ for $v > 0$, we can reduce order to obtain

$$\psi' = C\psi^{\sqrt{\delta}-1}. \tag{9.3}$$

We look for non-trivial solutions ψ that are defined globally. Hence the right-hand side of (9.3) must grow at most linearly, which forces

$$\sqrt{\delta} - 1 \leq 1 \quad \Rightarrow \quad \delta \leq 4.$$

1. If $-1 < \sqrt{\delta} - 1 < 1$, i.e. $0 < \delta < 4$, then we find $\psi(v) = C_1 v^{\frac{1}{2-\sqrt{\delta}}}$, among other possible solutions of this type (note that there is no uniqueness here). Then, using equation $\psi' = \sqrt{\varphi'}$, we find

$$\varphi(v) = cv^{\frac{2}{2-\sqrt{\delta}}-1} = cv^{p-1}, \quad p := \frac{2}{2-\sqrt{\delta}},$$

covering the classical power-type test function discussed in the beginning of this remark.

2. If $\delta = 4$, then ODE (9.3) becomes linear. Taking, for instance, $C = \frac{1}{2}$ in (9.3) yields the non-trivial solution

$$\psi(v) = 2e^{\frac{v}{2}}, \quad \varphi(v) = e^v.$$

This is essentially the test function used to treat the critical threshold $\delta = 4$, see Remark 5.1. Although $\psi(0) \neq 0$, which fails the original requirement $\psi(0) = 0$ needed above to apply the form-boundedness of b, one can overcome this by working with test functions $\varphi(v) = e^v - 1$ or $\varphi(v) = e^v - e^{-v}$ multiplied by a cutoff function, or working on torus instead of $\mathbb{R}^d$ [K3, KS5].

A moment of reflection, after inspecting the test function $\varphi(v) = e^v - 1$, suggests that the the theory of the Kolmogorov equation in the critical regime $\delta = 4$ in Remark 5.1 should be viewed as the limit $p \to \infty$ of the *asymptotic* L^p theory, namely, with the non-standard L^p test function

$$\varphi(v) = \left(1 + \frac{v}{p}\right)^p - 1.$$

Accordingly, one needs to study solutions v of the Kolmogorov backward equation *around* 1. We will address this point in subsequent paper.

Remark 9.3 (On time-inhomogeneous drifts)**.** Let us outline an alternative to Theorem 5.1, earlier approach to constructing Feller semigroup for $-\Delta + b \cdot \nabla$, $b \in \mathbf{F}_\delta$ [KoS, K2, KS8]. Using gradient bounds in (5.10) (or, rather, a bound obtained by interpolating between the first and the third terms in the left-hand side of (5.10)) and running parabolic Moser's iterations for solutions v_n of $(\partial_t - \Delta + b_n \cdot \nabla)v_n$, $v_n|_{t=0} = f$, one obtains the following inequality with constant $C = C(d, T, \|f\|_\infty, \|\nabla f\|_{W^{1,r}})$ independent of n, m:

$$\|v_n - v_m\|_{L^\infty([0,T]\times\mathbb{R}^d)} \leq C\|v_n - v_m\|^\gamma_{L^2([0,T]\times\mathbb{R}^d))}, \quad n, m \geq 1, \tag{9.4}$$

for certain fixed r. Crucially, γ does not depend on n, m and is strictly positive. It is relatively easy to show that the right-hand side of (9.4) converges to zero as $n, m \to \infty$.

In this way, (5.10) allows to constructs the Feller semigroup $e^{-t\Lambda}f := \lim_n v_n(t)$ for $\Lambda \supset -\Delta + b \cdot \nabla$ directly. See [K2], see also recent improvements in [KS8]; these works extended and simplified the pioneer work of Kovalenko-Semënov [KoS] that went via elliptic arguments.

Compared to the present paper, there is an important disadvantage to this construction, namely, the use of gradient bounds (5.10) introduces a constraint on form-bound δ of b of the form

$$\delta < \frac{c}{d^2}, \quad c \ll 1 \quad \text{(compare this with } \delta < 4 \text{ in Theorem 5.1)},$$

so e.g. no applications to particle systems. There are, however, some advantages to this approach:

1) Estimate (9.4) opens up a way for obtaining quantitative estimates on the rate of convergence of the approximating Feller semigroups, as long as one can estimate the rate of convergence in L^2 which is, in principle, a much simpler task.

   In contrast, the proof of Theorem 5.1(i) uses compactness arguments, so there is no way to control the rate of convergence.

2) It does not use any elliptic results such as Trotter's approximation theorem (Theorem 10.1), and works for time-inhomogeneous form-bounded drifts: for a.e. $t \geq 0$,

$$\langle |b(t)|^2, \varphi^2 \rangle \leq \delta \|\nabla \varphi\|_2^2 + g(t)\|\varphi\|_2^2, \quad \varphi \in W^{1,2}(\mathbb{R}^d),$$

for a function $0 \leq g = g_\delta \in L^1_{\rm loc}(\mathbb{R})$ that determines how irregular in time b is. (For instance, Ladyzhenskaya-Prodi-Serrin class

$$|b| \in L^q([0,T], L^p(\mathbb{R}^d)), \quad \frac{d}{p} + \frac{2}{q} \leq 1$$

satisfies the above condition.) One thus obtains the corresponding to $-\Delta + b(t,x) \cdot \nabla$ Feller propagator on C_∞.

In view of the counterexamples (a), (b) in the introduction we are rather satisfied with condition $\delta < 4$ in Theorems 5.1 and 5.2(i). However, in the context of gradient bounds (as in Theorem 5.2($viii$)), the question of what is the optimal condition on δ is still far from settled.

10. Proof of Theorem 5.1

Let us write, to shorten notations,

$$\Lambda_{n,m} := \Lambda(b_n, q_m) \equiv -\Delta + (b_n + q_m) \cdot \nabla, \quad D(\Lambda_{n,m}) = (1-\Delta)^{-1} C_\infty.$$

(***i***) This assertion will follow from the Trotter approximation theorem (see e.g. [Ka, IX.2.5]). Applied to contraction semigroups $\{e^{-t\Lambda_{n,m}}\}_{n,m \geq 1}$ in C_∞, this theorem is stated as follows:

Theorem 10.1 (Trotter's approximation theorem)**.** *Assume that there exists constant $\mu_0 > 0$ independent of n, m such that*

1°) $\sup_{n,m \geq 1} \|(\mu + \Lambda_{n,m})^{-1} f\|_\infty \leq \mu^{-1} \|f\|_\infty$, $\mu \geq \mu_0$.

2°) *there exists s-C_∞-$\lim_n \lim_m (\mu + \Lambda_{n,m})^{-1}$ for some $\mu \geq \mu_0$.*

3°) *$\mu(\mu + \Lambda_{n,m})^{-1} \to 1$ in C_∞ as $\mu \uparrow \infty$ uniformly in n, m.*

Then there exists a contraction strongly continuous semigroup $e^{-t\Lambda}$ on C_∞ such that

$$e^{-t\Lambda} = s\text{-}C_\infty\text{-}\lim_n \lim_m e^{-t\Lambda_{n,m}}$$

locally uniformly in $t \geq 0$.

10.0.1. *Key PDE results.*

Proposition 10.1 (Embedding property)**.** *Let $w_{n,m}$ is the classical solution to elliptic equation*

$$\big(\mu - \Delta + (b_n + q_m) \cdot \nabla\big) w_{n,m} = (b_n^i + q_m^i) f, \quad f \in \mathcal{S}. \tag{10.1}$$

where we have fixed $1 \leq i \leq d$ and have denoted $q_m^i = \sum_{j=1}^d \nabla_j Q_m^{ij}$. Put

$$A_m := |Q_m^i||\nabla f| + (1 + |Q_m^i|)|f| \quad \text{where } Q_m^i \text{ denotes the } i\text{-th row of } Q_m.$$

Then, for every $p > \frac{2}{2-\sqrt{\delta}}$, $p \geq 2$ and $1 < \theta < \frac{d}{d-2}$, there exist constants $\mu_0 > 0$, $0 < \beta < 1$ and K_j, $j = 1, 2$, independent of n, m, such that

$$\begin{aligned} \|w_{n,m}\|_\infty &\leq K_1 (\mu - \mu_0)^{-\frac{\beta}{p}} \|A_m\|_{p\theta'} \\ &\quad + K_2 (\mu - \mu_0)^{-\frac{1}{p\theta}} \|A_m\|_{p\theta}, \end{aligned} \tag{10.2}$$

for all $\mu > \mu_0$.

We prove Proposition 10.1 in Section 16.

Set $u_{n,m}$ be the classical solutions to elliptic equation

$$\big(\mu - \Delta + (b_n + q_m) \cdot \nabla\big) u_{n,m} = f, \quad \mu > 0. \tag{10.3}$$

Proposition 10.2 (A priori Hölder continuity)**.** *For every $\mu > 0$, $\{u_{n,m}\}$ are locally Hölder continuous uniformly in n, m.*

That is, $u_{n,m}$ are Hölder continuous, in every unit ball, with constants that do not depend on n, m or the center of the ball. These constants are, however, allowed to depend on $\|f\|_\infty$. We prove Proposition 10.2 in Section 17.

Proposition 10.3 (Convergence)**.** *There exists $\mu_0 > 0$ such that for every $\mu \geq \mu_0$, for all $p > \frac{2}{2-\sqrt{\delta}}$, $p \geq 2$, and any $x \in \mathbb{R}^d$, there exists the limit*

$$u := L^p_{\rho_x}\text{-}\lim_n \lim_m u_{n,m},$$

where $\rho_x(y) := \rho(y - x)$, provided that constant σ in the definition of weight ρ (this is (2.1)) is chosen sufficiently small (independently of x).

We prove Proposition 10.3 in Section 19.

Remark 10.1. The proof of Proposition 10.3 can be extended to show the existence of the limit $L^p\text{-}\lim_{n,m} u_{n,m}$, but at expense of imposing additional assumptions on drifts b or q, such as form-bound δ of b being strictly less than 1, or the stream matrix Q of q having entries in VMO. See Remark 19.2 in the end of the proof of Proposition 10.3.

Proposition 10.4 (Separation property/local maximum principle)**.** *Fix some $1 < \theta < \frac{d}{d-2}$ and $p > \frac{2}{2-\sqrt{\delta}}$, $p \geq 2$. There exists constants K, $\mu_0 > 0$ and σ (in the definition of weight ρ) independent of n, $m$such that for all $\mu \geq \mu_0$, for every $x \in \mathbb{R}^d$,*

$$\sup_{B_{\frac{1}{2}}(x)} |u_{n,m}| \leq K\bigg(\langle |f|^{p\theta} \rho_x \rangle^{\frac{1}{p\theta}} + \big\langle |f|^{p\theta'} \mathbf{1}_{B_1(x)} \big\rangle^{\frac{1}{p\theta'}}\bigg). \tag{10.4}$$

We prove Proposition 10.4 in Section 18.

We are in position to verify conditions of Trotter's theorem for $u_{n,m} = (\mu + \Lambda_{n,m})^{-1} f$:

10.0.2. *Proof of* 1°). This condition is a direct consequence of the fact that, by the classical theory, $e^{-t\Lambda_{n,m}}$ are L^∞ contractions.

10.0.3. *Proof of* 2°). By 1°), it suffices to verify the existence of the limit for all f belonging to a countable dense subset of C_c^∞. Proposition 10.2 and the Arzelà-Ascoli theorem yield: for every $r > 0$, $\{u_{n,m}\}$ is relatively compact in $C(\bar{B}_r)$. Proposition 10.3 allows to further conclude that $\{u_{n,m}\}$ converges in $C(\bar{B}_r)$, for every $r > 0$:

$$u \restriction \bar{B}_r = s\text{-}C(\bar{B}_r)\text{-}\lim_n \lim_m u_{m,n} \restriction \bar{B}_r. \tag{10.5}$$

Remark 10.2. We need Proposition 10.3 that any two partial limits of $u_{n,m}$ in $C(\bar{B}_r)$ coincide. The choice of the topology, and thus the weight, is secondary.

We need to improve (10.5) to global uniform convergence:

$$u = s\text{-}C_\infty\text{-}\lim_n \lim_m u_{n,m}. \tag{10.6}$$

To this end, we combine Proposition 10.4 and convergence (10.5). Namely, since $f \in C_c^\infty$ and weight ρ_x vanishes at infinity, it follows from (10.4) that solution $u_{n,m}$ is small uniformly in n, m when considered far away from the support of f. (Hence the name "separation property" for (10.4).)

We have verified condition 2°) of Trotter's theorem.

Remark 10.3. Using Proposition 10.4, it is easy to obtain the preservation of probability, i.e. that

$$e^{-t\Lambda(b_n,q_m)}(1-\mathbf{1}_{B_R}) \to 0 \quad \text{as } R\uparrow\infty \text{ uniformly in } n,\, m.$$

From here it follows easily that $e^{-t\Lambda}1 = 1$. (We use the fact that $e^{-t\Lambda(b_n,q_m)}$, $e^{-t\Lambda(b,q)}$ are semigroups of integral operators, so the expressions $e^{-t\Lambda(b_n,q_m)}(1-\mathbf{1}_R)$, $e^{-t\Lambda(b,q)}1$ are well-defined.)

10.0.4. *Proof of* 3°). In view of 1°), it suffices to verify 3°) on a dense subset of C_∞, e.g. C_c^∞. Fix $g \in C_c^\infty$. By the resolvent identity,

$$\begin{aligned}\mu(\mu+\Lambda_{n,m})^{-1}g - \mu(\mu-\Delta)^{-1}g &= \mu(\mu+\Lambda_{n,m})^{-1}(b_n+q_m)\cdot\nabla(\mu-\Delta)^{-1}g \\ &= (\mu+\Lambda_{n,m})^{-1}(b_n+q_m)\cdot\mu(\mu-\Delta)^{-1}\nabla g.\end{aligned}$$

Since $\mu(\mu-\Delta)^{-1}g \to g$ uniformly on $\mathbb{R}^d$ as $\mu\to\infty$, it suffices to show the convergence

$$\|(\mu+\Lambda_{n,m})^{-1}(b_n+q_m)\cdot\mu(\mu-\Delta)^{-1}\nabla g\|_\infty \to 0 \text{ as } \mu\to\infty \quad \text{uniformly in } n,\, m. \tag{10.7}$$

To that end, we apply Proposition 10.1 to $w_{n,m} := (\mu+\Lambda_{n,m})^{-1}(b_n^i+q_m^i)f$ with f taken to be

$$f := \mu(\mu-\Delta)^{-1}\nabla_i g. \tag{10.8}$$

Our goal is thus to prove

$$\|w_{n,m}\|_\infty \to 0 \quad \text{as } \mu\to\infty \text{ uniformly in } n,\, m. \tag{10.9}$$

If we can obtain bound

$$\sup_{\mu\geq 1,m}\|A_m\|_{p\theta},\ \sup_{\mu\geq 1,m}\|A_m\|_{p\theta'} < \infty \quad \text{for } f \text{ given by (10.8)}, \tag{10.10}$$

then the convergence (10.9) will follow thanks to the factors $(\mu-\mu_0)^{-\frac{\beta}{p}}$, $(\mu-\mu_0)^{-\frac{1}{p\theta}}$ in (10.2). The only slightly non-trivial aspect of proving (10.10) is that $|Q_m^i|$ can grow at infinity. But since the entries of Q_m^i are BMO functions, this growth cannot be arbitrary, see Lemma 10.2.

Proof of (10.10). So, from now on, let f be given by (10.8) where, recall, g has compact support (we will need this when we apply Lemma 10.1). In what follows, for brevity, $s=p\theta$ or $s=p\theta'$. We have

$$\begin{aligned}\||Q_m^i||f|\|_s^s &= \langle |Q_m^i|^s|\mu(\mu-\Delta)^{-1}\nabla_i g|^s\rangle \\ &\quad (\text{write } 1 = (1+|x|)^{-(d+\epsilon_0)s}(1+|x|)^{(d+\epsilon_0)s}) \\ &\leq \big\langle |Q_m^i|^{s\nu}(1+|x|)^{-(d+\epsilon_0)s\nu}\big\rangle^{\frac{1}{\nu}}\big\langle (1+|x|)^{(d+\epsilon_0)s\nu'}|\mu(\mu-\Delta)^{-1}\nabla_i g|^{s\nu'}\big\rangle^{\frac{1}{\nu'}}, \quad \nu>1.\end{aligned} \tag{10.11}$$

We will need the following elementary lemma.

Lemma 10.1. *For every $r > 0$,*

$$|(1 + |x|^r)\mu(\mu - \Delta)^{-1}\nabla_i g(x)| \leq c_R\mu(c\mu - \Delta)^{-1}|\nabla_i g|(x), \tag{10.12}$$

for some positive constants c, c_R that depend only on d, r and R, i.e. the radius of a fixed ball that contains the support of $|\nabla_i g|$.

(For reader's convenience, we prove Lemma 10.1 in Appendix A.) We will also need

Lemma 10.2 (see e.g. [Gr, Prop. 7.1.5])**.** *For every $f \in \mathrm{BMO}(\mathbb{R}^d)$, for all $1 \leq r < \infty$,*

$$\langle |f - (f)_{B_1}|^r \rho^r\rangle \leq C_{d,r,\varepsilon}\|f\|^r_{\mathrm{BMO}}, \tag{10.13}$$

where $\rho = \rho_\varepsilon$ is defined by (2.1). Hence

$$\langle |f|^r \rho^r\rangle \leq C'_{d,r,\varepsilon}\big(\|f\|^r_{\mathrm{BMO}} + \|f\mathbf{1}_{B_1}\|^r_{L^1}\big) < \infty. \tag{10.14}$$

By Lemma 10.1 (with $r := d+\epsilon_0$), we have pointwise estimate $(1+|x|)^{(d+\epsilon_0)}|\mu(\mu-\Delta)^{-1}\nabla_i g(x)| \leq c_R\mu(c\mu-\Delta)^{-1}|\nabla_i g|(x)$ with constants c_R, c independent of $\mu \geq 1$, so we can estimate in the second multiple in (10.11):

$$\big\langle (1 + |x|)^{(d+\epsilon_0)s\nu'}|\mu(\mu - \Delta)^{-1}\nabla_i g|^{s\nu'}\big\rangle \leq C\|\mu(c\mu - \Delta)^{-1}|\nabla_i g|\|^{s\nu'}_{s\nu'}.$$

In turn, by the contractivity of the heat semigroup in Lebesgue spaces, for all $\mu \geq 1$,

$$\|\mu(c\mu - \Delta)^{-1}|\nabla_i g|\|_{s\nu'} \leq c^{-1}\|\nabla_i g\|_{s\nu'}.$$

Thus, the second second multiple in (10.11) can be estimated as follows: for all $\mu \geq 1$,

$$\big\langle (1 + |x|)^{(d+\epsilon_0)s\nu'}|\mu(\mu - \Delta)^{-1}\nabla_i g|^{s\nu'}\big\rangle \leq C\|\nabla_i g\|_{s\nu'},$$

where C depends on the support of g, but does not depend on μ.

On the other hand, regarding the first multiple in (10.11), we have by Lemma 10.2 (that is, (10.14) with $r = s\nu$) after taking into account that the BMO semi-norms of Q^i_m are uniformly in m bounded,

$$\sup_m \big\langle |Q^i_m|^{s\nu}(1 + |x|)^{-(d+\epsilon_0)s\nu}\big\rangle < \infty.$$

We can thus conclude from (10.11):

$$\sup_{\mu\geq 1, m} \||Q^i_m||f|\|_s < \infty.$$

In the same way, since $\nabla f = \mu(\mu - \Delta)^{-1}\nabla_i\nabla g$ and the entries of ∇g have compact supports,

$$\sup_{\mu\geq 1, m} \||Q^i_m||\nabla f|\|_s < \infty.$$

The last two bounds yield (10.10). In view of the previous discussion, condition 3°) of Trotter's theorem is thus verified.

Now, Trotter's theorem applies and gives us assertion (i) of Theorem 5.1.

(***ii***) This is immediate from the existence of the limit in (i) and the fact that $[b]$, $[q]$ are closed with respect to passing to a sub-sequence.

(***iii***) It suffices for us to replace Proposition 10.3 with the following: for all sufficiently large μ and every $f \in \mathcal{S}$, $u_n = (\mu + \Lambda(b_n, q))^{-1}f$ converge to the same limit:

$$u_n \to u = (\mu + \Lambda(b, q))^{-1}f \quad \text{ in } L^2_\rho. \tag{10.15}$$

The rest repeats the proof of assertion (i).

Proof of (10.15). Since $b \in \mathbf{F}_\delta$, $\delta<1$ and $\operatorname{div} q=0$ provide, via Cauchy-Schwarz inequality and the compensatd compactness estimate, coercivity and boundedness of the quadratic form of $-\Delta+(b+q)\cdot\nabla$ in L^2, it is readily seen that function $u=(\mu+\Lambda(b,q))^{-1}f$, $f\in C_\infty\cap L^2$, constructed via approximation in (i), is a weak solution to elliptic equation $(\mu-\Delta+(b+q)\cdot\nabla)u=f$, i.e. $\mu\langle u,\varphi\rangle+\langle\nabla u,\nabla\varphi\rangle+\langle\nabla u,b\varphi\rangle+\langle\nabla u,q\varphi\rangle=\langle f,\varphi\rangle$ for all $\varphi\in W^{1,2}$.

Put $h_n:=u_n-u$. Our goal is to show convergence $h_n\to 0$ in L^2_ρ. Notice that, in contrast to the proof of assertion (i), we already have the limiting object u, which greatly simplifies the analysis.

Step 1. Proposition 10.4 ("separation property") extends to u, u_n and $f\in\mathcal{S}$ and yields: for a fixed $1<\theta<\frac{d}{d-2}$ there exist constants K and $\mu_0>0$ independent of n such that for all $\mu\geq\mu_0$, for every $x\in\mathbb{R}^d$,

$$\sup_{n\geq 1}\sup_{B_{\frac{1}{2}}(x)}|h_n|\leq K\biggl(\langle|f|^{2\theta}\rho_x\rangle\rangle^{\frac{1}{2\theta}}+\bigl\langle|f|^{2\theta'}\mathbf{1}_{B_1(x)}\bigr\rangle^{\frac{1}{2\theta'}}\biggr),$$

provided constant σ in the definition of ρ is fixed sufficiently small. Since $f\in\mathcal{S}$, it follows that for every $\varepsilon>0$ there exists sufficiently large $R>0$ such that

$$\sup_n\sup_{x\in\mathbb{R}^d\setminus B_R}|h_n(x)|<\varepsilon. \tag{10.16}$$

Step 2. The difference $h_n=u_n-u$ satisfies

$$\mu\langle h_n,\varphi\rangle+\langle\nabla h_n,\nabla\varphi\rangle+\langle b_n\cdot\nabla h_n,\varphi\rangle+\langle q\cdot\nabla h_n,\varphi\rangle=\langle(b-b_n)\cdot\nabla u,\varphi\rangle\qquad\text{for all }\varphi\in W^{1,2}.$$

Hence, taking $\varphi=h_n\rho$ and repeating the proof of the energy inequality of Proposition 15.1(ii) (take there $s=2$, which is possible since now b, b_n have form-bound $\delta<1$), with σ in the definition of ρ fixed sufficiently small, we obtain

$$(\mu-\mu_0)\langle|h_n|^2\rho\rangle+C_1\langle|\nabla h_n|^2\rho\rangle\leq\langle(b-b_n)\cdot\nabla u,h_n\rho\rangle,$$

where $C_1>0$ is independent of n. Thus, it suffices to show that $\langle(b-b_n)\cdot\nabla u,h_n\rho\rangle\to 0$ as $n\to\infty$.

Step 3. We represent

$$\langle(b-b_n)\cdot\nabla u,h_n\rho\rangle=\langle\mathbf{1}_{\mathbb{R}^d\setminus B_R}(b-b_n)\cdot\nabla u,h_n\rho\rangle+\langle\mathbf{1}_{B_R}(b-b_n)\cdot\nabla u,h_n\rho\rangle \tag{10.17}$$
$$=: I_1+I_2 \tag{10.18}$$

By Step 1 and $\rho\leq\sqrt{\rho}$, term I_1 can be estimated as follows: for every $\varepsilon>0$, for all $R=R(\varepsilon)>0$ sufficiently large

$$|I_1|=|\langle\mathbf{1}_{\mathbb{R}^d\setminus B_R}(b-b_n)\cdot\nabla u,h_n\rho\rangle|\leq K_b\|\nabla u\|_2\varepsilon,$$

$$K_b^2:=2\sup_n\langle|b_n|^2\rho\rangle\vee\langle|b|^2\rho\rangle<\infty\text{ by (15.15) (this is the place where we need weight }\rho),$$

where $\|\nabla u\|_2<\infty$ by the energy inequality (Proposition 15.1(i)).

Let R be as above (some some small $\varepsilon>0$). Let us deal with the second term $\langle\mathbf{1}_{B_R}(b-b_n)\cdot\nabla u,h_n\rho\rangle$ in (10.17) . Here we work over a compact set, so the weight ρ plays no role. By the energy inequality of Proposition 15.1(i), $\sup_n\|\nabla u_n\|_2<\infty$, and thus $\sup_n\|\nabla h_n\|_2<\infty$. Therefore, by the Rellich-Kondrashov theorem, there is a subsequence of $\{h_n\}$ (without loss of generality, $\{h_n\}$ itself) such that

$$h_n\restriction B_R\to g\text{ in }L^2(B_R)\text{ for some }g\in(L^2\cap L^\infty)(B_R).$$

(Here we have used a priori estimate $\|h_n\|_\infty \leq 2\mu^{-1}\|f\|_\infty$.) So,

$$|I_2| = |\langle \mathbf{1}_{B_R}(b-b_n)\cdot\nabla u, h_n\rho\rangle| \leq |\langle \mathbf{1}_{B_R}(b-b_n)\cdot\nabla u, g\rho\rangle| + |\langle \mathbf{1}_{B_R}(b-b_n)\cdot\nabla u, (h_n-g)\rho\rangle|,$$

where

$$\langle \mathbf{1}_{B_R}(b-b_n)\cdot\nabla u, g\rho\rangle \to 0 \quad \text{since } \mathbf{1}_{B_R}(\nabla u)g\rho \in L^2 \text{ and } b_n \to b \text{ weakly in } L^2,$$

and

$$\begin{aligned}|\langle \mathbf{1}_{B_R}(b-b_n)\cdot\nabla u, (h_n-g)\rho\rangle| &\leq \|\mathbf{1}_{B_R}(b-b_n)\sqrt{\rho}\|_2\|(h_n-g)\nabla u\|_2 \quad \text{(we have used } \rho \leq 1)\\ &\leq K_b\|\mathbf{1}_{B_R}\nabla u\|_{2+\varepsilon_1}\|\mathbf{1}_{B_R}(h_n-g)\|_{\frac{2(2+\varepsilon_1)}{\varepsilon_1}}, \qquad \varepsilon_1>0.\end{aligned}$$

Next, we apply

Lemma 10.3. $\|\nabla u\|_{2+\varepsilon_1} < \infty$ *provided that* $\varepsilon_1 > 0$ *is sufficiently small.*

We prove Lemma 10.3, which is of interest on its own, in Section 20. Its proof uses Gehring-Giaquinta-Modica's lemma.

In turn,

$$\|\mathbf{1}_R(h_n-g)\|_{\frac{2(2+\varepsilon_1)}{\varepsilon_1}} \to 0 \text{ as } n\to\infty$$

follows by interpolating between $\mathbf{1}_R(h_n-g) \to$ in L^2 and $\sup_n \|\mathbf{1}_R(h_n-g)\|_\infty \leq 4\mu^{-1}\|f\|_\infty < \infty$.

Combining the above estimate on I_1 and the convergence $I_2 \to 0$ as $n\to\infty$, we obtain convergence (10.15), which ends the proof of assertion (iii).

$(\boldsymbol{iv})$ Since $e^{-t\Lambda(b,q)}$ is a strongly continuous Feller semigroup on C_∞, there exist probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ on the canonical space $\mathbf{D}_T$ of càdlàg trajectories ω such that

$$e^{-t\Lambda(b,q)}f(x) = \mathbb{E}_{\mathbb{P}_x}[f(\omega_t)], \quad f\in C_\infty$$

and, for every $v \in D\big(\Lambda(b,q)\big)$, the process

$$t \mapsto v(\omega_t) - x + \int_0^t \Lambda(b,q)v(\omega_s)ds$$

is a martingale with respect to $\mathbb{P}_x$, see e.g. [RY, Ch. VII, §1]. Since both $e^{-t\Lambda(b_n,q_m)}$, $e^{-t\Lambda(b,q)}$ are strongly continuous Feller semigroups, the convergence of their finite-dimensional distributions, provided by assertion (i), yields

$$\mathbb{P}_x = w\text{-}\mathcal{P}(\mathbf{D})\text{-}\lim_n\lim_m \mathbb{P}_x^{n,m},$$

see e.g. [EK, Ch.4, Theorem 2.5]. Since $\mathbf{C}$ is closed in $\mathbf{D}$ and $\mathbb{P}_x^{n,m}(\mathbf{C}) = 1$, it follows the weak convergence that $\mathbb{P}_x(\mathbf{C}) = 1$. Thus, probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ are concentrated on $\mathbf{C}$, and we have

$$\mathbb{P}_x = w\text{-}\mathcal{P}(\mathbf{C})\text{-}\lim_n\lim_m \mathbb{P}_x^{n,m}, \tag{10.19}$$

as claimed.

$(\boldsymbol{v})$ We argue as in [HZ]. Set $\mathbb{P}_{\nu_0} := \int_{\mathbb{R}^d}\mathbb{P}_x\nu_0(x)dx$. Define in the same way $\mathbb{P}_{\nu_0}^{n,m}$. Then, in view of (10.19),

$$\mathbb{P}_{\nu_0} = w\text{-}\mathcal{P}(\mathbf{C})\text{-}\lim_n\lim_m \mathbb{P}_{\nu_0}^{n,m}.$$

By the Skorohod representation theorem, there exists a probability space $\mathfrak{F}' = \{\Omega', \mathcal{F}', \{\mathcal{F}'_t\}_{t\geq 0}, \mathbf{P}'\}$ and continuous processes $X_t^{n,m}$, X_t defined on this space such that $\mathbb{P}^{n,m}_{\nu_0}$, $\mathbb{P}_{\nu_0}$ are the laws of $X_t^{n,m}$, X_t and

$$X_t^{n,m}(\omega') \to X_t(\omega'), \quad (t \geq 0,\ \omega' \in \Omega') \tag{10.20}$$

(see e.g. [Bil, Ch. 1, Sect. 6]). In particular, $\mathbf{P}'(X_0^{n,m})^{-1}, \mathbf{P}' X_0^{-1}$ have density ν_0.

Fix $1 \leq i \leq d$. Our goal is to show that

$$\lim_{n_1,n_2} \lim_{m_1,m_2} \mathbf{E}'\biggl|\int_0^t \bigl(b^i_{n_1} + q^i_{m_1} - (b^i_{n_2} + q^i_{m_2})\bigr)(X_r)dr\biggr|^2 = 0. \tag{10.21}$$

It suffices to show that

$$\lim_{n_1,n_2} \lim_{m_1,m_2} \mathbf{E}'\biggl|\int_0^t \bigl(b^i_{n_1} + q^i_{m_1} - (b^i_{n_2} + q^i_{m_2})\bigr)(X_r^{n,m})dr\biggr|^2 = 0 \quad \text{uniformly in } n,\ m. \tag{10.22}$$

Indeed, having (10.22), we can appeal to (10.20) and the Dominated convergence theorem to show that, for any fixed n_1, m_1, n_2, m_2,

$$\lim_n \lim_m \mathbf{E}'\biggl|\int_0^t \bigl(b^i_{n_1} + q^i_{m_1} - (b^i_{n_2} + q^i_{m_2})\bigr)(X_r^{n,m})dr - \int_0^t \bigl(b^i_{n_1} + q^i_{m_1} - (b^i_{n_2} + q^i_{m_2})\bigr)(X_r)dr\biggr|^2 = 0,$$

which then yields (10.21).

So, let us prove (10.22). Put for brevity $F := b^i_{n_1} + q^i_{m_1} - (b^i_{n_2} + q^i_{m_2})$, so (10.22) becomes

$$\lim_{n_1,n_2} \lim_{m_1,m_2} \mathbf{E}'\biggl|\int_0^t F(X_r^{n,m})dr\biggr|^2 \to 0. \tag{10.23}$$

Let us rewrite the expression under the limit signs as follows. Let u be the classical solution to the terminal-value problem

$$\bigl(\partial_s + \Delta - (b_n + q_m)\cdot\nabla\bigr)u(s) = -F, \quad s < t, \quad u(t,\cdot) = 0. \tag{10.24}$$

Then

$$\begin{aligned}
\mathbf{E}'\biggl|\int_0^t F(X_r^{n,m})dr\biggr|^2 &= 2\mathbf{E}'\int_0^t F(X_s^{n,m})\int_s^t F(X_r^{n,m})drds \\
&= 2\mathbf{E}'\int_0^t F(X_s^{n,m})\mathbf{E}'\Bigl[\int_s^t F(X_r^{n,m})dr \mid \mathcal{F}'_s\Bigr]ds \\
&= 2\mathbf{E}'\int_0^t F(X_s^{n,m})u(s, X_s^{n,m})ds = 2\int_0^t \langle Fu, \nu\rangle ds,
\end{aligned}$$

where ν is the probability density of $X^{n,m}$, i.e. solution to Cauchy problem for the Fokker-Planck equation

$$\partial_t \nu(t) - \Delta\nu(t) - \operatorname{div}\bigl((b_n + q_m)\nu(t)\bigr) = 0, \quad t > 0, \quad \nu(0,\cdot) = \nu_0. \tag{10.25}$$

Now it is seen that convergence (10.23) will follow once we prove the next lemma.

Lemma 10.4.

$$I := \int_0^T \langle (b^i_{n_1} - b^i_{n_2})u, \nu\rangle ds \to 0, \quad J := \int_0^T \langle (q^i_{m_1} - q^i_{m_2})u, \nu\rangle ds \to 0$$

as $n_1, n_2, m_1, m_2 \to \infty$ uniformly in n, m.

Proof. Below we will use the following bound: provided that $r>1$ is chosen sufficiently close to 1, we have

$$\sup_{n,m,n_1,n_2,m_1,m_2}\left[\int_0^T\langle|\nabla u|^2\rangle ds,\int_0^T\langle u^4\rangle ds,\int_0^T\langle \nu^{2r}\rangle ds,\int_0^T\langle|\nabla \nu|^2\rangle ds\right]<\infty. \tag{10.26}$$

For the last two terms this bound follows right away from Corollary 15.2 with $s=2r$, $s=2$, respectively (since δ can be close to 1, $s=2r$ in Corollary 15.2 must be close to 2, hence the condition that r must be close to 1). For the first two terms this bound is obtained as follows. Since b_n, q_m have supports in the same ball B_R independent of n, m, we can rewrite F as

$$F=\big[b^i_{n_1}+q^i_{m_1}-(b^i_{n_2}+q^i_{m_2})\big]f,\quad f\in C_c^\infty \text{ is identically 1 on } B_R.$$

We now invoke the energy inequality of Proposition 15.2(i) for $s=2$, $s=4$, and then estimate $\langle F,u|u|^{s-2}\rangle$ in exactly the same way as in Step 1 of the proof of Proposition 10.1. (In fact, strictly speaking, in F there we have b^i_n, q^i_m instead of $b^i_{n_1}-b^i_{n_2}$ and $q^i_{m_1}-q^i_{m_2}$, but what matters in the proof is that the form-bound of $b^i_{n_1}-b^i_{n_2}$ can be chosen independently of n_1, n_2, and that the BMO semi-norm of $Q^i_{m_1}-Q^i_{m_2}$ is bounded from above by a constant independent of m_1, m_2, which is obviously true.) From here the bound on the first two terms in (10.26) follows.

Armed with (10.26), we estimate

$$I\leq\left(\int_0^T\|b^i_{n_1}-b^i_{n_2}\|_2^2ds\right)^{\frac12}\left(\int_0^T\|u\|_{2r'}^{2r'}ds\right)^{\frac{1}{2r'}}\left(\int_0^T\|\nu\|_{2r}^{2r}ds\right)^{\frac{1}{2r}},$$

with $1<r<\infty$ selected close to 1. The first term tends to 0 as $n_1,n_2\to\infty$, while the other two terms are uniformly (in n,m,n_1,n_2,m_1,m_2) bounded by (10.26). In turn,

$$J=-\int_0^T\langle(Q^i_{m_1}-Q^i_{m_2}),(\nabla u)\nu\rangle ds-\int_0^T\langle(Q^i_{m_1}-Q^i_{m_2}),u\nabla\nu\rangle ds=:J_1+J_2,$$

where the stream matrices $Q_{m.}$ can and will be chosen to satisfy

$$|Q_{m.}|\leq C(1+|x|)^{-d+2}\quad\forall|x|\geq 2R\gg1, \tag{10.27}$$

where R is chosen so that $\operatorname{sprt} q_{m.}\subset B_R(0)$ (Appendix B). The constant C does not depend on m_1, m_2. We have

$$|J_1|^2\leq\int_0^T\langle|\nabla u|^2\rangle ds\Big(T\langle|Q^i_{m_1}-Q^i_{m_2}|^{2r'}\rangle\Big)^{\frac{1}{r'}}\left(\int_0^T\langle\nu^{2r}\rangle ds\right)^{\frac1r}. \tag{10.28}$$

The first and the last factors are uniformly bounded in view of (10.26). Therefore, since $|Q^i_{m_1}-Q^i_{m_2}|\to0$, in particular, in $L^{2r'}_{\rm loc}$, we obtain

$$\langle|Q^i_{m_1}-Q^i_{m_2}|^{2r'}\rangle\to0\quad\text{as } m_1,m_2\to\infty. \tag{10.29}$$

Next,

$$|J_2|^2\leq\Big(T\langle|Q^i_{m_1}-Q^i_{m_2}|^4\rangle\Big)^{\frac12}\left(\int_0^T\langle u^4\rangle ds\right)^{\frac12}\int_0^T\langle|\nabla\nu|^2\rangle ds,$$

where, by the same argument as above, $\langle|Q^i_{m_1}-Q^i_{m_2}|^4\rangle\to0$ as $m_1,m_2\to\infty$ (note that, taking into account the above estimate on the polynomial rate of vanishing of $Q_{m.}$ at infinity, we have $4(d-2)>d$ even if $d=3$, so the integrals are finite). The other two factors are uniformly bounded. This ends the proof of Lemma 10.4. □

Remark 10.4. It is not difficult to remove the compact support assumption on b and q by inserting identity $\rho\rho^{-1}=1$ in the definitions of I and J, i.e.

$$I=\int_0^T\langle(b^i_{n_1}-b^i_{n_2})u\rho,\nu\rho^{-1}\rangle ds,\quad J=\int_0^T\langle(q^i_{m_1}-q^i_{m_2})u\rho,\nu\rho^{-1}\rangle ds$$

and then arguing as above, but using the global weighted L^2 convergence result of Lemma 15.1 and replacing the energy inequalities of Section 15 (i.e. Proposition 15.2 and Corollary 15.2) by their weighted counterparts with the weights ρ and ρ^{-1} (the last weight is not discussed in Section 15, but the arguments do not change since $|\nabla\rho^{-1}|$ is majorated by ρ^{-1}, see Section 2). This comes at expense of imposing condition $\langle\nu^{2r}\rho^{-\alpha}\rangle<\infty$ for $1<r<\frac{1}{\sqrt{\delta}}$ and some $\alpha>0$.

(***vi***) The proof of the dispersion estimates uses Nash's argument, see e.g. [KS3, proof of Theorem 4.2]. The construction of the semigroup follows closely [KS3, proof of Theorem 4.2]. In fact, we basically construct this semigroup in the proof of Proposition 10.3. The proof of the uniqueness of the weak solution follows the classical Lions's argument and the compensated compactness estimate (Proposition 2.1), see [QX] for details.

(***vii***) This is an immediate consequence of Proposition 10.4.

11. Proof of Theorem 5.2

For reader's convenience, we will prove assertion (vii). Put

$$R^n_\mu f(x):=\mathbb{E}_{\mathbb{P}_x(\tilde{b}_n)}\int_0^\infty e^{-\mu s}f(\omega_s)ds\quad\Big(=(\mu+\Lambda(\tilde{b}_n))^{-1}f(x)\Big),\quad f\in C_c^\infty,$$

$$R^Q_\mu f(x):=\mathbb{E}_{\mathbb{Q}_x}\int_0^\infty e^{-\mu s}f(\omega_s)ds,\quad\mu>0.$$

It suffices for us to show that $(\mu+\Lambda(b))^{-1}f(x)=R^Q_\mu f(x)$ for all $x\in\mathbb{R}^d$ and all $\mu>0$ sufficiently large, where $\Lambda(b)$ is the Feller realization of $-\Delta+b\cdot\nabla$ in C_∞ constructed in (i). This will imply that $\{\mathbb{Q}_x\}_{x\in\mathbb{R}^d}=\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$.

Step 1:

$$R^n_\mu f(x)\to R^Q_\mu f(x)\text{ as }n\to\infty\quad\forall\,x\in\mathbb{R}^d,$$

as follows right away from $\mathbb{Q}_x=w\text{-}\mathcal{P}(\mathbf{C})\text{-}\lim_n\mathcal{P}_x(\tilde{b}_n)$.

Step 2:

$$\|R^Q_\mu f\|_2\leqslant(\mu-\mu_0)^{-1}\|f\|_2\quad\text{ for all }\mu>\mu_0,$$

for some μ_0 independent of n. Indeed, by the elliptic energy inequality (see e.g. Proposition 15.1), $\|R^n_\mu f\|_2\leqslant(\mu-\mu_0)^{-1}\|f\|_2$ for all n. Now 2) follows from 1) by a weak compactness argument in L^2.

By Step 2, operators R^Q_μ admits unique extensions by continuity to L^2, which we denote by $R^Q_{\mu,2}$.

On the other hand, operators $(\mu+\Lambda(b))^{-1}|_{C_c^\infty\cap L^2}$ are bounded on L^2 and, in fact, constitute the resolvent Λ_2 of the generator of a strongly continuous semigroup in L^2, i.e.

$$(\mu+\Lambda_2(b))^{-1}:=\Big[(\mu+\Lambda(b))^{-1}|_{C_c^\infty\cap L^2}\Big]^{\rm clos}_{L^2\to L^2}.$$

One can also construct $\Lambda_2(b)$ directly (using e.g. quadratic forms), see [KS3].

Step 3:

$$\|b \cdot \nabla(\mu - \Delta)^{-1}\|_{2\to 2} \leqslant \delta,$$

which follows right away from $b \in \mathbf{F}_\delta$ and $\|\nabla(-\Delta)^{-\frac{1}{2}}\|_{2\to 2} = 1$.

Step 4:

$$(\mu + \Lambda_2(b))^{-1} f = (\mu - \Delta)^{-1}\big(1 + b \cdot \nabla(\mu - \Delta)^{-1}\big)^{-1} f \text{ in } L^2. \tag{11.1}$$

Indeed, since by our assumptions we have $\delta < 1$, in view of 3) the right-hand side of the previous formula is well defined. Now, we have to appeal for a moment to a "good" approximation $\{b_n\} \in [b]$, i.e. an approximation that actually does converge to b (recall that we do not require from $\{\tilde{b}_n\}$ any kind of convergence to b).

The sought identity (11.1) holds for every b_n. This follows by rearranging the usual Neumann series representation for $(\mu + \Lambda_2(b_n))^{-1}$ while taking into account the estimate of Step 3. So,

$$(\mu + \Lambda_2(b_n))^{-1} f = (\mu - \Delta)^{-1}\big(1 + b_n \cdot \nabla(\mu - \Delta)^{-1}\big)^{-1} f.$$

It remains to pass to the limit in n. In the left-hand side one has $(\mu+\Lambda_2(b_n))^{-1}f \to (\mu+\Lambda_2(b))^{-1}f$ in L^2 (see [KS2], but it is not difficult to prove this directly, see e.g. the proof of Proposition 10.3; here we need a simpler version of this in L^2). In the right-hand side the denominator of the geometric series $b_n \cdot \nabla(\mu - \Delta)^{-1} g \to b \cdot \nabla(\mu - \Delta)^{-1}$ in L^2 for every $g \in L^2$. This is immediate on $g \in C_c^\infty$, so it remains to apply the estimate of Step 3. So, we can pass to the limit in the right-hand side, arriving at the identity (11.1).

Step 5:

$$(\mu + \Lambda(b))^{-1} f = R^Q_\mu f \quad \text{ a.e. on } \mathbb{R}^d.$$

Indeed, since, by our assumptions, $\mathbb{Q}_x$ is a weak solution of the SDE (1.1), we have by Itô's formula

$$(\mu - \Delta)^{-1} h = R^Q_\mu\big[\big(1 + b \cdot \nabla(\mu - \Delta)^{-1}\big)h\big], \quad h \in C_c^\infty.$$

Since $1 + b \cdot \nabla(\mu - \Delta)^{-1} \in \mathcal{B}(L^2)$ by Step 3, we have, in view of Step 2,

$$(\mu - \Delta)^{-1} g = R^Q_{\mu,2}\big[\big(1 + b \cdot \nabla(\mu - \Delta)^{-1}\big)g\big], \quad g \in L^2.$$

Take $g = \big(1+b\cdot\nabla(\mu-\Delta)^{-1}\big)^{-1} f$, $f \in C_c^\infty$, which is possible by Step 3 and $\delta < 1$. Then, by Step 4, $(\mu+\Lambda_2(b))^{-1}f = R^Q_{\mu,2} f$. By the consistency property $(\mu+\Lambda(b))^{-1}|_{C_c^\infty \cap L^2} = (\mu+\Lambda_2(b))^{-1}|_{C_c^\infty \cap L^2}$ and the result follows.

Step 6: Fix some $r > 2 \vee (d-2)$ in the interval $[2, \frac{2}{\sqrt{\delta}}[$ (here we use our hypothesis on δ, which must be sufficiently small so that r can be large enough). Since $R^n_\mu f = (\mu + \Lambda(\tilde{b}_n))^{-1} f$, we obtain by assertion (xii) that for all $\mu > \mu_0$

$$\|\nabla R^n_\mu f\|_{\frac{rd}{d-2}} \leqslant K_2(\mu - \mu_0)^{-\frac{1}{2}+\frac{1}{r}} \|f\|_r.$$

By a weak compactness argument in L^{rj}, in view of Step 1, we have $\nabla R^Q_\mu f \in [L^{rj}]^d$, and there is a subsequence of $\{R^n_\mu f\}$ (without loss of generality, it is $\{R^n_\mu f\}$ itself) such that

$$\nabla R^n_\mu f \xrightarrow{w} \nabla R^Q_\mu f \quad \text{ in } [L^{rj}]^d.$$

By Mazur's lemma, there is a sequence of convex combinations of the elements of $\{\nabla R^n_\mu f\}_{n=1}^\infty$ that converges to $\nabla R^Q_\mu f$ strongly in $[L^{rj}]^d$, i.e.

$$\sum_\alpha c_\alpha \nabla R^{n_\alpha}_\mu f \xrightarrow{s} \nabla R^Q_\mu f \quad \text{ in } [L^{rj}]^d.$$

Now, in view of the latter, Step 1 and the Sobolev embedding theorem, we have $\sum_\alpha c_\alpha R_\mu^{n_\alpha} f \xrightarrow{s} R_\mu^Q f$ in C_∞. Therefore, by Step 5, $(\mu+\Lambda(b))^{-1}f(x) = R^Q_\mu f(x)$ for all $x \in \mathbb{R}^d$, $f \in C_c^\infty$, as claimed.

12. Proof of Theorem 6.1

(***i***) We modify the proof of Theorem 5.1, i.e. we verify conditions of the Trotter's approximation theorem, but now for Feller generators

$$\Lambda(a_n, b_n, q_m) := -a_n \cdot \nabla^2 + (b_n + q_m) \cdot \nabla, \quad D\big(\Lambda(a_n, b_n, q_m)\big) = (1-\Delta)^{-1} C_\infty.$$

Condition 1°) of Trotter's theorem is obvious.

Let us verify conditions 2°) and 3°). To this end, we note that Propositions 10.2, 10.4 and 10.1, i.e. a priori Hölder continuity of solutions, separation and embedding properties, are still valid for operators $\Lambda(a_m, b_n, q_m)$ (in fact, under more general condition $\delta < 4\xi^2$) since we can put these operators in divergence form

$$\Lambda(a_n, b_n, q_m) = -\nabla \cdot a_n \cdot \nabla + (\tilde{b}_n + q_m) \cdot \nabla, \quad \tilde{b}_n = \nabla a_n + b_n \in \mathbf{F}_\delta,$$

so De Giorgi's method applies.

Remark 12.1. Condition $\delta < 4\xi^2$ is seen from the following calculation, which we have to repeat several times (also, with the cutoff function η) when extending Propositions 10.2, 10.4 and 10.1 to include matrix fields a_n. We mutiply elliptic equation $(\mu - \nabla \cdot a_n \cdot \nabla + (\tilde{b}_n + q_m) \cdot \nabla)u = 0$ by $u^{p-1} \geq 0$ and integrate by parts, obtaining, after taking into accout $\operatorname{div} q_m = 0$,

$$\mu\langle u^p\rangle + \frac{4(p-1)}{p^2}\langle a_n \cdot \nabla u^{\frac{p}{2}}, \nabla u^{\frac{p}{2}}\rangle + \frac{2}{p}\langle \tilde{b}_n \cdot \nabla u^{\frac{p}{2}}, u^{\frac{p}{2}}\rangle = 0.$$

Since $a_n \in H_\xi$,

$$\mu\langle u^p\rangle + \frac{4(p-1)}{p^2}\xi\langle |\nabla u^{\frac{p}{2}}|^2\rangle + \frac{2}{p}\langle \tilde{b}_n \cdot \nabla u^{\frac{p}{2}}, u^{\frac{p}{2}}\rangle = 0,$$

Applying the quadratic inequality in the last term, we arrive at

$$\mu\langle u^p\rangle + \frac{4(p-1)}{p^2}\xi\langle |\nabla u^{\frac{p}{2}}|^2\rangle \leq 2\bigg(\alpha\langle |\tilde{b}_n|^2, v^p\rangle + \frac{1}{4\alpha}\langle |\nabla v^{\frac{p}{2}}|^2\rangle\bigg).$$

Now, using $\tilde{b}_n \in \mathbf{F}_\delta$ and selecting $\alpha = \frac{1}{2\sqrt{\delta}}$, we obtain

$$\mu\langle u^p\rangle + \left[\frac{4(p-1)}{p^2}\xi - \frac{2}{p}\sqrt{\delta}\right]\langle |\nabla u^{\frac{p}{2}}|^2\rangle \leq 0.$$

So, $\delta < 4\xi^2$ is exactly the condition that ensures that $\frac{4(p-1)}{p^2}\xi - \frac{2}{p}\sqrt{\delta} > 0$ for some finite $p \geq 2$ and hence gives us an energy inequality; that is, we need $p > \frac{2}{2-\xi^{-1}\sqrt{\delta}}$.

Proposition 10.3 is replaced by a simpler convergence result

$$u := L^2_{\rm loc}\text{-}\lim_n \lim_m u_{n,m}, \tag{12.1}$$

where $u_{n,m} = (\mu + \Lambda(a_n, b_n, q_m))^{-1} f$, $f \in C_c^\infty$. (As is explained in the proof of Theorem 5.1(i), we need convergence in *some* topology to establish the approximation uniqueness.) Let us prove (12.1). Due to our more restrictive assumption $\delta < \xi^2$ we can work in L^2 rather than L^p. By the

Steps 1-3 in the proof of Proposition 10.3, which extend easily to a_n for each fixed n, the limit $u_n := L^2\text{-}\lim_m u_{n,m}$ exists and satisfies the identity

$$\mu\langle u_n, \varphi\rangle - \langle a_n \cdot \nabla u_n, \nabla\varphi\rangle + \langle(\nabla a_n + b_n + q)\cdot\nabla u_n, \varphi\rangle = \langle f, \varphi\rangle, \quad \varphi \in C_c^\infty \tag{12.2}$$

with μ independent of n. The standard energy inequality argument (cf. Section 15) and the compensated compactness estimate of Proposition 2.1 allow us to extend (12.2) to test functions $\varphi \in W^{1,2}$, i.e. u_n is the standard weak solution to the elliptic equation $(\mu - \nabla\cdot a_n\cdot\nabla + (\nabla a_n + b_n + q)\cdot\nabla)u_n = f$ in L^2. Now, the convergence $\nabla a_n + b_n \to \nabla a + b$ in $L^2_{\rm loc}$ and the convergence $a_n \to a$ a.e. on $\mathbb{R}^d$, applied in the standard weak compactness argument in L^2, give us, via the uniqueness of the weak solution in L^2, the sought convergence $u_n \to u$ in $L^2_{\rm loc}$, where u satisfies (12.2) with a, b instead of a_n, b_n, moreover, this identity extends in the same way to $\varphi \in W^{1,2}$, so u is the standard weak solution $(\mu - \nabla\cdot a\cdot\nabla + (\nabla a + b + q)\cdot\nabla)u = f$ in L^2. In the case $q = 0$, this is essentially how the divergence form operator $-\nabla\cdot a\cdot\nabla + b\cdot\nabla$ with form-bounded b was treated in [KS3, Theorem 4.3]

Now, armed with the above analogues of Propositions 10.2, 10.3 and 10.1 for $\Lambda(a_m, b_n, q_m)$, we verify condition 2°) of Trotter's theorem in the same way as in the proof of Theorem 5.1.

Condition 3°) requires a comment. Fix $g \in C_c^\infty$. By the resolvent identity,

$$\begin{aligned}\mu(\mu + \Lambda(a_m, b_n, q_m))^{-1}g - \mu(\mu - \Delta)^{-1}g &= \mu(\mu + \Lambda(a_m, b_n, q_m))^{-1}(a_m - I)\cdot\nabla^2(\mu - \Delta)^{-1}g \\ &\quad + \mu(\mu + \Lambda(a_m, b_n, q_m))^{-1}(b_n + q_m)\cdot\nabla(\mu - \Delta)^{-1}g.\end{aligned}$$

Since $\mu(\mu - \Delta)^{-1}g \to g$ uniformly on $\mathbb{R}^d$ as $\mu \to \infty$, it suffices to show convergence

$$\|(\mu + \Lambda(a_m, b_n, q_m))^{-1}(a_m - I)\cdot\mu(\mu - \Delta)^{-1}\nabla^2 g\|_\infty \to 0 \tag{12.3}$$

and

$$\|(\mu + \Lambda(a_m, b_n, q_m))^{-1}(b_n + q_m)\cdot\mu(\mu - \Delta)^{-1}\nabla g\|_\infty \to 0 \tag{12.4}$$

$$\text{as } \mu \to \infty \quad \text{uniformly in } n,\ m.$$

To prove (12.4) we argue as in the proof of Theorem 5.1(i) and apply the discussed above analogue of Proposition 10.1 to $w_{n,m} := (\mu + \Lambda(a_m, b_n, q_m))^{-1}(b_n^i + q_m^i)f$ with f chosen as $f := \mu(\mu - \Delta)^{-1}\nabla_i g$. (Let us note in passing that Proposition 10.1 is valid for b_n^i and q_n^i in the RHS of the equation for $w_{n,m}$ replaced by the i-th components of general vector fields in $\mathbf{F}_\nu$, $\nu < \infty$, and $\mathbf{BMO}^{-1}$, i.e. the proof of Proposition 10.1 does not exploit any cancellations between the RHS of the equation and the drift term.)

The proof of (12.3) is even easier. Indeed, we can apply a straightforward analogue of Proposition 10.1 to $w_{n,m} := (\mu + \Lambda(a_m, b_n, q_m))^{-1}f$ with bounded f chosen as $f := ((a_m)_{ij} - \delta_{ij})\mu(\mu - \Delta)^{-1}\nabla_i\nabla_j g$, use the uniform in m boundedness of a_m on $\mathbb{R}^d$ and then argue as in the proof of Theorem 5.1(i).

(***iii***) The proof of the relaxed approximation uniqueness basically does not change. Since $\delta < \xi^2$, we continue to work in the standard setting of weak solutions in L^2. We get an extra term in Step 2: the difference $h_n = u_n - u$ satisfies

$$\mu\langle h_n, \varphi\rangle + \langle a_n\cdot\nabla h_n, \nabla\varphi\rangle + \langle b_n\cdot\nabla h_n, \varphi\rangle + \langle q\cdot\nabla h_n, \varphi\rangle = \langle\nabla\cdot(a_m - a)\cdot\nabla u, \nabla\varphi\rangle + \langle(b - b_n)\cdot\nabla u, \varphi\rangle$$

for all $\varphi \in W^{1,2}$, for all μ greater than some μ_0 independent of n. Taking $\varphi = h_n\rho$ and repeating the proof of the energy inequality of Proposition 15.1(ii) for $s = 2$, we obtain

$$(\mu - \mu_0)\langle|h_n|^2\rho\rangle + C_1\langle|\nabla h_n|^2\rho\rangle \leq \langle\nabla\cdot(a_n - a)\cdot\nabla u, \nabla(h_n\rho)\rangle + \langle(b - b_n)\cdot\nabla u, h_n\rho\rangle,$$

where the last term tends to zero as $n \to \infty$ by the argument in the proof of Theorem 5.1(ii). The term $\langle \nabla \cdot (a_n - a) \cdot \nabla u, (\nabla h_n)\rho + h_n \nabla \rho \rangle \to 0$ tends to zero by an even simpler argument:

$$|\langle \nabla \cdot (a_n - a) \cdot \nabla u, (\nabla h_n)\rho \rangle| \leq \| |a_n - a| |\nabla u| \sqrt{\rho} \|_2 \| (\nabla h_n) \sqrt{\rho} \|_2,$$

where the second multiple is bounded uniformly in n due to the energy inequality, and the first multiple tends to 0 by the Dominated convergence theorem (since $|\nabla u|\sqrt{\rho} \in L^2$, also by the energy inequality). (The proof that $\langle \nabla \cdot (a_n - a) \cdot \nabla u, h_n \nabla \rho \rangle \to 0$ as $n \to \infty$ is easier since $|\nabla \rho|$ is majorated by ρ.)

(***iv***) The proof repeats the proof of Theorem 5.1(iv).

(***v***) We obtain in the same way as in the proof of Theorem 5.1(v)

$$X_t^{n,m}(\omega') \to X_t(\omega'), \quad t \geq 0, \quad \omega' \in \Omega, \tag{12.5}$$

where

$$X_t^{n,m} = X_0 - \int_0^t (b_n(X_s^{n,m}) + q_m(X_s^{n,m}))ds + \int_0^t \sigma_n(X_s^{n,m})dB_s, \quad n, m = 1, 2, \dots$$

We only need to supplement the proof of Theorem 5.1(v) by the convergence

$$\int_0^t \sigma_n^{ij}(X_s^{n,m})dB_s \to \int_0^t \sigma^{ij}(X_s)dB_s \quad \text{in } L^2(\Omega'), \quad \text{for all } t \geq 0$$

Form now on, we drop index ij to lighten notations. Since $a_n \to a$ a.e. on $\mathbb{R}^d$, we have convergence of their square roots: $\sigma_n \to \sigma$ a.e. By Itô's isometry, our task is to show that

$$\mathbf{E}' \int_0^t |\sigma_n(X_s^{n,m}) - \sigma(X_s)|^2 ds \to 0 \quad \text{as } n, m \to \infty.$$

In turn, this convergence follows from:

$$\mathbf{E}' \int_0^t |\sigma_n(X_s^{n_0,m_0}) - \sigma(X_s^{n_0,m_0})|^2 ds \to 0 \text{ as } n \to \infty \text{ uniformly in } n_0,\, m_0 \geq 1,$$

and

$$\mathbf{E}' \int_0^t |\sigma(X_s^{n,m}) - \sigma(X_s)|^2 ds \to 0 \quad \text{as } n, m \to \infty.$$

The latter is immediate from (12.5) via the Dominated convergence theorem, and the former follows right away from the strong Feller property of the resolvents (assertion (vii)):

$$\begin{aligned}
\mathbf{E}' \int_0^t |\sigma_n(X_s^{n_0,m_0}) - \sigma(X_s^{n_0,m_0})|^2 ds &= \int_{\mathbb{R}^d} \nu_0(dx) \mathbb{E}_{\mathbb{P}_x^{n_0,m_0}} \int_0^t |\sigma_n(\omega_s) - \sigma(\omega_s)|^2 ds \\
&= \int_{\mathbb{R}^d} \int_0^t \big(e^{-s\Lambda(a_{n_0}, b_{n_0}, q_{m_0})} |\sigma_n - \sigma|^2\big)(x) ds \nu_0(x) dx \\
&= \int_{\mathbb{R}^d} \int_0^t e^{\mu s} e^{-\mu s} \big(e^{-s\Lambda} |\sigma_n - \sigma|^2\big)(x) ds \nu_0(x) dx \\
&\leq e^{\mu t} \int_{\mathbb{R}^d} \int_0^\infty e^{-\mu s} \big(e^{-s\Lambda} |\sigma_n - \sigma|^2\big)(x) ds \nu_0(x) dx.
\end{aligned}$$

The last term is, modulo $e^{\mu t}$, which is bounded anyway, is

$$\int_{\mathbb{R}^d} (\mu - \Lambda)^{-1} |\sigma_n - \sigma|^2(x) \nu_0(x) dx \leq C \int_{\mathbb{R}^d \setminus B_R} \nu_0(x) dx + \| \mathbf{1}_{B_R} (\mu - \Lambda))^{-1} |\sigma_n - \sigma|^2 \|_\infty,$$

where $C=\sup_n\|\sigma_n\|_\infty+\|\sigma\|_\infty$. The first integral can be made as small as needed by selecting R sufficiently large. To estimate the second term, we invoke the strong Feller property (vii):

$$\begin{aligned}&\|\mathbf{1}_{B_R}(\mu-\Lambda))^{-1}|\sigma_n-\sigma|^2\|_\infty\\&\leq K\sup_{x\in\frac{1}{2}\mathbb{Z}^d\cap B_R}\bigg[\langle|\sigma_n-\sigma|^{2p\theta}\rho_x\rangle^{\frac{1}{p\theta}}+\langle|\sigma_n-\sigma|^{2p\theta'}\rho_x\rangle^{\frac{1}{p\theta'}}\bigg].\end{aligned}$$

It remains to apply the Dominated convergence theorem in n. (Strictly speaking, we are applying Feller resolvent to discontinuous functions, but since the former is a family of integral operators, a standard limiting argument addresses this.)

13. Proof of Theorem 8.1

The proof of Theorem 8.1 follows closely the proof of Theorem 5.1, with only one calculation done differently. We stay at the level of a priori estimates, so b, q are additionally bounded and smooth. Assuming for simplicity that $c_\delta=c_{\delta_+}=0$ in the conditions on b, $(\operatorname{div} b)_+$, we consider Cauchy problem $(\partial_t-\Delta+(b+q)\cdot\nabla)v=0$, $v|_{t=0}=v_0\in C_c^\infty$ (without loss of generality, $v_0\geq 0$), multiply the parabolic equation by v^{p-1} and integrate by parts:

$$\frac{1}{p}\langle\partial_t v^p\rangle+\frac{4(p-1)}{p^2}\langle|\nabla v^{\frac{p}{2}}|^2\rangle+\frac{2}{p}\langle b\cdot\nabla v^{\frac{p}{2}},v^{\frac{p}{2}}\rangle=0,$$

where we have used $\operatorname{div} q=0$. Thus,

$$\partial_t\langle v^p\rangle+\frac{4(p-1)}{p}\langle|\nabla v^{\frac{p}{2}}|^2\rangle=-2\langle b\cdot\nabla v^{\frac{p}{2}},v^{\frac{p}{2}}\rangle.$$

In turn, $-\langle b\cdot\nabla v^{\frac{p}{2}},v^{\frac{p}{2}}\rangle=\langle bv^{\frac{p}{2}},\nabla v^{\frac{p}{2}}\rangle+\langle\operatorname{div} b,v^p\rangle$, hence $\langle\partial_t v^p\rangle+\frac{4(p-1)}{p}\langle|\nabla v^{\frac{p}{2}}|^2\rangle=\langle\operatorname{div} b,v^p\rangle$, and so

$$\langle\partial_t v^p\rangle+\frac{4(p-1)}{p}\langle|\nabla v^{\frac{p}{2}}|^2\rangle\leq\langle(\operatorname{div} b)_+,v^p\rangle,$$

Now, applying $\operatorname{div} b\in\mathbf{F}_{\delta_+}$, we obtain energy inequality

$$\langle\partial_t v^p\rangle+\left(\frac{4(p-1)}{p}-\delta_+\right)\langle|\nabla v^{\frac{p}{2}}|^2\rangle\leq 0$$

with $\frac{4(p-1)}{p}-\delta_+>0$ provided that $p>\frac{4}{4-\delta_+}$.

14. Proof of Theorem 8.3(iii)

(ii) We repeat the proof of Theorem 5.1(i):

Proposition 10.2 (“A priori Hölder continuity”) is replaced by [K6, Theorem 5].

Proposition 10.3 (“Convergence”) is replaced by [K6, Theorem 3(v)].

Proposition 10.4 (“Separation property”) is replaced by [K6, Propositions 5 and 6] in the proof of [K6, Theorem 5].

Proposition 10.1 (“Embedding property”) is replaced by [K6, Theorem 5].

15. Auxiliary results used in the proofs of Propositions 10.2-10.1

15.1. **Energy inequalities.** In the next two propositions we assume that

$$\begin{cases} b \in \mathbf{F}_\delta \text{ with } \delta < 4, \\ q \in \mathbf{BMO}^{-1} \end{cases} \tag{15.1}$$

(that is, we are in the assumptions of Theorem 5.1), and fix some $\{b_n\} \in [b]$, $\{q_m\} \in [q]$.

Proposition 15.1. *Assume that hypothesis (15.1) holds. Let $u = u_{n,m}$ denote the classical solution of the elliptic equation*

$$\big(\mu - \Delta + (b_n + q_m) \cdot \nabla\big) u = f, \quad \mu \geq 0, \quad f \in C_c^\infty. \tag{15.2}$$

Fix some $s > \frac{2}{2-\sqrt{\delta}}$, $s \geq 2$. Then the following are true:

(i) There exist positive constants μ_0, C_1 independent of n, m such that

$$(\mu - \mu_0)\langle |u|^s \rangle + C_1 \langle |\nabla |u|^{\frac{s}{2}}|^2 \rangle \leq \langle f, u|u|^{s-2} \rangle \tag{15.3}$$

for all $\mu \geq \mu_0$.

(ii) [Weighted variant] *Provided that constant $\sigma > 0$ in weight $\rho(y) = (1 + \sigma |y|^2)^{-\frac{d+\epsilon_0}{2}}$ is fixed sufficiently small, there exist positive constants μ_0, C_1 independent of n, m such that, for all $x \in \mathbb{R}^d$,*

$$(\mu - \mu_0)\langle |u|^s \rho_x \rangle + C_1 \langle |\nabla |u|^{\frac{s}{2}}|^2 \rho_x \rangle \leq \langle f, u|u|^{s-2} \rho_x \rangle, \quad \rho_x(y) = \rho(x-y), \tag{15.4}$$

for all $\mu \geq \mu_0$.

Remark 15.1. We have added condition $s \geq 2$ to save ourselves some efforts since to cover the values of the form-bound δ close to 4 we need to select s large anyway. But, generally speaking, $s \in]1, 2[$ does not pose a substantial difficulty, see e.g. [KS3, proof of Theorem 4.2].

In the proof of assertion (ii) we will need to control a term resulting from the interaction between the weight ρ and the stream matrix Q of drift q. This will be done using Proposition 2.2 (a compensated compactness type estimate) and Lemma 2.1 on $\mathrm{BMO}(\mathbb{R}^d)$ multipliers.

Proof of Proposition 15.1. We will only prove (ii). We write for brevity $b = b_n$, $q = q_m$. In the proof we will need estimates

$$|\nabla \rho_x(y)| \leq \frac{d+\epsilon_0}{2} \frac{2\sigma|x-y|}{1+\sigma|x-y|^2} \rho_x(y) \leq \frac{d+\epsilon_0}{2} \sqrt{\sigma} \rho_x(y). \tag{15.5}$$

We multiply equation (15.2) by $u|u|^{s-2}\rho_x$ and integrate by parts:

$$\mu\langle |u|^s \rangle + \frac{4(s-1)}{s^2} \langle |\nabla |u|^{\frac{s}{2}}|^2 \rho_x \rangle + \frac{2}{s} \langle \nabla |u|^{\frac{s}{2}}, |u|^{\frac{s}{2}} \nabla \rho_x \rangle$$
$$+ \frac{2}{s} \langle b \cdot \nabla |u|^{\frac{2}{s}}, |u|^{\frac{2}{s}}, \rho_x \rangle + \frac{2}{s} \langle q \cdot \nabla |u|^{\frac{s}{2}}, |u|^{\frac{s}{s}} \rho_x \rangle = \langle f, u|u|^{s-2} \rho_x \rangle,$$

where, taking into account anti-symmetry of the stream matrix Q of q,

$$\begin{aligned}\frac{2}{s}|\langle q\cdot\nabla|u|^{\frac{s}{2}},|u|^{\frac{s}{s}}\rho_x\rangle &= -\frac{2}{s}\langle Q\cdot\nabla|u|^{\frac{s}{2}},|u|^{\frac{s}{2}}\nabla\rho_x\rangle\\ &= \frac{2}{s}\langle Q\cdot\nabla(|u|^{\frac{s}{2}}\sqrt{\rho_x}),|u|^{\frac{s}{2}}\sqrt{\rho_x}\frac{\nabla\rho_x}{\rho_x}\rangle\\ &\text{(we apply Proposition 2.2))}\\ &\leq \frac{2C}{s}\|\frac{\nabla\rho_x}{\rho_x}\cdot Q\|_{\mathrm{BMO}}\|\nabla(|u|^{\frac{s}{2}}\sqrt{\rho_x})\|_2\|u^{\frac{s}{2}}\sqrt{\rho_x}\|_2,\end{aligned}$$

Applying Lemma 2.1 to $\frac{\nabla_i\rho(y)}{\rho(y)}=\frac{2\sigma y_i}{1+\sigma|y|^2}$, we obtain

$$K:=\sup_{x\in\mathbb{R}^d}\|\frac{\nabla\rho_x}{\rho_x}\cdot Q\|_{\mathrm{BMO}}<\infty,$$

so we can conclude the previous estimate as

$$\frac{2}{s}|\langle q\cdot\nabla|u|^{\frac{s}{2}},|u|^{\frac{s}{s}}\rho_x\rangle\leq\frac{2CK}{s}\big(\langle|\nabla|u|^{\frac{s}{2}}\sqrt{\rho_x}|^2\rangle+\langle|u|^s\rho_x\rangle\big)^{\frac{1}{2}}\langle|u|^s\rho_x\rangle^{\frac{1}{2}},\tag{15.6}$$

and then use Cauchy-Schwarz and the second inequality in (15.5) to estimate $\langle|\nabla|u|^{\frac{s}{2}}\sqrt{\rho_x}|^2\rangle\leq(1+\varepsilon_1)\langle|\nabla|u|^{\frac{s}{2}}|^2\rho_x\rangle+C_\sigma(1+\varepsilon_1^{-1})\langle|u|^s\rho_x\rangle$. Here we can fix any positive ε_1 because, going back to (15.6), in the end we will apply Cauchy-Schwart inequality which will allow us to make the constant in front of the term $\langle|\nabla|u|^{\frac{s}{2}}|^2\rho_x\rangle$ in the resulting upper bound on $\frac{2}{s}|\langle q\cdot\nabla|u|^{\frac{s}{2}},|u|^{\frac{s}{s}}\rho_x\rangle$ as small as we want.

Next,

$$\begin{aligned}\frac{2}{s}|\langle b\cdot\nabla|u|^{\frac{s}{2}},|u|^{\frac{2}{s}}\rho_x\rangle| &\leq\frac{2}{s}\bigg(\beta\langle|b|^2,|u|^s\rho_x\rangle+\frac{1}{4\beta}\langle|\nabla|u|^{\frac{s}{2}}|^2\rho_x\rangle\bigg)\\ &\text{(use } b\in\mathbf{F}_\delta)\\ &\leq\frac{2}{s}\bigg(\beta(\delta\|\nabla(|u|^{\frac{s}{2}}\sqrt{\rho_x})\|_2^2+c_\delta\langle|u|^s\rho_x\rangle)+\frac{1}{4\beta}\langle|\nabla|u|^{\frac{s}{2}}|^2\rho_x\rangle\bigg).\end{aligned}\tag{15.7}$$

Take $\beta=\frac{1}{2\sqrt{\delta}}$ and then apply inequality in the end of the previous paragraph, but with ε_1 chosen small.

Applying the last inequality in (15.5), we can replace in the previous inequalities all occurrences of $|\nabla\rho_x|$ by $C\sqrt{\sigma}\rho_x$. That way, we arrive at the inequality (15.4) with constant

$$C_1=\frac{4(s-1)}{s^2}-\frac{2}{s}\sqrt{\delta}-\text{(constant terms proportional to }\sqrt{\sigma}\text{ and }\varepsilon_1),$$

where μ_0 is given in terms of c_δ, $\frac{CK}{s}$ and ε_1^{-1}. Since $\frac{4(s-1)}{s^2}-\frac{2}{s}\sqrt{\delta}>0$ ($\Leftrightarrow s>\frac{2}{2-\sqrt{\delta}}$) by our choice of s, we can fix σ and ε_1 sufficiently small so that $C_1>0$. This ends the proof of Proposition 15.1. □

Corollary 15.1. *In the assumptions and notations of Proposition 15.1, we also have*

(i) There exist positive constants μ_0, C_1, C_2 independent of n, m such that

$$(\mu-\mu_0)\langle|u|^s\rangle+C_1\langle|\nabla|u|^{\frac{s}{2}}|^2\rangle\leq C_2\langle|f|^s\rangle\tag{15.8}$$

for all $\mu\geq\mu_0$.

(*ii*) [Weighted variant] *Provided that constant σ in the definition of weight ρ is fixed sufficiently small, there exist positive constants μ_0, C_1, C_2 independent of n, m such that, for all $x \in \mathbb{R}^d$,*

$$(\mu-\mu_0)\langle |u|^s\rho_x\rangle + C_1\langle |\nabla |u|^{\frac{s}{2}}|^2\rho_x\rangle \leq C_2\langle |f|^s\rho_x\rangle \tag{15.9}$$

for all $\mu \geq \mu_0$.

Proof. In the above proof of Proposition 15.1 we can take one step further and apply Young's inequality in order to estimate

$$\langle f, u|u|^{s-2}\rho_x\rangle \leq \frac{\varepsilon^s}{s}\langle |u|^s\rho_x\rangle + \frac{1}{\varepsilon^{s'}s'}\langle |f|^s\rho_x\rangle.$$

□

The previous energy inequalities have their parabolic counterparts:

Proposition 15.2. *Assume that hypothesis (15.1) is satisfied. Let $v = v_{n,m}$ be the classical solution of Cauchy problem*

$$\big(\mu + \partial_t - \Delta + (b_n + q_m)\cdot\nabla\big)v = 0, \quad v(0) = f \in C_c^\infty, \quad \mu \geq 0. \tag{15.10}$$

Fix some $s > \frac{2}{2-\sqrt{\delta}}$, $s \geq 2$. Then there exists $\mu_0 \geq 0$ independent of n, m such that for all $\mu \geq \mu_0$ the following are true:

(*i*)

$$(\mu-\mu_0)\int_0^t \langle |v|^s\rangle + \frac{1}{s}\sup_{r\in[0,t]}\langle |v(r)|^s\rangle + C_1\int_0^t\langle |\nabla |v|^{\frac{s}{2}}|^2\rangle \leq \frac{2}{s}\langle |f|^s\rangle \tag{15.11}$$

for constant $C_1 > 0$ independent of n, m.

(*ii*) [Weighted variant] *Provided that σ is the definition of weight ρ is chosen sufficiently small, we have*

$$(\mu-\mu_0)\int_0^t \langle |v|^s\rho_x\rangle + \frac{1}{s}\sup_{r\in[0,t]}\langle |v(r)|^s\rho_x\rangle + C_1\int_0^t\langle |\nabla |v|^{\frac{s}{2}}|^2\rho_x\rangle \leq \frac{2}{s}\langle |f|^s\rho_x\rangle \tag{15.12}$$

for constant $C_1 > 0$ independent of n, m and $x \in \mathbb{R}^d$.

Proof of Proposition 15.2. Let us prove (*ii*). We multiply the parabolic equation in (15.10) by $v|v|^{s-2}\rho_x$ and integrate over $[0,r]\times\mathbb{R}^d$. All the terms in the resulting integral identity, except the one containing $\partial_t v$, are dealt with as in the proof of the previous proposition. In turn, the term containing $\partial_t v$ is evaluated as follows:

$$\int_0^r \langle \partial_t v, v|v|^{s-2}\rho_x\rangle = \frac{1}{s}\int_0^r\langle \partial_t|v|^s\rho_x\rangle = \frac{1}{s}\big(\langle |v(r)|^s\rho_x\rangle - \frac{1}{s}\langle |f|^s\rho_x\rangle\big).$$

This gives us

$$(\mu-\mu_0)\int_0^r \langle |v|^s\rho_x\rangle + \frac{1}{s}\langle |v(r)|^s\rho_x\rangle + C_1\int_0^r\langle |\nabla |v|^{\frac{s}{2}}|^2\rho_x\rangle \leq \frac{1}{s}\langle |f|^s\rho_x\rangle. \tag{15.13}$$

for appropriate $\mu_0 \geq 0$ and $C > 0$, both independent of n, m. Let $\mu \geq \mu_0$. We have, in particular,

$$\frac{1}{s}\langle |v(r)|^s\rho_x\rangle \leq \frac{1}{s}\langle |f|^s\rho_x\rangle, \quad (\mu-\mu_0)\int_0^r\langle |v|^s\rho_x\rangle + C_1\int_0^r\langle |\nabla |v|^{\frac{s}{2}}|^2\rho_x\rangle \leq \frac{1}{s}\langle |f|^s\rho_x\rangle.$$

We can pass to the supremum in r in both inequalities since their right-hand side does not depend on r. Adding up the resulting inequalities, we arrive at (15.12). □

In the proof of assertion (v) of Theorem 5.1 we use energy inequality for the Fokker-Planck equation:

Corollary 15.2. *Assume that hypothesis (15.1) is satisfied with $\delta<1$. Let $\nu=\nu_{n,m}$ denote the classical solution to Cauchy problem*

$$\mu\nu+\partial_t\nu-\Delta\nu+\mathrm{div}\big[(b_n+q_m)\nu\big]=0,\quad \nu(0)=\nu_0\in C_c^\infty,\quad \mu\geq 0. \tag{15.14}$$

Fix some $2\leq s<\frac{2}{\sqrt{\delta}}$. Then there exist positive constants μ_0, C_1 independent of n, m such that

$$(\mu-\mu_0)\int_0^t\langle|\nu|^s\rangle+\frac{1}{s}\sup_{r\in[0,t]}\langle|\nu(r)|^s\rangle+C_1\int_0^t\langle|\nabla|\nu|^{\frac{s}{2}}|^2\rangle\leq\frac{2}{s}\langle|\nu_0|^s\rangle$$

for all $\mu\geq\mu_0$.

One also has a straightforward weighted counterpart of this inequality as in assertion (ii) of Proposition 15.2.

We will be applying Corollary 15.2 in the case when $\nu_0\geq 0$, $\langle\nu_0\rangle=1$.

Proof. All the terms in the corresponding integral identity, except the next one, are dealt with in the same way as in the proof of Proposition 15.1. Let $b=b_n$. Then

$$\langle\mathrm{div}(b\nu),\nu|\nu|^{s-2}\rangle=(s-1)\langle b\nu,|\nu|^{s-2}\nabla\nu\rangle=\frac{2}{s}(s-1)\langle b\cdot\nabla|\nu|^{\frac{s}{2}},|\nu|^{\frac{s}{2}}\rangle.$$

Hence we arrive at the counterpart of (15.13) with the coefficient of the dispersion term

$$C=\frac{4(s-1)}{s^2}-\frac{2}{s}(s-1)\sqrt{\delta}-(\text{constant terms proportional to }\varepsilon_1)$$

that must be positive. We can fix ε_1 as small as needed. What matters is the value of δ that ensures that $\frac{4(s-1)}{s^2}-\frac{2}{s}(s-1)\sqrt{\delta}>0$. The latter is equivalent to $s<\frac{2}{\sqrt{\delta}}$, which is satisfied by our assumptions. $\square$

15.2. **Global weighted L^2 summability of a form-bounded drift.** A form-bounded vector field $b\in\mathbf{F}_\delta$ is a priori only local summable: $|b|\in L^2_{\rm loc}$. In fact, condition $b\in\mathbf{F}_\delta$ implies global square summability of $|b|$, but with respect to weight ρ_x. Indeed, selecting $\sqrt{\rho_x}$ as the test function in the definition of $b\in\mathbf{F}_\delta$ and using (2.2), we obtain

$$\begin{aligned}\langle|b|^2\rho_x\rangle&\leq\frac{\delta}{4}\langle\frac{|\nabla\rho_x|^2}{\rho_x}\rangle+c_\delta\langle\rho_x\rangle\\&\leq\Big(\frac{\delta}{4}\frac{(d+\varepsilon_0)^2}{4}\sigma+c_\delta\Big)\langle\rho\rangle<\infty,\end{aligned} \tag{15.15}$$

where we have used $\langle\rho_x\rangle=\langle\rho\rangle$. Moreover, we have the following global convergence result:

Lemma 15.1. *Let $b\in\mathbf{F}_\delta$, $\{b_n\}\in[b]$. Then, for every $x\in\mathbb{R}^d$, $\langle|b_{n_1}-b_{n_2}|^2\rho_x\rangle\to 0$ as n_1, $n_2\to\infty$.*

Proof. First, let us show that

$$\lim_{R\to\infty}\langle|b_{n_1}-b_{n_2}|^2\mathbf{1}_{\mathbb{R}^d\setminus B_{R+1}}\rho_x\rangle\to 0\quad\text{ uniformly in }n_1,\,n_2. \tag{15.16}$$

Indeed, replacing $\mathbf{1}_{\mathbb{R}^d\setminus B_{R+1}}$ by greater function η_R^2, where

$$\eta_R(y):=\xi_R(|y|),\quad \xi_R(r):=\begin{cases}0, & 0\leq r<R,\\ r-R, & R\leq r\leq R+1,\\ 1, & r>R+1,\end{cases}$$

and noting that $|\nabla \eta_R(y)| \leq \mathbf{1}_{R<|y|<R+1}$, we estimate

$$\begin{aligned}\langle |b_{n_1}-b_{n_2}|^2 \mathbf{1}_{\mathbb{R}^d\setminus B_{R+1}}\rho_x\rangle &\leq \langle |b_{n_1}-b_{n_2}|^2\eta_R^2\rho_x\rangle \\ &\leq 4\delta\|\nabla(\eta_R\sqrt{\rho_x})\|_2^2+4c_\delta\|\eta_R\sqrt{\rho_x}\|_2^2,\end{aligned}$$

where $\|\eta_R\sqrt{\rho_x}\|_2\to 0$ as $R\to\infty$, and so, in view of the second estimate in (15.5), $\|\eta_R\nabla(\sqrt{\rho_x})\|_2\to 0$ as $R\to\infty$. Also, taking without loss of generality $x=0$,

$$\|(\nabla\eta_R)\sqrt{\rho}\|_2^2=\langle \mathbf{1}_{R\leq|\cdot|\leq R+1}\rho\rangle = CR^{-d-\varepsilon}R^d=R^{-\varepsilon}\to 0$$

as $R\to\infty$. This yields (15.16).

In turn, inside the ball B_R we have $\langle |b_{n_1}-b_{n_2}|^2\mathbf{1}_{B_R}\rho_x\rangle\to 0$ as $n_1,n_2\to\infty$ since $b_{n_1}-b_{n_2}\to 0$ in $L^2_{\rm loc}$. This ends the proof. □

16. Proof of Proposition 10.1 (Embedding property)

We will need the following lemma.

Lemma 16.1 ([G, Lemma 7.1]). *If $\{z_i\}_{i=0}^\infty\subset\mathbb{R}_+$ is a sequence of positive real numbers such that*

$$z_{i+1}\leq NC_0^iz_i^{1+\alpha}$$

for some $C_0>1$, $\alpha>0$, and

$$z_0\leq N^{-\frac{1}{\alpha}}C_0^{-\frac{1}{\alpha^2}}.$$

Then $\lim_i z_i=0$.

Throughout the proof, we write for brevity

$$b=b_n,\quad q=q_m,\quad Q^i=Q^i_m,\quad w=w_{n,m}.$$

It suffices to estimate the positive part of w:

$$\begin{aligned}\sup_{\mathbb{R}^d} w_+ &\leq K_1(\mu-\mu_0)^{-\frac{\beta}{p}}\big\langle |Q^i|^{p\theta'}(|\nabla f|^{p\theta'}+|f|^{p\theta'})\big\rangle^{\frac{1}{p\theta'}} \\ &\quad + K_2(\mu-\mu_0)^{-\frac{1}{p\theta}}\langle |Q^i|^{p\theta}(|\nabla f|^{p\theta}+|f|^{p\theta})\rangle^{\frac{1}{p\theta}}. \qquad (16.1)\end{aligned}$$

Step 1. First, we prove the following energy inequality: there exist generic constants $\mu_0\geq 0$ and C_0, C such that, for every $k\geq 0$ and all $\mu>\mu_0$, the positive part $v:=(w-k)_+$ of $w-k$ satisfies

$$(\mu-\mu_0)\|v^{\frac{p}{2}}\|_2^2+C_0\|\nabla v^{\frac{p}{2}}\|_2^2\leq C\big[\langle |Q^i|^p|\nabla f|^p\mathbf{1}_{v>0}\rangle+\langle(1+|Q^i|^p)|f|^p\mathbf{1}_{v>0}\rangle\big]. \qquad (16.2)$$

Proof of (16.2). We obtain from equation (10.1), using that both μ and k are non-negative, $(\mu-\Delta+(b+q)\cdot\nabla)(w-k)\leq(b^i+q^i)f$. We now basically apply the energy inequality of Proposition 15.1(i) with $s=p$. The fact that we have an elliptic differential inequality instead of an elliptic equation does not change anything since we multiply it by a non-negative function v^{p-1}. So, we integrate, apply $b\in\mathbf{F}_\delta$ and use anti-symmetry of the stream matrix Q of drift q:

$$(\mu-\mu_0)\langle v^p\rangle+C_1\langle|\nabla v^{\frac{p}{2}}|^2\rangle\leq\langle(b^i+q^i)f,v^{p-1}\rangle. \qquad (16.3)$$

Let us estimate the terms in the right-hand side. First,

$$
\begin{aligned}
|\langle b^i f, v^{p-1}\rangle| &\le \varepsilon\langle |b^i|^2, v^p\rangle + \frac{1}{4\varepsilon}\langle |f|^2 v^{p-2}\rangle \\
&\le \varepsilon\Big(\delta\langle |\nabla v^{\frac{p}{2}}|^2\rangle + c_\delta\langle v^p\rangle\Big) \\
&\quad + \frac{1}{4\varepsilon}\Big(\frac{p-2}{p}\langle v^p\rangle + \frac{2}{p}\langle |f|^p \mathbf{1}_{v>0}\rangle\Big).
\end{aligned}
$$

Second,

$$
\begin{aligned}
|\langle q^i f, v^{p-1}\rangle| &= |\langle \operatorname{div} Q^i, f v^{p-1}\rangle| \\
&\le |\langle Q^i, (\nabla f) v^{p-1}\rangle| + \frac{2(p-1)}{p}|\langle Q^i, f v^{\frac{p}{2}-1}\nabla v^{\frac{p}{2}}\rangle| =: K_1 + \frac{2(p-1)}{p}K_2,
\end{aligned}
$$

where

$$
\begin{aligned}
K_1 &\le \frac{1}{2}\langle |Q^i|^2 |\nabla f|^2 v^{p-2}\rangle + \frac{1}{2}\langle v^p\rangle \\
&\le \frac{1}{2}\Big(\frac{2}{p}\langle |Q^i|^p |\nabla f|^p \mathbf{1}_{v>0}\rangle + \frac{p-2}{p}\langle v^p\rangle\Big) + \frac{1}{2}\langle v^p\rangle
\end{aligned}
$$

and

$$
\begin{aligned}
K_2 &\le \varepsilon\langle |\nabla v^{\frac{p}{2}}|^2\rangle + \frac{1}{4\varepsilon}\langle |Q^i|^2 f^2 v^{p-2}\rangle \\
&\le \varepsilon\langle |\nabla v^{\frac{p}{2}}|^2\rangle + \frac{1}{4\varepsilon}\Big(\frac{2}{p}\langle |Q^i|^p |f|^p \mathbf{1}_{v>0}\rangle + \frac{p-2}{p}\langle v^p\rangle\Big).
\end{aligned}
$$

Thus,

$$
|\langle q^i f, v^{p-1}\rangle| \le \frac{2(p-1)}{p}\varepsilon\langle |\nabla v^{\frac{p}{2}}|^2\rangle + c_1\big(\langle |Q^i|^p |\nabla f|^p \mathbf{1}_{v>0}\rangle + \langle |Q^i|^p |f|^p \mathbf{1}_{v>0}\rangle\big) + c_2\langle v^p\rangle,
$$

where c_1, c_2 depend on ε and p.

Substituting the resulting estimates in (16.3) and selecting ε sufficiently small, we obtain estimate (16.2).

Step 2. In what follows, we will be selecting $k > 0$. Then $|\{v > 0\}| < \infty$. We obtain from (16.2), using the Sobolev embedding theorem,

$$
(\mu - \mu_0)\|v\|_p^p + C_S\|v\|^p_{\frac{pd}{d-2}} \le C\big\langle\big(|Q^i|^p|\nabla f|^p + (1+|Q^i|^p)|f|^p\big)\mathbf{1}_{v>0}\big\rangle. \tag{16.4}
$$

We estimate the left-hand side of (16.4) using the interpolation inequality:

$$
(\mu - \mu_0)^\beta \|v\|^p_{p\theta_0} \le \beta(\mu - \mu_0)\|v\|_p^p + (1-\beta)\|v\|^p_{L^{\frac{pd}{d-2}}}, \quad 0 < \beta < 1, \quad \frac{1}{p\theta_0} = \beta\frac{1}{p} + (1-\beta)\frac{d-2}{pd},
$$

where $1 < \theta_0 < \frac{d}{d-2}$. So,

$$
(\mu - \mu_0)^\beta \|v\|^p_{p\theta_0} \le C_2\big[\big\langle\big(|Q^i|^p|\nabla f|^p + (1+|Q^i|^p)|f|^p\big)\mathbf{1}_{v>0}\big\rangle\big]
$$

Let us fix β small enough so that we have $\theta_0 > \theta$. (Recall that $1 < \theta < \frac{d}{d-2}$ was fixed in the statement of the proposition.) Applying Hölder's inequality, we obtain

$$
(\mu - \mu_0)^\beta \|v\|^p_{p\theta_0} \le C_3 H |\{v > 0\}|^{\frac{1}{\theta}},
$$

where

$$H := \left\langle \left(|Q^i|^{p\theta'}|\nabla f|^{p\theta'} + (1+|Q^i|^{p\theta'})|f|^{p\theta'}\right)\mathbf{1}_{v>0}\right\rangle^{\frac{1}{\theta'}}.$$

On the other hand, again by Hölder's inequality,

$$\|v\|_{p\theta}^{p\theta} \leq \|v\|_{p\theta_0}^{p\theta}|\{v>0\}|^{1-\frac{\theta}{\theta_0}}.$$

Therefore, we obtain

$$\|v\|_{p\theta}^{p\theta} \leq \tilde{C}(\mu-\mu_0)^{-\beta\theta}H^\theta|\{v>0\}|^{2-\frac{\theta}{\theta_0}}.$$

Step 3. Now, put $v_m := (w-k_m)_+$, $k_m := \xi(1-2^{-m}) \uparrow \xi$, where constant $\xi > 0$ will be chosen later.

Remark 16.1. We have $k_0 = 0$, but we will not encounter the volume of $\{u > 0\}$ in the proof (clearly, $|\{u>0\}|$ can be infinite).

So,

$$\frac{1}{\xi^{p\theta}}\|v_{m+1}\|_{p\theta}^{p\theta} \leq \tilde{C}\frac{1}{\xi^{p\theta}}(\mu-\mu_0)^{-\beta\theta}H^\theta|\{u>k_{m+1}\}|^{2-\frac{\theta}{\theta_0}}.$$

From now on, we require constant ξ to satisfy $\xi^p \geq (\mu-\mu_0)^{-\beta}H$, so

$$\frac{1}{\xi^{p\theta}}\|v_{m+1}\|_{L^{p\theta}(B^{m+1})}^{p\theta} \leq \tilde{C}|\{u>k_{m+1}\}|^{2-\frac{\theta}{\theta_0}}.$$

Now,

$$\begin{aligned}|\{w>k_{m+1}\}| &= \left|\left\{\left(\frac{w-k_m}{k_{m+1}-k_m}\right)^{p\theta} > 1\right\}\right| \\ &\leq (k_{m+1}-k_m)^{-p\theta}\langle v_m^{p\theta}\rangle = \xi^{-p\theta}2^{p\theta(m+1)}\|v_m\|_{p\theta}^{p\theta},\end{aligned}$$

so, applying the previous two inequalities, we obtain

$$\frac{1}{\xi^{p\theta}}\|v_{m+1}\|_{p\theta}^{p\theta} \leq C2^{p\theta m(2-\frac{\theta}{\theta_0})}\left(\frac{1}{\xi^{p\theta}}\|v_m\|_{L^{p\theta}(B^m)}^{p\theta}\right)^{2-\frac{\theta}{\theta_0}}.$$

Step 4. Denote $z_m := \frac{1}{\xi^{p\theta}}\|v_m\|_{p\theta}^{p\theta}$. Then

$$z_{m+1} \leq C\gamma^m z_m^{1+\alpha}, \quad m = 0,1,2,\dots, \quad \alpha := 1-\frac{\theta}{\theta_0}, \ \ \gamma := 2^{p\theta(2-\frac{\theta}{\theta_0})}$$

and $z_0 = \frac{1}{\xi^{p\theta}}\langle w_+^{p\theta}\rangle \leq C^{-\frac{1}{\alpha}}\gamma^{-\frac{1}{\alpha^2}}$ provided that we fix ξ by

$$\xi^{p\theta} := C^{\frac{1}{\alpha}}\gamma^{\frac{1}{\alpha^2}}\langle w_+^{p\theta}\rangle + (\mu-\mu_0)^{-\beta\theta}H^\theta$$

(so that it also satisfies the previous requirement $\xi^p \geq (\mu-\mu_0)^{-\beta}H$). Hence, by Lemma 16.1, $z_m \to 0$ as $m \to \infty$. Therefore, $w_+ \leq \xi$. Thus, we obtain inequality

$$\sup_{\mathbb{R}^d} w_+ \leq K\left(\langle w_+^{p\theta}\rangle^{\frac{1}{p\theta}} + (\mu-\mu_0)^{-\frac{\beta}{p}}\left\langle\left(|Q^i|^{p\theta'}|\nabla f|^{p\theta'} + (1+|Q^i|^{p\theta'})|f|^{p\theta'}\right)\mathbf{1}_{v>0}\right\rangle^{\frac{1}{p\theta'}}\right). \tag{16.5}$$

Step 5. It remains to estimate $\langle w_+^{p\theta}\rangle^{\frac{1}{p\theta}}$. We already did this in (16.2) (use $p\theta > p$):

$$(\mu-\mu_0)^{\frac{1}{p\theta}}\|w_+\|_{p\theta} \leq C^{\frac{1}{p\theta}}\left[\langle|Q^i|^{p\theta}|\nabla f|^{p\theta}\rangle + \langle(1+|Q^i|^{p\theta})f^{p\theta}\rangle\right]^{\frac{1}{p\theta}}.$$

This inequality, applied in (16.5), yields (16.1) and thus ends the proof of Proposition 10.1. □

17. Proof of Proposition 10.2 (Hölder continuity)

The following is a well-known consequence of the John-Nirenberg inequality:

Proposition 17.1. *There exists constant $C = C(d)$ such that, for every $g \in \mathrm{BMO}$,*

$$\sup_{x\in\mathbb{R}^d, R>0} \frac{1}{|B_R(x)|} \langle e^{\frac{C|g-(g)_{B_R(x)}|}{\|g\|_{\mathrm{BMO}}}} \mathbf{1}_{B_R(x)} \rangle \leq C. \tag{17.1}$$

In particular, for every $1 \leq s < \infty$, we have $g \in L^s_{\mathrm{loc}}$ and

$$\sup_{x\in\mathbb{R}^d, R>0} \frac{1}{|B_R(x)|} \langle |g-(g)_{B_R(x)}|^s \mathbf{1}_{B_R(x)} \rangle \leq C(d,s)\|g\|^s_{\mathrm{BMO}}. \tag{17.2}$$

Put for brevity $b = b_n$, $q = q_m$, $Q = Q_m$ and $u = u_{n,m}$. So,

$$\big(\mu - \Delta + (b+q)\cdot\nabla\big)u = f.$$

Let us fix some $1 < \theta < \frac{d}{d-2}$. Let $p > \frac{2}{2-\sqrt{\delta}}$. By our assumption $\delta < 4$, so such p exist.

Lemma 17.1 (Caccioppoli's inequality)**.** *Let $v := (u-k)_+$, $k \in \mathbb{R}$. For all $0 < r < R \leq 1$, we have*

$$\|\nabla v^{\frac{p}{2}}\|^2_{L^2(B_r)} \leq \frac{K_1}{(R-r)^2}|B_R|^{\frac{1}{\theta'}}(1+\|Q\|^2_{\mathrm{BMO}})\|v^{\frac{p}{2}}\|^2_{L^{2\theta}(B_R)} + K_2\||f-\mu u|^{\frac{p}{2}}\mathbf{1}_{v>0}\|^2_{L^2(B_R)} \tag{17.3}$$

for constants K_1, K_2 independent of k, r, R and n, m.

Proof. Let $\{\eta = \eta_{r,R}\}$ be a family of $[0,1]$-valued smooth cut-off functions satisfying

$$\eta = 1 \text{ in } B_r, \quad \eta = 0 \text{ in } \mathbb{R}^d \setminus B_R, \tag{17.4}$$

$$|\nabla\eta| \leq \frac{c}{R-r}\mathbf{1}_{B_R}, \quad \frac{|\nabla\eta|^2}{\eta} \leq \frac{c}{(R-r)^2}\mathbf{1}_{B_R} \tag{17.5}$$

with constant c independent of r, R. We rewrite the equation for u in the form $\big(-\Delta+(b+q)\cdot\nabla\big)u = f - \mu u$, multiply it by $v^{p-1}\eta$ and integrate, obtaining

$$\begin{aligned}\frac{4(p-1)}{p^2}&\langle\nabla v^{\frac{p}{2}}, (\nabla v^{\frac{p}{2}})\eta\rangle + \frac{2}{p}\langle\nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\nabla\eta\rangle \\ &\leq -\frac{2}{p}\langle b\cdot\nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\eta\rangle - \frac{2}{p}\langle q\cdot\nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\eta\rangle + \langle f-\mu u, v^{p-1}\eta\rangle.\end{aligned}$$

Hence, by Cauchy-Schwarz,

$$\begin{aligned}\left(\frac{4(p-1)}{p} - \frac{4}{p}\epsilon\right)\langle|\nabla v^{\frac{p}{2}}|^2\eta\rangle &\leq \frac{p}{4\epsilon}\langle v^p\frac{|\nabla\eta|^2}{\eta}\rangle - 2\langle b\cdot\nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\eta\rangle - 2\langle q\cdot\nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\eta\rangle + p\langle f-\mu u, v^{p-1}\eta\rangle \\ &=: I_1 + I_2 + I_3 + I_4.\end{aligned} \tag{17.6}$$

Let us estimate terms I_1-I_4. We start with the term I_3 containing the distribution-valued vector field $q = \nabla Q$. The other terms I_1, I_2 and I_4 will be estimated in such a way as to fit the estimate on I_3. The following argument was used in [H]. Set $\tilde{Q}^{ij} := Q^{ij} - (Q^{ij})_R$, where, recall, $(Q^{ij})_R = (Q^{ij})_{B_R}$ is the average of Q^{ij} over ball B_R. We have

$$
\begin{aligned}
I_3 = -2\langle q \cdot \nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\eta\rangle &= -\sum_{i=1}^d \langle q^i \nabla_i v^p, \eta\rangle \\
&\text{(use identity } q^i = \sum_{j=1}^d \nabla_j Q^{ij} = \sum_{j=1}^d \nabla_j \tilde{Q}^{ij} \text{ and integrate by parts)} \\
&= \sum_{i,j=1}^d \langle \tilde{Q}^{ij} \nabla_j \nabla_i v^p, \eta\rangle + \sum_{i,j=1}^d \langle \tilde{Q}^{ij} \nabla_i v^p, \nabla_j \eta\rangle \\
&= \sum_{i,j=1}^d \langle \tilde{Q}^{ij} \nabla_j \nabla_i v^p, \eta\rangle + 2\sum_{i,j=1}^d \langle \tilde{Q}^{ij} \nabla_i v^{\frac{p}{2}}, v^{\frac{p}{2}} \nabla_j \eta\rangle.
\end{aligned}
$$

Due to the anti-symmetry of Q, the first sum on the right hand side is zero, so $I_3 = 2\langle \tilde{Q} \cdot \nabla v^{\frac{p}{2}}, v^{\frac{p}{2}} \nabla \eta\rangle$. Hence

$$
|I_3| \leq \varepsilon_1 \langle |\nabla v^{\frac{p}{2}}|^2 \eta\rangle + \frac{1}{\varepsilon_1} \langle |\tilde{Q}|^2 v^p \frac{|\nabla \eta|^2}{\eta}\rangle. \tag{17.7}
$$

The second term in the RHS of (17.7) is bounded as follows, using (17.5):

$$
\begin{aligned}
\langle |\tilde{Q}|^2 v^p \frac{|\nabla \eta|^2}{\eta}\rangle &\leq \frac{c}{(R-r)^2} \left\langle |\tilde{Q}|^{2\theta'} \mathbf{1}_{B_R} \right\rangle^{\frac{1}{\theta'}} \langle v^{p\theta} \mathbf{1}_{B_R}\rangle^{\frac{1}{\theta}} \\
&\text{(use (17.2))} \\
&\leq \frac{C}{(R-r)^2} |B_R|^{\frac{1}{\theta'}} \|Q\|^2_{\mathrm{BMO}} \langle v^{p\theta} \mathbf{1}_{B_R}\rangle^{\frac{1}{\theta}}.
\end{aligned} \tag{17.8}
$$

Thus, to summarize, we have

$$
|I_3| \leq \varepsilon_1 \langle |\nabla v^{\frac{p}{2}}|^2 \eta\rangle + \frac{1}{\varepsilon_1} \frac{C}{(R-r)^2} |B_R|^{\frac{1}{\theta'}} \|Q\|^2_{\mathrm{BMO}} \langle v^{p\theta} \mathbf{1}_{B_R}\rangle^{\frac{1}{\theta}}.
$$

Remark 17.1. If the entries of Q are in L^∞, then we can take $\theta = 1$. Indeed, in this case we can obtain (17.8) directly: $\langle |\tilde{Q}|^2 v^p \frac{|\nabla \eta|^2}{\eta}\rangle \leq \frac{4c}{(R-r)^2} \|Q\|^2_\infty \langle v^p \mathbf{1}_{B_R}\rangle$.

Now, we estimate the remaining terms I_1, I_2 and I_4. By (17.5),

$$
I_1 \leq \frac{cp}{4\epsilon(R-r)^2} |B_R|^{\frac{1}{\theta'}} \langle v^{p\theta} \mathbf{1}_{B_R}\rangle^{\frac{1}{\theta}}.
$$

Next,

$$\frac{1}{2}|I_2| \leq \langle |b||\nabla v^{\frac{p}{2}}|, v^{\frac{p}{2}}\eta\rangle \leq \alpha\langle|\nabla v^{\frac{p}{2}}|\eta\rangle + \frac{1}{4\alpha}\langle |b|^2, v^p\eta\rangle, \quad \alpha = \frac{\sqrt{\delta}}{2}$$

(use $b \in \mathbf{F}_\delta$)

$$\leq \frac{\sqrt{\delta}}{2}\langle|\nabla v^{\frac{p}{2}}|^2\eta\rangle + \frac{1}{2\sqrt{\delta}}\Big(\delta\langle|\nabla(v^{\frac{p}{2}}\sqrt{\eta})|^2\rangle + c_\delta\langle v^p\eta\rangle\Big)$$

$$\leq \frac{\sqrt{\delta}}{2}\langle|\nabla v^{\frac{p}{2}}|^2\eta\rangle + \frac{\sqrt{\delta}}{2}\left((1+\varepsilon_0)\langle|\nabla v^{\frac{p}{2}}|^2\eta\rangle + \frac{1}{4}\big(1+\frac{1}{\varepsilon_0}\big)\langle v^p\frac{|\nabla\eta|^2}{\eta}\rangle\right) + \frac{c_\delta}{2\sqrt{\delta}}\langle v^p\eta\rangle$$

$$\leq \frac{\sqrt{\delta}}{2}(2+\varepsilon_0)\langle|\nabla v^{\frac{p}{2}}|\eta\rangle + \left(\frac{\sqrt{\delta}}{8}\big(1+\frac{1}{\varepsilon_0}\big)\frac{c}{(R-r)^2} + \frac{c_\delta}{2\sqrt{\delta}}\right)|B_R|^{\frac{1}{\theta'}}\langle v^{p\theta}\mathbf{1}_{B_R}\rangle^{\frac{1}{\theta}}.$$

Finally, by Young's inequality, for every $\varepsilon_2 > 0$,

$$|I_4| = |\langle f-\mu u, v^{p-1}\eta\rangle| \leq \frac{1}{p\varepsilon_2^p}\langle|f-\mu u|^p\mathbf{1}_{\{v>0\}}\mathbf{1}_{B_R}\rangle + \frac{\varepsilon_2^{p'}}{p'}|B_R|^{\frac{1}{\theta'}}\langle v^{p\theta}\mathbf{1}_{B_R}\rangle^{\frac{1}{\theta}}.$$

Now, applying these estimates on I_1-I_4 in (17.6), we obtain:

$$\left(\frac{4(p-1)}{p} - \frac{4}{p}\varepsilon - (2+\varepsilon_0)\sqrt{\delta} - \varepsilon_1\right)\langle|\nabla v^{\frac{p}{2}}|^2\eta\rangle \leq \frac{C_1}{(R-r)^2}|B_R|^{\frac{1}{\theta'}}(1+\|Q\|^2_{\mathrm{BMO}})\langle v^{p\theta}\mathbf{1}_{B_R}\rangle^{\frac{1}{\theta}} + C_2\||f-\mu u|^{\frac{p}{2}}\mathbf{1}_{v>0}\mathbf{1}_{B_R}\|_2^2,$$

where ε, ε_0, ε_1 are fixed sufficiently small so that the expression in the brackets in the LHS is strictly positive. The latter is possible since $\delta < 4$ and $p > \frac{2}{2-\sqrt{\delta}}$. This ends the proof of Lemma 17.1. □

Lemma 17.2. *Fix $\alpha > 0$ by $\alpha(\alpha+1) = 1 - \frac{\theta(d-2)}{d}$. Then, for all $0 < r < R \leq 1$,*

$$\sup_{B_{\frac{R}{2}}} u \leq C\left(\frac{1}{|B_R|}\langle u^{p\theta}\mathbf{1}_{B_R\cap\{u>0\}}\rangle\right)^{\frac{1}{p\theta}}\left(\frac{|B_R\cap\{u>0\}|}{|B_R|}\right)^{\frac{\alpha}{p\theta}} + R^{\frac{2}{p}} \tag{17.9}$$

for C independent of n, m, r and R.

Proof. Step 1. Fix a family of cut-off functions $\eta = \eta_{r,R} \in C_c^\infty$ such that

$$\eta = 1 \quad \text{on } B_r \qquad \eta = 0 \quad \text{on } \mathbb{R}^d \setminus \overline{B}_{\frac{R+r}{2}}, \tag{17.10}$$

and

$$|\nabla\eta| \leq \frac{c}{R-r}\mathbf{1}_{B_{\frac{r+R}{2}}}, \quad \frac{|\nabla\eta|^2}{\eta} \leq \frac{c}{(R-r)^2}\mathbf{1}_{B_{\frac{r+R}{2}}} \tag{17.11}$$

with constant c independent of r, R. Set $v = (u-k)_+$, where $k \in \mathbb{R}$ will be chosen later. Using Sobolev's embedding theorem, we have

$$\langle v^{\frac{pd}{d-2}}\mathbf{1}_{B_r}\rangle^{\frac{d-2}{d}} \leq \langle (v^{\frac{p}{2}}\eta^{\frac{1}{2}})^{\frac{2d}{d-2}}\mathbf{1}_{B_{\frac{R+r}{2}}}\rangle^{\frac{d-2}{d}}$$
$$\leq C_S^2\langle|\nabla(v^{\frac{p}{2}}\eta^{\frac{1}{2}})|^2\mathbf{1}_{B_{\frac{R+r}{2}}}\rangle$$
$$\leq 2C_S^2\left(\langle|\nabla v^{\frac{p}{2}}|^2\eta\mathbf{1}_{B_{\frac{R+r}{2}}}\rangle + \langle v^p\frac{|\nabla\eta|^2}{\eta}\mathbf{1}_{B_{\frac{R+r}{2}}}\rangle\right).$$

We rewrite the latter, after applying Hölder's inequality in the second term, as follows:

$$\|v^{\frac{p}{2}}\mathbf{1}_{B_r}\|^2_{\frac{2d}{d-2}} \leq C_1\left(\|(\nabla v^{\frac{p}{2}})\mathbf{1}_{B_{\frac{R+r}{2}}}\|_2^2 + \frac{|B_R|^{\frac{1}{\theta'}}}{(R-r)^2}\|v^{\frac{p}{2}}\mathbf{1}_{B_{\frac{R+r}{2}}}\|^2_{2\theta}\right).$$

Next, we apply Lemma 17.1 to the first term in the RHS, obtaining

$$\|v^{\frac{p}{2}}\mathbf{1}_{B_r}\|^2_{\frac{2d}{d-2}} \leq C_2\bigg(\frac{|B_R|^{\frac{1}{\theta'}}}{(R-r)^2}\|v^{\frac{p}{2}}\mathbf{1}_{B_R}\|^2_{2\theta} + \||f-\mu u|^{\frac{p}{2}}\mathbf{1}_{v>0}\mathbf{1}_{B_R}\|_2^2\bigg).$$

Next, applying Hölder's inequality in the second term in the RHS, we estimate

$$\||f-\mu u|^{\frac{p}{2}}\mathbf{1}_{v>0}\mathbf{1}_{B_R}\|_2^2 \leq \|f-\mu u\|^p_\infty |B_R\cap\{v>0\}|^{\frac{1}{\theta}}|B_R|^{\frac{1}{\theta'}},$$

which gives us, upon noting that $\|f-\mu u\|_\infty \leq 2\|f\|_\infty$,

$$\|v^{\frac{p}{2}}\mathbf{1}_{B_r}\|^2_{\frac{2d}{d-2}} \leq C_2\bigg(\frac{|B_R|^{\frac{1}{\theta'}}}{(R-r)^2}\|v^{\frac{p}{2}}\mathbf{1}_{B_R}\|^2_{2\theta} + |B_R|^{\frac{1}{\theta'}}2^p\|f\|^p_\infty|B_R\cap\{u>k\}|^{\frac{1}{\theta}}\bigg).$$

Representing $|B_R|^{\frac{1}{\theta'}} = |B_R|^{\frac{d-2}{d}+\frac{2}{d}-\frac{1}{\theta}}$, and dividing both sides of the previous inequality by $|B_R|^{\frac{d-2}{d}}$, we have

$$\begin{aligned}\frac{1}{|B_R|^{\frac{d-2}{d}}}\|v^{\frac{p}{2}}\mathbf{1}_{B_r}\|^2_{\frac{2d}{d-2}} \leq C_2\bigg(&|B_R|^{\frac{2}{d}}\frac{1}{(R-r)^2}\frac{1}{|B_R|^{\frac{1}{\theta}}}\|v^{\frac{p}{2}}\mathbf{1}_{B_R}\|^2_{2\theta} \\ &+ |B_R|^{\frac{2}{d}}2^p\|f\|^p_\infty\left(\frac{|B_R\cap\{u>k\}|}{|B_R|}\right)^{\frac{1}{\theta}}\bigg).\end{aligned} \tag{17.12}$$

Next, note that if $h<k$, then $(u-k)_+ \leq (u-h)_+$. Therefore, by Chebyshev's inequality,

$$|B_R\cap\{u>k\}|^{\frac{1}{\theta}} \leq \frac{1}{(k-h)^p}\langle (u-h)_+^{p\theta}\mathbf{1}_{B_R}\rangle^{\frac{1}{\theta}}.$$

Recalling that $v=(u-k)_+$ and applying the last inequality in (17.12), we obtain

$$\left(\frac{\langle (u-k)_+^{\frac{pd}{d-2}}\mathbf{1}_{B_r}\rangle}{|B_R|}\right)^{\frac{d-2}{d}} \leq C_3|B_R|^{\frac{2}{d}}\left(\frac{1}{(R-r)^2}+\frac{1}{(k-h)^p}\right)\left(\frac{\langle (u-h)_+^{p\theta}\mathbf{1}_{B_R}\rangle}{|B_R|}\right)^{\frac{1}{\theta}}.$$

Now, using Hölder's inequality and then applying the previous estimate, we get

$$\begin{aligned}\frac{\langle (u-k)_+^{p\theta}\mathbf{1}_{B_r}\rangle}{|B_R|} &\leq \left(\frac{\langle (u-k)_+^{\frac{pd}{d-2}}\mathbf{1}_{B_r}\rangle}{|B_R|}\right)^{\frac{\theta(d-2)}{d}}\left(\frac{|B_r\cap\{u>k\}|}{|B_R|}\right)^{1-\frac{\theta(d-2)}{d}} \\ &\leq C_3^\theta|B_R|^{\frac{2\theta}{d}}\left(\frac{1}{(R-r)^2}+\frac{1}{(k-h)^p}\right)^\theta\left(\frac{\langle (u-h)_+^{p\theta}\mathbf{1}_{B_R}\rangle}{|B_R|}\right)\left(\frac{|B_R\cap\{u>k\}|}{|B_R|}\right)^{1-\frac{\theta(d-2)}{d}}\end{aligned}$$

Multiplying this inequality by $\left(\frac{|B_r\cap\{u>k\}|}{|B_R|}\right)^\alpha\left[\leq \frac{1}{(k-h)^{p\theta\alpha}}\left(\frac{\langle (u-h)_+^{p\theta}\mathbf{1}_{B_R}\rangle}{|B_R|}\right)^\alpha\right]$, and selecting $\alpha>0$ as in the statement of the lemma, i.e. such that $\alpha(\alpha+1)=1-\frac{\theta(d-2)}{d}$, we get

$$\begin{aligned}&\frac{\left\langle (u-k)_+^{p\theta}\mathbf{1}_{B_r}\right\rangle}{|B_R|}\left(\frac{|B_r\cap\{u>k\}|}{|B_R|}\right)^\alpha \\ &\leq C_3^\theta|B_R|^{\frac{2\theta}{d}}\left(\frac{1}{(R-r)^2}+\frac{1}{(k-h)^p}\right)^\theta\frac{1}{(k-h)^{p\theta\alpha}}\left(\frac{\langle (u-h)_+^{p\theta}\mathbf{1}_{B_R}\rangle}{|B_R|}\left(\frac{|B_R\cap\{u>k\}|}{|B_R|}\right)^\alpha\right)^{1+\alpha}\end{aligned}$$

At the next step, we are going to iterate this inequality.

Step 2. Now, define

$$r_m := \frac{R}{2}\left(1 + \frac{1}{2^m}\right), \quad B_m := B_{r_m}$$

$$k_m := \xi(1 - 2^{-m}),$$

for a positive constant ξ to be determined later. Setting

$$z_m \equiv z(k_m, r_m) := \frac{\left\langle (u-k_m)_+^{p\theta}\mathbf{1}_{B_m}\right\rangle}{|B_R|}\left(\frac{|B_m \cap \{u > k_m\}|}{|B_R|}\right)^\alpha,$$

we rewrite the previous inequality, upon selecting $k := k_{m+1}$ and $h := k_m$ there, as

$$z_{m+1} \leq C_3^\theta \frac{|B_R|^{\frac{2\theta}{d}}}{R^{2\theta}}\left(2^{2m} + 2^{mp}\frac{R^2}{\xi^p}\right)^\theta \frac{2^{mp\theta\alpha}}{\xi^{p\theta\alpha}} z_m^{1+\alpha}. \tag{17.13}$$

In what follows, we restrict out choice of constant ξ to those satisfying

$$\xi^p \geq R^2. \tag{17.14}$$

Then, since $p \geq 2$,

$$z_{m+1} \leq \big(\frac{C_4}{\xi^{p\alpha}}\big)^\theta 2^{mp\theta(1+\alpha)} z_m^{1+\alpha}.$$

Therefore, setting $C_0 = 2^{p\theta(1+\alpha)}$ and $N = (\frac{C_4}{\xi^{p\alpha}})^\theta$, we have the first inequality in Lemma 16.1, i.e. $z_{m+1} \leq NC_0^i z_m^{1+\alpha}$. To apply this lemma, we need to verify the second inequality there, i.e.

$$z_0 \leq N^{-\frac{1}{\alpha}} C_0^{-\frac{1}{\alpha^2}},$$

where, recall, $z_0 = \frac{\langle u^{p\theta}\mathbf{1}_{B_R \cap \{u>0\}}\rangle}{|B_R|}\big(\frac{|B_R \cap \{u>0\}|}{|B_R|}\big)^\alpha$. The previous inequality holds, by the definition of N, if we select ξ satisfying

$$\xi \geq 2^{\frac{1+\alpha}{\alpha^2}} C_4^{\frac{1}{p\alpha}} z_0^{\frac{1}{p\theta}} \tag{17.15}$$

We combine (17.14) and (17.15) by taking $\xi = 2^{\frac{1+\alpha}{\alpha^2}} C_4^{\frac{1}{p\alpha}} z_0^{\frac{1}{p\theta}} + R^{\frac{2}{p}}$. Now Lemma 16.1 yields $z(\xi, \frac{R}{2}) = 0$, i.e. $\sup_{B_{\frac{R}{2}}} u \leq \xi$. Hence, $\sup_{B_{\frac{R}{2}}} u \leq C z_0^{\frac{1}{p\theta}} + R^{\frac{2}{p}}$, as claimed. The proof of Lemma 17.2 is completed. □

Set

$$\operatorname{osc}(u, R) := \sup_{y,y' \in B_R} |u(y) - u(y')|.$$

Lemma 17.3. *Fix k_0 by $2k_0 = M(2R) + m(2R) := \sup_{B_{2R}} u + \inf_{B_{2R}} u$. Assume that $|B_R \cap \{u > k_0\}| \leq \gamma |B_R|$ for some $\gamma < 1$. If*

$$\operatorname{osc}(u, 2R) \geq 2^{n+1} C R^{\frac{2}{p}}, \tag{17.16}$$

then, for $k_n := M(2R) - 2^{-n-1}\operatorname{osc}(u, 2R)$,

$$\frac{|B_R \cap \{u > k_n\}|}{|B_R|} \leq c n^{-\frac{d}{2(d-1)}}.$$

Proof of Lemma 17.3. Let $h \in (k_0, k)$. Define

$$w := \begin{cases} (u-h)^{\frac{p}{2}} & \text{if } h < u < k, \\ (k-h)^{\frac{p}{2}} & \text{if } u \geq k, \\ 0 & \text{if } u \leq h. \end{cases}$$

Note that $w = 0$ in $B_R \setminus (B_R \cap \{u > k_0\})$. The measure of this set is greater than $\gamma |B_R|$, so the Sobolev embedding theorem yields

$$\begin{aligned}(k-h)^{\frac{p}{2}} \left|B_R \cap \{u > k\}\right|^{\frac{d-1}{d}} &\leq c_1 \langle w^{\frac{d}{d-1}} \mathbf{1}_{B_R} \rangle^{\frac{d-1}{d}} \\ &\leq c_2 \langle |\nabla w| \mathbf{1}_{\Delta} \rangle \\ &\leq c_2 |\Delta|^{\frac{1}{2}} \langle |\nabla (u-h)^{\frac{p}{2}}|^2 \mathbf{1}_{B_R \cap \{u > h\}} \rangle^{\frac{1}{2}},\end{aligned}$$

where $\Delta := B_R \cap \{u > h\} \setminus (B_R \cap \{u > k\})$. On the other hand, repeating the proof of Lemma 17.1, but estimating the term I_3 there via (17.7) rather than going all the way to (17.8), we obtain

$$\begin{aligned}\langle |\nabla (u-h)^{\frac{p}{2}}|^2 \mathbf{1}_{B_R \cap \{u>h\}} \rangle &\leq \frac{C_1}{R^2} \langle (u-h)^p \mathbf{1}_{B_{2R} \cap \{u>h\}} \rangle \\ &+ \frac{C_2}{R^2} \langle |Q - (Q)_R|^2 (u-h)^p \mathbf{1}_{B_{2R} \cap \{u>h\}} \rangle \\ &+ C_3 \langle |f - \mu u|^p \mathbf{1}_{B_{2R} \cap \{u>h\}} \rangle\end{aligned}$$

Applying the John-Nirenberg inequality (17.2) in the second term, we get

$$\begin{aligned}\langle |\nabla (u-h)^{\frac{p}{2}}|^2 \mathbf{1}_{B_R \cap \{u>h\}} \rangle &\leq C_4 R^{d-2} (M(2R) - h)^p + \frac{C_5}{R^2} (M(2R)-h)^p \cdot R^d \|Q\|^2_{\mathrm{BMO}} \\ &+ C_6 \|f - \mu u\|^p_\infty R^d \\ &\leq C R^{d-2} (M(2R) - h)^p + C R^d,\end{aligned}$$

(we have used $\|f - \mu u\|_\infty \leq 2\|f\|_\infty$). For $h \leq k_n$, we have $M(2R) - h \geq M(2R) - k_n = 2^{-n-1} \mathrm{osc}(u, 2R) \geq C R^{\frac{2}{p}}$, where we have used (17.16), in which case

$$(k-h)^{\frac{p}{2}} |B_R \cap \{u > k\}|^{\frac{d-1}{d}} \leq C |\Delta|^{\frac{1}{2}} R^{\frac{d-2}{2}} (M(2R) - h)^{\frac{p}{2}}.$$

Now, choosing increasing finite sequence $k = k_i := M(2R) - 2^{-i-1} \mathrm{osc}(u, 2R)$ for $i \in \{1, 2, \cdots, n\}$ and $h = k_{i-1}$. Then

$$M(2R) - h = 2^{-i} \mathrm{osc}(u, 2R), \quad |k - h| = 2^{-i-1} \mathrm{osc}(u, 2R)$$

so

$$|B_R \cap \{u > k_n\}|^{\frac{2(d-1)}{d}} \leq |B_R \cap \{u > k_i\}|^{\frac{2(d-1)}{d}} \leq C |\Delta_i| R^{d-2},$$

where $\Delta_i = B_R \cap \{u > k_i\} \setminus (B_R \cap \{u > k_{i-1}\})]$. Summing up over i, we obtain

$$n \left|B_R \cap \{u > k_n\}\right|^{\frac{2(d-1)}{d}} \leq C R^{d-2} \left|B_R \cap \{u > k_0\}\right| \leq C' R^{2(d-1)},$$

and the claimed inequality follows. This ends the proof of Lemma 17.3. □

We are in position to end the proof of Proposition 10.2. Fix $k_0 = \frac{1}{2}(M(2R) + m(2R))$. Without loss of generality, $|B_R \cap \{u > k_0\}| \leq \frac{1}{2}|B_R|$ (otherwise replace u by $-u$). Set $k_n = M(2R) - 2^{-n-1} \mathrm{osc}(u, 2R) > k_0$. By Lemma 17.2 applied to $u - k_n$ (k_n adds constant term μk_n in the equation, but since $\{k_n\}$ is bounded, the constant in Lemma 17.2 can be chosen to be independent of n), we have

$$\sup_{B_{\frac{R}{2}}}(u-k_n) \le C_1\left(\frac{1}{|B_R|}\langle (u-k_n)^{p\theta}\mathbf{1}_{B_R\cap\{u>k_n\}}\right)^{\frac{1}{p\theta}}\left(\frac{|B_R\cap\{u>k_n\}|}{|B_R|}\right)^{\frac{\alpha}{p\theta}}+R^{\frac{2}{p}}$$
$$\le C_1\sup_{B_R}(u-k_n)\left(\frac{|B_R\cap\{u>k_n\}|}{|B_R|}\right)^{\frac{1+\alpha}{p\theta}}+R^{\frac{2}{p}}. \tag{17.17}$$

Fix n by $cn^{-\frac{d}{2(d-1)}} \le (\frac{1}{2C_1})^{\frac{p\theta}{1+\alpha}}$. We consider two cases. First, let $\operatorname{osc}(u,2R)\ge 2^{n+1}R^{\frac{2}{p}}$. Then by Lemma 17.3 (with, say, $C=1$) applied in the RHS of (17.17),

$$M(R/2)-k_n \le \frac{1}{2}(M(2R)-k_n)+R^{\frac{2}{p}}$$

so

$$M(R/2)\le M(2R)-\frac{1}{2^{n+1}}\operatorname{osc}(u,2R)+\frac{1}{2}\frac{1}{2^{n+1}}\operatorname{osc}(u,2R)+R^{\frac{2}{p}},$$

which yields

$$M(R/2)-m(R/2)\le M(2R)-m(2R)-\frac{1}{2}\frac{1}{2^{n+1}}\operatorname{osc}(u,2R)+R^{\frac{2}{p}}$$
$$=\left(1-\frac{1}{2^{n+2}}\right)\operatorname{osc}(u,2R)+R^{\frac{2}{p}}.$$

Next, if $\operatorname{osc}(u,2R)\le 2^{n+1}R^{\frac{2}{p}}$, then

$$\operatorname{osc}(u,R/2)\le(1-\frac{1}{2^{n+2}})\operatorname{osc}(u,2R)+\frac{1}{2}R^{\frac{2}{p}}.$$

This result provides the desired Hölder continuity of u by applying the next Lemma 17.4 with $\tau=\frac{1}{4}$, $\delta=\log_\tau(1-2^{-n-1})$, and $0<\beta<\frac{2-p}{p}\wedge\delta$. Note that the second inequality in Lemma 17.3 is satisfied when $q=1$ and φ is non-decreasing, which is our situation. $\square$

Lemma 17.4 ([G, Lemma 7.3])**.** *Let $\varphi(t)$ be a positive function, and assume that there exists a constant q and a number $0<\tau<1$ such that for every $0<R<R_0$,*

$$\varphi(\tau R)\le\tau^\delta\varphi(R)+BR^\beta$$

with $0<\beta<\delta$, and

$$\varphi(t)\le q\varphi(\tau^k R)$$

for every t in the interval $(\tau^{k+1}R,\tau^k R)$. Then, for every $0<\rho<R<R_0$, we have

$$\varphi(\rho)\le C\left(\frac{\rho}{R}\right)^\beta\varphi(R)+B\rho^\beta$$

with a constant C that depends only on q, τ, δ, and β.

18. Proof of Proposition 10.4 (Separation property)

Step 1. Without loss of generality, $x=0$. We will need the following local estimate on solution $u=u_{n,m}$ of (10.3): for all $\mu \geq \mu_0>0$,

$$\sup_{B_{\frac{1}{2}}} |u| \leq K\biggl(\langle |u|^{p\theta}\mathbf{1}_{B_1}\rangle^{\frac{1}{p\theta}}+\langle |f|^{p\theta'}\mathbf{1}_{B_1}\rangle^{\frac{1}{p\theta'}}\biggr). \tag{18.1}$$

In fact, it suffices to prove the previous estimate for $\sup_{B_{\frac{1}{2}}} u_+$ in the LHS.

To that end, we first establish the following Caccioppoli's inequality for $v:=(u-k)_+$, $k\geq 0$:

$$\|v\|^p_{L^{\frac{pd}{d-2}}(B_r)} \leq C\biggl[(R-r)^{-2}|B_R|^{\frac{1}{\theta'}}\|v\|^p_{L^{p\theta}(B_R)}+\|f\mathbf{1}_{u>k}\|^p_{L^p(B_R)}\biggr]. \tag{18.2}$$

To prove (18.2), we argue as in the proof of Lemma 17.1, but treat the term μu differently: since μ and k are non-negative, we have

$$\mu(u-k)-\Delta(u-k)+(b+q)\cdot\nabla(u-k)\leq f.$$

Therefore, multiplying the previous inequality by $v^{p-1}\eta$, with the cutoff function η defined by (17.4), (17.5), and repeating the proof of Lemma 17.1, we obtain, for all $0<r<R\leq 1$,

$$\|v^{\frac{p}{2}}\|^2_{W^{1,2}(B_r)} \leq C_1\biggl[\frac{1}{(R-r)^2}|B_R|^{\frac{1}{\theta'}}\|v^{\frac{p}{2}}\|_{L^{2\theta}(B_R)}+\|f\mathbf{1}_{u>k}\|^p_{L^p(B_R)}\biggr].$$

The Sobolev embedding theorem now yields (18.2).

Set

$$R_m:=\frac{1}{2}+\frac{1}{2^{m+1}}, \quad m\geq 0,$$

so $B^m:=B_{R_m}$ is a decreasing sequence of balls converging to the ball of radius $\frac{1}{2}$. For the purposes of this proof, we can estimate $|B_R|^{\frac{1}{\theta'}}\leq 1$, which will make the iterations below converge slower, but will not change the sought estimate (18.1). Estimate (18.2) gives us

$$\begin{aligned}\|v\|^p_{L^{\frac{pd}{d-2}}(B_r)} &\leq C_1 2^{2m}\|v\|^p_{L^{p\theta}(B^m)}+C_2\|f\mathbf{1}_{u>k}\|^p_{L^p(B^m)}\\ &\leq C_1 2^{2m}\|v\|^p_{L^{p\theta}(B^m)}+C_2 H|B^m\cap\{v>0\}|^{\frac{1}{\theta}},\end{aligned} \tag{18.3}$$

where $H:=\langle |f|^{p\theta'}\mathbf{1}_{B^0}\rangle^{\frac{1}{\theta'}}$ ($B^0=B_1$ is the ball of radius 1). On the other hand, by Hölder's inequality,

$$\|v\|^{p\theta}_{L^{p\theta}(B^{m+1})}\leq \|v\|^{p\theta}_{L^{\frac{pd}{d-2}}(B^{m+1})}\biggl(|B^m\cap\{v>0\}|\biggr)^{1-\frac{(d-2)\theta}{d}}.$$

Applying (18.3) in the first multiple in the RHS, we obtain

$$\|v\|^{p\theta}_{L^{p\theta}(B^{m+1})}\leq C\biggl(2^{2\theta m}\|v\|^{p\theta}_{L^{p\theta}(B^m)}+H^\theta|B^m\cap\{v>0\}|\biggr)\biggl(|B^m\cap\{v>0\}|\biggr)^{1-\frac{(d-2)\theta}{d}}.$$

Put

$$v_m:=(u-k_m)_+, \quad k_m:=\xi(1-2^{-m})\uparrow\xi,$$

where constant $\xi > 0$ will be chosen later. Using $2^{2\theta m} \le 2^{p\theta m}$ and dividing by $\xi^{p\theta}$, we obtain

$$
\begin{aligned}
&\frac{1}{\xi^{p\theta}}\|v_{m+1}\|^{p\theta}_{L^{p\theta}(B^{m+1})} \\
&\le C\bigg(\frac{2^{p\theta m}}{\xi^{p\theta}}\|v_{m+1}\|^{p\theta}_{L^{p\theta}(B^m)} + \frac{1}{\xi^{p\theta}}H^\theta|B^m \cap \{u > k_{m+1}\}|\bigg)\big(|B^m \cap \{u > k_{m+1}\}|\big)^{1-\frac{(d-2)\theta}{d}}.
\end{aligned}
$$

From now on, we require that constant ξ satisfies $\xi^p \ge H$, so

$$
\begin{aligned}
&\frac{1}{\xi^{p\theta}}\|v_{m+1}\|^{p\theta}_{L^{p\theta}(B^{m+1})} \\
&\le C\bigg(\frac{2^{p\theta m}}{\xi^{p\theta}}\|v_{m+1}\|^{p\theta}_{L^{p\theta}(B^m)} + |B^m \cap \{u > k_{m+1}\}|\bigg)\big(|B^m \cap \{u > k_{m+1}\}|\big)^{1-\frac{(d-2)\theta}{d}}.
\end{aligned}
\tag{18.4}
$$

Now,

$$
\begin{aligned}
|B^m \cap \{u > k_{m+1}\}| &= \bigg|B^m \cap \bigg\{\bigg(\frac{u-k_m}{k_{m+1}-k_m}\bigg)^{2\theta} > 1\bigg\}\bigg| \\
&\le (k_{m+1}-k_m)^{-p\theta}\langle v_m^{p\theta}\mathbf{1}_{B^m}\rangle = \xi^{-p\theta}2^{p\theta(m+1)}\|v_m\|^{p\theta}_{L^{p\theta}(B^m)},
\end{aligned}
$$

so using $\|v_{m+1}\|_{L^{p\theta}(B^m)} \le \|v_m\|_{L^{p\theta}(B^m)}$ in (18.4) and applying the previous inequality, we obtain

$$
\frac{1}{\xi^{p\theta}}\|v_{m+1}\|^{p\theta}_{L^{p\theta}(B^{m+1})} \le C2^{p\theta m(2-\frac{\theta}{\theta_0})}\bigg(\frac{1}{\xi^{p\theta}}\|v_m\|^{p\theta}_{L^{p\theta}(B^m)}\bigg)^{2-\frac{(d-2)\theta}{d}}.
$$

Denote $z_m := \frac{1}{\xi^{p\theta}}\|v_m\|^{p\theta}_{L^{p\theta}(B^m)}$. Then

$$
z_{m+1} \le C\gamma^m z_m^{1+\alpha}, \quad m = 0, 1, 2, \dots, \quad \alpha := 1 - \frac{(d-2)\theta}{d}, \quad \gamma := 2^{p\theta(2-\frac{(d-2)\theta}{d})}
$$

and $z_0 = \frac{1}{\xi^{p\theta}}\langle u_+^{p\theta}\mathbf{1}_{B^0}\rangle \le C^{-\frac{1}{\alpha}}\gamma^{-\frac{1}{\alpha^2}}$ (recall: $B^0 := B_{R_0} \equiv B_1$) provided that we fix c by

$$
\xi^{p\theta} := C^{\frac{1}{\alpha}}\gamma^{\frac{1}{\alpha^2}}\langle u_+^{p\theta}\mathbf{1}_{B^0}\rangle + H^\theta.
$$

Hence, by Lemma 16.1, $z_m \to 0$ as $m \to \infty$. It follows that $\sup_{B_{1/2}} u_+ \le \xi$, and the claimed inequality follows.

Step 2. Next, we bound $\langle |u|^{p\theta}\mathbf{1}_{B_1}\rangle^{\frac{1}{p\theta}}$ using Lemma 15.1 with $s = p\theta$, which allows us to conclude that if σ (in the definition of weight ρ) is fixed sufficiently small, then for all $x \in \mathbb{R}^d$

$$
\sup_{B_{\frac{1}{2}}(x)} u_+ \le K\bigg(\langle |f|^{p\theta}\rho_x\rangle^{\frac{1}{p\theta}} + \big\langle |f|^{p\theta'}\mathbf{1}_{B_1(x)}\big\rangle^{\frac{1}{p\theta'}}\bigg).
$$

This ends the proof of Proposition 10.4. $\square$

19. Proof of Proposition 10.3 (Convergence)

The proof given below is close to [KS3, Proof of Theorem 4.3]. This argument (in L^2) can, in principle, be replaced by an argument based on the Lions' variational approach that handles well both $b \in \mathbf{F}_\delta$ and $q \in \mathbf{BMO}^{-1}$ (regarding the latter, see [QX]).

In order to prove the existence of the limit, we need to construct first intermediate semigroups $e^{-t\Lambda(b_n,q)}$ in L^p. Here b_n are bounded and smooth, but q can be singular.

Step 1. At the first step, we construct $e^{-t\Lambda(0,q)}$ in L^2. Here we work over complex numbers (in this regard, see Remark 19.1). Define sesquilinear form

$$\tau[v,w] := \langle \nabla v, \nabla w\rangle + \langle q \cdot \nabla v, w\rangle, \quad D(\tau) = W^{1,2},$$

where $\langle v, w\rangle = \langle v\bar{w}\rangle$.

a) τ is bounded: $|\tau[v,w]| \leq C\|\nabla v\|_2 \|\nabla w\|_2$. Indeed, by the compensated compactness estimate (Proposition 2.1), $|\langle q \cdot \nabla v, w\rangle| \leq C'\|\nabla v\|_2\|\nabla w|_2$.

b) τ is a sectorial form:

$$\operatorname{Im}\tau[v,v] \leq K \operatorname{Re}\tau[v,v], \quad v \in D(\tau)$$

for some constant $K > 0$. Indeed, writing $v = r + ie$, where r, e are real-valued elements of $W^{1,2}$, we have

$$\tau[v,v] = \langle |\nabla r|^2 + |\nabla e|^2\rangle + \langle q \cdot \big[(\nabla r)r + (\nabla e)e\big]\rangle + i\langle q \cdot [(\nabla e)r - (\nabla r)e]\rangle,$$

so, taking into that the second term vanishes due to the anti-symmetry of Q, we have

$$\operatorname{Re}\tau[v,v] = \langle |\nabla r|^2 + |\nabla e|^2\rangle, \quad \operatorname{Im}\tau[v,v] = \langle q \cdot [(\nabla e)r - (\nabla r)e]\rangle.$$

Now, invoking again the compensated compactness, we obtain $\operatorname{Im}\tau[v,v] \leq C\|Q\|_{\mathrm{BMO}}\operatorname{Re}\tau[u,u]$.

c) $\operatorname{Re}\tau[v,v]$ is a closed form, i.e. if $v_k \to v$ in L^2 and $\operatorname{Re}\tau[v_k - v_l] \to 0$ as $k, l \to \infty$, then $\tau[v_k - v] \to 0$ (we only need to look at the real part of τ due to its sectoriality). But in our case the latter is just a re-statement that $W^{1,2}$ is a complete space.

Therefore, there exists a unique (m-sectorial) operator $\Lambda(0,q)$ such that

$$\langle \Lambda(0,q)v, w\rangle = \tau[v,w], \quad v \in D(\Lambda(0,q)) \subset W^{1,2}, \quad w \in D(\tau) = W^{1,2},$$

see [Ka, Ch. VI, §2]. This operator, being m-sectorial, generates a holomorphic semigroup $e^{-t\Lambda(0,q)}$ in L^2.

Remark 19.1. A property of holomorphic semigroups that we will need at Step 5 is as follows: for every $f \in L^2$, $t > 0$, $e^{-t\Lambda(0,q)}f$ belongs to the domain $D(\Lambda(0,q))$.

Step 2. We need to construct $e^{-t\Lambda(b_n,q)}$, i.e. to add the drift term $b_n \cdot \nabla$. So, we re-do what we did above for the sesquilinear form

$$\tau[v,w] := \langle \nabla v, \nabla w\rangle + \langle q \cdot \nabla v, w\rangle + \langle b_n \cdot \nabla v, w\rangle. \tag{19.1}$$

In particular, applying Cauchy-Schwarz' inequality to $\langle b_n \cdot \nabla v, w\rangle$, we obtain

$$\operatorname{Im}\tau[v,v] \leq K\big(\operatorname{Re}\tau[v,v] + C\langle v, v\rangle\big), \tag{19.2}$$

where $C = C(\|b_n\|_\infty) \geq 0$. The cited results apply to the case (19.2), and we get (m-sectorial) generator $\Lambda(b_n,q)$ of a holomorphic semigroup in L^2 such that

$$\langle \Lambda(b_n,q)v, w\rangle = \tau[v,w] \quad \text{for } \tau \text{ defined by (19.1)}, \tag{19.3}$$

where $v \in D(\Lambda(b_n,q)) \subset W^{1,2}$, $w \in D(\tau) = W^{1,2}$. (Or we could appeal to the Hille perturbation theorem and work with the algebraic sum $\Lambda(b_n,q) := \Lambda(0,q) + b_n \cdot \nabla$, $D(\Lambda(b_n,q)) = D(\Lambda(0,q))$, that still generates a holomorphic semigroup in L^2.)

Step 3. Let us now note that we have the following quasi-contraction estimate for $e^{-t\Lambda(b_n,q_m)}$, $n,m = 1,2,\dots$, in L^p, $p > \frac{2}{2-\sqrt{\delta}}$, $p \geq 2$:

$$\|e^{-t\Lambda(b_n,q_m)} f\|_p \leq e^{\omega t}\|f\|_p, \quad t \geq 0, \quad f \in C_c^\infty. \tag{19.4}$$

for some ω independent of n,m. Indeed, setting $u_{n,m} = e^{-t(\omega+\Lambda(b_n,q_m))}f$, we multiply the corresponding parabolic equation $(\omega + \partial_t - \Delta + (b_n+q_m)\cdot\nabla)u_{n,m} = 0$ by $u_{n,m}|u_{n,m}|^{p-2}$ and repeat the proof of Lemma 15.1 (without the weight there) with the obvious modifications for the time derivative term (see Step 5 below for details). The constant $\omega > 0$ needs to be chosen to account e.g. for the constant c_δ resulting from the use of condition $b \in \mathbf{F}_\delta$.

If we do include the weight ρ_x, then, again arguing as in the proof of Lemma 15.1, we obtain a weighted quasi contraction estimate

$$\|e^{-t\Lambda(b_n,q_m)} f\|_{L^p_{\rho_x}} \leq e^{\omega' t}\|f\|_{L^p_{\rho_x}}, \quad t \geq 0, \tag{19.5}$$

for any $x \in \mathbb{R}^d$, with ω' independent of n, m or x. (Since $\rho_x \leq 1$, the result of the next section implies that

$$u_{m,n} = (\mu + \Lambda(b_n,q_m))^{-1}f \to u_n = (\mu + \Lambda(b_n,q))^{-1}f \quad \text{in } L^p_{\rho_x}. \tag{19.6}$$

Moreover, (19.5) ensures that the resulting semigroup $e^{-t\Lambda(b_n,q)}$ is strongly continuous in $L^p_{\rho_x}$.)

Step 4. From now on, we work over reals. Let us show convergence

$$e^{-t\Lambda(b_n,q_m)} \to e^{-t\Lambda(b_n,q)} \text{ in } L^2 \text{ loc. uniformly in } t \geq 0$$

as $m \to \infty$. Here n is fixed.

Since n is fixed, it is easily seen that the operator norms of the resolvents $\|(\mu+\Lambda(b_n,q_m))^{-1}\|_{2\to 2}$ are uniformly in m bounded, provided $\mu = \mu(\|b_n\|_\infty)$ is fixed sufficiently large. It suffices for us (see [Ka, Ch. XI, §5]) to show the convergence of the resolvents

$$(\mu + \Lambda(b_n,q_m))^{-1}f \to (\mu + \Lambda(b_n,q))^{-1}f \quad \text{in } L^2 \text{ as } m \to \infty$$

for all $f \in L^2_c$ (subscript c means compact support).

The standard argument yields that $u_n = (\mu + \Lambda(b_n,q))^{-1}f$ is the unique weak solution to the elliptic equation $(\mu - \Delta + (b_n+q)\cdot\nabla)u_n = f$, where the former means that

$$\mu\langle u_n,\varphi\rangle + \langle \nabla u_n, \nabla\varphi\rangle + \langle (b_n+q)\cdot\nabla u_n, \varphi\rangle = \langle f,\varphi\rangle, \quad \varphi \in C_c^\infty.$$

(The compensated compactness estimate $|\langle q\cdot\nabla u_n,\varphi\rangle| = |\langle Q\cdot\nabla u_n,\nabla\varphi\rangle| \leq C'\|\nabla u_n\|_2\|\nabla\varphi\|_2$ of Proposition 2.1 allows us to pass to test functions $\varphi \in W^{1,2}$.) In turn, this uniqueness and the usual weak compactness argument shows that $u_{n,m} = (\mu + \Lambda(b_n,q_m))^{-1}f$, i.e. solutions to the approximating elliptic equations $(\mu - \Delta + (b_n+q_m)\cdot\nabla)u_{n,m} = f$, converge weakly in $W^{1,2}$ to the same limit u_n. Now we can appeal to the Rellich-Kondrashov theorem to obtain $u_{n,m} \to u_n$ in L^2_{loc}, and further to upgrade this convergence to $u_{n,m} \to u_n$ in L^2 by "cutting tails" of $u_{n,m}$ at infinity uniformly in m using the upper Gaussian bound on the heat kernel of $-\Delta + (b_n+q_m)\cdot\nabla$, see [QX], and taking into account that f has compact support. (The constants in the upper Gaussian heat kernel bound, which yields the bound on the integral kernel of the resolvents, will depend on n. Since each b_n is bounded, adding the drift term in $b_n\cdot\nabla$ does not affect the proof of the upper bound in [QX] which employs Moser's iterations and the Davies device. In fact, one can account for $b_n \cdot \nabla$ by once again introducing a constant term in the operator that will absorb the contribution from $b_n \cdot \nabla$ in Moser's method.)

Now, since $e^{-t\Lambda(b_n,q_m)}$ are L^∞ contractions, we obtain by interpolation that

$$e^{-t\Lambda(b_n,q_m)} \to e^{-t\Lambda(b_n,q)} \text{ in } L^p \text{ loc. uniformly in } t \geq 0$$

for all $p \geq 2$. This implies the convergence of the resolvents: as $m \to \infty$,

$$u_{n,m} = (\mu + \Lambda(b_n, q_m))^{-1} f \to u_n = (\mu + \Lambda(b_n, q))^{-1} f \quad \text{ in } L^p \tag{19.7}$$

with μ independent of n, m.

Step 5. Having constructed the intermediate semigroups $e^{-t\Lambda(b_n,q)}$, our goal now is to show that they converge as $n \to \infty$. For reader's convenience, at this step we give a proof that assumes additionally that $b_n \to b$ in L^2. In this case we obtain that, every $f \in C_c^\infty$,

$$\{v_n(t) := e^{-t(\omega+\Lambda(b_n,q))} f\}_{n=1}^\infty \quad \text{ is a Cauchy sequence in } L^\infty([0,1], L^p) \tag{19.8}$$

for

$$p > \frac{2}{2-\sqrt{\delta}}, \quad p \geq 2, \text{ for some fixed } \omega.$$

Taking into account Remark 19.1 about $v(t) \in D(\Lambda(b_n, q))$ for all $t > 0$ and the identity (19.3), we can write

$$\langle \partial_t v_n, \psi \rangle + \omega \langle v_n, \psi \rangle + \langle \nabla v_n, \nabla \psi \rangle + \langle b_n \cdot \nabla v_n, \psi \rangle - \langle Q \cdot \nabla v_n, \nabla \psi \rangle = 0, \tag{19.9}$$

for all $\psi(t, \cdot) \in W^{1,2}$, where, recall, $\nabla Q = q$. Set

$$h := v_{n_1} - v_{n_2}.$$

Subtracting identities (19.9) for v_{n_1} and v_{n_2} from each other, we obtain

$$\langle \partial_t h, \psi \rangle + \omega \langle h, \psi \rangle + \langle \nabla h, \nabla \psi \rangle + \langle b_{n_1} \cdot \nabla h, \psi \rangle - \langle Q \cdot \nabla h, \nabla \psi \rangle = -(b_{n_1} - b_{n_2}) \cdot \nabla v_{n_2}.$$

We are basically in the setting of Proposition 15.2(i) with the only difference that Q is no longer smooth and we are dealing with weak solutions of the parabolic equations rather than classical solutions. However, the latter does not pose a difficulty: since $p \geq 2$ the standard result on the composition of Lipschitz functions with the elements of Sobolev spaces yield that $h(t)|h(t)|^{p-2}$, $h(t)|h(t)|^{\frac{p}{2}-1}$, $|h(t)|^{\frac{p}{2}} \in W^{1,2}$. Therefore, we can take $\psi = h|h|^{p-2}$, obtaining

$$\begin{aligned} \frac{1}{p} \partial_t \langle |h|^p \rangle + \omega \langle |h|^p \rangle + \frac{4(p-1)}{p^2} \langle |\nabla |h|^{\frac{p}{2}}|^2 \rangle + \frac{2}{p} \langle \langle b_{n_1} \cdot \nabla |h|^{\frac{p}{2}}, |h|^{\frac{p}{2}} \rangle \\ \leq |\langle (b_{n_1} - b_{n_2}) \cdot \nabla v_{n_2}, h|h|^{p-2} \rangle|, \end{aligned} \tag{19.10}$$

where we have used anti-symmetry of Q and $h(0) = 0$. We handle the term containing b_{n_1} as in the proof of Proposition 15.1, i.e. first applying quadratic inequality and then $b_{n_1} \in \mathbf{F}_\delta$. Now, handling the time derivative term as in the proof of Proposition 15.2, we obtain

$$\begin{aligned} \frac{1}{p} \sup_{s \in [0,t]} \langle |h(s)|^p \rangle + \left[\omega - 2\frac{c_\delta}{\sqrt{\delta}}\right] \int_0^t \|h\|_p^p ds + \left[\frac{4(p-1)}{p^2} - \frac{2}{p}\sqrt{\delta}\right] \int_0^t \langle |\nabla |h|^{\frac{p}{2}}|^2 \rangle ds \\ \leq 2 \int_0^t \|b_{n_1} - b_{n_2}\|_2 \|\nabla v_{n_2}\|_2 \|h\|_\infty^{p-1} ds. \end{aligned} \tag{19.11}$$

Take $\frac{1}{2}\omega := 2\frac{c_\delta}{\sqrt{\delta}}$. Since $p > \frac{2}{2-\sqrt{\delta}}$, the expressions in square brackets are strictly positive. In the right-hand side of (19.11), $\|b_{n_1} - b_{n_2}\|_2 \to 0$ as $n_1, n_2 \to \infty$ and $\|h(s)\|_\infty^{p-1} \leq 2^{p-1} \|f\|_\infty^{p-1}$,

$s \in [0,t]$, for all n_1, n_2. It remains to note that $\int_0^t \|\nabla v_{n_2}\|_2 ds$ is uniformly in n_2 bounded. Indeed, by (19.9) with $n = n_2$ and $\psi = v_{n_2}$ upon noting that

$$\begin{aligned} &|\langle b_{n_2} \cdot \nabla v_{n_2}, v_{n_2}\rangle| + |\langle Q \cdot \nabla v_{n_2}, \nabla v_{n_2}\rangle| \\ &= |\langle b_{n_2} \cdot \nabla v_{n_2}, v_{n_2}\rangle| \leq \frac{1}{2}\|b_{n_2}\|_2^2 \|v_{n_2}\|_\infty^2 + \frac{1}{2}\|\nabla v_{n_2}\|_2^2, \end{aligned}$$

where we have used again the anti-symmetry of Q. Now we use $\sup_{n_2} \|b_{n_2}\|^2 < \infty$ and the obvious a priori estimate $\|v_{n_2}(t)\|_\infty \leq \|f\|_\infty$, $t \geq 0$. Thus, the right-hand side of (19.11) tends to 0 as n_1, $n_2 \to \infty$. Hence $h \to 0$ in $L^\infty([0,1], L^p)$, and (19.8) follows. The latter and the contractivity estimate (19.4) yield that the limit $v = L^p\text{-}\lim_n v_n$ (loc. uniformly in $t \geq 0$) determines a strongly continuous semigroup in L^p, say, $v(t) =: e^{-t\Lambda(b,q)} f$. In turn, the convergence of the semigroups yields the convergence of the resolvents

$$u_n = (\mu + \Lambda(b_n, q))^{-1} f \to u = (\mu + \Lambda(b,q))^{-1} f \quad \text{in } L^p$$

with μ independent of n (proportional to ω). In view of (19.7), the existence of the limit $u = L^p\text{-}\lim_n u_n = L^p\text{-}\lim_n \lim_m u_{n,m}$ follows.

Step 6. Finally, we prove convergence in the general case, i.e. we do not assume global convergence of b_n to b in L^2. We show that, for any $x \in \mathbb{R}^d$, for every $f \in C_c^\infty$,

$$\{v_n(t) := e^{-t(\mu + \Lambda(b_n,q))} f\}_{n=1}^\infty \quad \text{is a Cauchy sequence in } L^\infty([0,1], L^p_{\rho_x}) \tag{19.12}$$

for $p > \frac{2}{2-\sqrt{\delta}}$, $p \geq 2$, for some fixed ω. Then, repeating the argument in the end of Step 5 but using (19.5) and (19.6), we will obtain the claimed in Proposition 10.3 existence of the limit

$$u \left(= L^p_{\rho_x}\text{-}\lim_n u_n \right) = L^p_{\rho_x}\text{-}\lim_n \lim_m u_{n,m}.$$

Let us prove (19.12). This time, taking $\psi = h|h|^{p-2}\rho_x$, we obtain

$$\begin{aligned} \sup_{s \in [0,t]} \langle |h(s)|^p \rho_x\rangle + C_1 \int_0^t \langle |h|^p \rho_x\rangle ds + C_2 \int_0^t \langle |\nabla |h|^{\frac{p}{2}}|^2 \rho_x\rangle ds& \\ \leq C_3 \int_0^t |\langle (b_{n_1} - b_{n_2}) \cdot \nabla v_{n_2}, h|h|^{p-2}\rho_x\rangle| ds& \end{aligned} \tag{19.13}$$

for constants C_1-C_3 independent of n, m, constant C_1 being strictly positive provided that σ in the definition of ρ is fixed sufficiently small. Arguing as at the previous step, we obtain

$$\int_0^1 \sup_{n_2} \langle |\nabla v_{n_2}(s)|^2 \rho_x\rangle ds < \infty.$$

Therefore,

$$\begin{aligned} \int_0^T |\langle (b_{n_1} - b_{n_2}) \cdot \nabla v_{n_2}, h|h|^{p-2}\rho_x\rangle| ds &\leq \langle |b_{n_1} - b_{n_2}|^2 \rho_x\rangle \|h\|_\infty^{p-1} \int_0^1 \langle |\nabla v_{n_2}|^2 \rho_x\rangle ds \\ &\quad (\text{use } \|h\|_\infty \leq 2\|f\|_\infty \text{ and apply Lemma 15.1}) \\ &\to 0 \quad \text{as } n_1, n_2 \to \infty. \end{aligned}$$

Combining this with (19.13), we obtain the claimed convergence (19.12). □

Remark 19.2. If we try to obtain a stronger result about the existence of the limit s-L^p- $\lim_{n,m} u_{m,n}$ by extending the proof of Lemma 10.3 to the sequence $h := u_{n_1,m_1} - u_{n_2,m_2}$, then we get an extra term in the right-hand side of (19.10): $|\langle (Q_{m_1} - Q_{m_2}) \cdot \nabla u_{n_2,m_2}, \nabla (h|h|^{p-2})\rangle|$. It can be dealt with in two ways:

(a) We can estimate

$$|\langle (Q_{m_1} - Q_{m_2}) \cdot \nabla u_{n_2,m_2}, \nabla (h|h|^{p-2})\rangle| \leq \|Q_{m_1} - Q_{m_2}\|_s \|\nabla u_{n_2,m_2}\|_{s'} (p-1)(2\|f\|_\infty)^{p-2} \|\nabla h\|_2,$$

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{2}.$$

So, assuming for the illustration purposes that we have global convergence $\|Q_{m_1} - Q_{m_2}\|_s \to 0$ as $m_1, m_2 \to \infty$, we need a bound on $\|\nabla u_{n_2,m_2}\|_{s'}$ for a $s' > 2$. In principle, s' can be chosen to be close to 2. To obtain such an estimate, we can use Gehring-Giaquinta-Modica's lemma as in the proof of Theorem 8.3, but this, at least in the present form of the argument, requires us to consider the equation in L^2, hence we need to require $\delta < 1$ rather than $\delta < 4$ as in (5.3).

(b) Another option is to use the estimate

$$|\langle (Q_{m_1} - Q_{m_2}) \cdot \nabla u_{n_2,m_2}, \nabla (h|h|^{p-2})\rangle| \leq \|Q_{m_1} - Q_{m_2}\|_{\mathrm{BMO}} \|\nabla u_{n_2,m_2}\|_2 (2\|f\|_\infty)^{p-2} \|\nabla h\|_2,$$

where $\|\nabla u_{n_2,m_2}\|_2$ can be estimated as in the proof of Theorem 10.3, but now to have convergence $\|Q_{m_1} - Q_{m_2}\|_{\mathrm{BMO}} \to 0$ as $m_1, m_2 \to \infty$ we need a stronger hypothesis of the matrix field Q and thus on q, namely, that Q has entries in $\mathrm{VMO}(\mathbb{R}^d)$.

$\square$

20. Proof of Lemma 10.3

It suffices to carry out the proof for b_n and q_m, and then use the convergence result of Proposition 10.3. Thus, our goal is to show that

$$\sup_{n,m} \|\nabla u_{n,m}\|_{2+\varepsilon} < \infty \tag{20.1}$$

for some $\varepsilon > 0$ independent of n, m.

Put for brevity $b = b_n$, $q = q_m$ and $u = u_{n,m}$, so

$$\big(\mu - \Delta + (b+q) \cdot \nabla\big) u = f.$$

Let us fix some $1 < \theta < \frac{d}{d-2}$.

By Lemma 17.1 (with $p = 2$ there, which is admissible since $\delta < 1$), the function $v := (u-k)_+$ ($k \in \mathbb{R}$) satisfies Caccioppoli's inequality: for all $x \in \mathbb{R}^d$, $0 < r < R < \frac{1}{2}$,

$$\|\nabla v\|^2_{L^2(B_r(x))} \leq \frac{K_1}{(R-r)^2} |B_R|^{\frac{1}{\theta'}} (1 + \|Q\|^2_{\mathrm{BMO}}) \|v\|^2_{L^{2\theta}(B_R)} + K_2 \|(f - \mu u)\mathbf{1}_{v>0}\|^2_{L^2(B_R)}, \tag{20.2}$$

for constants K_1, K_2 independent of k, r, R and n, m. (There is some abuse of notation: our R here is not the radius of a fixed large ball in Lemma 10.3, but this should not cause a confusion.) We will obtain the sought bound (20.1) by applying a corollary of this Caccioppoli's inequality in the Gehring-Giaquinta-Modica lemma:

Lemma 20.1. *Assume that there exist constants $K \geq 1$, $1 < \nu < \infty$ such that, for given $0 \leq g \in L^q_{\rm loc}$, $0 \leq h \in L^\nu_{\rm loc} \cap L^\infty$ we have, for all $x \in \mathbb{R}^d$,*

$$\biggl(\frac{1}{|B_R|}\langle g^\nu \mathbf{1}_{B_R(x)}\rangle\biggr)^{\frac{1}{\nu}} \leq \frac{K}{|B_{2R}|}\langle g\mathbf{1}_{B_{2R}(x)}\rangle + \biggl(\frac{1}{|B_{2R}|}\langle h^\nu \mathbf{1}_{B_{2R}(x)}\rangle\biggr)^{\frac{1}{\nu}}$$

for all $0 < R < \frac{1}{2}$. Then $g \in L^s_{\rm loc}$ for some $s > \nu$ and, for all $x \in \mathbb{R}^d$,

$$\biggl(\frac{1}{|B_R|}\langle g^s \mathbf{1}_{B_R(x)}\rangle\biggr)^{\frac{1}{s}} \leq C_1\biggl(\frac{1}{|B_{2R}|}\langle g^\nu \mathbf{1}_{B_{2R}(x)}\rangle\biggr)^{\frac{1}{\nu}} + C_2\biggl(\frac{1}{|B_{2R}|}\langle h^s \mathbf{1}_{B_{2R}(x)}\rangle\biggr)^{\frac{1}{s}}.$$

Remark 20.1. The authors of [KrS] proved Lemma 20.1 with explicit constants independent of the dimension d.

We are in position to prove (20.1). Put, for brevity, $x = 0$.

Step 1. Set $(u_n)_{B_{2R}} := \frac{1}{|B_{2R}|}\langle u_n \mathbf{1}_{B_{2R}}\rangle$. Applying (20.2) to the positive and the negative parts of $u_n - (u_n)_{B_{2R}}$, we obtain

$$\langle |\nabla u_n|^2 \mathbf{1}_{B_R}\rangle \leq \frac{K_1}{|B_{2R}|^{\frac{2}{d}}}|B_{2R}|^{\frac{1}{\theta'}}(1+\|Q\|^2_{\rm BMO})\langle |u_n-(u_n)_{B_{2R}}|^{2\theta}\mathbf{1}_{B_{2R}}\rangle^{\frac{1}{\theta}} + K_2\langle |f-\mu u_n|^2\mathbf{1}_{B_{2R}}\rangle, \quad 0<R<\frac{1}{2}. \tag{20.3}$$

By the Sobolev-Poincaré inequality,

$$\biggl(\frac{1}{|B_{2R}|}\langle (u_n-(u_n)_{B_{2R}})^{2\theta}\mathbf{1}_{B_{2R}}\rangle\biggr)^{\frac{1}{2\theta}} \leq C|B_R|^{\frac{1}{d}}\biggl(\frac{1}{|B_{2R}|}\langle |\nabla u_n|^{\frac{2\theta d}{d+2\theta}}\mathbf{1}_{B_{2R}}\rangle\biggr)^{\frac{d+2\theta}{2\theta d}}, \tag{20.4}$$

i.e.

$$\langle (u_n-(u_n)_{B_{2R}})^{2\theta}\mathbf{1}_{B_{2R}}\rangle^{\frac{1}{\theta}} \leq C^2|B_R|^{\frac{2}{d}+\frac{1}{\theta}}\biggl(\frac{1}{|B_{2R}|}\langle |\nabla u_n|^{\frac{2\theta d}{d+2\theta}}\mathbf{1}_{B_{2R}}\rangle\biggr)^{\frac{d+2\theta}{\theta d}}.$$

Plug in above estimate in (20.4), and divide both side by $|B_R|$ then for appropriate constants C_1 and c were C_1 depends on $\|Q\|_{\rm BMO}$

$$\frac{1}{|B_R|}\langle |\nabla u_n|^2\mathbf{1}_{B_R}\rangle \leq C_1\biggl(\frac{1}{|B_{2R}|}\langle |\nabla u_n|^{\frac{2\theta d}{d+2\theta}}\mathbf{1}_{B_{2R}}\rangle\biggr)^{\frac{d+2\theta}{\theta d}} + \frac{c}{|B_{2R}|}\langle |f-\mu u_n|^2\mathbf{1}_{B_{2R}}\rangle,$$

Then the condition of the Gehring-Giaquinta-Modica lemma is verified with $g = |\nabla u_n|^{\frac{2\theta d}{d+2\theta}}$, $g^\nu = |\nabla u_n|^2$ (so $\nu = \frac{d+2\theta}{\theta d}$) and $h = (c^{\frac{1}{2}}|f-\mu u_n|)^{\frac{2\theta d}{d+2\theta}}$, $h^\nu = c|f-\mu u_n|^2$. Hence there exists $s > \frac{d+2\theta}{\theta d}$ such that

$$\biggl(\frac{1}{|B_R|}\langle |\nabla u_n|^{s\frac{2\theta d}{d+2\theta}}\mathbf{1}_{B_R}\rangle\biggr)^{\frac{1}{s}} \leq C_1\biggl(\frac{1}{|B_{2R}|}\langle |\nabla u_n|^2\mathbf{1}_{B_{2R}}\rangle\biggr)^{\frac{\theta d}{d+2\theta}} + C_2\biggl(\frac{1}{|B_{2R}|}\langle |f-\mu u_n|^{s\frac{2\theta d}{d+2\theta}}\mathbf{1}_{B_{2R}}\rangle\biggr)^{\frac{1}{s}},$$

where all constants are independent of n, or

$$\frac{1}{|B_R|}\langle |\nabla u_n|^{s\frac{2\theta d}{d+2\theta}}\mathbf{1}_{B_R}\rangle \leq C_1'\biggl(\frac{1}{|B_{2R}|}\langle |\nabla u_n|^2\mathbf{1}_{B_{2R}}\rangle\biggr)^{s\frac{\theta d}{d+2\theta}} + C_2'\frac{1}{|B_{2R}|}\langle |f-\mu u_n|^{s\frac{2\theta d}{d+2\theta}}\mathbf{1}_{B_{2R}}\rangle.$$

Fix some R, say, $R = 1$. We consider equally spaced grid $\frac{1}{2}\mathbb{Z}^d$ in $\mathbb{R}^d$ so that the smaller balls centered at the nodes of the grid cover $\mathbb{R}^d$, apply the previous estimate on each ball, and then sum up. We obtain a global estimate

$$\|\nabla u_n\|^{s\frac{2\theta d}{d+2\theta}}_{s\frac{2\theta d}{d+2\theta}} \leq C_3\sum_{x\in c\mathbb{Z}^d}\biggl(\frac{1}{|B_2|}\langle |\nabla u_n|^2\mathbf{1}_{B_2(x)}\rangle\biggr)^{s\frac{\theta d}{d+2\theta}} + C_4\|f-\mu u_n\|^{s\frac{2\theta d}{d+2\theta}}_{s\frac{2\theta d}{d+2\theta}}. \tag{20.5}$$

To deal with the first term in the right-hand side, we split the grid into two parts: $I := \{x \in c\mathbb{Z}^d \mid \frac{1}{|B_2|}\langle |\nabla u_n|^2 \mathbf{1}_{B_2(x)}\rangle > 1\}$ and its complement I^c. For the nodes in the complement we have, taking into account that $s\frac{\theta d}{d+2\theta} > 1$,

$$\sum_{x \in I^c} \biggl(\frac{1}{|B_2|}\langle |\nabla u_n|^2 \mathbf{1}_{B_2(x)}\rangle\biggr)^{s\frac{\theta d}{d+2\theta}} \leq \sum_{x \in c\mathbb{Z}^d} \frac{1}{|B_2|}\langle |\nabla u_n|^2 \mathbf{1}_{B_2(x)}\rangle \leq C_5 \langle |\nabla u_n|^2\rangle.$$

In turn, there are only finitely many nodes in I. In fact, the cardinality of I can be estimated in terms of $\langle |\nabla u_n|^2\rangle$:

$$|I| < \sum_{x \in I} \frac{1}{|B_2|}\langle |\nabla u_n|^2 \mathbf{1}_{B_2(x)}\rangle \leq C_6 \langle |\nabla u_n|^2\rangle.$$

So,

$$\sum_{x \in I} \biggl(\frac{1}{|B_2|}\langle |\nabla u_n|^2 \mathbf{1}_{B_2(x)}\rangle\biggr)^{s\frac{\theta d}{d+2\theta}} \leq \sum_{x \in I} \biggl(\frac{1}{|B_2|}\langle |\nabla u_n|^2\rangle\biggr)^{s\frac{\theta d}{d+2\theta}} \leq C_7 \langle |\nabla u_n|^2\rangle^{1+s\frac{\theta d}{d+2\theta}}.$$

We arrive at a global estimate

$$\|\nabla u_n\|_{s\frac{2\theta d}{d+2\theta}}^{s\frac{2\theta d}{d+2\theta}} \leq C_8 \biggl(\|\nabla u_n\|_2^2 + \|\nabla u_n\|_2^{2+s\frac{2\theta d}{d+2\theta}}\biggr) + C_4 \|f - \mu u_n\|_{s\frac{2\theta d}{d+2\theta}}^{s\frac{2\theta d}{d+2\theta}}.$$

Step 2. Let us show that in the right-hand side of the estimate of Step 1 we have $\sup_n \|\nabla u_n\|_2^2 < \infty$. To this end, we multiply $(\mu - \Delta + b_n \cdot \nabla)u_n = f$ by u_n and integrate, obtaining $\mu \|u_n\|_2^2 + \|\nabla u_n\|_2^2 + \langle b_n \cdot \nabla u_n, u_n\rangle = \langle f, u_n\rangle$, where

$$\langle b_n \cdot \nabla u_n, u_n\rangle = -\frac{1}{2}\langle \operatorname{div} b_n, u_n^2\rangle \geq -\frac{1}{2}\langle (\operatorname{div} b_n)_+, u_n^2\rangle.$$

Hence, by our form-boundedness assumption on $(\operatorname{div} b_n)_+$,

$$\biggl(\mu - \frac{c_{\delta_+}}{2}\biggr)\|u_n\|_2^2 + \biggl(1 - \frac{\delta_+}{2}\biggr)\|\nabla u_n\|_2^2 \leq \langle f, u_n\rangle. \tag{20.6}$$

So, applying the quadratic inequality in the right-hand side, we arrive at $(\mu - \frac{c_{\delta_+}}{2} - \frac{1}{2})\|u_n\|_2^2 + (1 - \frac{\delta_+}{2})\|\nabla u_n\|_2^2 \leq \frac{1}{2}\|f\|_2^2$. Since $\delta_+ < 2$, $\sup_n \|\nabla u_n\|_2^2 < \infty$ for $\mu \geq \mu_0 := \frac{c_{\delta_+}}{2} + \frac{1}{2}$.

Step 3. Next, $\|u_n\|_2 \leq C\|f\|_2$ and a priori bound $\|u_n\|_\infty \leq \|f\|_\infty$ yield $\sup_n \|u_n\|_{s\frac{2\theta d}{d+2\theta}} < \infty$. Hence $\sup_n \|f - \mu u_n\|^2_{s\frac{2\theta d}{d+2\theta}} < \infty$.

Steps 1-3 give us the sought gradient bound $\sup_n \|\nabla u_n\|_{s\frac{2\theta d}{d+2\theta}} < \infty$, which thus ends the proof. $\square$

Appendix A. Proof of Lemma 10.1

We estimate

$$(1+|x|^r)\mu(\mu-\Delta)^{-1}\nabla_i g(x)=\mu\int_0^\infty\int_{\mathbb{R}^d}e^{-\mu s}(4\pi s)^{-\frac{d}{2}}(1+|x|^r)e^{-\frac{|x-y|^2}{4s}}\nabla_i g(y)dyds \tag{A.1}$$

as follows. For all $\mu\geq 1$, provided that $1\leq s<\infty$,

$$e^{-\mu s}(4\pi s)^{-\frac{d}{2}}\big|(1+|x|^r)e^{-\frac{|x-y|^2}{4s}}\nabla_i g(y)\big|\leq C_R e^{-c_1\mu s}(4\pi s)^{-\frac{d}{2}}e^{-\frac{|x-y|^2}{c_2 s}}|\nabla_i g(y)| \tag{A.2}$$

for some $c_2>4$, $0<c_1<1$. To see this, it suffices to show that

$$e^{-\mu s}(4\pi s)^{-\frac{d}{2}}(1+|x|^r)e^{-\frac{|x|^2}{4s}}\leq Ce^{-c_1\mu s}(4\pi s)^{-\frac{d}{2}}e^{-\frac{|x|^2}{c_2 s}},\quad x\in\mathbb{R}^d, s\geq 1, \tag{A.3}$$

since y varies only in B_R, and the sought estimate is non-trivial when $|x|\gg R$. Inequality (A.3) reduces to

$$e^{-\gamma_1\mu s}(1+|x|^r)e^{-\frac{|x|^2}{\gamma_2 s}}\leq C$$

for $0<\gamma_1<1$ and $\gamma_2>4$ ($\gamma_1=1-c_1$, $\frac{1}{\gamma_2}=\frac14-\frac{1}{c_2}$). So, denoting $x'=\frac{x}{\sqrt{s}}$, we obtain for $\mu\geq 1$, for all $x'\in\mathbb{R}^d$, $s\geq 1$,

$$\begin{aligned}e^{-\gamma_1\mu s}(1+|x|^r)e^{-\frac{|x|^2}{\gamma_2 s}}&\leq e^{-\gamma_1\mu s}s^{\frac{r}{2}}(1+|x'|^{d+1})e^{-\frac{|x'|^2}{\gamma_2}}\\&\leq Ce^{-\frac{\gamma_1}{2}\mu s}e^{-\frac{|x'|^2}{2\gamma_2}}\leq C,\end{aligned}$$

which gives us the previous estimate and hence (A.3). We thus have (A.2) for all $1\leq s<\infty$. Note that (A.2) is trivial for $0<s<1$. So, armed with (A.2) for all $0<s<\infty$, we estimate (A.1):

$$\mu\int_0^\infty\int_{\mathbb{R}^d}e^{-\mu s}(4\pi s)^{-\frac{d}{2}}(1+|x|^r)e^{-\frac{|x-y|^2}{4s}}|\nabla_i g(y)|dyds\leq c_R\mu(c\mu-\Delta)^{-1}|\nabla_i g|(x),\quad c=4c_1c_2^{-1},$$

i.e. we have obtained (10.12). ☐

Appendix B. Vanishing of stream matrix at infinity

Lemma B.1. *Assume that $q\in\mathbf{BMO}^{-1}$ has compact support in $B_1(0)$. Then we can find a stream matrix $Q=-Q^\top\in[\mathrm{BMO}]^{d\times d}$ for q, i.e. $q=\nabla Q$, that decays polynomially at infinity:*

$$|Q(x)|\leq C_R|x|^{-d+2}\quad\forall\,|x|\geq R>2.$$

It suffices carry out the proof for scalar distributions. Recall that a tempered distribution $h\in\mathcal{S}'$ belongs to BMO^{-1} if and only if

$$\sup_{x\in\mathbb{R}^d,R>0}\frac{1}{|B_R|}\int_{B_R(x)}\int_0^{R^2}|e^{t\Delta}h|^2dtdy<\infty.$$

Now, given $h\in\mathrm{BMO}^{-1}$, one can find a vector field $H=(H_j)_{j=1}^d\in[\mathrm{BMO}]^d$ such that

$$h=\operatorname{div}H$$

by arguing as follows (see [KT]). Put

$$h_{kj}:=\nabla_k\nabla_j(-\Delta)^{-1}h.$$

By [KT, Lemma 4.1], $\|h_{kj}\|_{\mathrm{BMO}^{-1}} \leq \|h\|_{\mathrm{BMO}^{-1}}$. For each fixed $1 \leq j \leq d$, we have $\nabla_r h_{kj} = \nabla_k h_{rj}$ for all $1 \leq k, r \leq d$, i.e. $h_{\cdot j}$ is curl-free. Therefore, there exists H_j such that $\nabla H_j = h_{\cdot j}$, e.g. take

$$H_j = (-\Delta)^{-1} \mathrm{div}\, h_{\cdot j}. \tag{B.1}$$

This $H_j \in \mathrm{BMO}$ by the Carleson's characterization of BMO (Section 2). We have $\nabla_j H_j = h_{jj}$, hence

$$\mathrm{div}\, H \equiv \sum_{j=1}^d \nabla_j H_j = \sum_{k=1}^d h_{jj} = h.$$

Proof of Lemma B.1. Assume that $h \in \mathrm{BMO}^{-1}$ has compact support in $B_1(0)$. We will show that, by following the above procedure, we obtain a "primitive" vector field H for h which decays at infinity as $|x|^{-d+2}$. Indeed, since h has compact support, $h_{kj}(x) = O((1+|x|)^{-d})$ as $|x| \to \infty$, and so $\mathrm{div}\, h_{\cdot j}(x) = O((1+|x|)^{-d-1})$. Therefore, by (B.1),

$$|H_j| \leq C(-\Delta)^{-1}(1+|\cdot|)^{-d-1}.$$

Invoking the Sobolev embedding property of $(-\Delta)^{-1}$, we obtain that

$$(-\Delta)^{-1}(1+|\cdot|)^{-d-1} \in L^{\frac{d}{d-2}}.$$

So, given that $(-\Delta)^{-1}(1+|\cdot|)^{-d-1}$ is a bounded rotationally-invariant function, we obtain the sought polynomial rate of decay of $(-\Delta)^{-1}(1+|\cdot|)^{-d-1}$ and hence of H_j. (It is also not difficult to estimate the rate of decay of $(-\Delta)^{-1}(1+|\cdot|)^{-d-1}$ directly.) □

Appendix C. Adams' estimates

The following are special cases of estimates proved by D.R. Adams. Let $V \geq 0$.

Lemma C.1 ([A1, Theorem 7.3])**.** *Let $0 < \alpha < d$, $1 < q < \infty$. Let $s > 1$. If*

$$\sup_{x \in \mathbb{R}^d, r>0} r^{\alpha q} \biggl(\frac{1}{|B_r|} \langle V^s \mathbf{1}_{B_r(x)} \rangle \biggr)^{\frac{1}{s}} < \infty,$$

then, for all $\varphi \in \mathcal{S}$,

$$\|V^{\frac{1}{q}} \varphi\|_q \leq C \|(-\Delta)^{\frac{\alpha}{2}} \varphi\|_q.$$

Lemma C.2 ([A2])**.** *Let $1 < p < q < \infty$, $p < d$. Then*

$$\|V^{\frac{1}{q}} \varphi\|_q \leq C \|(-\Delta)^{\frac{1}{2}} \varphi\|_p,$$

if and only if

$$\sup_{x \in \mathbb{R}^d, r>0} r^{-q\left(\frac{d}{p}-1\right)} \langle V \mathbf{1}_{B_r(x)} \rangle < \infty.$$

Thus, Morrey class is responsible for the $L^p(\mathbb{R}^d, dx) \to L^q(\mathbb{R}^d, V dx)$, $p < q$, estimates. It is only sufficient for the $L^q(\mathbb{R}^d, dx) \to L^q(\mathbb{R}^d, V dx)$ estimates (already the larger Chang-Wilson-Wolff class shows that Morrey class is not necessary, see Section 3). In turn, the $L^q(\mathbb{R}^d, dx) \to L^q(\mathbb{R}^d, V dx)$ estimates can be used e.g. to prove weak well-posedness of SDEs via (7.6), see Section 7.

Appendix D. Multiplicative form-boundedness and Morrey class M_1

The following result was proved in [M, Theorem 1.4.7]. We reproduce the proof for reader's convenience.

Proposition D.1. *Let $b \in [L^1_{\rm loc}]^d$. Then*

$$\langle |b|\varphi,\varphi\rangle \leq \delta \|\nabla \varphi\|_2 \|\varphi\|_2 \quad \forall\, \varphi \in C_c^\infty(\mathbb{R}^d) \tag{D.1}$$

if and only if

$$\langle |b|\mathbf{1}_{B_r(x)}\rangle \leq K r^{d-1} \tag{D.2}$$

for some constant K independent of $r>0$ and $x \in \mathbb{R}^d$ (then K is proportional to δ).

In other words, $b \in \mathbf{MF}_\delta$ with $c_\delta = 0$ if and only if $|b|$ belongs to the Morrey class M_1.

Proof of Proposition D.1. Define weighted Lebesgue norm over ball $B_r(x)$:

$$\|\varphi\|^q_{L^q(B_r(x),|b|)} := \int_{B_r(x)} |\varphi(y)|^q |b(y)| dy.$$

Define analogously $L^q(\mathbb{R}^d,|b|)$. D.R. Adams' Lemma C.2 will play a crucial role in the proof.

Proof of (D.2) $\Rightarrow$ (D.1). So, we assume that (D.2) holds, and prove the multiplicative inequality (D.1).

Lemma D.1. *For every $\varphi \in C(\overline{B_r})$,*

$$\|\varphi\|_{L^2(B_r,|b|)} \leq C_2 \left(r^{\frac{1}{2}} \|\nabla \varphi\|_{L^2(B_r)} + r^{-\frac{1}{2}}\|\varphi\|_{L^2(B_r)}\right), \tag{D.3}$$

for constant C_2 independent of φ or r.

Proof. Any function $\varphi \in C^\infty(\overline{B_1})$ can be extended to a function in $C^1_c(B_2)$, which we still denote, with some abuse of notation, by φ, so that

$$\|\nabla\varphi\|_{L^2(B_2)} \leq C_1\big(\varphi\|_{L^2(B_1)} + \|\nabla \varphi\|_{L^2(B_1)}\big),$$

where C_1 does not depend on φ. After rescaling, we obtain

$$\|\nabla\varphi\|_{L^2(B_{2r})} \leq C\left(\|\nabla\varphi\|_{L^2(B_r)} + r^{-1}\|\varphi\|_{L^2(B_r)}\right). \tag{D.4}$$

Now that the integrals over B_r do not depend on the choice of the extension of φ. The integrals over B_{2r}which do depend on the choice of the extension, like the one in the left-hand side of the previous inequality, will appear below, but only at the intermediate steps, and will not appear in the final estimates.

Now, put $q = 2\frac{d-1}{d-2}$. Then our hypothesis $b \in M_1$ allows us to apply Lemma C.2 with thus selected q and $p=2$. For $p=2$, we have $\|(-\Delta)^{\frac{1}{2}}u\|_2 = \|\nabla u\|_2$. So,

$$\Big(\|\varphi\|_{L^q(B_r,|b|)} \leq\Big) \quad \|\varphi\|_{L^q(B_{2r},|b|)} \leq C\|\nabla\varphi\|_{L^2(B_{2r})}, \tag{D.5}$$

where, of course, it is crucial that φ has support in the open ball B_{2r}, so integrating over $\mathbb{R}^d$ is equivalent to integrating over B_{2r}.

By Hölder's inequality,

$$\|\varphi\|_{L^2(B_r,|b|)} \leq \langle |b|\mathbf{1}_{B_r}\rangle^{\frac{1}{2}-\frac{1}{q}} \|\varphi\|_{L^q(B_r,|b|)}, \tag{D.6}$$

where $\frac{1}{2} - \frac{1}{q} = \frac{1}{2}\frac{1}{d-1}$.

Finally, assemblying estimates (D.6)-(D.4) and using our hypothesis (D.2) on $|b|$, we arrive at

$$\|\varphi\|_{L^2(B_r,|b|)} \leq \langle |b|\mathbf{1}_{B_r}\rangle^{\frac{1}{2}-\frac{1}{q}}\|\varphi\|_{L^q(B_r,|b|)} \leq CK^{\frac{1}{2}\frac{1}{d-1}}r^{\frac{1}{2}}\|\nabla\varphi\|_{L^2(B_{2r})}$$
$$\leq C_2\left(r^{\frac{1}{2}}\|\nabla\varphi\|_{L^2(B_r)}+r^{-\frac{1}{2}}\|\varphi\|_{L^2(B_r)}\right),$$

i.e. we have proved (D.3). □

We are in position to complete the proof of implication (D.2) ⇒ (D.1). Fix some $r_0>0$, and do the following for every $x\in\mathbb{R}^d$:

– If $r_0\|\nabla\varphi\|_{L^2(B_{r_0}(x))}\geq\|\varphi\|_{L^2(B_{r_0}(x))}$, then from (D.3)

$$\|\varphi\|_{L^2(B_{r_0}(x),|b|)} \leq C_2\left(r_0^{\frac{1}{2}}\|\nabla\varphi\|_{L^2(B_{r_0}(x))}+\|\nabla\varphi\|^{\frac{1}{2}}_{L^2(B_{r_0}(x))}\|\varphi\|^{\frac{1}{2}}_{L^2(B_{r_0}(x))}\right)$$
$$\leq C_2\left(r_0^{\frac{1}{2}}\|\nabla\varphi\|_{L^2(B_{r_0}(x))}+\|\nabla\varphi\|^{\frac{1}{2}}_{L^2(B_{r_0}(x))}\|\varphi\|^{\frac{1}{2}}_{L^2(B_{r_0}(x))}\right). \tag{D.7}$$

- Otherwise, increase r until one reaches equality $r\|\nabla\varphi\|_{L^2(B_r(x))}=\|\varphi\|_{L^2(B_r(x))}$, in which case (D.3) yields

$$\|\varphi\|_{L^2(B_r(x),|b|)} \leq C_2\left(\|\nabla\varphi\|^{\frac{1}{2}}_{L^2(B_r(x))}\|\varphi\|^{\frac{1}{2}}_{L^2(B_r(x))}+\|\nabla\varphi\|^{\frac{1}{2}}_{L^2(B_r(x))}\|\varphi\|^{\frac{1}{2}}_{L^2(B_r(x))}\right)$$
$$=2C_2\|\nabla\varphi\|^{\frac{1}{2}}_{L^2(B_r(x))}\|\varphi\|^{\frac{1}{2}}_{L^2(B_r(x))}. \tag{D.8}$$

In any case, we obtain, for every $x\in\mathbb{R}^d$ and $r=r(x)$ (either r_0 or increased r, depending on x):

$$\|\varphi\|_{L^2(B_r(x),|b|)} \leq 2C_2\bigg(\|\nabla\varphi\|^{\frac{1}{2}}_{L^2(B_r(x))}\|\varphi\|^{\frac{1}{2}}_{L^2(B_r(x))}+r_0^{\frac{1}{2}}\|\nabla\varphi\|_{L^2(B_{r_0}(x))}\bigg). \tag{D.9}$$

Our goal now is to turn local estimate (D.9) into the multiplicative inequality (D.1). So, we cover $\operatorname{supp}|b|$ with balls $\{B_{r(x)}(x)\}$. Using [M, Theorem 1.2.1], i.e. a variant of the Besicovich covering theorem, one can extract at most countable subcover of finite multiplicity $M=M(d)$. We emphasize that while the covering depends on φ, the multiplicity M does not. Let $\{B^i\}$ be this sub-cover. Squaring (D.9) (considered over each ball B^i in the sub-cover), summing up in i, and using Hölder inequality, we have:

$$\|\varphi\|^2_{L^2(\mathbb{R}^d,|b|)} \leq M\sum_i\|\varphi\|^2_{L^2(B^i,|b|)}$$
$$\leq C_3\left(\sum_i\|\nabla\varphi\|_{L^2(B^i)}\|\varphi\|_{L^2(B^i)}+r_0\sum_i\|\nabla\varphi\|^2_{L^2(B^i)}\right)$$
$$\leq C_3\left(\left(\sum_i\|\nabla\varphi\|^2_{L^2(B^i)}\right)^{\frac{1}{2}}\left(\sum_i\|\varphi\|^2_{L^2(B^i)}\right)^{\frac{1}{2}}+r_0\sum_i\|\nabla\varphi\|^2_{L^2(B^i)}\right).$$

By Jensen's inequality,

$$\big(\sum_i\|\nabla\varphi\|^2_{L^2(B^i)}\big)^{1/2}\leq\sum_i\|\nabla\varphi\|_{L^2(B^i)}\leq M\|\nabla\varphi\|_2$$

and similarly for $\|\varphi\|_2$, which gives us

$$\|\varphi\|^2_{L^2(\mathbb{R}^d,|b|)}\leq M^2C_3(\|\nabla\varphi\|_2\|\varphi\|_2+r_0\|\nabla\varphi\|_2^2)$$

Since r_0 was fixed arbitrarily, letting $r_0\downarrow 0$ yields the multiplicative inequality (D.1).

The reverse direction (D.1) $\Rightarrow$ (D.2) is easier. Choosing in (D.1) test functions $\varphi = \varphi_r \in C_c^\infty(B_r(x))$ such that $\|\nabla\varphi\|_2 \leq cr^{\frac{d-2}{2}}$, $\|\varphi\|_2 \leq cr^{d/2}$, we obtain $\langle |b|, \mathbf{1}_{B_r(x)}\rangle \leq Cr^{d-1}$, i.e. $b \in M_1$. This completes the proof. $\square$

References

[A1] D. R. Adams, Weighted nonlinear potential theory, Trans. Amer. Math. Soc., **297** (1986), 73-94.

[A2] D. R. Adams, A trace inequality for generalized potentials. Stud. Math. **48** (1973) 99-105.

[AX] D. Adams and J. Xiao, Morrey spaces potentials capacities with some PDE applications, J. London Math. Soc. (2024), DOI: 10.1112/jlms.70131.

[ABK] S. Armstrong, A. Bou-Rabee and T. Kuusi, Superdiffusive central limit theorem for a Brownian particle in a critically-correlated incompressible random drift, arXiv:2404.01115.

[B] R. Bass, Diffusions and Elliptic Operators, Springer, 1997.

[BC] R. Bass and Z.-Q. Chen, Brownian motion with singular drift. Ann. Probab., 31 (2003), 791-817.

[Bil] P. Billingsley, Convergence of Probability Measures. Second Edition. Wiley, 1999.

[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.

[BK] S. E. Boutiah and D. Kinzebulatov, Heat kernels of particles with strong attracting interactions, in preparation.

[BJW] D. Bresch, P.-E. Jabin and Z. Wang, Mean field limit and quantitative estimates with singular attractive kernels, Duke Math. J., **172** (2023), 2591-2641 (arXiv:2011.08022).

[CC] G. Cannizzaro and K. Chouk. Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential, Ann. Probab. **46**(3) (2018), 1710-1763.

[C] P. Cattiaux, Entropy on the path space and application to singular diffusions and mean-field models, arXiv:2404.09552.

[CP] P. Cattiaux and L. Pédèches, The 2-D stochastic Keller-Segel particle model: existence and uniqueness, ALEA, Lat. Am. J. Probab. Math. Stat., **13** (2016), 447-463.

[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.-E. Chaudru de Raynal and S. Menozzi, On multidimensional stable-driven stochastic differential equations with Besov drift. Electron. J. Probab. **27** (2022), 1-52.

[CJM] P.E. 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.

[CJM2] P.E. Chaudru de Raynal, J.-F. Jabir and S. Menozzi, Multidimensional stable-driven McKean-Vlasov SDEs with distributional interaction kernel: critical thresholds and related models, arXiv:2302.09900.

[CE] A. Cheskidov and T. Eguchi, Global well-posedness of the Navier-Stokes equations for small initial data in frequency localized Koch-Tataru's space, arXiv:2503.11642.

[Ch] A. S. Cherny, On the uniqueness in law and the pathwise uniqueness for stochastic differential equations, Theory of Probability and its Applications, **46**(3) (2002), 406-419.

[CF] F. Chiarenza and M. Frasca, A remark on a paper by C. Fefferman, Proc. Amer. Math. Soc., **108** (1990), 407-409.

[CLMS] R. Coifman, P.-L. Lions, Y. Meyer and S. Semmes, Compensated compactness and Hardy spaces, J. Math. Pures Appl. **72** (1992) 247-286.

[CPZ] L. Corrias, B. Perthame and H. Zaag, Global solutions of some chemotaxis and angiogenesis systems in high space dimensions, Milan J. Math., **72** (2004), 1-28.

[DD] F. Delarue and R. Diel, Rough paths and 1d SDE with a time dependent distributional drift: application to polymers, Probab. Theory Related Fields, **165** (2016), 1-63.

[EK] S. N. Ethier and T.G. Kurtz, Markov Processes. Characterization and Convergence. Wiley, 2005.

[Fe] C. Fefferman, The uncertainty principle, Bull. Amer. Math. Soc., **9** (1983), 129-206.

[FMT] V. Felli, E. M. Marchini and S. Terracini, On Schrödinger operators with multipolar inverse-square potentials, J. Funct. Anal., **250** (2007), 265-316.

[Fl] F. Flandoli, Regularization by additive noise, in Random Perturbation of PDEs and Fluid Dynamic Models, Lecture Notes in Mathematics, **2015**, Springer (2011), 49-84.

[FGP] F. Flandoli, M. Gubinelli and E. Priola, Well-posedness of the transport equation by stochastic perturbation, Invent. Math., **180** (2010), 1-53.

[FIR] F. Flandoli, E. Issoglio and F. Russo, Multidimensional stochastic differential equations with distributional drift. Trans. Amer. Math. Soc, **369** (2017), 1665–1688.

[FR] F. Flandoli and M. Romito, Markov selections and their regularizing effect for the stochastic three-dimensional Navier–Stokes equations, Probab. Theory Related Fields, **145** (2009), 271-315.

[F] N. Fournier, Stochastic particles for the Keller-Segel equation, CIRM, 2024.

[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.

[FT] N. Fournier and Y. Tardy, A simple proof of non-explosion for measure solutions of the Keller-Segel equation, Kinetic and Related Models, **16**(2) (2023), 178-186 (arXiv:2202.03508).

[GS] M. Glazkov and T. Shilkin, On the L^1-stability for parabolic equations with a supercritical drift term, arXiv:2411.03816.

[GP] L. Gräfner and N. Perkowski, Weak well-posedness of energy solutions to singular SDEs with supercritical distributional drift, arXiv:2407.09046 (2024)

[G] E. Giusti, Direct Methods in the Calculus of Variations, World Scientific, 2003.

[Gr] L. Grafakos, Modern Fourier Analysis. Second Edition. Springer, 2009.

[HZ] Z. Hao and X. Zhang, SDEs with supercritical distributional drifts, arXiv:2312.11145.

[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.

[HL] T. Hoffmann-Ostenhof and A. Laptev, Hardy inequalities with homogeneous weights, J. Funct. Anal., **268** (2015), 3278-3289.

[HHLT] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, A. Laptev and J. Tidblom, Many-particle Hardy inequalities, J. Lond. Math. Soc., **77** (2008), 99-114.

[JL] W. Jager and S. Luckhaus, On explosions of solutions to a system of partial differential equations modelling chemotaxis, Trans. Amer. Math. Soc., **329** (1992), 819-824.

[Ka] T. Kato, "Perturbation Theory for Linear Operators", Springer-Verlag Berlin Heidelberg, 1995.

[KSa] R. Kerman and E. Sawyer, The trace inequality and eigenvalue estimates for Schrödinger operators, Ann. Inst. Fourier **36**(4) (1986), 207-228.

[K1] 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).

[K2] D. Kinzebulatov, Feller evolution families and parabolic equations with form-bounded vector fields, Osaka J. Math., **54** (2017), 499-516 (arXiv:1407.4861).

[K3] D. Kinzebulatov, Laplacian with singular drift in a critical borderline case, Math. Nachr., to appear (arXiv:2309.04436).

[K4] D.Kinzebulatov, Form-boundedness and SDEs with singular drift, INdAM Meeting 2022: Kolmogorov Operators and Their Applications (arXiv:2305.00146).

[K5] D. Kinzebulatov, Parabolic equations and SDEs with time-inhomogeneous Morrey drift, arXiv:2301.13805.

[K6] D. Kinzebulatov, On particle systems and critical strengths of general singular interactions, Ann. Inst. Henri Poincaré (B) Probab. Stat., to appear (arXiv:2402.17009).

[K7] D. Kinzebulatov, Regularity theory of Kolmogorov operator revisited, Canadian Bull. Math. **64** (2021), 725-736 (arXiv:1807.07597).

[K8] D. Kinzebulatov, Feller generators with measure-valued drifts, Potential Anal., **48** (2018), 207-222.

[K9] D. Kinzebulatov, Non-local parabolic equations with singular (Morrey) time-inhomogeneous drift, La Matematica, the volume dedicated to the memory of D.R. Adams, to appear (arXiv:2405.08652).

[KM1] D. Kinzebulatov and K.R. Madou, Strong solutions of SDEs with singular (form-bounded) drift via Roeckner-Zhao approach, Stochastics and Dynamics, to appear (arXiv:2306.04825).

[KM2] D.Kinzebulatov and K.R.Madou, Stochastic equations with time-dependent singular drift, J. Differential Equations, **337** (2022), 255-293 (arXiv:2105.07312).

[KM3] 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).

[KM4] D. Kinzebulatov and K.R. Madou, On strong solutions of SDEs with critical discontinuities in diffusion coefficients, *in preparation*.

[KS1] D. Kinzebulatov and Yu.A. Semënov, Brownian motion with general drift, Stoch. Proc. Appl., **130** (2020), 2737-2750 (arXiv:1710.06729).

[KS2] D. Kinzebulatov and Yu. A. Semënov, Sharp solvability for singular SDEs, Electr. J. Probab., **28** (2023), article no. 69, 1-15. (arXiv:2110.11232).

[KS3] 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).

[KS4] D. Kinzebulatov and Yu. A. Semënov, Heat kernel bounds for parabolic equations with singular (form-bounded) vector fields, Math. Ann., **384** (2022), 1883-1929.

[KS5] D. Kinzebulatov and Yu. A. Semënov, Regularity for parabolic equations with singular non-zero divergence vector fields, J. Differential Equations, to appear.

[KS6] D. Kinzebulatov and Yu. A. Semënov, Feller generators with drifts in the critical range, J. Differential Equations, to appear (arXiv:2405.12332).

[KS7] 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).

[KS8] D. Kinzebulatov and Yu.A. Semënov, Remarks on parabolic Kolmogorov operator, Prob. Theory. Appl., to appear (arXiv:2303.03993).

[KSS] D. Kinzebulatov, Yu. A. Semënov and R. Song, Stochastic transport equation with singular drift, Ann. Inst. Henri Poincaré (B) Probab. Stat. **60**(1) (2024), 731-752 (arXiv:2102.10610).

[KV] D. Kinzebulatov and R. Vafadar, On divergence-free (form-bounded type) drifts, Discrete Contin. Dyn. Syst. Ser. S., **17** (2024), 2083-2107 (arXiv:2209.04537).

[KT] H. Koch and D. Tataru, Well-posedness for the Navier-Stokes equations, Advances in Mathematics, **157**(1) (2001), 22-35.

[KoS] 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 $L_w^{1/2}(R^l)$-potentials, J. Math. Phys., **22** (1981), 1033-1044.

[KrS] J. Kristensen and B. Stroffolini, The Gehring lemma: dimension free estimates, Nonlinear Analysis **177**(B) (2018), 601-610.

[Kr1] N.V. Krylov, Once again on weak solutions of time inhomogeneous Itô's equations with VMO diffusion and Morrey drift, Electr. J. Probab. **29**, paper 95, 1-19 (2024) (arXiv:2304.04634).

[Kr2] N.V. Krylov, On strong solutions of Itô's equations with $D\sigma$ and b in Morrey classes containing L^d, Ann. Probab. **51**(5) (2023) 1729-1751 (arXiv:2111.13795).

[Kr3] N.V. Krylov, On parabolic equations in Morrey spaces with VMO a and Morrey b, c, NoDEA, to appear.

[Kr4] N.V. Krylov, Time-homogeneous Stochastic Itô Equations and Second-Order PDEs with Singularities, book.

[Kr5] N.V. Krylov, Essentials of Real Analysis and Sobolev-Morrey Spaces for Second-order Elliptic and Parabolic PDEs with Singular Lower-order Coefficients, book.

[Kr6] 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).

[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.

[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.

[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.

[M] V. G. Mazya, Sobolev Spaces, Springer 1985.

[M2] V.G. Mazya, Capacitary criteria for the boundedness of linear operators in function spaces, Soviet Math. Dokl. **11** (1970), 133-137.

[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.

[MV2] V. G. Mazya and I. E. Verbitsky, Infinitesimal form boundedness and Trudinger's subordination for the Schrödinger operator, Invent. Math. **162** (2005), 81-136.

[NY] E. Nakai and K. Yabuta, Pointwise multipliers for functions of bounded mean oscillation, J. Math. Soc. Japan **37**(2) (1985), 207-218.

[N] J. Nash, Continuity of solutions of parabolic and elliptic equations, Amer. Math. J. **80** (1) (1958), 931-954.

[ORT] A. Ohashi, F. Russo, and A. Teixeira. SDEs for Bessel processes in low dimension and path-dependent extensions. ALEA **20** (2023), 1111-1138.

[QX] Z. Qian and G. Xi, Parabolic equations with singular divergence-free drift vector fields, J. London Math. Soc. **100** (1) (2019), 17-40.

[QX2] Z. Qian and G. Xi, Parabolic equations with divergence-free drift in space $L^l_t L^q_x$, Indiana Univ. Math. J. **68**(3) (2019), 761-797.

[RY] D. Revuz and M. Yor, Continuous Martingales and Browinan Motion, Springer, 1991.

[RZ] M. Röckner and G. Zhao, SDEs with critical time dependent drifts: weak solutions, Bernoulli, **29** (2023), 757-784 (arXiv:2012.04161).

[RZ2] M. Röckner and G. Zhao, SDEs with critical time dependent drifts: strong solutions, arXiv:2103.05803.

[S] Yu. A. Semënov, Regularity theorems for parabolic equations, J. Funct. Anal., **231** (2006), 375-417.

[SSŠZ] G. Seregin, L. Silvestre, V. Šverak and A. Zlatoš, On divergence-free drifts, J. Differential Equations, **252**(1) (2012), 505-540.

[T] Y. Tardy, Weak convergence of the empirical measure for the Keller-Segel model in both sub-critical and critical cases, Electr. J. Probab., to appear (arXiv:2205.04968).

[V] A. Yu. Veretennikov, Strong solutions and explicit formulas for solutions of stochastic integral equations, Mat. Sb. **111** (3) (1980) 434-452; in Russian, English translation in Math. USSR-Sbornik **39** (1981) 387-403.

[HWY] J. L. Wei, J. Hu and C. Yuan, Stochastic differential equations with critically low regularity growing drift, https://ssrn.com/abstract=5025237

[Za] Q. S. Zhang, Gaussian bounds for the fundamental solutions of $\nabla(A\nabla u) + B\nabla u - u_t = 0$, Mauscripta Math. **93** (1997), 381-390.

[Za2] Q. S. Zhang, A strong regularity result for parabolic equations, Comm. Math. Phys. **244** (2004), 245-260.

[Za3] Q. S. Zhang, Local estimates of two linear parabolic equations with singular coefficients, Pacific J. Math. **223** (2006), 367-396.

[Z] X. Zhang, Stochastic differential equations with Sobolev coefficients and applications, Ann. Appl. Probab. **26** (2016), 2697–2732.

[ZZ1] X. Zhang and G. Zhao, Stochastic Lagrangian paths for Leray solutions of $3D$ Naiver-Stokes equations, Comm. Math. Phys., **381**(2) (2021), 491-525.

[ZZ2] X. Zhang and G. Zhao, Heat kernels and ergodicity of SDEs with distributional drifts. arXiv:1710.10537, 2017.

[Zh] G. Zhao, Stochastic Lagrangian flows for SDEs with rough coefficients, arXiv:1911.05562, 2019.

Email address: `damir.kinzebulatov@mat.ulaval.ca`
Email address: `reihaneh.vafadar-seyedi-nasl.1@ulaval.ca`

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