

# ON PARTICLE SYSTEMS AND CRITICAL STRENGTHS OF GENERAL SINGULAR INTERACTIONS

D. KINZEBULATOV

ABSTRACT. For finite interacting particle systems with strong repulsing-attracting or general interactions, we prove global weak well-posedness almost up to the critical threshold of the strengths of attracting interactions (independent of the number of particles), and establish other regularity results, such as a heat kernel bound in the regions where strongly attracting particles are close to each other. Our main analytic instruments are a variant of De Giorgi's method in  $L^p$  and an abstract desingularization theorem.

## CONTENTS

1. Introduction	2
1.1. Notations	5
2. Main results	5
2.1. Particle systems	5
2.2. SDEs with general singular drifts	16
2.3. Embedding theorem and Hölder continuity of solutions	20
2.4. Gradient bounds	21
3. Smooth approximation of form-bounded vector fields	22
4. Proofs of Lemmas 1 and 2	24
5. Caccioppoli's inequality	26
6. Proof of Theorem 8	29
7. Proof of Theorem 9	32
8. Proof of Theorem 5	35
9. Proof of Theorem 6	40
10. Proof of Theorems 1 and 2	41
11. Proof of Theorem 3	42
12. Proof of Theorem 4	44
13. Proof of Theorem 10	45
14. Proof of Theorem 7	48
Appendix A. A desingularization theorem from [33]	52
References	54

---

UNIVERSITÉ LAVAL, DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, QUÉBEC, QC, CANADA  
*E-mail address:* [damir.kinzebulatov@mat.ulaval.ca](mailto:damir.kinzebulatov@mat.ulaval.ca).

*Key words and phrases.* Interacting particle systems, singular stochastic equations, form-boundedness.  
 The research of the author is supported by the NSERC (grant RGPIN-2017-05567).

## 1. INTRODUCTION

The paper is devoted to study of well-posedness and other properties of multi-particle system

$$dX_i = M_i(X_i)dt - \frac{1}{N} \sum_{j=1, j \neq i}^N K_{ij}(X_i - X_j)dt + \sqrt{2}dB_i, \quad X_i(0) = x_i \in \mathbb{R}^d \quad (1.1)$$

( $i = 1, \dots, N$ ), where  $[0, \infty[ \ni t \mapsto X_i(t)$  is the trajectory of the  $i$ -th particle in  $\mathbb{R}^d$ ,  $\{B_i(t)\}_{t \geq 0}$  are  $d$ -dimensional independent Brownian motions, under broad assumptions on singular drifts and interaction kernels  $M_i, K_{ij} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  that can be repulsing-attracting or have general form, e.g.  $|K|$  in  $L^d$  or weak  $L^d$  (Definitions 2-4). Concrete examples include critical-order point singularities

$$K(y) = \pm \sqrt{\kappa} \frac{d-2}{2} \frac{y}{|y|^2}, \quad y \in \mathbb{R}^d \quad (1.2)$$

(defined to be zero at the origin) or hypersurface singularities

$$|K(y)|^2 = C \frac{\mathbf{1}_{\{\frac{1}{2} \leq |y| \leq \frac{3}{2}\}}}{||y| - 1|(-\ln ||y| - 1|)^\beta}, \quad \beta > 1 \quad (1.3)$$

(so  $|K| \notin L_{\text{loc}}^{2+\epsilon}$  for any  $\epsilon > 0$ ), or their sums. Throughout the paper, dimension  $d \geq 3$ . Important case  $d = 2$  requires separate study which we plan to carry out elsewhere.

In Theorems 1-4 we cover half-critical or critical (at least in high dimensions  $d$ ) ranges of the strength of attracting interactions, establish, in particular, global weak well-posedness of stochastic particle system (1.1) and, for the model singular attracting kernel in (1.2), prove a necessarily non-Gaussian upper bound on the heat kernel of (1.1) which we believe to be optimal in the regions where the particles are close to each other. To this end, we use a variant of De Giorgi's method and an abstract desingularization theorem obtained earlier in [33].

To illustrate the notion of the critical strength of interactions, let us consider particle system (1.1) with the model singular attracting kernel (1.2)

$$dX_i = -\frac{1}{N} \sum_{j=1, j \neq i}^N \sqrt{\kappa} \frac{d-2}{2} \frac{X_i - X_j}{|X_i - X_j|^2} dt + \sqrt{2}dB_i. \quad (1.4)$$

As the strength of attraction  $\kappa$  increases, the behaviour of trajectories  $X_i$  undergoes qualitative changes, e.g. the particles start to agglomerate and one can prove for (1.4) only local existence until the moment the particles collide. The strength of attraction  $\kappa_* = 16(\frac{d}{d-2})^2$  is a critical threshold for system (1.4)<sup>1</sup>. Informally, this can be seen by inspecting the density of the formal invariant measure of (1.4):

$$\psi(x) = \prod_{1 \leq i < j \leq N} |x_i - x_j|^{-\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N}}, \quad x = (x_1, \dots, x_N) \in \mathbb{R}^{Nd},$$

which is locally summable if and only if  $\kappa < \kappa_*$  (also, as  $\kappa$  reaches and surpasses  $\kappa_{**} = 16$ ,  $\psi$  ceases to be in  $W_{\text{loc}}^{2,1}$ ).

In the case  $N = 2$ , one can show that if  $\kappa > \kappa_*$  and  $X_1(0) = X_2(0)$ , then system (1.4) has no weak solution, and if  $\kappa > \kappa_{**}$ , then particles collide in finite time with positive probability<sup>2</sup>. However, if  $\kappa < \kappa_{**}$ , then (1.4) has global in time weak solution.

<sup>1</sup>This normalization of the coefficient in (1.4) is due to the use of Hardy's inequality, see Example 1(b).

<sup>2</sup> $Z = (X_1 - X_2)/\sqrt{2}$  satisfies  $dZ = -\frac{\sqrt{\kappa}}{2} \frac{d-2}{2} Z|Z|^{-2} dt + \sqrt{2}dB$  in  $\mathbb{R}^d$ , so a well-known counterexample applies, see e.g. [7].

Regarding the case when the number of particles  $N$  is large, we refer to [9, 13] for the proofs of weak well-posedness of the two-dimensional counterpart of (1.4) and analysis of the phase transition effects as the strength of attraction  $\kappa$  increases. The well-posedness and blow-up effects for the two-dimensional McKean-Vlasov SDE arising as the limit of the particle system as  $N \rightarrow \infty$  (the corresponding McKean-Vlasov PDE is the famous Keller-Segel model of chemotaxis) is studied in [13] and in recent papers [14, 47] where the authors handle the entire ranges of the strengths of attracting interactions as well as the critical threshold. The proofs in the cited papers exploit the special structure of the interaction kernel  $K(y) = c|y|^{-2}y$ .

Many applications require one to handle more general than  $K(y) = c|y|^{-2}y$  singular interactions in dimension  $d = 2, 3$  and higher. There is extensive literature on singular interaction kernels having special form, e.g. gradient form, which includes many kernels arising in Statistical Physics. We refer, in particular, to [3, 40] where the authors proved for such interaction kernels weak and strong well-posedness of particle system (1.1), however, excluding the singularities of (1.2). We also refer to [17, 47] where the authors prove, as a part of their results on the propagation of chaos, strong well-posedness of the particle system (1.1) for  $K$  in the sub-critical Ladyzhenskaya-Prodi-Serrin class. Applied to (1.1), their condition reads as  $|K| \in L^p + L^\infty$ ,  $p > d$ , as is needed to use Girsanov's transform, which, again, excludes (1.2). We also mention article [8] where the authors work at the PDE level on the torus, consider interaction kernels of gradient form with the interaction potential pointwise comparable to  $\sqrt{\kappa} \frac{d-2}{2} \log|x|$  (which thus includes the attracting kernel in (1.2)) and obtain quantitative estimates on the propagation of chaos for the McKean-Vlasov PDE for all  $\kappa < \kappa_*$ .

In the present paper we address the problem of well-posedness of stochastic particle system (1.1) directly, by rewriting (1.1) as SDE

$$dZ = b(Z)dt + \sqrt{2}dB, \quad B \text{ is a Brownian motion in } \mathbb{R}^{Nd} \tag{1.5}$$

with  $Z = (X_1, \dots, X_N)$  and drift  $b = (b_1, \dots, b_N) : \mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$  defined by

$$b_i(x) := \frac{1}{N} \sum_{j=1, j \neq i}^N K_{ij}(x_i - x_j), \quad x = (x_1, \dots, x_N) \in \mathbb{R}^{Nd}, \quad 1 \leq i \leq N \tag{1.6}$$

and then applying results on well-posedness for SDEs with general drifts, in particular, our Theorem 5 below. Until recently, the results on general singular SDEs could not compete, in terms of the admissible point singularities of the drift, with the aforementioned results on particle systems with model interactions (1.2). In the past few years, however, there was a substantial progress in proving weak and strong well-posedness of SDE (1.5) with general drift  $b$ , which now can have critical-order singularities, see [27, 28, 25, 24], [37, 38, 39], [43, 44]. That said, to apply these results to (1.1) in a way that would allow to control the strength of interactions (measured, in the case of (1.2), by constant  $\kappa$ ), one needs to keep track of the strength of the singularities of the drift  $b$  (measured by “form-bounds”, see below). To the best of our knowledge, the only paper that covered all (at least in high dimensions) admissible values of the form-bound of general drift  $b$  was [28]. In a number of ways, the present paper continues [28].

More specifically, below we consider interaction kernels satisfying the following dimension-free conditions. Let  $\langle \cdot, \cdot \rangle$  denote the integration over  $\mathbb{R}^d$  and  $\|\cdot\|_2$  the  $L^2$  norm. For a general  $K$  we require that the following quadratic form inequality holds:

$$\langle |K|^2 \varphi, \varphi \rangle \leq \kappa \|\nabla \varphi\|_2^2 + c_\kappa \|\varphi\|_2^2 \quad \forall \varphi \in C_c^\infty. \tag{1.7}$$

In the case  $K$  has an attraction component, we require

$$\begin{cases} \langle K|\varphi, \varphi \rangle \leq \kappa_0 \|\nabla\varphi\|_2 \|\varphi\|_2 + c_{\kappa_0} \|\varphi\|_2^2 \\ \langle (\operatorname{div}K)_+\varphi, \varphi \rangle \leq \kappa_+ \|\nabla\varphi\|_2^2 + c_{\kappa_+} \|\varphi\|_2^2, \end{cases} \quad (1.8)$$

where constant  $\kappa_+$  measures the strength of attraction;  $\kappa_0$  and  $c_\kappa, c_{\kappa_0}, c_{\kappa_+}$  are any finite constants. These are broad assumptions on the interaction kernel  $K$  that allow it to have critical-order singularities. See the next section for examples which include interaction kernels (1.2) and (1.3). The constants  $\kappa, \kappa_0$  and  $\kappa_+$  are called form-bounds. Importantly, the vector field  $b$  defined by (1.6) satisfies the same conditions but with the integration over  $\mathbb{R}^{Nd}$  and essentially the same form-bounds (i.e. obtained from  $\kappa, \kappa_+$  by dividing the latter by constants that tend to 1 as the number of particles increases). Thus, we obtain our results for the interacting particle system (1.1) (Theorems 1-4) right away from our results on the general singular SDE (1.5). These are Theorem 5-7. For this to work, in Theorem 5 we need to impose dimension-free assumptions on the form-bounds of the drift  $b$  in  $\mathbb{R}^{Nd}$  (so that the assumptions on the strengths of interactions  $\kappa, \kappa_+$  would not depend on the number of particles). This is achieved, as in [28], by means of De Giorgi's method ran in  $L^p$  which allows to "decouple" the proof of the tightness estimate needed to establish the existence of a martingale solution (cf. (2.17)) from any strong gradient bounds on solutions of the corresponding elliptic or parabolic equations (that, in turn, introduce dependence on the dimension in the assumptions on the form-bounds of  $b$ ). Let us add that running De Giorgi's method in  $L^p$  with  $p$  large (rather than with the usual  $p = 2$ ) allows to maximize admissible values of the form bounds/strengths of interactions.

Theorem 5-7 on the general singular SDE (1.5) are of interest on their own. The existence of a martingale solution part of Theorem 5 under condition  $(\mathbb{A}_1)$  (i.e. analogue of (1.7)) was established in [28], but we included this statement in Theorem 5 anyway for the sake of completeness. The novelty of Theorem 5 is related to condition (1.8), the strong Markov and Feller properties. See comments after Theorem 3.

The analytic core of the paper are Theorems 8 and 9 from which Theorem 5 and a number of other results follow. Theorem 8 is proved by showing that solution of the elliptic Kolmogorov equation  $u$  belongs to appropriate  $L^p$  De Giorgi's classes and then following De Giorgi's method. These De Giorgi classes, however, are somewhat different from the  $L^p$  De Giorgi classes found in the literature, i.e. they contain the integrals of

$$|\nabla(u - k)_+^{p/2}|^2, \quad k \in \mathbb{R},$$

rather than the integrals of  $|\nabla(u - k)_+|^p$ , the latter arising naturally in the problems of optimization. Theorem 9, i.e. the embedding theorem, plays a key role in the proof of Theorem 5, but also has applications outside of the present paper. In order to establish Caccioppoli's inequality under condition (1.8) (as is needed to prove Theorems 8 and 9) we extend to the non-homogeneous  $L^p$  setting the iteration procedure introduced in earlier paper [34].

De Giorgi's method was used earlier in the context of singular SDEs in [43, 51, 53]. There the authors considered singular drifts arising in the study of 3D Navier-Stokes equations.

Let us also mention recent results in [11] on interacting particle systems and McKean-Vlasov SDEs with the interaction kernels and drifts in Besov's spaces of distributions, see also references therein. However, these assumptions are somewhat orthogonal to the present work and, at least at the moment, do not include the model interaction kernels (1.2) (while including other highly irregular distributional kernels) or keep track of the strength of interactions  $\kappa$ . We also comment in Remark 9 on the existing literature on SDEs with supercritical drifts.

1.1. **Notations.** Put

$$\langle f \rangle := \int_{\mathbb{R}^d} f(y) dy, \quad \langle f, g \rangle := \langle fg \rangle$$

(all functions in this paper are real-valued). Let  $L^p = L^p(\mathbb{R}^d, dy)$  be the Lebesgue spaces endowed with the norm  $\|\cdot\|_p$ . Let  $W^{1,p}$  denote the corresponding Sobolev spaces. Denote by  $[L^p]^d$  the space of vector fields  $\mathbb{R}^d \rightarrow \mathbb{R}^d$  with entries in  $L^p$ . We denote by  $\|\cdot\|_{p \rightarrow q}$  the  $L^p \rightarrow L^q$  operator norm. Let  $C_\infty$  denote the space of continuous functions on  $\mathbb{R}^d$  vanishing at infinity, endowed with the sup-norm. Let  $B_R(y) \subset \mathbb{R}^d$  be the open ball of radius  $R$  centered at  $y \in \mathbb{R}^d$ ,  $|B_R(x)|$  denotes its volume. Set  $B_R := B_R(0)$ . Given a function  $f$ , we denote its positive and negative parts by

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

Set

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

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 Friedrichs mollifier of a function  $h \in L^1_{\text{loc}}$  (or a vector field with entries in  $L^1_{\text{loc}}$ ) by  $E_\varepsilon h := \gamma_\varepsilon * h$ .

## 2. MAIN RESULTS

2.1. **Particle systems.** Let us first consider particle system

$$X_i(t) = x_i - \frac{1}{N} \sum_{j=1, j \neq i}^N \int_0^t K_{ij}(X_i(s) - X_j(s)) ds + \sqrt{2} B_i(t), \quad 1 \leq i \leq N, \quad t \in [0, T], \quad (2.1)$$

where  $x = (x_1, \dots, x_N) \in \mathbb{R}^{Nd}$ ,  $N \geq 2$ . That is, we exclude the drift terms  $M(X_i)$  from (1.1). However, we will explain in Remark 5 below how to put the drifts back there.

Let  $e_t : C([0, T], \mathbb{R}^{Nd}) \rightarrow \mathbb{R}^{Nd}$  be defined by  $e_t(\omega) := \omega_t$ .

**DEFINITION 1.** A probability measure  $\mathbb{P}_x$  ( $x \in \mathbb{R}^{Nd}$ ) on the canonical space of continuous trajectories  $\omega = (\omega^1, \dots, \omega^N)$  in  $\mathbb{R}^{Nd}$  is called a martingale solution to particle system (2.1) on  $[0, T]$  if

- 1)  $\mathbb{P}_{x,0} = \delta_x$ , where  $\mathbb{P}_{x,t} := \mathbb{P} \circ e_t^{-1}$  (on  $\mathbb{R}^{Nd}$ );
- 2)  $\mathbb{E}_x \sum_{i=1}^N \sum_{j=1, j \neq i}^N \int_0^T |K_{ij}(\omega_t^i - \omega_t^j)| dt < \infty$ ;
- 3) for every  $\phi \in C_c^2(\mathbb{R}^{Nd})$  the process

$$[0, T] \ni r \mapsto \phi(\omega_r) - \phi(x) + \int_0^r (-\Delta_y \phi(\omega_t) + \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N K_{ij}(\omega_t^i - \omega_t^j) \cdot \nabla_{y_i} \phi(\omega_t)) dt$$

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

We start with the general interaction kernels.

**DEFINITION 2.** A vector field  $K \in [L^2_{\text{loc}}]^d$  is said to be form-bounded if it satisfies quadratic form inequality

$$\langle |K|^2 \varphi, \varphi \rangle \leq \kappa \|\nabla \varphi\|_2^2 + c_\kappa \|\varphi\|_2^2 \quad (2.2)$$

for all  $\varphi \in W^{1,2}$  for some constants  $\kappa$  (“form-bound”) and  $c_\kappa$ . This is written as  $K \in \mathbf{F}_\kappa$ .

The class of form-bounded vector fields  $\mathbf{F}_\kappa$  is closed with respect to addition and multiplication by functions from  $L^\infty$  (up to change of  $\kappa$ ).

The form-boundedness condition  $b$  appears already in the Lax-Milgram theorem. It provides coercivity in  $L^2$  of the Dirichlet form of  $-\nabla \cdot a \cdot \nabla + b \cdot \nabla$  with a uniformly elliptic matrix  $a$ .

EXAMPLE 1. (a) If  $|K| \in L^p$  for some  $d \leq p \leq \infty$ , then  $K$  is in  $\mathbf{F}_\kappa$  with arbitrarily small form-bound  $\kappa$  (in this sense the class  $|K| \in L^d$  is sub-critical). Indeed, if e.g.  $|K| \in L^d$ , then, for every  $\varepsilon > 0$ , we can represent  $K = K_1 + K_2$ , where  $\|K_1\|_d < \varepsilon$  and  $\|K_2\|_\infty < \infty$ . So, we obtain, using the Sobolev embedding theorem,

$$\|K\varphi\|_2^2 \leq 2\|K_1\|_d^2 \|\varphi\|_{\frac{2d}{d-2}}^2 + 2\|K_2\|_\infty^2 \|\varphi\|_2^2 \leq C_S 2\|K_1\|_d^2 \|\nabla\varphi\|_2^2 + 2\|K_2\|_\infty^2 \|\varphi\|_2^2,$$

hence  $K \in \mathbf{F}_\kappa$  with  $\kappa = C_S 2\varepsilon$  and  $c_\kappa = 2\|K_2\|_\infty^2$ .

(b) (Critical point singularities) The model singular interaction kernel

$$K(y) = \pm \sqrt{\kappa} \frac{d-2}{2} \frac{y}{|y|^2}, \quad y \in \mathbb{R}^d, \quad (2.3)$$

is in  $\mathbf{F}_\kappa$  with  $c_\kappa = 0$  (but it is not in any  $\mathbf{F}_{\kappa'}$  with  $\kappa' < \kappa$ , regardless of the value of  $c_{\kappa'}$ ). This is the well known Hardy inequality.

(c) (Weak  $L^d$  class interaction kernels) More generally, vector fields  $K$  in  $L^{d,\infty}$ , i.e. such that

$$\|K\|_{d,\infty} := \sup_{s>0} s |\{y \in \mathbb{R}^d : |K(y)| > s\}|^{1/d} < \infty \quad (2.4)$$

are in  $\mathbf{F}_\kappa$  with  $\sqrt{\kappa} = \|K\|_{d,\infty} |B_1(0)|^{-\frac{1}{d}} \frac{2}{d-2}$ , see [35]. This example includes (2.3).

(d) (Morrey class interaction kernels) The scaling-invariant Morrey class  $M_{2+\varepsilon}$ , with  $\varepsilon > 0$  fixed arbitrarily small, consists of vector fields  $K \in [L_{\text{loc}}^{2+\varepsilon}]^d$  such that

$$\|K\|_{M_{2+\varepsilon}} := \sup_{r>0, y \in \mathbb{R}^d} r \left( \frac{1}{|B_r(y)|} \int_{B_r(y)} |K|^{2+\varepsilon} dy \right)^{\frac{1}{2+\varepsilon}} < \infty. \quad (2.5)$$

By [12], if  $K \in M_{2+\varepsilon}$ , then  $K \in \mathbf{F}_\kappa$  with  $\kappa = c\|K\|_{M_{2+\varepsilon}}$  for a constant  $c = c(\varepsilon)$ . This sufficient condition for form-boundedness can be further refined by considering the Chang-Wilson-Wolff class [10].

On the other hand, a simple argument with cutoff functions shows that the class of form-bounded vector fields  $\mathbf{F}_\kappa$  is contained in the Morrey class  $M_2$ .

(e) (Hypersurface singularities) Any interaction kernel  $K$  of the form (1.3) is form-bounded, which can be seen from the previous example by arguing locally.

An example of different kind is given by the weighted Hardy inequality of [19]. Fix  $0 \leq \Phi \in L^q(S^{d-1})$  for some  $q \geq \frac{2(d-2)^2}{2(d-1)} + 1$ , where  $S^{d-1}$  is the unit sphere in  $\mathbb{R}^d$ . If

$$|K(y)|^2 \leq \kappa \frac{(d-2)^2}{4} c \frac{\Phi(y/|y|)}{|y|^2},$$

where  $c := \frac{|S^{d-1}|^{\frac{1}{q}}}{\|\Phi\|_{L^q(S^{d-1})}}$ , then  $K \in \mathbf{F}_\kappa$  with  $c_\kappa = 0$ .

Put

$$\mathbf{F} := \{K \mid K \in \mathbf{F}_\kappa \text{ for some } \kappa < \infty\}.$$

**Theorem 1.** *Assume that the interaction kernels  $K_{ij}$  in particle system (2.1) satisfy*

$$K_{ij} \in \mathbf{F}_\kappa \quad \text{with } \kappa < 4 \left( \frac{N}{N-1} \right)^2 \quad (2.6)$$

*Then the following are true:*

- (i) *There exists a strong Markov family of martingale solutions  $\{\mathbb{P}_x\}_{x \in \mathbb{R}^{Nd}}$  of particle system (2.1).*
- (ii) *The function*

$$u(x) := \mathbb{E}_{\mathbb{P}_x} \int_0^\infty e^{-\lambda s} f(\omega_s^1, \dots, \omega_s^N) ds, \quad x \in \mathbb{R}^{Nd}, \quad f \in C_c^\infty(\mathbb{R}^{Nd}), \quad (2.7)$$

*where  $\lambda$  is assumed to be sufficiently large, is a locally Hölder continuous weak solution to elliptic Kolmogorov equation*

$$\left( \lambda - \Delta + \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N K_{ij}(x_i - x_j) \cdot \nabla_{x_i} \right) u = f, \quad x = (x_1, \dots, x_N), \quad (2.8)$$

*see definitions in Remark 8 where we also discuss the uniqueness of  $u$ .*

Using a standard argument, we can further show that the martingale solutions in assertion (i) are weak solutions. We recall from the discussion in the introduction that if  $\kappa$  is taken to be too large then a weak solution to the particle system (2.1) ceases to exist. So, in Theorem 1(i) we are dealing with the critical scale of the strength of interactions.

Having at hand Theorem 1, we are now in position to ask what additional assumptions on  $K_{ij}$  provide other regularity properties of the particle system. Indeed, in Theorems 3 and 4 we will make the assertion of Theorem 1 and of its counterpart in the case  $K_{ij}$  have repulsion-attraction structure – Theorem 2 – more detailed, at expense of imposing some additional assumptions on the interaction kernels  $K_{ij}$ .

In Remark 3 we comment on the regularity theory of the backward Kolmogorov equation for particle system (2.1) in the case when  $\kappa = 4 \left( \frac{N}{N-1} \right)^2$ .

We now turn to the interaction kernels having repulsion-attraction structure.

**DEFINITION 3.** A vector field  $K \in [L_{\text{loc}}^1]^d$  is said to be multiplicatively form-bounded if

$$\langle |K| \varphi, \varphi \rangle \leq \kappa_0 \|\nabla \varphi\|_2 \|\varphi\|_2 + c_{\kappa_0} \|\varphi\|_2^2 \quad (2.9)$$

for all  $\varphi \in W^{1,2}$  for some constants  $\kappa_0$  (“multiplicative form-bound”) and  $c_{\kappa_0}$ . We abbreviate this as  $K \in \mathbf{MF}_{\kappa_0}$ .

Every form-bounded vector field is multiplicatively form-bounded, but not vice versa, see Remark 6. But we will additionally require from the positive part  $(\text{div} K)_+$  of  $\text{div} K$ , which can be viewed as describing attraction between the particles, to be a form-bounded “potential”:

**DEFINITION 4.**  $(\text{div} K)_+ \in L_{\text{loc}}^1$  is said to be form-bounded if there exists constant  $\kappa_+$  such that

$$\langle (\text{div} K)_+ \varphi, \varphi \rangle \leq \kappa_+ \|\nabla \varphi\|_2^2 + c_{\kappa_+} \|\varphi\|_2^2 \quad (2.10)$$

for some  $c_{\kappa_+}$ . We abbreviate this, with slight abuse of notation, as  $(\text{div} K)_+^{\frac{1}{2}} \in \mathbf{F}_{\kappa_+}$ .

**EXAMPLE 2.** (a) (Critical point singularities) The model singular interaction kernel (2.3) is multiplicatively form-bounded, see Remark 6. In turn,

$$\text{div} K = \pm \sqrt{\kappa} \frac{(d-2)^2}{2} |y|^{-2},$$

so, if there is plus in front of  $\sqrt{\kappa}$  (attraction), then by Hardy's inequality  $(\operatorname{div} K)_+^{\frac{1}{2}} \in \mathbf{F}_{\kappa_+}$ ,  $\kappa_+ = 2\sqrt{\kappa}$ ,  $c_{\kappa_+} = 0$ .

- (b) (Essentially largest possible scaling-invariant Morrey class) The class  $\mathbf{MF}_{\kappa_0}$  contains the largest possible up to the strict inequality in  $\varepsilon > 0$  scaling-invariant Morrey class  $M_{1+\varepsilon}$ , i.e. if

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

then  $M_{1+\varepsilon} \subset \mathbf{MF}_{\kappa_0}$  with  $\kappa_0 = c(d, \varepsilon) \|K\|_{M_{1+\varepsilon}}$ , see Remark 6. Similarly, if the following Morrey class condition is satisfied

$$\sup_{r>0, y \in \mathbb{R}^d} r^2 \left( \frac{1}{|B_r(y)|} \int_{B_r(y)} |(\operatorname{div} K)_+|^{1+\varepsilon} dy \right)^{\frac{1}{1+\varepsilon}} < \infty$$

then  $(\operatorname{div} K)_+^{\frac{1}{2}} \in \mathbf{F}_{\kappa_+}$  with appropriate  $\kappa_+$ .

The class  $b \in \mathbf{MF}_\delta$  under additional hypothesis  $\operatorname{div} b = 0$  appeared in [45] as a broad sufficient condition for two-sided Gaussian bounds on the heat kernel of  $-\nabla \cdot a \cdot \nabla + b \cdot \nabla$ , proved using a variant of Nash's method.

Put

$$\mathbf{MF} := \{K \mid K \in \mathbf{MF}_\kappa \text{ for some } \kappa < \infty\}.$$

**Theorem 2.** *Assume that the interaction kernels  $K_{ij}$  in particle system (2.1) satisfy*

$$K_{ij} \in \mathbf{MF}, \quad \begin{cases} (\operatorname{div} K_{ij})_- \in L^1 + L^\infty, \\ (\operatorname{div} K_{ij})_+^{\frac{1}{2}} \in \mathbf{F}_{\kappa_+} \text{ with } \kappa_+ < 4\frac{N}{N-1} \end{cases} \quad |K_{ij}|^{\frac{1+\alpha}{2}} \in \mathbf{F} \quad (2.11)$$

for some  $\alpha > 0$  fixed arbitrarily close to zero. Then all assertions of Theorem 1 are valid for such interaction kernels as well.

**Theorem 3.** (i) *For the interaction kernels*

$$K_{ij}(y) = \sqrt{\kappa} \frac{d-2}{2} |y|^{-2} y + K_{0,ij}(y), \quad y \in \mathbb{R}^d, \quad (2.12)$$

if the strength of attraction

$$\kappa < 16$$

and  $K_{0,ij}$  satisfy (2.6) or (2.11) with sufficiently small form-bounds, then all assertions of Theorems 1 and 2 remain valid.

- (ii) *Furthermore, for the model attracting interaction kernel  $K(y) = \sqrt{\kappa} \frac{d-2}{2} |y|^{-2} y$ ,  $\kappa < 16$ , the heat kernel  $e^{-t\Lambda}(x, z)$  of particle system (2.1) (which exists by Theorem 4(i)) satisfies, up to modification on a measure zero set, the heat kernel bound*

$$e^{-t\Lambda}(x, z) \leq C t^{-\frac{Nd}{2}} \prod_{1 \leq i < j \leq N} \eta(t^{-\frac{1}{2}} |z_i - z_j|), \quad t \in ]0, T],$$

for some  $C = C_T$ , for all  $x \in \mathbb{R}^{Nd}$ ,  $z = (z_1, \dots, z_N) \in \mathbb{R}^{Nd}$  provided  $z_i \neq z_j$  ( $i \neq j$ ), for a fixed function  $1 \leq \eta \in C^2(]0, \infty[)$  such that

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



REMARK 1 (On Theorems 1, 2 and their proofs). It is not difficult to modify the proofs of Theorems 1 and 2 to extend them to the sums of the interaction kernels satisfying (2.6) and (2.11), under properly adjusted assumptions on the form-bounds.

We prove Theorems 1 and 2 by embedding the multi-particle system (2.1) in the general SDE (2.30) considered in  $\mathbb{R}^{Nd}$ , with drift  $b = (b_1, \dots, b_N) : \mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$  defined by

$$b_i(x) := \frac{1}{N} \sum_{j=1, j \neq i}^N K_{ij}(x_i - x_j), \quad x = (x_1, \dots, x_N) \in \mathbb{R}^{Nd}, \quad 1 \leq i \leq N. \quad (2.13)$$

**Lemma 1.** *If  $K_{ij} \in \mathbf{F}_\kappa(\mathbb{R}^d)$ , then  $b$  defined by (2.13) satisfies*

$$\begin{cases} b \in \mathbf{F}_\delta(\mathbb{R}^{Nd}) \\ \text{with } \delta = \frac{(N-1)^2}{N^2} \kappa, \quad c_\delta = \frac{(N-1)^2}{N} c_\kappa. \end{cases}$$

**Lemma 2.** *If  $K_{ij} \in \mathbf{MF}_\kappa(\mathbb{R}^d)$ ,  $(\operatorname{div} K_{ij})_+^{\frac{1}{2}} \in \mathbf{F}_{\kappa_+}(\mathbb{R}^d)$ ,  $|K_{ij}|^{\frac{1+\alpha}{2}} \in \mathbf{F}_\sigma(\mathbb{R}^d)$ , then  $b$  defined by (2.13) satisfies*

$$\begin{cases} b \in \mathbf{MF}_\delta(\mathbb{R}^{Nd}) \\ \text{with } \delta = \frac{N-1}{\sqrt{N}} \kappa, \quad c_\delta = (N-1)c_\kappa, \end{cases} \quad (2.14)$$

$$\begin{cases} (\operatorname{div} b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+}(\mathbb{R}^{Nd}), \\ \text{with } \delta_+ = \frac{N-1}{N} \kappa_+, \quad c_{\delta_+} = (N-1)c_{\kappa_+}, \end{cases} \quad (2.15)$$

$$\begin{cases} |b|^{\frac{1+\alpha}{2}} \in \mathbf{F}_\chi(\mathbb{R}^{Nd}), \\ \text{with } \chi = \frac{(N-1)^{1+\alpha}}{N^{1+\alpha}} \sigma 2^\alpha, \quad c_\chi = \frac{(N-1)^{1+\alpha}}{N^\alpha} c_\sigma 2^\alpha. \end{cases} \quad (2.16)$$

Lemmas 1, 2 allow us to obtain the existence of a strong Markov family of martingale solutions to (2.1) in Theorem 1(i) and Theorem 2(i) from Theorem 5(i) for SDE (2.30). Theorem 5, and other results in Section 2.2 dealing with general singular drifts, are of interest on their own.

In Theorem 5 the family of martingale solutions for (2.30) is constructed by applying a tightness argument where the central role belongs to the estimate

$$\mathbf{E} \int_{t_0}^{t_1} |b_\varepsilon(Y_\varepsilon(s))| ds \leq C(t_1 - t_0)^{\frac{\gamma}{1+\gamma}}, \quad t_0, t_1 \in [0, T] \quad (2.17)$$

(this is (8.5)), where  $b_\varepsilon$  is a regularization of  $b$  that does not increase form-bounds  $\delta$ ,  $\delta_+$  (see Definition 5) in Lemmas 1, 2, and  $Y_\varepsilon$  is the strong solution of (2.30) with drift  $b_\varepsilon$ . Constants  $C$ ,  $\gamma > 0$  are independent of  $\varepsilon$ .

To prove (2.17) and, furthermore, to prove the strong Markov property, we establish regularity results for non-homogeneous elliptic PDEs (2.38) and (2.43). These are Theorems 8 and 9, obtained via De Giorgi's method ran in  $L^p$ , where  $p$  depends on the values of form-bounds  $\delta$  and  $\delta_+$ . Theorems 8 and 9 are the analytic core of the present paper.

We prove Theorem 8 by showing that  $u$  belongs to  $L^p$  De Giorgi's classes and then following the arguments in [15, Ch. 7], that is, applying De Giorgi's method. As was mentioned in the introduction, our  $L^p$  De Giorgi classes are somewhat different from the  $L^p$  De Giorgi classes found in the literature (cf. [15]).

The observation that working in  $L^p$  for large  $p$  allows to relax the assumptions on the form-bounds goes back to [36] where the authors dealt with Kolmogorov operator

$$-\Delta + b \cdot \nabla, \quad b \in \mathbf{F}_\delta, \quad \delta < 4$$

at the level of the semigroup theory.

In [28], the authors proved, using De Giorgi's iterations in  $L^p$ , that the general SDE (2.30) with  $b \in \mathbf{F}_\delta$ ,  $\delta < 4$  has a martingale solution for every initial point. This result yields the existence of a martingale solution part of Theorem 5 under condition  $(\mathbb{A}_1)$  on  $b$ , which we included in Theorem 5 for the sake of completeness. In what concerns  $(\mathbb{A}_1)$ , in the present paper we make the next step and also prove the strong Markov property.

One of the main observations of the present paper is related to condition  $(\mathbb{A}_2)$  of Theorem 5. This condition dictates the assumption (2.11) on the interaction kernel  $K$  when the latter has repulsion-attraction structure. In  $(\mathbb{A}_2)$ , we relax the a priori condition  $|b| \in L^2_{\text{loc}}$  as in  $(\mathbb{A}_1)$  to  $|b| \in L^{1+\alpha}_{\text{loc}}$  for  $\alpha > 0$  fixed arbitrarily small, aiming at stronger hypersurface singularities of  $b$  (and thus of  $K$ ). To achieve this, we once again need to work in  $L^p$  for  $p$  large. In fact, when dealing with the right-hand side of non-homogeneous equation

$$(\mu - \Delta + b \cdot \nabla)u = |b|f \quad (f \in C_c^\infty),$$

as is needed to prove weak well-posedness of the general SDE (2.30), we require

$$p' = \frac{p}{p-1} \geq 1 + \alpha$$

(cf. Theorem 9). If we were to consider this non-homogeneous equation in  $L^2$ , we would have to take  $\alpha = 1$ , and so  $(\mathbb{A}_2)$  and (2.11) would force the old form-boundedness assumption on drift  $b$ , i.e. as in  $(\mathbb{A}_1)$ . Thus, in the present paper we find *another compelling reason to work in  $L^p$  with  $p$  large, not related to the values of form-bounds*.

Another technical novelty of the paper is Theorem 9, i.e. the embedding theorem, which has applications beyond this paper.

The proof of Caccioppoli's inequality (Proposition 1) under assumption (2.11) uses an iteration procedure introduced in [34]. In [34], the authors worked in  $L^2$  and used Moser's method to prove the Harnack inequality for positive solutions of  $(-\nabla \cdot a \cdot \nabla + b \cdot \nabla)u = 0$  with measurable uniformly elliptic matrix  $a$  and  $b \in \mathbf{MF}_\delta$ ,  $\delta < \infty$ , provided that the form-bounds of the positive and the negative parts of  $\text{div } b$  satisfy some sub-critical assumptions.

The additional right-most condition on  $K_{ij}$  in (2.11) is essentially much weaker than the left-most condition (informally, the former treats  $|K|$  as a potential, while a proper "potential analogue" of the drift perturbation  $K \cdot \nabla$  would be  $|K|^2$ ). For instance, if we were to state condition (2.11) on the scale of  $L^p$  spaces, then it would become

$$|K| \in L^d + L^\infty, \quad \begin{cases} (\text{div } K)_- \in L^1 + L^\infty, \\ (\text{div } K)_+ \in L^{\frac{d}{2}} + L^\infty \end{cases} \quad |K| \in L^{\frac{d}{2}(1+\alpha)} + L^\infty,$$

where, recall,  $\alpha > 0$  is fixed arbitrarily small, i.e. the right-most condition follows from the left-most one. The same would happen if we were working on the scale of scaling-invariant Morrey spaces (cf. Example 1(d)).

If in Theorem 2  $(\text{div } K_{ij})_+ = 0$ , i.e. there is no attraction, then we essentially impose only a condition on  $|K_{ij}|$ , which can have any finite multiplicative form-bound. In fact, in a purely repulsing situation considered in [3] one can relax the assumptions on the interaction kernels  $K_{ij}$  of gradient form even further.

**REMARK 2** (On Theorem 3 and its proof). The improvement of the assumptions on  $\kappa$  in Theorem 3(i), compared to Theorems 1 and 2, is due to a refinement of Lemma 2 by means of the multi-particle Hardy inequality of [20]: for  $d \geq 3$ , all  $N \geq 2$ ,

$$C_{d,N} \sum_{1 \leq i < j \leq N} \int_{\mathbb{R}^{Nd}} \frac{|\varphi(x)|^2}{|x_i - x_j|^2} dx \leq \int_{\mathbb{R}^{Nd}} |\nabla \varphi(x)|^2 dx, \quad x = (x_1, \dots, x_N), \quad (2.18)$$

for all  $\varphi \in W^{1,2}(\mathbb{R}^{Nd})$ , where

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

In the proof of Theorem 3(ii) we replace constant  $C_{d,N}$  with smaller constant  $\frac{(d-2)^2}{N}$ . However, the maximum for large  $N$  and  $d \leq 6$  is actually attained on the second argument. So, constant 16 in Theorem 3 can be somewhat improved for  $d \leq 6$ . See also [16] regarding further improvements of  $C_{d,N}$ .

The heat kernel bound in Theorem 3(ii) is not unexpected (although we could not find it in the literature). Indeed, an elementary calculation shows that

$$\psi(x) := \prod_{1 \leq i < j \leq N} |x_i - x_j|^{-\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N}}.$$

is a Lyapunov function of the formal adjoint of  $\Lambda = -\Delta_x - \sqrt{\kappa} \frac{d-2}{2} \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{x_i - x_j}{|x_i - x_j|^2} \cdot \nabla_{x_i}$ , i.e. the following identity holds:

$$-\Delta_x \psi + \sqrt{\kappa} \frac{d-2}{2} \frac{1}{N} \sum_{i=1}^N \nabla_{x_i} \left( \sum_{j=1, j \neq i}^N \frac{x_i - x_j}{|x_i - x_j|^2} \psi \right) = 0.$$

One can expect that such Lyapunov function will appear as a multiple in the heat kernel bounds. That said, the question of how to prove such an estimate is non-trivial due to singularities in the drift. An interesting aspect of Theorem 3(ii) is its proof, which uses an abstract desingularization result from [33], see Appendix A.

In Theorem 3(ii), we expect to have two-sided bound

$$C_1 e^{-\frac{|x-y|^2}{c_2 t}} \varphi_t(y) \leq e^{-t\Lambda}(x, y) \leq C_3 e^{-\frac{|x-y|^2}{c_4 t}} \varphi_t(y), \quad (2.19)$$

where

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

as is suggested by the analogous results for Kolmogorov operator  $-\Delta - \sqrt{\kappa}|x|^{-2}x \cdot \nabla$ ,  $0 < \kappa < 4$  on  $\mathbb{R}^d$ , see [41]. Moreover, there should be an analogous to Theorem 3(ii) and (2.19) result in the case of attracting interactions, see [41] and [32] regarding  $-\Delta + \sqrt{\kappa}|x|^{-2}x \cdot \nabla$ ,  $0 < \kappa < \infty$ . ([32, 33] deal with the fractional Laplacian  $(-\Delta)^{\alpha/2}$  perturbed by the model singular drift term  $c|x|^{-\alpha}x \cdot \nabla$ ,  $1 < \alpha < 2$ .)

In the next theorem we show that if one is willing to replace the multiplicative form-boundedness condition (2.11) by more restrictive condition (2.21), or even restrict admissible values of the strength of the interactions, then more can be said about the martingale solutions  $\{\mathbb{P}_x\}_{x \in \mathbb{R}^{Nd}}$  constructed in Theorems 1 and 2.

We will need the following definition. Let  $K$  satisfy (2.2), let  $\{K_\varepsilon\}$  be some sequence of vector fields.

DEFINITION 5. Let us say that  $K_\varepsilon$  do not increase the form-bounds of  $K$  if

$$\|K_\varepsilon \varphi\|_2^2 \leq \kappa \|\nabla \varphi\|_2^2 + c_\kappa \|\varphi\|_2^2 \quad \forall \varphi \in W^{1,2}(\mathbb{R}^d), \quad \forall \varepsilon > 0,$$

i.e.  $\{K_\varepsilon\}$  satisfy (2.2) with the same constants as  $K$ .

The previous definition extends naturally to  $K$  satisfying (2.9), (2.10) or (2.21) below. In all these cases, in Section 3 we show that the vector fields  $K_\varepsilon$  defined by

$$K_\varepsilon := E_\varepsilon K, \quad \varepsilon \downarrow 0, \quad E_\varepsilon \text{ is the Friedrichs mollifier,} \quad (2.20)$$

are bounded, smooth and do not increase the form-bounds of  $K$ .

**Theorem 4.** *Assume that the interaction kernels  $K_{ij}$  in particle system (2.1) satisfy either condition (2.6) in Theorem 1, i.e.*

$$K_{ij} \in \mathbf{F}_\kappa \quad \text{with } \kappa < 4 \left( \frac{N}{N-1} \right)^2,$$

or a more restrictive condition than (2.11) in Theorem 2:

$$K_{ij} \in \mathbf{F}, \quad \begin{cases} (\operatorname{div} K_{ij})_- \in L^1 + L^\infty, \\ (\operatorname{div} K_{ij})_+^{\frac{1}{2}} \in \mathbf{F}_{\kappa_+} \text{ with } \kappa_+ < 4 \frac{N}{N-1}. \end{cases} \quad (2.21)$$

If  $K_{ij}$  satisfy (2.6), fix  $p > \frac{2}{2 - \frac{N-1}{N}\sqrt{\kappa}}$ . If  $K_{ij}$  satisfy (2.21), fix  $p > \frac{4}{4 - \frac{N-1}{N}\kappa_+}$ . Let  $\{\mathbb{P}_x\}_{x \in \mathbb{R}^{Nd}}$  be the strong Markov family of martingale solutions of particle system (2.1) constructed in Theorems 1 and 2.

(i) The family of operators  $\{P_t\}_{t \geq 0}$  defined by

$$P_t f(x) := \mathbb{E}_{\mathbb{P}_x}[f(\omega_t^1, \dots, \omega_t^N)], \quad f \in C_c^\infty(\mathbb{R}^{Nd}),$$

admits extension by continuity to a strongly continuous quasi contraction Markov semigroup on  $L^p$  of integral operators, say  $P_t =: e^{-t\Lambda_p}$ , such that

$$\|e^{-t\Lambda_p}\|_{p \rightarrow q} \leq c\omega^{\omega t} t^{-\frac{Nd}{2}(\frac{1}{p} - \frac{1}{q})}, \quad p \leq q \leq \infty \quad (2.22)$$

for appropriate constants  $c$  and  $\omega$ . If  $p = 2$ , then we have

$$\Lambda_2 \supset -\Delta + \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N K_{ij}(x_i - x_j) \cdot \nabla_{x_i} \upharpoonright C_c^\infty(\mathbb{R}^{Nd}).$$

This semigroup is unique in the following sense: for any sequence of bounded smooth interaction kernels  $K_{ij}^n \rightarrow K_{ij}$  in  $[L_{\text{loc}}^2(\mathbb{R}^d)]^d$  that do not increase the form-bounds of  $K$ , for every  $f \in C_c^\infty(\mathbb{R}^{Nd})$  solutions  $\{v_n\}$  to

$$(\partial_t - \Delta + \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N K_{ij}^n(x_i - x_j) \cdot \nabla_{x_i}) v_n = 0, \quad v_n(0) = f$$

converge to the same limit  $e^{-t\Lambda_p} f$  in  $L^p(\mathbb{R}^{Nd})$  loc. uniformly in  $t \geq 0$ .

(ii) If

$$K \in \mathbf{F}_\kappa \quad \text{with } \kappa < \frac{1}{(N-1)^2 d^2},$$

then, for every  $x \in \mathbb{R}^{Nd}$ , martingale solution  $\mathbb{P}_x$  satisfies for a given  $q \in ]Nd, \frac{N}{N-1}\kappa^{-\frac{1}{2}}[$  Krylov-type bounds

$$\begin{cases} \mathbb{E}_x \int_0^T |h(s, X_1(s), \dots, X_N(s))| ds \leq c \|h\|_{L^q([0, T] \times \mathbb{R}^{Nd})} & \text{and} \\ \mathbb{E}_x \int_0^T |b(X_1(s), \dots, X_N(s))| |h(\tau, X_1(s), \dots, X_N(s))| ds \leq c \|b\|_{L^{\frac{q}{2}}([0, T] \times \mathbb{R}^{Nd})}^{\frac{2}{q}} \end{cases} \quad (2.23)$$

for all  $h \in C_c([0, T] \times \mathbb{R}^{Nd})$ , for some constant  $c > 0$ , where vector field  $b : \mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$  is defined by (2.13).

Moreover,  $\mathbb{P}_x$  is the only martingale solution to (2.1) that satisfies (2.23).

(iii) Assertions (i), (ii) are also valid for interaction kernels of the form

$$K_{ij}(x_i, x_i - x_j) = \zeta(x_i) K_{ij}^0(x_i - x_j), \quad (2.24)$$

where

$$\|\zeta\|_\infty \leq 1, \quad K_{ij}^0 \in \mathbf{F}_\kappa \text{ with } \kappa < 4 \left( \frac{N}{N-1} \right)^2.$$

(iv) There exists constant  $C < 1$  such that if  $K_{ij}$  is of the form (2.24) with  $\zeta$  additionally having compact support and

$$\kappa < \frac{C}{(N-1)^2 d^2}, \quad (2.25)$$

then for every  $(x_1, \dots, x_N) \in \mathbb{R}^{Nd}$  particle system (2.1) has a strong solution on  $[0, T]$  that is unique among all strong solutions defined on the same probability space satisfying (2.23).

The proof of Theorem 4 uses results from [25, 26, 29, 36, 45]. Theorem 4 also extends to the interaction kernels considered in Theorem 3 with (2.11) replaced with (2.21).

REMARK 3 (Borderline strengths of interactions). One can reach the borderline values of the strengths of interactions

$$\begin{aligned} \kappa &= 4 \left( \frac{N}{N-1} \right)^2 && \text{if (2.6) holds, or} \\ \kappa_+ &= 4 \frac{N}{N-1} && \text{if (2.21) holds} \end{aligned}$$

although, at the moment, only on the torus and at the PDE level, by considering the corresponding to (2.1) Kolmogorov backward equation in the Orlicz space with gauge function  $\Phi = \cosh - 1$ . This space is situated between all  $L^p$  and  $L^\infty$ . We refer to [22] for details.

REMARK 4 (On Theorem 4 and its proof). Assertion (iv) follows, after applying Lemma 1, from the result in [25] whose proof, in turn, follows closely the method of Röckner-Zhao [44]. In [25], the authors needed a technical hypothesis that  $b$  has compact support, hence the condition in (iv) on the support of  $\zeta$ .

We are rather satisfied with Theorems 1-3 and assertion (i) of Theorem 4 where the assumptions on  $\kappa$  and  $\kappa_+$  stabilize to positive values as  $N \rightarrow \infty$ . In assertions (ii) and (iv) of Theorem 4, however, the admissible strength of interactions  $\kappa$  degenerates as the number of particles  $N$  goes to infinity, which is a by-product of our method of embedding particle system (2.1) in the general SDE (2.30). Another drawback of assertions (ii) and (iv) is the difficulty with taking into account the repulsion/attraction structure of the interaction kernel  $K$  simply by looking at the divergence of  $K$ . Regarding the conditional weak uniqueness as in (ii), in Theorem 7 we consider the general SDE (2.30) and propose another condition on the drift  $b$  that provides conditional weak uniqueness for (2.30) while taking into account the repulsion/attraction. We show in Example 3 that there is some truth to this condition: it is always satisfied in dimensions  $d \geq 4$  for the model repulsing drift  $b(x) = -\sqrt{\delta} \frac{d-2}{2} |x|^{-2} x$ , regardless of the value for the form-bound  $\delta > 0$ , as one would expect. This requires us to obtain gradient bounds in  $L^q$  starting with  $q > d - 2$ , hence the need to work in the elliptic setting. (In the parabolic setting

we would need  $q > d$ .) That said, this result, when applied via Lemma 2 to drift (2.13) with  $K_{ij}(y) = -\sqrt{\kappa} \frac{d-2}{2} |y|^{-2} y$ , leads to a condition on  $\kappa$  that still depends on the number of particles  $N$ . So, there is still work to be done to find a proper analogue of Theorem 7 for particle system (2.1).

REMARK 5 (Drifts). One can easily extend Theorems 1-4 to more general particle system

$$dX_i = M_i(X_i)dt - \frac{1}{N} \sum_{j=1, j \neq i}^N K_{ij}(X_i - X_j)dt + \sqrt{2}dB_i,$$

where, for every  $1 \leq i \leq N$ ,

$$M_i \in \mathbf{F}_\mu.$$

Let us discuss for simplicity the case when  $K_{ij}$  satisfy (2.6). We require that  $\mu, \kappa$  satisfy

$$\left(\sqrt{\mu} + \frac{N-1}{N}\sqrt{\kappa}\right)^2 < 4.$$

We only need to embed this particle system into (2.30), i.e. prove an analogue of Lemma 1 for vector field  $b = b^M + b^K$  with  $b^M, b^K : \mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$  having components

$$b_i^M(x) := M_i(x_i), \quad b_i^K(x) := \frac{1}{N} \sum_{j=1, j \neq i}^N K_{ij}(x_i - x_j), \quad 1 \leq i \leq N, \quad (2.26)$$

and then e.g. use Theorem 5 for the general SDE (2.30) as we do in the proof of Theorem 2. Repeating the proof of Lemma 1, we obtain right away that

$$b^K \in \mathbf{F}_{\delta^K}(\mathbb{R}^{Nd}) \quad \text{with} \quad \delta^K = \frac{(N-1)^2}{N^2}\kappa, \quad c_{\delta^K} = \frac{(N-1)^2}{N}c_\kappa$$

and

$$b^M \in \mathbf{F}_{\delta^M}(\mathbb{R}^{Nd}) \quad \text{with} \quad \delta^M = \mu, \quad c_{\delta^M} = Nc_\mu,$$

see Remark 10 in Section 4 for the proof. It remains to note that the sum of two form-bounded vector fields is form-bounded, i.e.  $b = b^M + b^K$  is in  $\mathbf{F}_\delta$  with  $\sqrt{\delta} = \sqrt{\delta^M} + \sqrt{\delta^K}$ , and  $\delta$  must be strictly less than 4, cf. Theorem 5.

Arguing similarly, one can treat general drifts  $M_i(X_1, \dots, X_N)$  ( $1 \leq i \leq N$ ) in the particle system (adjusting the hypothesis on the form-bound, i.e. now  $\delta^M = N\mu$ ).

REMARK 6 (Sufficient condition for multiplicative form-boundedness). A Borel measurable vector field  $K : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is said to belong to the class of weakly form-bounded vector fields  $\mathbf{F}_\kappa^{1/2}$  if  $|K| \in L_{\text{loc}}^1$  and

$$\| |K|^{\frac{1}{2}}(\lambda - \Delta)^{-\frac{1}{4}} \|_{2 \rightarrow 2} \leq \sqrt{\kappa} \quad (L^2 \rightarrow L^2 \text{ operator norm}) \quad (2.27)$$

for some  $\lambda > 0$ . We have

$$\mathbf{F}_\kappa^{1/2} \subset \mathbf{MF}_\kappa. \quad (2.28)$$

Indeed, if  $K \in \mathbf{F}_\kappa^{1/2}$ , then, arguing as in [45], we have

$$\begin{aligned} \langle |K|\varphi, \varphi \rangle &\leq \kappa \langle (\lambda - \Delta)^{\frac{1}{2}}\varphi, \varphi \rangle \leq \kappa \|(\lambda - \Delta)^{\frac{1}{2}}\varphi\|_2 \|\varphi\|_2 \\ &= \kappa \sqrt{\|\nabla\varphi\|_2^2 + \lambda\|\varphi\|_2^2} \|\varphi\|_2 \leq \kappa \|\nabla\varphi\|_2 \|\varphi\|_2 + \kappa\sqrt{\lambda}\|\varphi\|_2^2, \end{aligned}$$

i.e.  $K \in \mathbf{MF}_\kappa$ .

The class  $\mathbf{F}_\kappa^{1/2}$  (and therefore  $\mathbf{MF}_\kappa$ ) contains the largest possible up to the strict inequality in  $\varepsilon > 0$  scaling-invariant Morrey class  $M_{1+\varepsilon}$ , i.e. if

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

then  $M_{1+\varepsilon} \subset \mathbf{F}_\kappa^{1/2}$  with  $\kappa = c(d, \varepsilon) \|K\|_{M_{1+\varepsilon}} [1]$ .

It is easily seen that  $M_{1+\varepsilon}$  is larger than  $M_2$ , which, in turn, contains  $\mathbf{F}_\kappa$ . That said, we also need to control the form-bounds. In fact, we have

$$\mathbf{F}_\kappa \subset \mathbf{F}_{\sqrt{\kappa}}^{1/2}. \quad (2.29)$$

Indeed, rewriting  $K \in \mathbf{F}_\kappa$  as

$$\| |K|(\lambda - \Delta)^{-\frac{1}{2}} \|_{2 \rightarrow 2} \leq \sqrt{\kappa}$$

(with  $\lambda = c_\kappa/\kappa$ ), we obtain the required result by applying the Heinz inequality. In (2.29) we have a proper inclusion because the class of weakly form-bounded vector fields also contains the Kato class of vector fields  $\| |K|(\lambda - \Delta)^{-\frac{1}{2}} \|_\infty \leq \sqrt{\kappa}$  while  $\mathbf{F}_\kappa$  does not (see [29]).

**REMARK 7** (Stronger hypersurface singularities). We refer to [24] and [27] for a weak well-posedness theory of the general SDE (2.30) with  $b \in \mathbf{F}_\delta^{1/2}$  or with  $b$  in the time-inhomogeneous analogue of the Morrey class  $M_{1+\varepsilon}$  which allows to include some critical singularities of  $b$  in time. This allows to treat

$$b(x) = \pm \frac{cx}{\|x-1\|^{1-\gamma}} \eta(x),$$

for a fixed  $0 < \gamma < 1$ ,  $c \in \mathbb{R}$  and  $0 \leq \eta \in C_c^\infty$ , i.e. hypersurface singularities that are essentially twice more singular than (1.3). That said, in this result the assumption on the form-bound  $\delta$  is also dimension-dependent, so if we were to apply this result to (2.1), we would arrive at the assumption on the strength of interaction  $\kappa$  of the form  $\kappa < \frac{C}{(Nd)^2}$ .

**REMARK 8** (On the uniqueness of weak solution to elliptic Kolmogorov PDE). 1. Our most complete uniqueness result for the Kolmogorov elliptic equation in (2.8) with interaction kernels satisfying (2.6) or (2.21) is in fact proved in [22] on the torus, see Remark 3. Speaking of  $\mathbb{R}^d$ , let us first say a few words about the case of very sub-critical strengths of interactions.

**DEFINITION 6.** If  $K$  satisfies (2.6) with  $\kappa < (\frac{N}{N-1})^2$ , we say that  $u$  is a weak solution of (2.8) if  $u \in W^{1,2} \cap L^\infty$  and

$$\mu \langle u, \varphi \rangle + \langle \nabla u, \nabla \varphi \rangle + \frac{1}{N} \left\langle \sum_{i=1}^N \sum_{j=1, j \neq i}^N K_{ij}(x_i - x_j) \cdot \nabla_{x_i} u, \varphi \right\rangle = \langle f, \varphi \rangle$$

for every  $\varphi \in W^{1,2}$ . (Recall that in (2.8) the initial function is bounded, so the solution is bounded as well.)

**DEFINITION 7.** If  $K$  satisfies (2.21) with  $\kappa_+ < 2\frac{N}{N-1}$ , then  $u$  is a weak solution to (2.8) if  $u \in W^{1,2} \cap L^\infty$  and

$$\begin{aligned} & \mu \langle u, \varphi \rangle + \langle \nabla u, \nabla \varphi \rangle \\ & - \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \left[ \langle \operatorname{div} K_{ij}(x_i - x_j) u, \varphi \rangle + \langle K_{ij}(x_i - x_j) u, \nabla_{x_i} \varphi \rangle \right] = \langle f, \varphi \rangle \end{aligned}$$

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

In both cases the uniqueness of the weak solution follows upon applying Lemmas 1, 2 and the Lax-Milgram theorem in  $L^2$ , i.e. we can take  $p = 2$  in Theorem 4.

In the general case, we need to consider (2.8) in  $L^p$ , where  $p$  is as in Theorem 4. In this regard, we refer to [45] for the definition of weak solution and results on weak solutions of parabolic equations in  $L^p$ .

2. If  $K$  satisfies (2.11), then we can prove that  $u$  constructed in Theorem 2(i) is a weak solution of (2.8) e.g. in the following sense.

DEFINITION 8. If  $K$  satisfies (2.11), then we say that  $u$  is a weak solution of (2.8) if  $u \in W_{\text{loc}}^{1,2} \cap L^\infty$  and

$$\begin{aligned} & \mu \langle u, \varphi \rangle + \langle \nabla u, \nabla \varphi \rangle \\ & - \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \left[ \langle \text{div } K_{ij}(x_i - x_j)u, \varphi \rangle + \langle K_{ij}(x_i - x_j)u, \nabla_{x_i} \varphi \rangle \right] = \langle f, \varphi \rangle \end{aligned}$$

for all  $\varphi \in W_{\text{loc}}^{1,2} \cap L_c^\infty$  ( $L_c^\infty$  are bounded functions with compact supports).

The latter is a way to establish a link between function  $u$  defined by (2.7) and the formal elliptic equation (2.8). However, the proof of uniqueness of such a weak solution under condition (2.11) remains elusive. Still, we can prove that  $u$  given by (2.7) is unique among weak solutions that can be obtained via a reasonable regularization of  $K$ , cf. Theorem 6. Alternatively, we can restrict our attention to the subclass of weakly form-bounded vector fields, see Remark 6, and prove uniqueness via the Lax-Milgram theorem in the triple of Bessel potential spaces  $\mathcal{W}^{\frac{1}{2},2} \subset \mathcal{W}^{-\frac{1}{2},2} \subset \mathcal{W}^{-\frac{3}{2},2}$  (rather than the standard  $W^{1,2} \subset L^2 \subset W^{-1,2}$ ), see [31], although this comes at the cost of requiring that the weak form-bound of  $K$  (and therefore its multiplicative form-bound  $\kappa_0$ , cf. (2.28)) must be strictly less than 1.

**2.2. SDEs with general singular drifts.** Theorems 1, 2 and 3(i) are proved by embedding the particle system in a general SDE with singular drift, which we consider here, to lighten the notations, in  $\mathbb{R}^d$  instead of  $\mathbb{R}^{Nd}$ :

$$Y(t) = y - \int_0^t b(Y(s))ds + \sqrt{2}B(t), \quad t \in [0, T], \quad y \in \mathbb{R}^d, \quad (2.30)$$

where a priori  $b \in [L_{\text{loc}}^1]^d$ ,  $\{B(t)\}_{t \geq 0}$  is a Brownian motion in  $\mathbb{R}^d$ , and then applying Theorem 5 below.

Set

$$b_n := E_{\varepsilon_n} b, \quad \varepsilon_n \downarrow 0, \quad E_\varepsilon \text{ is the Friedrichs mollifier,} \quad \varepsilon_n \downarrow 0. \quad (2.31)$$

**Theorem 5.** *Assume that a Borel measurable vector field  $b$  in SDE (2.30) satisfies one of the following two conditions:*

$$b \in \mathbf{F}_\delta \quad \text{with } \delta < 4 \quad (\mathbb{A}_1)$$

or

$$b \in \mathbf{MF}, \quad \begin{cases} (\text{div } b)_- \in L^1 + L^\infty, \\ (\text{div } b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+} \text{ with } \delta_+ < 4, \end{cases} \quad |b|^{\frac{1+\alpha}{2}} \in \mathbf{F} \quad (\mathbb{A}_2)$$

for some  $\alpha > 0$  fixed arbitrarily small. Then the following are true:

(i) *There exists a strong Markov family  $\{\mathbb{P}_y\}_{y \in \mathbb{R}^d}$  of martingale solutions of SDE (2.30).*



(ii) The function

$$u(x) := \mathbb{E}_{\mathbb{P}_x} \int_0^\infty e^{-\lambda s} f(\omega_s) ds, \quad x \in \mathbb{R}^d, \quad f \in C_c^\infty(\mathbb{R}^d), \quad (2.32)$$

where  $\lambda$  is assumed to be sufficiently large, is a locally Hölder continuous weak solution to elliptic Kolmogorov equation  $(\lambda - \Delta + b \cdot \nabla)u = f$  (see Remark 8 for the definitions).

From now on, let us replace condition  $(\mathbb{A}_2)$  with somewhat more restrictive hypothesis

$$\begin{cases} b \in \mathbf{F}, \\ (\operatorname{div} b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+} \text{ with } \delta_+ < 4, \quad (\operatorname{div} b)_-^{\frac{1}{2}} \in L^1 + L^\infty. \end{cases} \quad (\mathbb{A}_3)$$

If  $b$  satisfies  $(\mathbb{A}_1)$ , fix  $p > \frac{2}{2-\sqrt{\delta}}$ . If  $b$  satisfies  $(\mathbb{A}_3)$ , fix  $p > \frac{4}{4-\delta_+}$ .

(iii) ([36, 45], see also [29]) The family of operators

$$P_t f(x) := \mathbb{E}_{\mathbb{P}_x}[f(\omega_t)], \quad t > 0, \quad f \in C_c^\infty$$

admits extension by continuity to a strongly continuous quasi contraction Markov semigroup on  $L^p$ , say,  $P_t =: e^{-t\Lambda_p}$ , such that

$$\|e^{-t\Lambda_p}\|_{p \rightarrow q} \leq c e^{\omega t} t^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}, \quad p \leq q \leq \infty$$

for appropriate constants  $c$  and  $\omega$ . The semigroup is unique in the sense that for any sequence of bounded smooth vector fields

$$b_n \rightarrow b \quad \text{in } [L_{\text{loc}}^2]^d$$

that do not increase the form-bounds on  $b$  in  $(\mathbb{A}_2)$  or  $(\mathbb{A}_3)$ , the classical solutions  $v_n$  to

$$(\partial_t - \Delta + b_n \cdot \nabla)v_n = 0, \quad v_n(0) = f \in C_c^\infty$$

converge to the same limit  $e^{-t\Lambda_p} f$  in  $L^p$  loc. uniformly in  $t \geq 0$ .

(iv) The resolvent  $(\mu + \Lambda_p(b))^{-1}$  has Feller property, i.e. for each  $\mu$  greater than some  $\mu_0 > 0$  it extends by continuity to a bounded linear operator on  $C_\infty$ :

$$R_\mu(b) := [(\mu + \Lambda_p(b))^{-1} \upharpoonright L^p \cap C_\infty]_{C_\infty \rightarrow C_\infty}^{\text{clos}} \in \mathcal{B}(C_\infty).$$

Moreover,

$$R_\mu(b_n) \rightarrow R_\mu(b) \quad \text{strongly in } C_\infty, \quad \mu \geq \mu_0,$$

where  $R_\mu(b_n)$  coincides with the resolvent of  $-\Delta + b_n \cdot \nabla$  on  $C_\infty$ ,  $n = 1, 2, \dots$

In (iii) we can consider  $b_n$  defined by (2.31).

The next result shows that, in principle, there is nothing pathological about condition  $(\mathbb{A}_2)$ . That said, its proof, compared to the proof of the approximation uniqueness in Theorem 5(iii) under condition  $(\mathbb{A}_3)$ , requires more detailed information about the solutions of the corresponding elliptic equation.

**Theorem 6** (On the approximation uniqueness). *Assume that a Borel measurable vector field  $b$  satisfies*

$$\begin{cases} b \in \mathbf{MF}, \\ (\operatorname{div} b)_- \in L^1 + L^\infty, \quad (\operatorname{div} b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+} \text{ with } \delta_+ < 2 \end{cases} \quad (2.33)$$

Then there exist constants  $\mu_0$  and  $0 < \varkappa < 1$  such that if, additionally,  $|b| \in L^{2-\varkappa}$ , then, for any sequence  $b_n$  of bounded smooth vector fields satisfying (2.33) with the same constants as  $b$  and

such that  $b_n \rightarrow b$  in  $L^{2-\alpha}$ , the sequence of the classical solutions  $u_n$  to  $(\mu - \Delta + b_n \cdot \nabla)u_n = f$ ,  $f \in C_c^\infty$ ,  $\mu \geq \mu_0$ , converge:

$$u_n \rightarrow u \quad \text{in } L^2.$$

Moreover, if  $\{\tilde{b}_n\}$  is some other sequence satisfying the same assumptions as  $\{b_n\}$ , then the corresponding solutions  $\tilde{u}_n$  converge in  $L^2$  to the same limit  $u$ .

Notice more restrictive than in Theorem 5 condition on  $\delta_+$  in (2.33). The proof of Theorem 6 follows the argument in [34] dealing with the approximation uniqueness of solution to the corresponding Dirichlet problem in a bounded domain. That is, it also uses Caccioppoli's inequality (Proposition 1) and Gehring's lemma.

Combining Theorem 6 with Theorem 8, we can show that the limit  $u$  is locally Hölder continuous. One can further show that  $u$  is a weak solution of  $(\mu - \Delta + b \cdot \nabla)u = f$  and construct the corresponding semigroup in  $L^p$ , but we will not pursue this here.

We need assertion (iv) of Theorem 5, obtained by means of Theorems 8 and 9, in the proof of the following uniqueness result.

**Theorem 7** (Krylov-type estimates and conditional uniqueness). *Assume that a Borel measurable vector field  $b$  satisfies one of the following conditions:*

$$b \in \mathbf{F}_\delta \text{ with } \delta < \left(\frac{2}{q}\right)^2 \wedge 1 \text{ for some } q > (d-2) \vee 2 \quad (\mathbb{B}_1)$$

or

$$\left\{ \begin{array}{l} b \in \mathbf{F}_\delta \cap [W_{\text{loc}}^{1,1}(\mathbb{R}^d)]^d \text{ for some finite } \delta, \text{ has symmetric Jacobian } Db := (\nabla_k b_i)_{k,i=1}^d, \\ \text{the normalized eigenvectors } e_j \text{ and eigenvalues } \lambda_j \geq 0 \text{ of the negative part of } Db - \frac{\text{div } b}{q} I \\ \text{for some } q > (d-2) \vee 2 \text{ satisfy } \sqrt{\lambda_j} e_j \in \mathbf{F}_{\nu_j} \text{ with } \nu := \sum_{j=1}^d \nu_j < \frac{4(q-1)}{q^2}. \end{array} \right. \quad (\mathbb{B}_2)$$

Then the following are true for the strong Markov family of martingale solutions of SDE (2.30) constructed in Theorem 5:

(i) For every  $y \in \mathbb{R}^d$ , martingale solution  $\mathbb{P}_y$  satisfies Krylov-type bound

$$\mathbb{E}_{\mathbb{P}_y} \int_0^\infty e^{-\lambda s} |\mathbf{g}f|(\omega_s) ds \leq C \|\mathbf{g}\|_2^{\frac{q}{2}} \|f\|_2^{\frac{2}{q}}, \quad \forall \mathbf{g} \in \mathbf{F}, \quad \forall f \in C_c, \quad (2.34)$$

for  $q > (d-2) \vee 2$  close to  $(d-2) \vee 2$ , for all  $\lambda$  sufficiently large.

(i')  $\{\mathbb{P}_y\}_{y \in \mathbb{R}^d}$  is the only Markov family of martingale solutions to (2.30) that satisfies Krylov-type bound in (i).

(ii) For every  $y \in \mathbb{R}^d$ ,  $\mathbb{P}_x$  satisfies Krylov bound:

$$\mathbb{E}_{\mathbb{P}_y} \int_0^\infty e^{-\lambda s} |f(\omega_s)| ds \leq C \|f\|_{\frac{qd}{d+q-2}}, \quad \forall f \in C_c \quad (2.35)$$

for all  $\lambda$  sufficiently large.

(ii') We make (2.35) more restrictive by selecting  $q$  close to  $(d-2) \vee 2$ , so that in (2.35)  $\frac{qd}{d+q-2} = \frac{d}{2-\varepsilon} \wedge \frac{2}{1-\varepsilon}$  for some  $\varepsilon > 0$  small. Let  $\{\mathbb{P}_y^2\}_{t \in \mathbb{R}^d}$  be another Markov family of martingale solutions for (2.30) that satisfies Krylov bound

$$\mathbb{E}_{\mathbb{P}_y^2} \int_0^\infty e^{-\lambda s} |f(\omega_s)| ds \leq C \|f\|_{\frac{d}{2-\varepsilon} \wedge \frac{2}{1-\varepsilon}}, \quad \forall f \in C_c \quad (2.36)$$

(one such family exists, it is  $\{\mathbb{P}_y\}_{y \in \mathbb{R}^d}$  from above). Assume additionally that, for some  $\varepsilon_1 \in ]\varepsilon, 1[$  we have

$$(1 + |x|^{-2})^{-\beta} |b|^{\frac{d}{2-\varepsilon_1} \vee \frac{2}{1-\varepsilon_1}} \in L^1$$

for some  $\beta > \frac{d}{2}$  fixed arbitrarily large, and either  $(\mathbb{B}_1)$  holds with  $\delta < \frac{4}{q_*} \wedge 1$ , where

$$q_* := \begin{cases} \frac{d-2}{\varepsilon_1 - \varepsilon} & \text{if } d \geq 4, \\ 2\left(\frac{1}{3(\varepsilon_1 - \varepsilon)} \vee 1\right) & \text{if } d = 3, \end{cases}$$

or  $(\mathbb{B}_2)$  holds with  $q = q_*$  and  $\nu < \frac{4(q_* - 1)}{q_*^2}$ . Then  $\{\mathbb{P}_y^2\}_{y \in \mathbb{R}^d}$  coincides with  $\{\mathbb{P}_y\}_{y \in \mathbb{R}^d}$  from above.

Some remarks are in order.

1. In the last assertion, the uniqueness class of martingale solutions satisfying Krylov bound (2.36), which depends on our choice of  $\varepsilon$ , determines the extra conditions that one needs to impose on  $b$ . Note that if in  $(\mathbb{B}_1)$  one has  $|b| \in L^d$ , or in  $(\mathbb{B}_2)$  the eigenvectors have entries in  $L^d$ , then the form-bounds  $\delta$  and  $\nu_j$  ( $j = 1, \dots, d$ ), respectively, can be chosen arbitrarily small, in which case these extra conditions on  $b$  are trivially satisfied.
2. In  $(\mathbb{B}_2)$  we require Jacobian  $Db$  to be symmetric, so  $b = \nabla V$  for some potential  $V$ .

Let us illustrate condition  $(\mathbb{B}_2)$  with the following example.

EXAMPLE 3. Let  $d \geq 4$ . Let

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

a drift that pushes solution  $Y_t$  of (2.30) away from the origin. Put for brevity  $c := \sqrt{\delta} \frac{d-2}{2} > 0$ . We have  $\operatorname{div} b = -c(d-2)|x|^{-2}$  and  $\nabla_j b_i = c[-|x|^{-2} \delta_{ij} + 2x_i x_j |x|^{-4}]$ . Therefore, for every  $\xi = (\xi_i) \in \mathbb{R}^d$ ,

$$\begin{aligned} \xi^\top \left( Db - \frac{\operatorname{div} b}{q} \right) \xi &= \sum_{i,j=1}^d \xi_j [(\nabla_j b_i) - \frac{1}{q} (\operatorname{div} b) I] \xi_i = c \left( \frac{d-2}{q} - 1 \right) |x|^{-2} |\xi|^2 + 2c |x|^{-4} (x \cdot \xi)^2 \\ &= B_+ - B_-, \end{aligned}$$

where  $B_+ \geq 0$  is the matrix with entries  $2cx_i x_j |x|^{-4}$ , and  $B_- := -c \left( \frac{d-2}{q} - 1 \right) |x|^{-2} I \geq 0$ . Thus, constant  $\nu$  in condition  $(\mathbb{B}_2)$  can be made as small as needed by selecting  $q > d-2$  sufficiently close to  $d-2$ , and so for this  $b$  condition  $(\mathbb{B}_2)$  can be satisfied for any strength of repulsion from the origin.

In the previous example it is crucial that we can select  $q$  as close to  $d-2$  as needed. By working in the parabolic setting we could obtain a stronger uniqueness results, i.e. for every fixed initial point  $x$ . However, the parabolic setting requires us to take  $q > d$  [26], and so the previous example becomes invalid: we have to require smallness of  $\delta$  even in the case of repulsion.

REMARK 9 (On some other classes of singular vector fields arising in the study of singular SDEs and PDEs). 1. A number of important results on the regularity theory of  $-\Delta + b \cdot \nabla$  was obtained in [49, 50] which considered supercritical form-boundedness type conditions on  $b$  (in the context of the study of 3D Navier-Stokes equations). These are conditions of the type: there exists  $\varepsilon \in ]0, 1[$  such that  $|b| \in L_{\text{loc}}^{1+\varepsilon}([0, \infty[ \times \mathbb{R}^d)$  and

$$\begin{aligned} \int_0^\infty \int_{\mathbb{R}^d} |b(t, \cdot)|^{1+\varepsilon} \xi^2(t, \cdot) dt &\leq \delta \int_0^\infty \|\nabla \xi(t, \cdot)\|_2^2 dt + \int_0^\infty g(t) \|\xi(t, \cdot)\|_2^2 dt \\ &\text{for all } \xi \in C_c^\infty([0, \infty[ \times \mathbb{R}^d) \end{aligned} \quad (2.37)$$

for some  $\delta > 0$  and  $0 \leq g \in L^1_{\text{loc}}([0, \infty[)$  under, necessarily, some assumptions on  $\text{div } b$  which cannot be too singular. Here supercriticality/criticality/subcriticality refer to how the assumptions on  $b$  behave under rescaling the equation. In the supercritical case one has to sacrifice a large portion of the regularity theory of  $-\Delta + b \cdot \nabla$  including the usual Harnack inequality and the Hölder continuity of solutions to the elliptic and parabolic equations. See also counterexample to the uniqueness in law for SDEs with supercritical drifts in [53]. However, some parts of the theory, such as the local boundedness of weak solutions, can be salvaged, see cited papers, see also recent developments in [4, 18]. Let us also note that if we were to specify (2.37) to the critical case when the usual regularity theory is still valid, then we would need to take  $\varepsilon = 1$ , i.e. we would obtain condition (2.21), but not more general condition (2.11).

2. As was noted in [28], after supplementing (2.37) with condition  $(\text{div } b)_+^{\frac{1}{2}} \in \mathbf{F}_\nu$  for some  $\nu < 4$ , one can still prove the existence of a martingale solution to SDE (2.30). In the present paper we work in the critical setting which allows to us preserve most of the important results in the regularity theory of elliptic equations that do not involve estimates on the second order derivatives of the solutions (which are destroyed by the form-boundedness assumptions), and thus allows to prove, e.g. the strong Markov property, approximation uniqueness or conditional weak uniqueness results for particle system (1.1) (see, however, [18] who constructed a Markov family of weak solutions in a supercritical setting using a selection procedure).

Let us also add that above supercriticality refers to the assumptions on  $b$ , but not on  $(\text{div } b)_+$ . In fact, as the counterexample to weak solvability of (1.5) with the model attracting drift shows, one cannot go beyond the form-boundedness assumption (critical) on  $(\text{div } b)_+$ .

**2.3. Embedding theorem and Hölder continuity of solutions.** To prove Theorem 5, we need the following regularity results for non-homogeneous elliptic equations (2.38), (2.43). In these results we assume additionally that the coefficients of (2.38), (2.43) are bounded and smooth. However, importantly, the constants in the regularity estimates are *generic*, i.e. they depend only on the structure parameters of the equation such as the dimension  $d$ , constant term  $\lambda$  and the form-bounds of the vector fields (but not on the smoothness or boundedness of the coefficients).

**Theorem 8** (Hölder continuity of solutions). *Let  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a bounded smooth vector field such that either  $(\mathbb{A}_1)$  (i.e.  $b \in \mathbf{F}_\delta$  with  $\delta < 4$ ) holds, or*

$$\begin{cases} b \in \mathbf{MF}, \\ (\text{div } b_+)^{\frac{1}{2}} \in \mathbf{F}_{\delta_+} \text{ with } \delta_+ < 4, \end{cases} \quad (\bar{\mathbb{A}}_2)$$

where  $\text{div } b = \text{div } b_+ - \text{div } b_-$  for some bounded smooth functions  $\text{div } b_\pm \geq 0$ . Then a classical solution  $u$  to non-homogeneous equation

$$(\lambda - \Delta + b \cdot \nabla)u = f \in C_c^\infty, \quad \lambda \geq 0, \quad (2.38)$$

is locally Hölder continuous with generic constants that also depend on  $\|f\|_\infty$ .

(The difference between  $(\bar{\mathbb{A}}_2)$  and  $(\mathbb{A}_2)$  is that in the former case we do not require  $\text{div } b_\pm$  to be positive and negative parts of  $\text{div } b$ , which are continuous but not necessarily smooth.)

The fact that the constants are generic is of course the main point of Theorem 8.

Define weight

$$\rho(y) = (1 + k|y|^2)^{-\frac{d}{2}-1}, \quad y \in \mathbb{R}^d, \quad (2.39)$$

where constant constant  $k > 0$  will be chosen sufficiently small. This weight has property

$$|\nabla \rho| \leq \left(\frac{d}{2} + 1\right) \sqrt{k} \rho. \quad (2.40)$$

For a fixed  $x \in \mathbb{R}^d$ , put  $\rho_x(y) := \rho(y - x)$ .

**Theorem 9** (Embedding theorem). *Let  $b, \mathbf{h} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be bounded smooth vector fields such that*

$$b \in \mathbf{F}_\delta \text{ with } \delta < 4, \quad \mathbf{h} \in \mathbf{F} \quad (2.41)$$

or

$$\begin{cases} b \in \mathbf{MF}, \\ (\operatorname{div} b_+)^{\frac{1}{2}} \in \mathbf{F}_{\delta_+} \text{ with } \delta_+ < 4, \end{cases} \quad |\mathbf{h}|^{\frac{1+\gamma}{2}} \in \mathbf{F} \text{ for some small } \gamma > 0, \quad (2.42)$$

where  $\operatorname{div} b = \operatorname{div} b_+ - \operatorname{div} b_-$  for some bounded smooth functions  $\operatorname{div} b_\pm \geq 0$ . In the former case, fix  $p > \frac{2}{2-\sqrt{\delta}}$ ,  $p \geq 2$ , and in the latter case fix  $p > \frac{4}{4-\delta_+}$ ,  $p' \leq 1 + \gamma$ ,  $p \geq 2$ .

Then, for a fixed  $1 < \theta < \frac{d}{d-2}$ , there exist generic constants  $\lambda_0, k$  (in  $\rho$ ),  $C$  and  $\beta \in ]0, 1[$  such that the classical solution  $u$  to non-homogeneous equation

$$(\lambda - \Delta + b \cdot \nabla)u = |\mathbf{h}f|, \quad f \in C_c^\infty \quad (2.43)$$

on  $\mathbb{R}^d$  satisfies

$$\begin{aligned} \|u\|_\infty \leq C \sup_{x \in \frac{1}{2}\mathbb{Z}^d} & \left( (\lambda - \lambda_0)^{-\frac{1}{p\theta}} \langle (\mathbf{1}_{|\mathbf{h}|>1} + |\mathbf{h}|^{p\theta} \mathbf{1}_{|\mathbf{h}|\leq 1}) |f|^{p\theta} \rho_x \rangle^{\frac{1}{p\theta}} \right. \\ & \left. + \lambda^{-\beta} \langle (\mathbf{1}_{|\mathbf{h}|>1} + |\mathbf{h}|^{p\theta'} \mathbf{1}_{|\mathbf{h}|\leq 1}) |f|^{p\theta'} \rho_x \rangle^{\frac{1}{p\theta'}} \right) \end{aligned}$$

for all  $\lambda \geq \lambda_0 \vee 1$ .

**2.4. Gradient bounds.** The following result is used in the proof of Theorem 7.

**Theorem 10** (Gradient bounds). *Assume that a bounded smooth vector field  $b$  satisfies either condition  $(\mathbb{B}_1)$  of Theorem 7 or*

$$\begin{cases} b \in \mathbf{F}_\delta \cap [W_{\text{loc}}^{1,1}(\mathbb{R}^d)]^d \text{ with finite } \delta \text{ and symmetric Jacobian } Db := (\nabla_k b_i)_{k,i=1}^d, \\ \text{and the negative part } B_- \text{ of matrix } Db - \frac{\operatorname{div} b}{q} I \text{ for some } q > (d-2) \vee 2 \\ \text{satisfies } \langle B_- h, h \rangle \leq \nu \langle |\nabla h|^2 \rangle + c_\nu \langle |h|^2 \rangle \text{ for some } \nu < \frac{4(q-1)}{q^2}. \end{cases} \quad (\bar{\mathbb{B}}_2)$$

Then the following are true:

(i) For every  $\mathbf{g} \in \mathbf{F}$  there exist generic constants  $\mu_0$  and  $K$  such that, for every  $\mu > \mu_0$ , the classical solution  $u$  to elliptic equation  $(\mu - \Delta - b \cdot \nabla)u = |\mathbf{g}|f$ ,  $f \in C_c^\infty$ , satisfies

$$\|\nabla |\nabla u|^{\frac{q}{2}}\|_2 \leq K(\mu - \mu_0)^{-\frac{1}{q}} \|\mathbf{g}|f|^{\frac{q}{2}}\|_2.$$

(ii) There exist generic constants  $\mu_0$  and  $K$  such that the classical solution  $u$  to elliptic equation  $(\mu - \Delta - b \cdot \nabla)u = f$ ,  $f \in C_c^\infty$ , satisfies, for all  $\mu > \mu_0$ ,

$$\|\nabla |\nabla u|^{\frac{q}{2}}\|_2 \leq K(\mu - \mu_0)^{-\frac{1}{q}} \|f\|_{\frac{qd}{d+q-2}}.$$

For example, condition  $(\bar{\mathbb{B}}_2)$  holds if condition  $(\mathbb{B}_2)$  of Theorem 7 is satisfied, see Lemma 11.

Assuming that  $b \in \mathbf{F}_\delta$ ,  $\delta < (\frac{2}{d-2})^2 \wedge 1$ , [36] proved estimate

$$\|\nabla|\nabla u|^{\frac{q}{2}}\|_2 \leq K(\mu - \mu_0)^{\frac{1}{q} - \frac{1}{2}} \|f\|_q, \quad q \in ](d-2) \vee 2, \frac{2}{\sqrt{\delta}}[ \quad (2.44)$$

for solution  $u$  to elliptic equation  $(\mu - \Delta - b \cdot \nabla)u = f$ . This estimate was used in [36] to construct the corresponding to  $-\Delta - b \cdot \nabla$  Feller semigroup via a Moser-type iteration procedure. The norm  $\|f\|_q$  in the right-hand side of (2.44) does not allow to obtain the uniqueness result in Theorem 7 from (2.44), unless  $b$  satisfies additional assumption  $|b| \in L^{(d-2)\vee 2}$ . Still, the argument of [36] can be modified to include a weaker norm of  $f$ , and this is what we do in the proof of Theorem 10. In particular, we use the test function

$$\phi = -\nabla \cdot (\nabla u |\nabla u|^{q-2}) \quad (2.45)$$

of [36]. In more recent literature one can find other test functions that give gradient bounds on  $u$  of the same type as in Theorem 10 (moreover, these test functions work for larger classes of equations). However, importantly, test function (2.45) yields the least restrictive assumptions on form-bounds  $\delta$  and  $\nu$ , which are in the focus of the present paper. In fact, one can argue that by multiplying the elliptic equation by test function (2.45) and integrating by parts, one differentiates the equation in the optimal direction  $\frac{\nabla u}{|\nabla u|}$ . We refer to [23] for more detailed discussion and references.

### 3. SMOOTH APPROXIMATION OF FORM-BOUNDED VECTOR FIELDS

The proofs of Lemmas 3-6 employ some arguments from [30]. Let  $b \in [L^1_{\text{loc}}(\mathbb{R}^d)]^d$ . Define

$$b_\varepsilon := E_\varepsilon b, \quad \varepsilon > 0,$$

where, recall,  $E_\varepsilon h$  denotes the Friedrichs mollifier of function (or vector field)  $h$ , see Section 1.1 for the definition.

**Lemma 3.** *If  $b \in \mathbf{F}_\delta$ , then the following is true:*

1.  $b_\varepsilon \in [L^\infty(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)]^d$ ,  $b_\varepsilon \rightarrow b$  in  $[L^2_{\text{loc}}(\mathbb{R}^d)]^d$  as  $\varepsilon \downarrow 0$ .
2.  $b_\varepsilon \in \mathbf{F}_\delta$  with the same constant  $c_\delta$  (thus, independent of  $\varepsilon$ ).

*Proof.* 1. The smoothness of  $b_\varepsilon$  and the convergence follow from the standard properties of Friedrichs mollifiers, so it remains to prove that  $|b_\varepsilon| \in L^\infty$ . By Hölder's inequality,

$$|b_\varepsilon(x)| \leq \sqrt{E_\varepsilon |b|^2(x)} = \sqrt{\gamma_\varepsilon(x - \cdot) |b(\cdot)|^2},$$

so

$$\begin{aligned} |b_\varepsilon(x)| &\leq \langle |b(\cdot)|^2 \gamma_\varepsilon(x - \cdot) \rangle^{\frac{1}{2}} \\ &\quad \text{(we apply the hypothesis } b \in \mathbf{F}_\delta) \\ &\leq (\delta \langle |\nabla \sqrt{\gamma_\varepsilon(x - \cdot)}|^2 \rangle + c_\delta)^{\frac{1}{2}} = (C\varepsilon^{-1} + c_\delta)^{\frac{1}{2}}. \end{aligned}$$

Hence  $|b_\varepsilon| \in L^\infty$  for each  $\varepsilon > 0$ .

2. Put  $\varphi_\varepsilon = \sqrt{E_\varepsilon |\varphi|^2}$ ,  $\varphi \in W^{1,2}$ . Then

$$\|b_\varepsilon \varphi\|_2^2 \leq \langle E_\varepsilon |b|^2, |\varphi|^2 \rangle = \|b_\varepsilon \varphi\|_2^2 \leq \delta \|\nabla \varphi_\varepsilon\|_2^2 + c(\delta) \|\varphi_\varepsilon\|_2^2,$$

where

$$\|\nabla\varphi_\varepsilon\|_2 = \left\| \frac{E_\varepsilon(|\varphi|\nabla\varphi)}{\sqrt{E_\varepsilon|\varphi|^2}} \right\|_2 \leq \|\sqrt{E_\varepsilon|\nabla\varphi|^2}\|_2 = \|E_\varepsilon|\nabla\varphi|^2\|_1^{\frac{1}{2}} \leq \|\nabla\varphi\|_2 \leq \|\nabla\varphi\|_2 \quad (3.1)$$

and, clearly,  $\|\varphi_\varepsilon\|_2 \leq \|\varphi\|_2$ .  $\square$

**Lemma 4.** *If  $b \in \mathbf{MF}_\delta$ , then the following is true:*

1.  $b_\varepsilon \in [L^\infty(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)]^d$ ,  $b_\varepsilon \rightarrow b$  in  $[L^1_{\text{loc}}(\mathbb{R}^d)]^d$ .
2.  $b_\varepsilon \in \mathbf{MF}_\delta$  with the same  $c_\delta$ .

*Proof.* 1. We only need to prove  $|b_\varepsilon| \in L^\infty$ . By  $b \in \mathbf{MF}_\delta$ , for all  $x \in \mathbb{R}^d$ ,

$$|b_\varepsilon(x)| \leq \langle |b(\cdot)|\gamma_\varepsilon(x - \cdot) \rangle \leq \delta \langle |\nabla\sqrt{\gamma_\varepsilon(x - \cdot)}|^2 \rangle^{\frac{1}{2}} + c_\delta = C\varepsilon^{-\frac{1}{2}} + c_\delta.$$

2. Let  $\varphi_\varepsilon = \sqrt{E_\varepsilon|\varphi|^2}$ ,  $\varphi \in W^{1,2}$ . We have

$$\langle |b_\varepsilon|\varphi, \varphi \rangle = \langle |b|E_\varepsilon|\varphi|^2 \rangle = \langle |b|\varphi_\varepsilon^2 \rangle \leq \delta \|\nabla\varphi_\varepsilon\|_2 \|\varphi_\varepsilon\|_2 + c_\delta \|\varphi_\varepsilon\|_2^2,$$

where, repeating the previous proof,  $\|\nabla\varphi_\varepsilon\|_2 \leq \|\nabla\varphi\|_2$ ,  $\|\varphi_\varepsilon\|_2 \leq \|\varphi\|_2$ .  $\square$

Assume that  $\text{div } b \in L^1_{\text{loc}}$ . We can represent  $\text{div } b_\varepsilon = E_\varepsilon \text{div } b$  as

$$\text{div } b_\varepsilon = \text{div } b_{\varepsilon,+} - \text{div } b_{\varepsilon,-},$$

where

$$\text{div } b_{\varepsilon,+} := E_\varepsilon(\text{div } b)_+, \quad \text{div } b_{\varepsilon,-} := E_\varepsilon(\text{div } b)_-.$$

Note that smooth functions  $\text{div } b_{\varepsilon,\pm} \geq 0$  are in general greater than the positive and the negative parts  $(\text{div } b_\varepsilon)_+ := \text{div } b_\varepsilon \vee 0$ ,  $(\text{div } b_\varepsilon)_- := -(\text{div } b_\varepsilon \wedge 0)$  of  $\text{div } b_\varepsilon$ .

**Lemma 5.** *If  $(\text{div } b)_+ \in \mathbf{F}_{\delta,+}$ ,  $(\text{div } b)_- \in L^1 + L^\infty$ , then the following is true:*

1.  $\text{div } b_{\varepsilon,+} \in L^\infty \cap C^\infty$ ,  $\text{div } b_{\varepsilon,+} \rightarrow (\text{div } b)_+$  in  $L^1_{\text{loc}}$  as  $\varepsilon \downarrow 0$ .
2.  $\text{div } b_{\varepsilon,+} \in \mathbf{F}_{\delta,+}$  with the same  $c_{\delta,+}$  as the one for  $b$ .

*Proof.* The first statement follows from the properties of Friedrichs mollifiers and the following estimate (we use notations from the previous proof): for every  $x \in \mathbb{R}^d$ ,

$$\text{div } b_{\varepsilon,+}(x) \leq \langle (\text{div } b)_+(\cdot)\gamma_\varepsilon(x - \cdot) \rangle \leq \delta_+ \langle |\nabla\sqrt{\gamma_\varepsilon(x - \cdot)}|^2 \rangle + c_{\delta,+} = C\varepsilon^{-1} + c_{\delta,+}$$

Let us prove the second statement:

$$\langle \text{div } b_{\varepsilon,+}\varphi, \varphi \rangle = \langle (\text{div } b)_+\varphi_\varepsilon^2 \rangle \leq \delta_+ \|\nabla\varphi_\varepsilon\|_2^2 + c_{\delta,+} \|\varphi_\varepsilon\|_2^2 \leq \delta_+ \|\nabla\varphi\|_2^2 + c_{\delta,+} \|\varphi\|_2^2.$$

$\square$

Finally, we will need

**Lemma 6.** *If  $|\mathbf{h}|^{\frac{1+\gamma}{2}} \in \mathbf{F}_\chi$  ( $\gamma > 0$ ), then the following is true:*

1.  $\mathbf{h}_\varepsilon := E_\varepsilon \mathbf{h} \in [L^\infty(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d)]^d$ ,  $\mathbf{h}_\varepsilon \rightarrow \mathbf{h}$  in  $[L^1_{\text{loc}}(\mathbb{R}^d)]^d$  as  $\varepsilon \downarrow 0$ ,
2.  $|\mathbf{h}_\varepsilon| \in \mathbf{F}_\chi$  with the same  $c_\chi$ .

*Proof.* By Hölder's inequality,  $|\mathbf{h}_\varepsilon|^{1+\gamma} \leq E_\varepsilon |\mathbf{h}|^{1+\gamma}$ , so  $\langle |\mathbf{h}_\varepsilon|^{1+\gamma}\varphi^2 \rangle \leq \langle |\mathbf{h}|^{1+\gamma}, \varphi_\varepsilon^2 \rangle$ , where, recall,  $\varphi_\varepsilon = \sqrt{E_\varepsilon|\varphi|^2}$ ,  $\varphi \in W^{1,2}$ . Now we apply  $|\mathbf{h}|^{\frac{1+\gamma}{2}} \in \mathbf{F}_\chi$  and use  $\|\nabla\varphi_\varepsilon\|_2 \leq \|\nabla\varphi\|_2$ ,  $\|\varphi_\varepsilon\|_2 \leq \|\varphi\|_2$ .  $\square$

## 4. PROOFS OF LEMMAS 1 AND 2

Recall:  $b = (b_1, \dots, b_N) : \mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$  is defined by

$$b_i(x) := \frac{1}{N} \sum_{j=1, j \neq i}^N K_{ij}(x_i - x_j), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^{Nd}, \quad 1 \leq i \leq N.$$

Below  $|\cdot|$  denotes, depending on the context, the Euclidean norm in  $\mathbb{R}^{Nd}$  or  $\mathbb{R}^d$ . In this section,  $\langle \cdot, \cdot \rangle$  is the integration over  $\mathbb{R}^{Nd}$ .

*Proof of Lemma 1.* We have

$$\begin{aligned} |b(x)|^2 &\leq \sum_{i=1}^N |b_i(x)|^2 \leq \sum_{i=1}^N \left( \frac{1}{N} \sum_{j=1, j \neq i}^N |K_{ij}(x_i - x_j)| \right)^2 \\ &\leq \sum_{i=1}^N \frac{N-1}{N^2} \sum_{j=1, j \neq i}^N |K_{ij}(x_i - x_j)|^2. \end{aligned}$$

Therefore,  $\langle |b|^2 \varphi^2 \rangle \leq \sum_{i=1}^N \frac{N-1}{N^2} \sum_{j=1, j \neq i}^N \langle |K_{ij}(x_i - x_j)|^2 \varphi^2 \rangle$ , where, denoting by  $\bar{x}$  vector  $x$  with component  $x_i$  removed, we estimate

$$\begin{aligned} \langle |K_{ij}(x_i - x_j)|^2 \varphi^2 \rangle &= \int_{\mathbb{R}^{(N-1)d}} \int_{\mathbb{R}^d} |K_{ij}(x_i - x_j)|^2 \varphi^2(x_i, \bar{x}) dx_i d\bar{x} \\ &\quad (\text{we use } K_{ij} \in \mathbf{F}_\kappa(\mathbb{R}^d) \text{ in } x_i \text{ variable}) \\ &\leq \kappa \int_{\mathbb{R}^{(N-1)d}} \int_{\mathbb{R}^d} |\nabla_{x_i} \varphi(x_i, \bar{x})|^2 dx_i d\bar{x} + c_\kappa \int_{\mathbb{R}^{Nd}} \varphi^2 dx \\ &= \kappa \langle |\nabla_{x_i} \varphi|^2 \rangle + c_\kappa \langle \varphi^2 \rangle. \end{aligned}$$

Hence  $\langle |b|^2 \varphi^2 \rangle \leq \frac{(N-1)^2}{N^2} \kappa \langle |\nabla \varphi|^2 \rangle + \frac{(N-1)^2}{N} c_\kappa \langle \varphi^2 \rangle$ , as claimed.  $\square$

*Proof of Lemma 2.* Let us first prove (2.14). We have

$$\langle |b| \varphi^2 \rangle \leq \sum_{i=1}^N \langle |b_i| \varphi^2 \rangle = \sum_{i=1}^N \frac{1}{N} \sum_{j=1, j \neq i}^N \langle |K_{ij}(x_i - x_j)| \varphi^2 \rangle. \quad (4.1)$$

Denoting by  $\bar{x}$  the variable  $x$  with component  $x_i$  removed, we estimate

$$\begin{aligned} \langle |K_{ij}(x_i - x_j)| \varphi^2 \rangle &= \int_{\mathbb{R}^{(N-1)d}} \int_{\mathbb{R}^d} |K_{ij}(x_i - x_j)| \varphi^2(x_i, \bar{x}) dx_i d\bar{x} \\ &\quad (\text{apply } K_{ij} \in \mathbf{MF}_\kappa(\mathbb{R}^d) \text{ in } x_i \text{ variable}) \\ &\leq \int_{\mathbb{R}^{(N-1)d}} \left[ \left( \int_{\mathbb{R}^d} |\nabla_{x_i} \varphi(x_i, \bar{x})|^2 dx_i \right)^{\frac{1}{2}} \left( \int_{\mathbb{R}^d} \varphi^2(x_i, \bar{x}) dx_i \right)^{\frac{1}{2}} + c_\kappa \int_{\mathbb{R}^d} \varphi^2(x_i, \bar{x}) dx_i \right] d\bar{x} \\ &\leq \kappa \langle |\nabla_{x_i} \varphi|^2 \rangle^{\frac{1}{2}} \langle \varphi^2 \rangle^{\frac{1}{2}} + c_\kappa \langle \varphi^2 \rangle. \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{i=1}^N \frac{1}{N} \sum_{j=1, j \neq i}^N \langle |K_{ij}(x_i - x_j)| \varphi^2 \rangle &\leq \sum_{i=1}^N \frac{N-1}{N} \left[ \kappa \langle |\nabla_{x_i} \varphi|^2 \rangle^{\frac{1}{2}} \langle \varphi^2 \rangle^{\frac{1}{2}} + c_\kappa \langle \varphi^2 \rangle \right] \\ &\leq \frac{N-1}{N} \sqrt{N} \kappa \langle |\nabla \varphi|^2 \rangle^{\frac{1}{2}} \langle \varphi^2 \rangle^{\frac{1}{2}} + (N-1) c_\kappa \langle \varphi^2 \rangle. \end{aligned}$$



Applying these estimates in (4.1), we obtain (2.14).

Next, we prove (2.15). We have  $\operatorname{div} b(x) = \sum_{i=1}^N \frac{1}{N} \sum_{j=1, j \neq i}^N (\operatorname{div} K_{ij})(x_i - x_j)$ . So,

$$(\operatorname{div} b)_+ = \sum_{i=1}^N \frac{1}{N} \sum_{j=1, j \neq i}^N (\operatorname{div} K_{ij})_+(x_i - x_j).$$

Hence, by  $(\operatorname{div} K_{ij})_+^{\frac{1}{2}} \in \mathbf{F}_{\kappa_+}(\mathbb{R}^d)$  (note that this condition is linear in  $(\operatorname{div} K_{ij})_+$ ),

$$\langle (\operatorname{div} b)_+, \varphi^2 \rangle \leq \frac{N-1}{N} \kappa_+ \langle |\nabla \varphi|^2 \rangle + (N-1) c_{\kappa_+} \langle \varphi^2 \rangle,$$

i.e. we have proved (2.15) for  $(\operatorname{div} b)_+$ .

Now, we prove (2.16). We have

$$\begin{aligned} |b|^{1+\alpha} &\leq \sum_{i=1}^N |b_i(x)|^{1+\alpha} \leq \sum_{i=1}^N \left( \frac{1}{N} \sum_{j=1, j \neq i}^N |K_{ij}(x_i - x_j)| \right)^{1+\alpha} \\ &\leq 2^\alpha \sum_{i=1}^N \frac{(N-1)^\alpha}{N^{1+\alpha}} \sum_{j=1, j \neq i}^N |K_{ij}(x_i - x_j)|^{1+\alpha}. \end{aligned}$$

Therefore, applying  $|K_{ij}|^{\frac{1+\alpha}{2}} \in \mathbf{F}_\sigma(\mathbb{R}^d)$ , we obtain

$$\langle |b|^{1+\alpha} \varphi^2 \rangle \leq \frac{(N-1)^{1+\alpha}}{N^{1+\alpha}} \sigma 2^\alpha \langle |\nabla \varphi|^2 \rangle + N \frac{(N-1)^{1+\alpha}}{N^{1+\alpha}} c_\sigma 2^\alpha \langle \varphi^2 \rangle,$$

which gives us (2.16).  $\square$

REMARK 10. In Remark 5 we promised to prove that vector field  $b^M : \mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$  defined by (2.26) is in  $\mathbf{F}_{\delta^M}(\mathbb{R}^{Nd})$  with  $\delta^M = \mu$ ,  $c_{\delta^M} = N c_\mu$ . Here is the proof:

$$|b^M(x)|^2 = \sum_{i=1}^N |M_i(x_i)|^2,$$

where (recall that  $\langle \cdot, \cdot \rangle$  is the integration over  $\mathbb{R}^{Nd}$ ,  $\bar{x}$  is vector  $x \in \mathbb{R}^{Nd}$  with component  $x_i \in \mathbb{R}^d$ )

$$\begin{aligned} \langle |M_i(x_i)| \varphi^2 \rangle &= \int_{\mathbb{R}^{(N-1)d}} \int_{\mathbb{R}^d} |M_i(x_i)| \varphi^2(x_i, \bar{x}) dx_i d\bar{x} \\ &\text{(we use } M_i \in \mathbf{F}_\mu(\mathbb{R}^d) \text{ in } x_i \text{ variable)} \\ &\leq \mu \int_{\mathbb{R}^{(N-1)d}} \int_{\mathbb{R}^d} |\nabla_{x_i} \varphi(x_i, \bar{x})|^2 dx_i d\bar{x} + c_\mu \int_{\mathbb{R}^{Nd}} \varphi^2 dx = \mu \langle |\nabla_{x_i} \varphi|^2 \rangle + c_\mu \langle \varphi^2 \rangle. \end{aligned}$$

So,

$$\langle |b^M(x)|^2 \varphi^2 \rangle = \sum_{i=1}^N \langle |M_i(x_i)|^2 \rangle \leq \mu \langle |\nabla \varphi|^2 \rangle + N c_\mu \langle \varphi^2 \rangle,$$

as claimed.

## 5. CACCIOPPOLI'S INEQUALITY

We use the next proposition in the proofs of Theorems 8 and 9.

**Proposition 1.** *Let  $b, h : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be bounded smooth vector fields such that either*

$$b \in \mathbf{F}_\delta \text{ with } \delta < 4, \quad h \in \mathbf{F}_\chi \text{ with } \chi < \infty \quad (5.1)$$

or

$$\begin{cases} b \in \mathbf{MF}_\delta \text{ for some } \delta < \infty, \\ (\operatorname{div} b_+)^{\frac{1}{2}} \in \mathbf{F}_{\delta_+} \text{ with } \delta_+ < 4, \end{cases} \quad |h|^{\frac{1+\gamma}{2}} \in \mathbf{F}_\chi \quad \text{with } \chi < \infty \quad (5.2)$$

for some  $\gamma > 0$  fixed arbitrarily small, where  $\operatorname{div} b = \operatorname{div} b_+ - \operatorname{div} b_-$  for some bounded smooth functions  $\operatorname{div} b_\pm \geq 0$ .

In the former case, fix  $p > \frac{2}{2-\sqrt{\delta}}$ ,  $p \geq 2$ , and in the latter case fix  $p > \frac{4}{4-\delta_+}$ ,  $p' \leq 1 + \gamma$ ,  $p \geq 2$ .

Let  $u$  be a classical solution to non-homogeneous equation

$$(\lambda - \Delta + b \cdot \nabla)u = |hf|, \quad f \in C_c^\infty, \quad \lambda \geq 0. \quad (5.3)$$

Fix  $R_0 \leq 1$ . Then, for all  $0 < r < R \leq R_0$  and every  $k \in \mathbb{R}$  the positive part  $v := (u - k)_+$  of  $u - k$  satisfies

$$\begin{aligned} \lambda \|v^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 + \|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_r}\|_2^2 &\leq \frac{K_1}{(R-r)^2} \|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2^2 \\ &\quad + K_2 \|(\mathbf{1}_{|h|>1} + |h|^{\frac{p}{2}} \mathbf{1}_{|h|\leq 1}) |f|^{\frac{p}{2}} \mathbf{1}_{u>c} \mathbf{1}_{B_R}\|_2^2 \end{aligned}$$

for generic constants  $K_1, K_2$  (so, independent of  $r, R$  and  $k$ ).

*Proof.* Let us first carry out the proof in the more difficult case when  $b$  and  $h$  satisfy condition (5.2). We attend to the case when  $b$  and  $h$  satisfy (5.1) in Remark 11.

We fix a family of  $[0, 1]$ -valued smooth cut-off functions  $\{\eta = \eta_{r_1, r_2}\}_{0 < r_1 < r_2 < R}$  on  $\mathbb{R}^d$  satisfying

$$\eta = \begin{cases} 1 & \text{in } B_{r_1}, \\ 0 & \text{in } \mathbb{R}^d - \bar{B}_{r_2}, \end{cases}$$

and

$$\frac{|\nabla \eta|^2}{\eta} \leq \frac{c}{(r_2 - r_1)^2} \mathbf{1}_{B_{r_2}}, \quad (5.4)$$

$$\sqrt{|\nabla \eta|} \leq \frac{c}{\sqrt{r_2 - r_1}} \mathbf{1}_{B_{r_2}}, \quad (5.5)$$

$$|\nabla \sqrt{|\nabla \eta|}| \leq \frac{c}{(r_2 - r_1)^{\frac{3}{2}}} \mathbf{1}_{B_{r_2}} \quad (5.6)$$

for some constant  $c$ . For instance, one can take, for  $r_1 \leq |y| \leq r_2$ ,

$$\eta(y) := 1 - \int_1^{1 + \frac{|y| - r_1}{r_2 - r_1}} \varphi(s) ds, \quad \text{where } \varphi(s) := C e^{-\frac{1}{4 - (s - \frac{3}{2})^2}}, \quad \operatorname{sprt} \varphi = [1, 2],$$

with constant  $C$  adjusted to  $\int_1^2 \varphi(s) ds = 1$ .

We multiply equation (5.3) by  $v^{p-1} \eta$  and integrate to obtain

$$\begin{aligned} \lambda \langle v^p \eta \rangle + \frac{4(p-1)}{p^2} \langle \nabla v^{\frac{p}{2}}, (\nabla v^{\frac{p}{2}}) \eta \rangle + \frac{2}{p} \langle \nabla v^{\frac{p}{2}}, v^{\frac{p}{2}} \nabla \eta \rangle \\ + \frac{2}{p} \langle b \cdot \nabla v^{\frac{p}{2}}, v^{\frac{p}{2}} \eta \rangle = \langle |h|f, v^{p-1} \eta \rangle. \end{aligned}$$

Then, applying quadratic inequality (fix some  $\epsilon > 0$ ), we have

$$\begin{aligned}
p\lambda\langle v^p\eta\rangle + \left(\frac{4(p-1)}{p} - \frac{4}{p}\epsilon\right)\langle|\nabla v^{\frac{p}{2}}|^2\eta\rangle &\leq \frac{p}{4\epsilon}\langle v^p\frac{|\nabla\eta|^2}{\eta}\rangle - 2\langle b\cdot\nabla v^{\frac{p}{2}}, v^{\frac{p}{2}}\eta\rangle + p\langle|h|f, v^{p-1}\eta\rangle \quad (5.7) \\
&\text{(we are integrating by parts)} \\
&\leq \frac{p}{4\epsilon}\langle v^p\frac{|\nabla\eta|^2}{\eta}\rangle + \langle bv^{\frac{p}{2}}, v^{\frac{p}{2}}\nabla\eta\rangle + \langle\text{div } b, v^p\eta\rangle + p\langle|h|f, v^{p-1}\eta\rangle =: I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

By (5.4),

$$I_1 \leq \frac{cp}{4\epsilon(r_2 - r_1)^2}\|v^{\frac{p}{2}}\mathbf{1}_{B_{r_2}}\|_2^2.$$

By (5.2),

$$\begin{aligned}
I_2 &\leq \langle|b|v^{\frac{p}{2}}, v^{\frac{p}{2}}|\nabla\eta\rangle \\
&\leq \delta\|\nabla(v^{\frac{p}{2}}\sqrt{|\nabla\eta|})\|_2\|v^{\frac{p}{2}}\sqrt{|\nabla\eta|}\|_2 + c_\delta\|v^{\frac{p}{2}}\sqrt{|\nabla\eta|}\|_2^2 \\
&\leq \delta\left(\|(\nabla v^{\frac{p}{2}})\sqrt{|\nabla\eta|}\|_2 + \|v^{\frac{p}{2}}\nabla\sqrt{|\nabla\eta|}\|_2\right)\|v^{\frac{p}{2}}\sqrt{|\nabla\eta|}\|_2 + c_\delta\|v^{\frac{p}{2}}\sqrt{|\nabla\eta|}\|_2^2.
\end{aligned}$$

Hence, using (5.5), (5.6), we obtain

$$\begin{aligned}
I_2 &\leq \delta c\left(\frac{1}{\sqrt{r_2 - r_1}}\|(\nabla v^{\frac{p}{2}})\mathbf{1}_{B_{r_2}}\|_2 + \frac{1}{(r_2 - r_1)^{\frac{3}{2}}}\|v^{\frac{p}{2}}\mathbf{1}_{B_{r_2}}\|_2\right)\frac{1}{\sqrt{r_2 - r_1}}\|v^{\frac{p}{2}}\mathbf{1}_{B_{r_2}}\|_2 \\
&\quad + \frac{c_\delta c}{r_2 - r_1}\|v^{\frac{p}{2}}\mathbf{1}_{B_{r_2}}\|_2^2.
\end{aligned}$$

Thus, since  $r_2 - r_1 < 1$ ,

$$I_2 \leq \frac{C_1}{r_2 - r_1}\|(\nabla v^{\frac{p}{2}})\mathbf{1}_{B_{r_2}}\|_2\|v^{\frac{p}{2}}\mathbf{1}_{B_{r_2}}\|_2 + C_1\left(1 + \frac{1}{(r_2 - r_1)^2}\right)\|v^{\frac{p}{2}}\mathbf{1}_{B_{r_2}}\|_2^2$$

for appropriate constant  $C_1$ .

Next, by (5.2),

$$\begin{aligned}
I_3 &\leq \langle\text{div } b_+, v^p\eta\rangle \leq \delta_+\|\nabla(v^{\frac{p}{2}}\sqrt{\eta})\|_2^2 + c_{\delta_+}\|v\sqrt{\eta}\|_2^2 \\
&= \delta_+\|(\nabla v^{\frac{p}{2}})\sqrt{\eta} + v\frac{\nabla\eta}{\sqrt{\eta}}\|_2^2 + c_{\delta_+}\|v\sqrt{\eta}\|_2^2 \\
&\leq \delta_+\left((1 + \epsilon')\|(\nabla v^{\frac{p}{2}})\sqrt{\eta}\|_2^2 + \left(1 + \frac{1}{\epsilon'}\right)\frac{1}{4}\|v^{\frac{p}{2}}\frac{\nabla\eta}{\sqrt{\eta}}\|_2^2\right) + c_{\delta_+}\|v^{\frac{p}{2}}\sqrt{\eta}\|_2^2 \quad (\epsilon' > 0) \\
&\leq \delta_+(1 + \epsilon')\|(\nabla v^{\frac{p}{2}})\sqrt{\eta}\|_2^2 + \frac{c_1}{(r_2 - r_1)^2}\|v^{\frac{p}{2}}\mathbf{1}_{B_{r_2}}\|_2^2, \quad c_1 := \delta_+\left(1 + \frac{1}{\epsilon'}\right)\frac{1}{4}c + c_{\delta_+}.
\end{aligned}$$

Again by (5.2),

$$\begin{aligned}
\frac{1}{p}I_4 &= \langle (|\mathbf{h}\mathbf{1}_{|\mathbf{h}|>1} + \mathbf{h}\mathbf{1}_{|\mathbf{h}|\leq 1})f, v^{p-1}\eta \rangle \\
&\quad (\text{we open the brackets and apply Young's inequality}) \\
&\leq \varepsilon'' \langle |\mathbf{h}|^{p'} \mathbf{1}_{|\mathbf{h}|>1} v^p \eta \rangle + \frac{1}{4\varepsilon''} \langle \mathbf{1}_{|\mathbf{h}|>1} |f|^p \eta \rangle \\
&\quad + \varepsilon'' \langle v^p \eta \rangle + \frac{1}{4\varepsilon''} \langle |\mathbf{h}|^p \mathbf{1}_{|\mathbf{h}|\leq 1} |f|^p \eta \rangle \\
&\quad (\text{we are applying } |\mathbf{h}|^{\frac{1+\gamma}{2}} \in \mathbf{F}_\chi \text{ in the first term (recall: } p' \leq 1 + \gamma)) \\
&\leq 2\varepsilon'' \chi \langle |\nabla v^{\frac{p}{2}}|^2 \eta \rangle + 2\varepsilon'' \chi \langle v^p \frac{|\nabla \eta|^2}{\eta} \rangle + \varepsilon'' (c_\chi + 1) \langle v^p \eta \rangle + \frac{1}{4\varepsilon''} \langle \Theta |f|^p \mathbf{1}_{v>0} \eta \rangle,
\end{aligned}$$

where  $\Theta := \mathbf{1}_{|\mathbf{h}|>1} + |\mathbf{h}|^p \mathbf{1}_{|\mathbf{h}|\leq 1}$ . Selecting  $\varepsilon''$  sufficiently small and applying the estimates on  $I_1$ - $I_4$  in (5.7), we obtain

$$\begin{aligned}
\lambda \|v^{\frac{p}{2}} \mathbf{1}_{B_{r_1}}\|_2^2 + \| |\nabla v^{\frac{p}{2}}| \mathbf{1}_{B_{r_1}} \|_2^2 &\leq \frac{C_1}{r_2 - r_1} \|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_{r_2}}\|_2 \|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2 \\
&\quad + C_2 \left(1 + \frac{1}{(r_2 - r_1)^2}\right) \|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2^2 + C_3 \|\Theta^{\frac{1}{2}} |f|^{\frac{p}{2}} \mathbf{1}_{v>0} \mathbf{1}_{B_R}\|_2^2. \quad (5.8)
\end{aligned}$$

Divide (5.8) by  $\|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2^2$ :

$$\frac{\lambda \|v^{\frac{p}{2}} \mathbf{1}_{B_{r_1}}\|_2^2 + \|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_{r_1}}\|_2^2}{\|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2^2} \leq \frac{C_1}{r_2 - r_1} \frac{\lambda \|v^{\frac{p}{2}} \mathbf{1}_{B_{r_2}}\|_2^2 + \|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_{r_2}}\|_2^2}{\|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2^2} \quad (5.9)$$

$$+ C_2 \left(1 + \frac{1}{(r_2 - r_1)^2}\right) + C_3 S^2, \quad (5.10)$$

where

$$S^2 := \frac{\|\Theta^{\frac{1}{2}} |f|^{\frac{p}{2}} \mathbf{1}_{v>0} \mathbf{1}_{B_R}\|_2^2}{\|v^{\frac{p}{2}} \mathbf{1}_{v>0} \mathbf{1}_{B_R}\|_2^2}.$$

Inequality (5.10) is the pre-Caccioppoli inequality that we will now iterate.

Put

$$a_n^2 := \frac{\lambda \|v^{\frac{p}{2}} \mathbf{1}_{B_{R-\frac{R-r}{2^{n-1}}}}\|_2^2 + \|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_{R-\frac{R-r}{2^{n-1}}}}\|_2^2}{\|v^{\frac{p}{2}} \mathbf{1}_{v>0} \mathbf{1}_{B_R}\|_2^2},$$

the inequality (5.10) yields

$$a_n^2 \leq C(R-r)^{-1} 2^n a_{n+1} + C^2(R-r)^{-2} 2^{2n} + C^2 S^2$$

with constant  $C$  independent of  $n$ . We multiply this inequality by  $(R-r)^2$  and divide by  $C^2 2^{2n}$ . Then, setting  $y_n := \frac{(R-r)a_n}{C^2 2^n}$ , we obtain

$$y_n^2 \leq y_{n+1} + 1 + (R-r)^2 S^2 \quad (5.11)$$

for all  $n = 1, 2, \dots$ . A priori, all  $a_n$ 's are bounded by a non-generic constant

$$(\lambda \|v^{\frac{p}{2}} \mathbf{1}_B\|_2^2 + \|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_R}\|_2^2) / \|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2 < \infty,$$

so  $\sup_n y_n < \infty$ . Therefore, we can iterate (5.11) and thus estimate all  $y_n$ ,  $n = 1, 2, \dots$ , via nested square roots  $1 + (R - r)^2 S^2 + \sqrt{1 + (R - r)^2 S^2 + \sqrt{\dots}}$ , obtaining

$$y_n^2 \leq 3 + 2(R - r)^2 S^2, \quad n = 1, 2, \dots$$

Taking  $n = 1$ , we arrive at  $a_1 \leq K_1(R - r)^{-2} + K_2 S^2$  for appropriate constants  $K_1$  and  $K_2$ , i.e.

$$\frac{\lambda \|v^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 + \|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_r}\|_2^2}{\|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2^2} \leq K_1(R - r)^{-2} + K_2 \frac{\|\Theta^{\frac{1}{2}} |f|^{\frac{p}{2}} \mathbf{1}_{v>0} \mathbf{1}_{B_R}\|_2^2}{\|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2^2},$$

as claimed.

REMARK 11. If  $b$  and  $\mathbf{h}$  satisfy condition (5.1), then we can work with somewhat simpler cutoff functions  $\eta \in C_c^\infty$ ,  $\eta = 1$  in  $B_{r_1}$ ,  $\eta = 0$  in  $\mathbb{R}^d \setminus B_{r_2}$ , i.e.  $|\nabla \eta| \leq c_1(r_2 - r_1)^{-1}$ ,  $|\Delta \eta| \leq c_2(r_2 - r_1)^{-2}$ , and we do not need to integrate by parts in order to estimate the second term in the RHS of (5.7). Instead, we can just apply quadratic inequality:

$$2|\langle b \cdot \nabla v^{\frac{p}{2}}, v^{\frac{p}{2}} \eta \rangle| \leq \alpha |\langle \nabla v^{\frac{p}{2}}, v^{\frac{p}{2}} \eta \rangle| + \frac{1}{4\alpha} \langle |b|^2, v^p \eta \rangle, \quad \alpha = \frac{2}{\sqrt{\delta}}.$$

Regarding the terms containing  $\mathbf{h}$ , we simply take  $\gamma = 1$ , which transforms condition  $|\mathbf{h}|^{\frac{1+\gamma}{2}} \in \mathbf{F}_\chi$ ,  $\chi < \infty$  from (5.2) into condition  $\mathbf{h} \in \mathbf{F}_\chi$  in (5.1), and argue as in the estimate on  $I_4$  above.

This ends the proof of Proposition 1.  $\square$

## 6. PROOF OF THEOREM 8

If  $b$  satisfies  $(\mathbb{A}_1)$ , then we fix throughout this proof  $p > \frac{2}{2-\sqrt{\delta}}$ ,  $p \geq 2$ . If  $b$  satisfies  $(\bar{\mathbb{A}}_2)$ , then we fix  $p > \frac{2}{4-\delta_+}$ ,  $p \geq 2$ . Let  $u$  be a classical solution to non-homogeneous equation

$$(\lambda - \Delta + b \cdot \nabla)u = f, \quad f \in C_c^\infty.$$

Set

$$v := (u - k)_+, \quad k \in \mathbb{R}.$$

Fix  $R_0 \leq 1$ . Here is a special case of Proposition 1 obtained by taking  $\mathbf{h} = 1$  and discarding the term containing  $\lambda$  there. (Strictly speaking, in Proposition 1 we have  $|f|$  in the RHS, but this does not affect the proof.)

**Proposition 2.** *For all  $0 < r < R \leq R_0$ ,*

$$\|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_r}\|_2^2 \leq \frac{K_1}{(R - r)^2} \|v^{\frac{p}{2}} \mathbf{1}_{B_R}\|_2^2 + K_2 \| |f|^{\frac{p}{2}} \mathbf{1}_{u>k} \mathbf{1}_{B_R} \|_2^2$$

for generic constants  $K_1, K_2$ .

**Lemma 7** ([15, Sect.7.2]). *If  $\{z_m\}_{m=0}^\infty \subset \mathbb{R}_+$  is a sequence of positive real numbers such that*

$$z_{m+1} \leq N C_0^m z_m^{1+\alpha}$$

for some  $C_0 > 1$ ,  $\alpha > 0$ , and

$$z_0 \leq N^{-\frac{1}{\alpha}} C_0^{-\frac{1}{\alpha^2}}.$$

Then  $\lim_m z_m = 0$ .

Below we follow closely De Giorgi's method as it is presented in [15, Ch.7], with appropriate modifications to account for our somewhat different definition of  $L^p$  De Giorgi's classes (i.e. functions satisfying the inequality in Proposition 2), see discussion in Remark 1.

**Proposition 3.** *For all  $0 < r < R \leq R_0$ ,*

$$\sup_{B_{\frac{R}{2}}} u \leq C_1 \left( \frac{1}{|B_R|} \langle u^p \mathbf{1}_{B_R \cap \{u > 0\}} \rangle \right)^{\frac{1}{p}} \left( \frac{|B_R \cap \{u > 0\}|}{|B_R|} \right)^{\frac{\alpha}{p}} + C_2 \|f\|_{\infty} R^{\frac{2}{p}}$$

for generic constants  $C_1, C_2$  that also depend on  $\|f\|_{\infty}$ , where  $\alpha > 0$  is fixed by  $\alpha(\alpha + 1) = \frac{2}{d}$ .

*Proof.* Without loss of generality,  $R_0 = 1$ . Let  $\frac{1}{2} < r < \rho \leq 1$ . Fix  $\eta \in C_c^{\infty}$ ,  $\eta = 1$  on  $B_r$ ,  $\eta = 0$  on  $\mathbb{R}^d \setminus \bar{B}_{\frac{r+\rho}{2}}$ ,  $|\nabla \eta| \leq \frac{4}{\rho-r}$ . Set  $\zeta := \eta v = \eta(u - k)_+$ ,  $k \in \mathbb{R}$ . Using Hölder's inequality and Sobolev's embedding theorem, we obtain

$$\begin{aligned} \|v^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 &\leq \|\zeta^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 \leq \langle \mathbf{1}_{B_r \cap \{u > k\}} \rangle^{\frac{2}{d}} \langle \zeta^{\frac{pd}{d-2}} \mathbf{1}_{B_{\frac{r+\rho}{2}}} \rangle^{\frac{d-2}{d}} \\ &\leq c_1 |B_r \cap \{u > k\}|^{\frac{2}{d}} \langle |\nabla \zeta^{\frac{p}{2}}|^2 \mathbf{1}_{B_{\frac{r+\rho}{2}}} \rangle \\ &= c_1 |B_r \cap \{u > k\}|^{\frac{2}{d}} \langle (|\nabla \eta^{\frac{p}{2}}| v^{\frac{p}{2}})^2 + \eta^{\frac{p}{2}} |\nabla v^{\frac{p}{2}}|^2 \mathbf{1}_{B_{\frac{r+\rho}{2}}} \rangle \end{aligned}$$

Hence

$$\|v^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 \leq c_2 |B_r \cap \{u > k\}|^{\frac{2}{d}} \left( \frac{1}{(\rho-r)^2} \|v^{\frac{p}{2}} \mathbf{1}_{B_{\frac{r+\rho}{2}}}\|_2^2 + \|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_{\frac{r+\rho}{2}}}\|_2^2 \right).$$

Proposition 2 yields:

$$\|(\nabla v^{\frac{p}{2}}) \mathbf{1}_{B_{\frac{r+\rho}{2}}}\|_2^2 \leq \frac{K_1}{(\rho-r)^2} \|v^{\frac{p}{2}} \mathbf{1}_{B_{\rho}}\|_2^2 + K_2 \|f\|_{\infty}^p |B_{\rho} \cap \{u > k\}|, \quad (6.1)$$

so

$$\begin{aligned} \|v^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 &\leq C |B_r \cap \{u > k\}|^{\frac{2}{d}} \left( \frac{1}{(\rho-r)^2} \|v^{\frac{p}{2}} \mathbf{1}_{B_{\rho}}\|_2^2 + \|f\|_{\infty}^p |B_{\rho} \cap \{u > k\}| \right) \\ &\leq \frac{C |B_{\rho} \cap \{u > k\}|^{\frac{2}{d}}}{(\rho-r)^2} \|v^{\frac{p}{2}} \mathbf{1}_{B_{\rho}}\|_2^2 + C \|f\|_{\infty}^p |B_{\rho} \cap \{u > k\}|^{1+\frac{2}{d}}. \end{aligned} \quad (6.2)$$

Now, returning from notation  $v$  to  $(u - k)_+$ , we note that if  $h < k$ , then  $\|(u - k)_+^{\frac{p}{2}} \mathbf{1}_{B_{\rho} \cap \{u > k\}}\|_2 \leq \|(u - h)_+^{\frac{p}{2}} \mathbf{1}_{B_{\rho} \cap \{u > h\}}\|_2$  and  $\|(u - h)_+^{\frac{p}{2}} \mathbf{1}_{B_{\rho} \cap \{u > h\}}\|_2^2 \geq (k - h)^p |B_{\rho} \cap \{u > h\}|$ . Therefore, we obtain from (6.2)

$$\begin{aligned} \|(u - k)_+^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 &\leq \frac{C}{(\rho-r)^2} \|(u - h)_+^{\frac{p}{2}} \mathbf{1}_{B_{\rho}}\|_2^2 |B_{\rho} \cap \{u > h\}|^{\frac{2}{d}} \\ &\quad + \frac{C \|f\|_{\infty}^p}{(k-h)^p} \|(u - h)_+^{\frac{p}{2}} \mathbf{1}_{B_{\rho}}\|_2^2 |B_{\rho} \cap \{u > h\}|^{\frac{2}{d}}. \end{aligned}$$

Multiplying this inequality by  $|B_r \cap \{u > k\}|^{\alpha}$  ( $\leq \frac{1}{(k-h)^{p\alpha}} \|(u - h)_+^{\frac{p}{2}} \mathbf{1}_{B_{\rho}}\|_2^{2\alpha}$ ) and using  $\alpha^2 + \alpha = \frac{2}{d}$ , we obtain

$$\begin{aligned} &\|(u - k)_+^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 |B_r \cap \{u > k\}|^{\alpha} \\ &\leq C \left[ \frac{1}{(\rho-r)^2} + \frac{\|f\|_{\infty}^p}{(k-h)^p} \right] \frac{1}{(k-h)^{p\alpha}} \left( \|(u - h)_+^{\frac{p}{2}} \mathbf{1}_{B_{\rho}}\|_2^2 |B_{\rho} \cap \{u > h\}|^{\alpha} \right)^{1+\alpha}. \end{aligned}$$

Now, take  $r := r_{i+1}$ ,  $\rho := r_i$ , where  $r_i := \frac{R}{2}(1 + \frac{1}{2^i})$  and  $k := k_{i+1}$ ,  $h := k_i$ , where  $k_i := \xi(1 - 2^{-i})$ , with constant  $\xi \geq R^{\frac{2}{p}}$  to be chosen later. Then, setting

$$z_i = z(k_i, r_i) := \|(u - k_i)_+^{\frac{p}{2}} \mathbf{1}_{B_{r_i}}\|_2^2 |B_{r_i} \cap \{u > k_i\}|^\alpha,$$

we have  $z_{i+1} \leq K[2^{2i} + \frac{2^{pi}R^2}{\xi^p}] \frac{1}{R^2} \frac{2^{pi\alpha}}{\xi^{p\alpha}} z_i^{1+\alpha}$  hence

$$z_{i+1} \leq 2^{p(1+\alpha)i} \frac{2K}{R^2} \frac{1}{d^{p\alpha}} z_i^{1+\alpha}.$$

We apply Lemma 7. In the notation of this lemma,  $C_0 = 2^p$  and  $N = \frac{2K}{R^2} \frac{1}{\xi^{p\alpha}}$ . We need

$$z_0 \leq N^{-\frac{1}{\alpha}} C_0^{-\frac{1}{\alpha^2}}$$

(where, recall,  $z_0 = \langle u^p \mathbf{1}_{B_R \cap \{u > 0\}} \rangle |B_R \cap \{u > 0\}|^\alpha$ ), which amounts to requiring  $\xi \geq C_1 R^{\frac{2}{p\alpha}} z_0^{\frac{1}{p}}$ .

Take  $\xi := R^{\frac{2}{p}} + C_1 R^{\frac{2}{p\alpha}} z_0^{\frac{1}{p}}$ . By Lemma 7,  $z(d, \frac{R}{2}) = 0$ , i.e.  $\sup_{\frac{R}{2}} u \leq \xi$ . The claimed inequality follows.  $\square$

Set  $\text{osc}(u, R) := \sup_{y', y \in B_R} |u(y) - u(y')|$ .

**Proposition 4.** 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

$$\text{osc}(u, 2R) \geq 2^{n+1} C_2 R^{\frac{2}{p}},$$

then, for  $k_n := M(2R) - 2^{-n-1} \text{osc}(u, 2R)$ ,

$$|B_R \cap \{u > k_n\}| \leq cn^{-\frac{d}{2(d-1)}} |B_R|.$$

*Proof.* For  $h \in ]k_0, k[$ , set  $w := (u - h)^{\frac{p}{2}}$  if  $h < u < k$ , set  $w := (k - h)^{\frac{p}{2}}$  if  $u \geq k$ , and  $w := 0$  if  $u \leq h$ . 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 applies and yields

$$\begin{aligned} (k - h)^{\frac{p}{2}} |B_R \cap \{u > k\}|^{\frac{d-1}{d}} &\leq c_1 \langle w^{\frac{d}{d-1}} \mathbf{1}_{B_R} \rangle \leq c_2 \langle |\nabla w| \mathbf{1}_\Delta \rangle \\ &\leq c_2 |\Delta|^{\frac{1}{2}} \langle |\nabla(u - h)^{\frac{p}{2}}|^2 \mathbf{1}_{B_R \cap \{u > h\}} \rangle^{\frac{1}{2}}, \end{aligned}$$

where  $\Delta := B_R \cap \{u > h\} \setminus (B_R \cap \{u > k\})$ . Now, it follows from Proposition 2 that

$$\begin{aligned} \langle |\nabla(u - h)^{\frac{p}{2}}|^2 \mathbf{1}_{B_R \cap \{u > h\}} \rangle &\leq \frac{C_3}{R^2} \langle (u - h)^p \mathbf{1}_{B_{2R} \cap \{u > h\}} \rangle + C_4 |B_{2R} \cap \{u > h\}| \\ &\leq C_3 R^{d-2} (M(2R) - h)^p + C_5 R^d. \end{aligned}$$

For  $h \leq k_n$  we have  $M(2R) - h \geq M(2R) - k_n \geq C_2 R^{\frac{2}{p}}$ . Therefore,

$$(k - h)^{\frac{p}{2}} |B_R \cap \{u > k\}|^{\frac{d-1}{d}} \leq c |\Delta|^{\frac{1}{2}} R^{\frac{d-2}{2}} (M(2R) - h)^{\frac{p}{2}}.$$

Select  $k = k_i := M(2R) - 2^{-i-1} \text{osc}(u, 2R)$ ,  $h = k_{i-1}$ . Then

$$M(2R) - h = 2^{-i} \text{osc}(u, 2R), \quad |k - h| = 2^{-i-1} \text{osc}(u, 2R),$$

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 in  $i$  from 1 to  $n$ , we obtain

$$n|B_R \cap \{u > k_n\}|^{\frac{2(d-1)}{d}} \leq CR^{d-2}|B_R \cap \{u > k_0\}| \leq C'R^{2(d-1)},$$

and the claimed inequality follows.  $\square$

*Proof of Theorem 8, completed.* Fix  $k_0$  by  $2k_0 = M(2R) - m(2R)$ . Without loss of generality,  $|B_R \cap \{u > k_0\}| \leq \frac{1}{2}|B_R|$  (otherwise we replace  $u$  by  $-u$ ). Set  $k_n := M(2R) - 2^{-n-1}\text{osc}(u, 2R) > k_0$ . By Proposition 3,

$$\begin{aligned} \sup_{B_R}(u - k_n) &\leq C_1 \left( \frac{1}{|B_R|} \langle (u - k_n)^p \mathbf{1}_{B_R \cap \{u > k_n\}} \rangle \right)^{\frac{1}{p}} \left( \frac{|B_R \cap \{u > k_n\}|}{|B_R|} \right)^{\frac{\alpha}{p}} + C_3 R^{\frac{2}{p}} \\ &\leq C_1 \sup_{B_R}(u - k_n) \left( \frac{|B_R \cap \{u > k_n\}|}{|B_R|} \right)^{\frac{1+\alpha}{p}} + C_3 R^{\frac{2}{p}} \end{aligned}$$

Fix  $n$  by  $cn^{-\frac{d}{2(d-1)}} \leq (\frac{1}{2})^{\frac{p}{1+\alpha}}$ . Then, if  $\text{osc}(u, 2R) \geq 2^{n+1}C_2R^{\frac{2}{p}}$ , we obtain from Proposition 4

$$M\left(\frac{R}{2}\right) - k_n \leq \frac{1}{2}(M(2R) - k_n) + C_3R^{\frac{2}{p}},$$

hence

$$\text{osc}\left(u, \frac{R}{2}\right) \leq \left(1 - \frac{1}{2^{n+1}}\right)\text{osc}(u, 2R) + C_3R^{\frac{2}{p}}. \quad (6.3)$$

If  $\text{osc}(u, 2R) \geq 2^{n+1}C_2R^{\frac{2}{p}}$ , then, clearly,

$$\text{osc}\left(u, \frac{R}{2}\right) \leq \left(1 - \frac{1}{2^{n+1}}\right)\text{osc}(u, 2R) + 2^{n+1}C_2R^{\frac{2}{p}}. \quad (6.4)$$

This yields the sought Hölder continuity of  $u$  via a standard algebraic lemma, see [15, Lemma 7.1].  $\square$

## 7. PROOF OF THEOREM 9

We will use Proposition 1. The assumptions of Theorem 9 are exactly those of Proposition 1.

**Proposition 5.** *There exists generic constants  $K$  and  $\beta \in ]0, 1[$  such that, for all  $\lambda \geq 1$ , the positive part  $u_+$  of solution  $u$  of non-homogeneous equation (2.43) satisfies*

$$\sup_{B_{\frac{1}{2}}} u_+ \leq K \left( \langle u_+^{p\theta} \mathbf{1}_{B_1} \rangle^{\frac{1}{p\theta}} + \lambda^{-\beta} \langle (\mathbf{1}_{|h|>1} + |h|^p \mathbf{1}_{|h|\leq 1})^{\theta'} |f|^{p\theta'} \mathbf{1}_{B_1} \rangle^{\frac{1}{p\theta'}} \right). \quad (7.1)$$

*Proof.* Proposition 1 yields

$$\begin{aligned} \lambda \|v^{\frac{p}{2}} \mathbf{1}_{B_r}\|_2^2 + \|v^{\frac{p}{2}}\|_{W^{1,2}(B_r)}^2 &\leq \tilde{K}_1 (R-r)^{-2} \|v\|_{L^p(B_R)}^p \\ &\quad + K_2 \|\Theta^{\frac{1}{p}} f \mathbf{1}_{u>k}\|_{L^p(B_R)}^p, \quad v := (u-k)_+, \quad k \in \mathbb{R}, \end{aligned}$$

where  $\Theta := \mathbf{1}_{|h|>1} + |h|^p \mathbf{1}_{|h|\leq 1}$  and  $\tilde{K}_1, K_2$  are generic constants. By the Sobolev embedding theorem,

$$\lambda \|v\|_{L^p(B_R)}^p + \|v\|_{L^{\frac{pd}{d-2}}(B_r)}^p \leq C_1 (R-r)^{-2} \|v\|_{L^p(B_R)}^p + C_2 \|\Theta^{\frac{1}{p}} f \mathbf{1}_{u>k}\|_{L^p(B_R)}^p.$$



By the interpolation inequality,

$$\lambda^{p\beta} \|v\|_{L^q(B_r)}^p \leq \beta \lambda \|v\|_{L^p(B_r)}^p + (1-\beta) \|v\|_{L^{\frac{pd}{d-2}}(B_r)}^p, \quad 0 < \beta < 1, \quad \frac{1}{q} = \beta \frac{1}{p} + (1-\beta) \frac{d-2}{pd}.$$

Put  $\theta_0 := \frac{q}{p}$ , so  $1 < \theta_0 < \frac{d}{d-2}$ . We fix  $\beta \in ]0, 1[$  sufficiently small so that  $\theta < \theta_0$  where, recall,  $0 < \theta < \frac{d}{d-2}$  was fixed earlier.

REMARK. We could take  $\beta = 0$ , in which case  $q = \frac{pd}{d-2}$  and  $\theta_0 = \frac{d}{d-2}$ . However, then factor  $\lambda^{p\beta}$  in the previous estimate becomes 1, which we prefer to avoid keeping in mind some future applications of Theorem 9.

Hence, taking into account that  $q = p\theta_0$ ,

$$\lambda^{p\beta} \|v\|_{L^{p\theta_0}(B_r)}^p \leq \tilde{C}_1 (R-r)^{-2} \|v\|_{L^p(B_R)}^p + \tilde{C}_2 \|\Theta^{\frac{1}{p}} f \mathbf{1}_{u>k}\|_{L^p(B_R)}^p.$$

Applying Hölder's inequality in the RHS, we obtain

$$\lambda^{p\beta} \|v\|_{L^{p\theta_0}(B_r)}^p \leq \tilde{C}_1 (R-r)^{-2} |B_R|^{\frac{\theta-1}{2\theta}} \|v\|_{L^{p\theta}(B_R)}^p + \tilde{C}_2 \|\Theta^{\frac{1}{p}} f \mathbf{1}_{u>k}\|_{L^p(B_R)}^p. \quad (7.2)$$

Set

$$R_m := \frac{1}{2} + \frac{1}{2^{m+1}}, \quad m \geq 0,$$

so  $B_m \equiv B_{R_m}$  is a decreasing sequence of balls converging to the ball of radius  $\frac{1}{2}$ . By (7.2),

$$\begin{aligned} \lambda^{p\beta} \|v\|_{L^{p\theta_0}(B_{m+1})}^p &\leq \hat{C}_1 2^{2m} \|v\|_{L^{p\theta}(B_m)}^p + \tilde{C}_2 \|\Theta^{\frac{1}{p}} f \mathbf{1}_{u>k}\|_{L^p(B_m)}^p \\ &\leq \hat{C}_1 2^{2m} \|v\|_{L^{p\theta}(B_m)}^p + \tilde{C}_2 H |B_m \cap \{v > 0\}|^{\frac{1}{\theta}}, \end{aligned} \quad (7.3)$$

where

$$H := \langle \Theta^{\theta'} |f|^{p\theta'} \mathbf{1}_{B_m} \rangle^{\frac{1}{\theta'}}.$$

On the other hand, by Hölder's inequality,

$$\|v\|_{L^{p\theta}(B_{m+1})}^{p\theta} \leq \|v\|_{L^{p\theta_0}(B_{m+1})}^{p\theta} \left( |B_m \cap \{v > 0\}| \right)^{1 - \frac{\theta}{\theta_0}}.$$

Applying (7.3) in the first multiple in the RHS, we obtain

$$\|v\|_{L^{p\theta}(B_{m+1})}^{p\theta} \leq \tilde{C} \lambda^{-p\beta\theta} \left( 2^{2\theta m} \|v\|_{L^{p\theta}(B_m)}^{p\theta} + H^\theta |B_m \cap \{v > 0\}| \right) \left( |B_m \cap \{v > 0\}| \right)^{1 - \frac{\theta}{\theta_0}}.$$

Now, put  $v_m := (u - k_m)_+$  where  $k_m := \xi(1 - 2^{-m}) \uparrow \xi$ , where constant  $\xi > 0$  will be chosen later. Then, using  $2^{2\theta m} \leq 2^{p\theta m}$  and dividing by  $\xi^{p\theta}$ ,

$$\begin{aligned} &\frac{1}{\xi^{p\theta}} \|v_{m+1}\|_{L^{p\theta}(B_{m+1})}^{p\theta} \\ &\leq \tilde{C} \lambda^{-p\beta\theta} \left( \frac{2^{p\theta m}}{\xi^{p\theta}} \|v_{m+1}\|_{L^{p\theta}(B_m)}^{p\theta} + \frac{1}{\xi^{p\theta}} H^\theta |B_m \cap \{u > k_{m+1}\}| \right) \left( |B_m \cap \{u > k_{m+1}\}| \right)^{1 - \frac{\theta}{\theta_0}}. \end{aligned}$$

Using that  $\lambda \geq 1$ , we further obtain

$$\begin{aligned} & \frac{1}{\xi^{p\theta}} \|v_{m+1}\|_{L^{p\theta}(B_{m+1})}^{p\theta} \\ & \leq \tilde{C} \left( \frac{2^{p\theta m}}{\xi^{p\theta}} \|v_{m+1}\|_{L^{p\theta}(B_m)}^{p\theta} + \frac{1}{\xi^{p\theta}} \lambda^{-p\beta\theta} H^\theta |B_m \cap \{u > k_{m+1}\}| \right) \left( |B_m \cap \{u > k_{m+1}\}| \right)^{1 - \frac{\theta}{\theta_0}}. \end{aligned}$$

From now on, we require that constant  $\xi$  satisfies  $\xi^p \geq \lambda^{-p\beta\theta} H$ , so

$$\begin{aligned} & \frac{1}{\xi^{p\theta}} \|v_{m+1}\|_{L^{p\theta}(B_{m+1})}^{p\theta} \\ & \leq \tilde{C} \left( \frac{2^{p\theta m}}{\xi^{p\theta}} \|v_{m+1}\|_{L^{p\theta}(B_m)}^{p\theta} + |B_m \cap \{u > k_{m+1}\}| \right) \left( |B_m \cap \{u > k_{m+1}\}| \right)^{1 - \frac{\theta}{\theta_0}}. \end{aligned} \tag{7.4}$$

Now,

$$\begin{aligned} |B_m \cap \{u > k_{m+1}\}| &= |B_m \cap \left\{ \left( \frac{u - k_m}{k_{m+1} - k_m} \right)^{2\theta} > 1 \right\}| \\ &\leq (k_{m+1} - k_m)^{-p\theta} \langle v_m^{p\theta} \mathbf{1}_{B_m} \rangle = \xi^{-p\theta} 2^{p\theta(m+1)} \|v_m\|_{L^{p\theta}(B_m)}^{p\theta}, \end{aligned}$$

so using in (7.4)  $\|v_{m+1}\|_{L^{p\theta}(B_{m+1})} \leq \|v_m\|_{L^{p\theta}(B_m)}$  and applying the previous inequality, we obtain

$$\frac{1}{\xi^{p\theta}} \|v_{m+1}\|_{L^{p\theta}(B_{m+1})}^{p\theta} \leq \tilde{C} 2^{p\theta m(2 - \frac{\theta}{\theta_0})} \left( \frac{1}{\xi^{p\theta}} \|v_m\|_{L^{p\theta}(B_m)}^{p\theta} \right)^{2 - \frac{\theta}{\theta_0}}.$$

Denote  $z_m := \frac{1}{\xi^{p\theta}} \|v_m\|_{L^{p\theta}(B_m)}^{p\theta}$ . Then

$$z_{m+1} \leq \tilde{C} \gamma^m z_m^{1+\alpha}, \quad m = 0, 1, 2, \dots, \quad \alpha := 1 - \frac{\theta}{\theta_0}, \quad \gamma := 2^{p\theta(2 - \frac{\theta}{\theta_0})}$$

and  $z_0 = \frac{1}{\xi^{p\theta}} \langle u_+^{p\theta} \mathbf{1}_{B_m} \rangle \leq \tilde{C}^{-\frac{1}{\alpha}} \gamma^{-\frac{1}{\alpha^2}}$  provided that we fix  $c$  by

$$\xi^{p\theta} := \tilde{C}^{\frac{1}{\alpha}} \gamma^{\frac{1}{\alpha^2}} \langle u_+^{p\theta} \mathbf{1}_{B_1} \rangle + \lambda^{-p\beta\theta} H^\theta.$$

Hence, by Lemma 7,  $z_m \rightarrow 0$  as  $m \rightarrow \infty$ . It follows that  $\sup_{B_{1/2}} u_+ \leq \xi$ , and the claimed inequality follows.  $\square$

To end the proof of Theorem 9, we need to estimate  $\langle u_+^{p\theta} \mathbf{1}_{B_1} \rangle^{1/p\theta}$  in the RHS of (7.1) in terms of  $\mathbf{h}$  and  $f$ . We will do it by estimating  $\langle u_+^{p\theta} \rho \rangle^{1/p\theta}$ , where, recall,  $\rho(x) = (1 + k|x|^2)^{-\frac{d}{2}-1}$ , and then applying inequality  $\rho \geq c \mathbf{1}_{B_1}$  for appropriate constant  $c = c_d$ .

**Proposition 6.** *There exist generic constants  $C$ ,  $k$  and  $\lambda_0 > 0$  such that for all  $\lambda \geq \lambda_0$ ,*

$$(\lambda - \lambda_0) \langle u^p \rho \rangle + \langle |\nabla u^{\frac{p}{2}}|^2 \rho \rangle \leq C \langle (\mathbf{1}_{|\mathbf{h}|>1} + |\mathbf{h}|^p \mathbf{1}_{|\mathbf{h}|\leq 1}) |f|^p \rho \rangle. \tag{7.5}$$

*Proof.* Let  $b$  satisfy condition (2.42). We may assume without loss of generality that  $p > \frac{2}{2-\delta_+}$  is rational with odd denominator. We multiply equation (2.43) by  $u^{p-1} \rho$  and integrate to obtain

$$\lambda \langle u^p \rho \rangle + \frac{4(p-1)}{p^2} \langle \nabla u^{\frac{p}{2}}, (\nabla u^{\frac{p}{2}}) \rho \rangle + \frac{2}{p} \langle \nabla u^{\frac{p}{2}}, u^{\frac{p}{2}} \nabla \rho \rangle + \frac{2}{p} \langle b \cdot \nabla u^{\frac{p}{2}}, u^{\frac{p}{2}} \rho \rangle = \langle |\mathbf{h}| f, u^{p-1} \rho \rangle.$$

Now we argue as in the proof of Proposition 1, but instead of the iterations we use a straightforward estimate  $|\nabla\rho| \leq (\frac{d}{2} + 1)\sqrt{k}\rho$  in order to get rid of  $\nabla\rho$  in the previous identity. We arrive at

$$\begin{aligned} p\lambda\langle u^p\rho\rangle &+ \left(\frac{4(p-1)}{p} - \frac{4}{p}\varepsilon\right)\langle|\nabla u^{\frac{p}{2}}|^2\rho\rangle \\ &\leq \frac{p}{4\varepsilon}\left(\frac{d}{2} + 1\right)^2 k\langle v^p\rho\rangle + \left(\frac{d}{2} + 1\right)\sqrt{k}\langle|b|u^p\rho\rangle + \langle\operatorname{div} b_+, u^p\rho\rangle \quad (\varepsilon, \varepsilon' > 0) \\ &+ p\left(2\varepsilon'\chi\langle|\nabla u^{\frac{p}{2}}|^2\rho\rangle + 2\varepsilon'\chi\left(\frac{d}{2} + 1\right)^2 k\langle u^p\rho\rangle + \varepsilon'(c_\chi + 1)\langle u^p\rho\rangle + \frac{1}{4\varepsilon'}\langle(\mathbf{1}_{|h|>1} + |h|^p\mathbf{1}_{|h|\leq 1})|f|^p\rho\rangle\right). \end{aligned}$$

The terms  $\langle|b|u^p\rho\rangle$ ,  $\langle(\operatorname{div} b)_+, u^p\rho\rangle$  are estimated by applying quadratic inequality and using condition (2.42). Selecting  $\varepsilon$ ,  $\varepsilon'$ ,  $k$  sufficiently small, we arrive at the sought inequality.

If  $b$  satisfies (2.41), then the proof is similar but easier (i.e. we do not need to integrate by parts, only apply quadratic inequality to  $\langle b \cdot \nabla u^{\frac{p}{2}}, u^{\frac{p}{2}}\rho\rangle$  and use form-boundedness of  $b$ ).  $\square$

*Proof of Theorem 9, completed.* By Proposition 5, for all  $\lambda \geq 1$ ,

$$\sup_{y \in B_{\frac{1}{2}}(x)} |u(y)| \leq K \left( \langle|u|^{p\theta}\rho_x\rangle^{\frac{1}{p\theta}} + \lambda^{-\beta}\langle(\mathbf{1}_{|h|>1} + |h|^{p\theta'}\mathbf{1}_{|h|\leq 1})|f|^{p\theta'}\rho_x\rangle^{\frac{1}{p\theta'}} \right),$$

where  $\rho_x(y) := \rho(y - x)$ , and constant  $C$  is generic, so

$$\|u\|_\infty \leq K \sup_{x \in \frac{1}{2}\mathbb{Z}^d} \left( \langle|u|^{p\theta}\rho_x\rangle^{\frac{1}{p\theta}} + \lambda^{-\beta}\langle(\mathbf{1}_{|h|>1} + |h|^{p\theta'}\mathbf{1}_{|h|\leq 1})|f|^{p\theta'}\rho_x\rangle^{\frac{1}{p\theta'}} \right).$$

Applying Proposition 6 to the first term in the RHS (with  $p\theta$  instead of  $p$ ), we obtain for all  $\lambda \geq \lambda_0 \vee 1$

$$\begin{aligned} \|u\|_\infty &\leq C \sup_{x \in \frac{1}{2}\mathbb{Z}^d} \left( (\lambda - \lambda_0)^{-\frac{1}{p\theta}} \langle(\mathbf{1}_{|h|>1} + |h|^{p\theta}\mathbf{1}_{|h|\leq 1})|f|^{p\theta}\rho_x\rangle^{\frac{1}{p\theta}} \right. \\ &\quad \left. + \lambda^{-\beta}\langle(\mathbf{1}_{|h|>1} + |h|^{p\theta'}\mathbf{1}_{|h|\leq 1})|f|^{p\theta'}\rho_x\rangle^{\frac{1}{p\theta'}} \right). \end{aligned}$$

This ends the proof of Theorem 9.  $\square$

## 8. PROOF OF THEOREM 5

(i) By the assumption of the theorem, a Borel measurable vector field  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfies either

$$b \in \mathbf{F}_\delta \quad \text{with } \delta < 4 \tag{A_1}$$

or

$$\left\{ \begin{array}{l} b \in \mathbf{MF}_\delta \text{ for some } \delta < \infty, \\ (\operatorname{div} b)_- \in L^1 + L^\infty, \\ (\operatorname{div} b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+} \text{ with } \delta_+ < 4, \\ |b|^{\frac{1+\alpha}{2}} \in \mathbf{F}_\chi \quad \text{for some } \alpha > 0 \text{ fixed arbitrarily small, and some } \chi < \infty. \end{array} \right. \tag{A_2}$$

We define a regularization of  $b$  as in Section 3:

$$b_\varepsilon := E_\varepsilon b, \quad \varepsilon \downarrow 0,$$

where  $E_\varepsilon$  is the Friedrichs mollifier. Then, recall,  $\{b_\varepsilon\}$  are bounded and smooth, preserve all form-bounds in  $(\mathbb{A}_1)$  or in  $(\mathbb{A}_2)$ , and converge to  $b$  in  $[L_{\text{loc}}^2]^d$  or in  $[L_{\text{loc}}^1]^d$ , respectively.

Step 1. By the classical theory, for every  $\varepsilon > 0$ , there exist unique strong solution  $Y_\varepsilon$  to SDE

$$Y_\varepsilon(t) = y - \int_0^t b_\varepsilon(Y_\varepsilon(s)) ds + \sqrt{2}B(t), \quad y \in \mathbb{R}^d,$$

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

Fix  $T > 0$ .

**Lemma 8.** *Let vector field  $\mathbf{g} \in [C_b(\mathbb{R}^d)]^d$  be such that:*

1. *If  $b$  satisfies condition  $(\mathbb{A}_2)$ , then*

$$\langle |\mathbf{g}|^{1+\alpha} \varphi, \varphi \rangle \leq \chi \|\nabla \varphi\|_2^2 + c_\chi \|\varphi\|_2^2, \quad \varphi \in W^{1,2}, \quad (8.1)$$

where constants  $\chi, c_\chi$  are from condition  $(\mathbb{A}_2)$ .

2. *If  $b$  satisfies condition  $(\mathbb{A}_1)$ , then (8.1) holds with  $\alpha = 1$ ,  $\chi = \delta$  and  $c_\chi = c_\delta$ .*

Fix  $\gamma > 0$  by  $1 + \alpha = (1 + \gamma)^2$ . Then

$$\mathbf{E} \int_{t_0}^{t_1} |\mathbf{g}(Y_\varepsilon(s))| ds \leq C_2 (t_1 - t_0)^{\frac{\gamma}{1+\gamma}}, \quad (8.2)$$

where constant  $C_2$  does not depend on  $\varepsilon, y$  or  $t_0, t_1$  (but it depends, by Theorem 9, on constants  $\chi, c_\chi$ ).

(We will be applying (8.2) with  $\mathbf{g} = b_\varepsilon$ .)

*Proof of Lemma 8.* First, let  $\mathbf{g} \in [C_c(\mathbb{R}^d)]^d$ . By Hölder's inequality,

$$\begin{aligned} \mathbf{E} \int_{t_0}^{t_1} |\mathbf{g}(Y_\varepsilon(s))| ds &= \mathbf{E} \int_{t_0}^{t_1} e^{\lambda t} e^{-\lambda t} |\mathbf{g}(Y_\varepsilon(s))| ds \\ &\leq e^{\lambda T} (t_1 - t_0)^{\frac{\gamma}{1+\gamma}} \left( \mathbf{E} \int_0^\infty e^{-(1+\gamma)\lambda t} |\mathbf{g}(Y_\varepsilon(s))|^{1+\gamma} ds \right)^{\frac{1}{1+\gamma}} \\ &= e^{\lambda T} (t_1 - t_0)^{\frac{\gamma}{1+\gamma}} u_\varepsilon(x)^{\frac{1}{1+\gamma}} \end{aligned} \quad (8.3)$$

where  $u_\varepsilon$  is the classical solution to non-homogeneous elliptic equation

$$[(1 + \gamma)\lambda - \Delta + b_\varepsilon \cdot \nabla] u_\varepsilon = |\mathbf{g}|^{1+\gamma}.$$

Note that, in view of the results of Section 3, condition  $(\mathbb{A}_2)$  implies the second condition  $(\bar{\mathbb{A}}_2)$  on  $b$  of Theorem 8 for  $b_\varepsilon$ . (If  $b$  satisfies condition  $(\mathbb{A}_1)$ , then  $b_\varepsilon$  satisfy the same condition in Theorem 8.) Further, we take in Theorem 9  $\mathbf{h} := |\mathbf{g}|^\gamma$  and  $f = 1$  in a neighbourhood of the support of  $\mathbf{g}$ . In view of  $1 + \alpha = (1 + \gamma)^2$  and (8.1),  $\mathbf{h}$  satisfies condition  $|\mathbf{h}|^{\frac{1+\gamma}{2}} \in \mathbf{F}_\chi$  of Theorem 9. Thus, Theorem 9 applies and yields

$$\|u_\varepsilon\|_\infty \leq C \sup_{x \in \frac{1}{2}\mathbb{Z}^d} \left( \langle (\mathbf{1}_{|\mathbf{g}|>1} + |\mathbf{g}|^{(1+\gamma)p\theta} \mathbf{1}_{|\mathbf{g}|\leq 1}) \rho_x \rangle^{\frac{1}{p\theta}} + \langle (\mathbf{1}_{|\mathbf{g}|>1} + |\mathbf{g}|^{(1+\gamma)p\theta'} \mathbf{1}_{|\mathbf{g}|\leq 1}) \rho_x \rangle^{\frac{1}{p\theta'}} \right), \quad (8.4)$$

where the right-hand side is finite (by our choice of  $\rho$ ) and clearly does not depend on  $\varepsilon$ . It is seen now that (8.2) follows from (8.3). Using Fatou's lemma, we can replace the requirement that  $\mathbf{g}$  has compact support by  $\mathbf{g} \in [C_b(\mathbb{R}^d)]^d$ .  $\square$

Inequality (8.2) yields, upon taking  $\mathbf{g} := b_\varepsilon$ ,

$$\mathbf{E} \int_{t_0}^{t_1} |b_\varepsilon(Y_\varepsilon(s))| ds \leq C_2(t_1 - t_0)^{\frac{\gamma}{1+\gamma}} \quad (8.5)$$

(note that  $|b_\varepsilon|^{1+\gamma}$  have independent of  $\varepsilon$  finite form-bound  $\chi$  and constant  $c_\chi$ , see Lemma 6). This gives us the next lemma. We will write  $Y_\varepsilon^y$  to emphasize the dependence of solution  $Y_\varepsilon$  on  $y$ .

**Lemma 9.** (i) For every  $\beta > 0$ ,

$$\sup_{\varepsilon > 0} \sup_{y \in \mathbb{R}^d} \mathbf{P} \left[ \sup_{t \in [0,1], \sigma' \in [0,\sigma]} |Y_\varepsilon^y(t + \sigma') - Y_\varepsilon^y(t)| > \beta \right] \leq \hat{C}H(\sigma), \quad (8.6)$$

where constant  $\hat{C}$  and function  $H$  are independent of  $\varepsilon$ , and  $H(\sigma) \downarrow 0$  as  $\sigma \downarrow 0$ .

(ii) For every  $y \in \mathbb{R}^d$ , the family of probability measures

$$\mathbb{P}_x^\varepsilon := (\mathbf{P} \circ Y_\varepsilon^y)^{-1}, \quad \varepsilon > 0,$$

is tight on the canonical space of continuous trajectories on  $[0, T]$ .

*Proof of Lemma 9.* The argument is standard. For reader's convenience, we include it below (we repeat more or less verbatim a part of [28]). Put for brevity  $T = 1$ . We have, for a stopping time  $0 \leq \tau \leq 1$ ,

$$Y_\varepsilon^y(\tau + \sigma) - Y_\varepsilon^y(\tau) = \int_\tau^{\tau+\sigma} b_n(s, Y_\varepsilon^y(s)) ds + \sqrt{2}(B(\tau + \sigma) - B(\tau)), \quad 0 < \sigma < 1. \quad (8.7)$$

Next, note that (8.5) yields

$$\mathbf{E} \int_\tau^{\tau+\sigma} |b_n(s, Y_\varepsilon^y(s))| ds \leq C_0 \sigma^\mu, \quad (8.8)$$

see Remark 1.2 in [51] (to show that (8.5)  $\Rightarrow$  (8.8), the authors of [51] use a decreasing sequence of stopping times  $\tau_m$  converging to  $\tau$  and taking values in  $S = \{k2^{-m} \mid k \in \{0, 1, 2, \dots\}\}$ , and note that the proof of estimate (8.8) with  $\tau_m$  in place of  $\tau$  can be reduced to applying (8.5) on intervals  $[t_0, t_1] := [c, c + \sigma]$ ,  $c \in S$ .) Thus, applying (8.8) in (8.7), one obtains

$$\mathbf{E} \sup_{\sigma' \in [0, \sigma]} |Y_\varepsilon^y(\tau + \sigma') - Y_\varepsilon^y(\tau)| \leq C_0 \sigma^{\frac{\gamma}{\gamma+1}} + C_1 \sigma^{\frac{1}{2}} =: H(\sigma).$$

Now, applying [52, Lemma 2.7], we obtain: there exists constant  $\hat{C}$  independent of  $\varepsilon$  such that

$$\sup_\varepsilon \sup_{x \in \mathbb{R}^d} \mathbf{E} \left[ \sup_{t \in [0,1], \sigma' \in [0,\sigma]} |Y_\varepsilon^y(t + \sigma') - Y_\varepsilon^y(t)|^{\frac{1}{2}} \right] \leq \hat{C}H(\sigma). \quad (8.9)$$

Applying Chebyshev's inequality in (8.9), since  $H(\sigma) \downarrow 0$  as  $\sigma \downarrow 0$ , we obtain the first assertion of the lemma. The second assertion follows from the first one, see [46, Theorem 1.3.2].  $\square$

Fix  $y \in \mathbb{R}^d$ . Let  $\mathbb{P}_y$  be a weak subsequential limit point of  $\{\mathbb{P}_x^\varepsilon\}$ ,

$$\mathbb{P}_y^{\varepsilon_k} \rightarrow \mathbb{P}_y \text{ weakly} \quad \text{for some } \varepsilon_k \downarrow 0. \quad (8.10)$$

Let us rewrite (8.2) as

$$\mathbb{E}_y^\varepsilon \int_{t_0}^{t_1} |\mathbf{g}(\omega_s)| ds \leq C_2(t_1 - t_0)^{\frac{\gamma}{1+\gamma}}.$$

Taking  $\mathbf{g} := b_{\varepsilon_m}$  and then applying (8.10), we obtain  $\mathbb{E}_y \int_{t_0}^{t_1} |b_{\varepsilon_m}(\omega_s)| ds \leq C_2(t_1 - t_0)^{\frac{\gamma}{1+\gamma}}$ , and hence, using e.g. Fatou's lemma,  $\mathbb{E}_y \int_{t_0}^{t_1} |b(\omega_s)| ds \leq C_2(t_1 - t_0)^{\frac{\gamma}{1+\gamma}} < \infty$ .

Step 2. Let us show that, for any fixed  $y \in \mathbb{R}^d$ , any subsequential limit point  $\mathbb{P}_y$  of  $\{\mathbb{P}_y^\varepsilon\}$  (say, (8.10) holds) is a solution to the martingale problem for SDE (2.30).

It suffices to show that  $\mathbb{E}_y[M_{t_1}^\varphi G] = \mathbb{E}_x[M_{t_0}^\varphi G]$  for every  $\mathcal{B}_{t_0}$ -measurable  $G \in C_b(C([0, T], \mathbb{R}^d))$ . We will do this by passing to the limit in  $k$  in

$$\mathbb{E}_y^{\varepsilon_k}[M_{t_1}^{\varphi, \varepsilon_k} G] = \mathbb{E}_y^{\varepsilon_k}[M_{t_0}^{\varphi, \varepsilon_k} G],$$

where

$$M_t^{\varphi, \varepsilon} = \varphi(\omega_t) - \varphi(\omega_0) + \int_0^t (-\Delta\varphi + b_\varepsilon \cdot \nabla\varphi)(\omega_s) ds, \quad \varphi \in C_c^2.$$

That is, we need to prove

$$\lim_k \mathbb{E}_y^{\varepsilon_k} \int_0^t (b_{\varepsilon_k} \cdot \nabla\varphi)(\omega_s) G(\omega) ds = \mathbb{E}_y \int_0^t (b \cdot \nabla\varphi)(\omega_s) G(\omega) ds, \quad (8.11)$$

*Proof of (8.11).* First, let us note that repeating the proof of (8.2), but this time selecting  $\mathbf{h} := \mathbf{g}|\mathbf{g}|^\gamma$ ,  $\mathbf{g} := b_{\varepsilon_{m_1}} - b_{\varepsilon_{m_2}}$ ,  $f := |\nabla\varphi|$ , we have

$$\begin{aligned} & \mathbb{E}_y^\varepsilon \int_{t_0}^{t_1} |b_{\varepsilon_{m_1}}(\omega_s) - b_{\varepsilon_{m_2}}(\omega_s)| |\nabla\varphi(\omega_s)| ds \\ & \leq C_3 \sup_{x \in \frac{1}{2}\mathbb{Z}^d} \left( \langle (\mathbf{1}_{|\mathbf{g}|>1} + |\mathbf{g}|^{(1+\gamma)p\theta} \mathbf{1}_{|\mathbf{g}|\leq 1}) |\nabla\varphi|^{p\theta} \rho_x \rangle^{\frac{1}{p\theta}} + \langle (\mathbf{1}_{|\mathbf{g}|>1} + |\mathbf{g}|^{(1+\gamma)p\theta'} \mathbf{1}_{|\mathbf{g}|\leq 1}) |\nabla\varphi|^{p\theta'} \rho_x \rangle^{\frac{1}{p\theta'}} \right)^{\frac{1}{1+\gamma}}, \end{aligned}$$

Since  $\varphi$  has compact support, the RHS converges to 0 as  $m_1, m_2 \rightarrow \infty$ . Now, it follows from the weak convergence (8.10) and Fatou's lemma that

$$\begin{aligned} & \mathbb{E}_y \int_{t_0}^{t_1} |b(\omega_s) - b_{\varepsilon_m}(\omega_s)| |\nabla\varphi(\omega_s)| ds \\ & \leq C_3 \sup_{x \in \frac{1}{2}\mathbb{Z}^d} \left( \langle (\mathbf{1}_{|b-b_{\varepsilon_m}|>1} + |b-b_{\varepsilon_m}|^{(1+\gamma)p\theta} \mathbf{1}_{|b-b_{\varepsilon_m}|\leq 1}) |\nabla\varphi|^{p\theta} \rho_x \rangle^{\frac{1}{p\theta}} \right. \\ & \quad \left. + \langle (\mathbf{1}_{|b-b_{\varepsilon_m}|>1} + |b-b_{\varepsilon_m}|^{(1+\gamma)p\theta'} \mathbf{1}_{|b-b_{\varepsilon_m}|\leq 1}) |\nabla\varphi|^{p\theta'} \rho_x \rangle^{\frac{1}{p\theta'}} \right)^{\frac{1}{1+\gamma}}, \end{aligned}$$

where the RHS converges to 0 as  $m \rightarrow \infty$ . We are in position to prove (8.11):

$$\begin{aligned} & \left| \mathbb{E}_y^{\varepsilon_{n_k}} \int_0^t (b_{\varepsilon_{n_k}} \cdot \nabla\varphi)(\omega_s) G(\omega) ds - \mathbb{E}_y \int_0^t (b \cdot \nabla\varphi)(\omega_s) G(\omega) ds \right| \\ & \leq \left| \mathbb{E}_y^{\varepsilon_{n_k}} \int_0^t |b_{\varepsilon_{n_k}} - b_{\varepsilon_m}| |\nabla\varphi|(\omega_s) |G(\omega)| ds \right| \\ & \quad + \left| \mathbb{E}_y^{\varepsilon_{n_k}} \int_0^t (b_{\varepsilon_m} \cdot \nabla\varphi)(\omega_s) G(\omega) ds - \mathbb{E}_y \int_0^t (b_{\varepsilon_m} \cdot \nabla\varphi)(\omega_s) G(\omega) ds \right| \\ & \quad + \left| \mathbb{E}_y \int_0^t |b_{\varepsilon_m} - b| |\nabla\varphi|(\omega_s) |G(\omega)| ds \right|, \end{aligned}$$

where the first and the third terms in the RHS can be made arbitrarily small using the estimates above and the boundedness of  $G$  by selecting  $m$ , and then  $n_k$ , sufficiently large. The second term can be made arbitrarily small in view of (8.10) by selecting  $n_k$  even larger. Thus, (8.11) follows.

Step 3. Let us now find a subsequence  $\varepsilon_k \downarrow 0$  that works for all  $y \in \mathbb{R}^d$  and yields a strong Markov family of probability measures  $\mathbb{P}_y$ ,  $y \in \mathbb{R}^d$ , solutions to the martingale problem for SDE (2.30). Denote  $R_\lambda^\varepsilon f := u_\varepsilon$ , where  $u_\varepsilon$  is the classical solution of  $(\lambda - \Delta + b_\varepsilon \cdot \nabla)u_\varepsilon = f$  in  $\mathbb{R}^d$ ,  $f \in C_c^\infty$ ,  $\lambda \geq \lambda_0 \vee 1$ ;

$$R_\lambda^\varepsilon f(y) = \mathbb{E}_{\mathbb{P}_y^\varepsilon} \int_0^\infty e^{-\lambda s} f(\omega_s) ds.$$

By Theorem 9,  $u_\varepsilon$  are uniformly in  $\varepsilon$  bounded on  $\mathbb{R}^d$ . By Theorem 8 applied to  $b_\varepsilon$ , solutions  $u_\varepsilon$  are Hölder continuous on every compact, also uniformly in  $\varepsilon > 0$ . By the Arzelà-Ascoli theorem and a standard diagonal argument there exists a subsequence  $\varepsilon_k \downarrow 0$  such that sequence  $\{R_\lambda^\varepsilon f\}$  converges locally uniformly on  $\mathbb{R}^d$ , for every  $f$  in a fixed dense subset of  $C_b$ . Let us denote the limit by  $R_\lambda f$ . The latter, and the uniform in  $\varepsilon$  estimate  $\|R_\lambda^\varepsilon f\|_\infty \leq \frac{1}{\lambda} \|f\|_\infty$  allow us to extend  $R_\lambda f$  to all  $f \in C_b$ . Thus,  $R_\lambda f \in C_b$ ,  $f \in C_b$ . Now, for this subsequence  $\varepsilon_k \downarrow 0$ , for any  $y_k \rightarrow y$ , any two subsequential limits  $\mathbb{P}^1, \mathbb{P}^2$  of  $\{\mathbb{P}_{y_k}^\varepsilon\}$  (we use (8.10)) have the same finite-dimensional distributions (see [5] for details, if needed) and therefore coincide:  $\mathbb{P}_y := \mathbb{P}^1 = \mathbb{P}^2$ . Hence  $\mathbb{E}_{\mathbb{P}_y} \int_0^\infty e^{-\lambda s} f(\omega_s) ds = R_\lambda f(y)$ . By what was proved above,  $\mathbb{P}_y$  is a martingale solution of (2.30). A simple argument (see [5]) now gives that, for every  $t > 0$ ,  $y \mapsto \mathbb{E}_{\mathbb{P}_y} f(X_t)$  is a continuous function. The latter, in turn, yields that  $\{\mathbb{P}_y\}_{y \in \mathbb{R}^d}$  is a strong Markov family (the proof can be found e.g. in [5] or [6, Sect. I.3]).

This completes the proof of assertion (i).

(ii) Let  $b_n$  be defined by (2.31), so that vector fields  $\{b_n\}$  do not increase the form-bounds of  $b$ . In the end of the proof of (i) we show that there exists a subsequence  $b_{n_k}$  (for brevity,  $\{b_n\}$  itself) such that, for every  $f \in C_c^\infty(\mathbb{R}^d)$ , the classical solutions  $\{u_n\}$  to elliptic equations

$$(\lambda - \Delta + b_n \cdot \nabla)u_n = f$$

converge locally uniformly on  $\mathbb{R}^{Nd}$  to

$$x \mapsto \mathbb{E}_{\mathbb{P}_x} \int_0^\infty e^{-\lambda s} f(\omega_s^1, \dots, \omega_s^N) ds, \quad x \in \mathbb{R}^{Nd}, \quad x \in \mathbb{R}^{Nd}. \quad (8.12)$$

where  $\lambda$  is assumed to be sufficiently large. This yields the local Hölder continuity of  $u$ . At the same time,  $u_n$  are weak solutions of (2.8) in the sense of Definitions 6 and 8. The possibility to pass to the limit  $\varepsilon \downarrow 0$  in these definitions follows from the standard compactness argument (for details, if needed, see e.g. [34]).

(iii) The proof goes by showing that  $v_n$  constitute a Cauchy sequence in  $L^\infty([0, 1], L^p(\mathbb{R}^d))$ , see [21], see also [29]. At the elliptic level this was done earlier in [36] using Trotter's theorem. The proof of the  $(L^p, L^q)$  estimate is due to [45]. (Strictly speaking, these papers did not consider condition  $(\mathbb{A}_3)$ , but it is easy to modify the proofs there to cover the case  $(\mathbb{A}_3)$  as well.)

(iv) It suffices to show that, for all  $\mu \geq \mu_0$ , for every  $f \in C_c^\infty$ ,

$$R_\mu^\varepsilon f \rightarrow (\mu + \Lambda_p)^{-1} f \quad \text{in } C_\infty \text{ as } \varepsilon \downarrow 0, \quad (8.13)$$

possibly after a modification of  $(\mu + \Lambda_p)^{-1} f$  on a measure zero set. The rest follows from estimates  $\|R_{\mu, \varepsilon} f\|_\infty \leq \mu^{-1} \|f\|_\infty$ ,  $\|(\mu + \Lambda_p)^{-1} f\|_\infty \leq \mu^{-1} \|f\|_\infty$  (an immediate consequence of the fact that the corresponding semigroups are  $L^\infty$  contractions) using a density argument.

Let us prove (8.13). Put  $u_\varepsilon := R_{\mu,\varepsilon}f$ , so  $u_\varepsilon$  is the classical solution to  $(\mu - \Delta + b \cdot \nabla)u_\varepsilon = f$  on  $\mathbb{R}^d$ . Then, by Propositions 5 and 6 (with  $\mathbf{h} = 1$ ), for all  $\mu(\geq 1 \vee \lambda_0) + 1$

$$\sup_{y \in B_{\frac{1}{2}}(x)} |u_\varepsilon(y)| \leq C \left( \langle |f|^{p\theta} \rho_x \rangle^{\frac{1}{p\theta}} + \langle |f|^{p\theta'} \rho_x \rangle^{\frac{1}{p\theta'}} \right)$$

for constant  $C$  independent of  $\varepsilon$ . It is seen now that for a fixed  $f \in C_c^\infty$ , for a given  $\varepsilon > 0$ , we can find  $R > 0$  such that

$$\sup_{y \in \mathbb{R}^d \setminus \bar{B}_R(0)} |u_\varepsilon(y)| < \varepsilon.$$

In turn, inside the closed ball  $\bar{B}_R(0)$ , the family of solutions  $\{u_\varepsilon\}_{\varepsilon > 0}$  is equicontinuous by Theorem 8. So, applying Arzelà-Ascoli theorem and using the convergence result for the semigroups in  $L^p$  from assertion (iii), we obtain (8.13).  $\square$

## 9. PROOF OF THEOREM 6

The proof is an application of Proposition 1 and Gehring's lemma:

**Lemma 10.** *Assume that there exist constants  $K \geq 1$ ,  $1 < q < \infty$  such that, for given  $0 \leq g \in L^q$ ,  $0 \leq h \in L^q \cap L^\infty$  we have*

$$\left( \frac{1}{|B_R|} \langle g^q \mathbf{1}_{B_R} \rangle \right)^{\frac{1}{q}} \leq \frac{K}{|B_{2R}|} \langle g \mathbf{1}_{B_{2R}} \rangle + \left( \frac{1}{|B_{2R}|} \langle h^q \mathbf{1}_{B_{2R}} \rangle \right)^{\frac{1}{q}}$$

for all  $0 < R < \frac{1}{2}$ . Then  $g \in L^s$  for some  $s > q$  and

$$\left( \frac{1}{|B_R|} \langle g^s \mathbf{1}_{B_R} \rangle \right)^{\frac{1}{s}} \leq C_1 \left( \frac{1}{|B_{2R}|} \langle g^q \mathbf{1}_{B_{2R}} \rangle \right)^{\frac{1}{q}} + C_2 \left( \frac{1}{|B_{2R}|} \langle h^s \mathbf{1}_{B_{2R}} \rangle \right)^{\frac{1}{s}}.$$

Let us prove Theorem 6.

Step 1. Applying Proposition 1 (condition (5.2)) with  $\mathbf{h} = 1$  and  $p = 2$  to  $u_n - (u_n)_{B_{2R}}$  and  $-u_n + (u_n)_{B_{2R}}$ , where  $(u_n)_{B_{2R}} := \frac{1}{|B_{2R}|} \langle u_n \mathbf{1}_{B_{2R}} \rangle$ , we obtain

$$\langle |\nabla u_n|^2 \mathbf{1}_{B_R} \rangle \leq \frac{K_1}{|B_{2R}|^{\frac{2}{d}}} \langle (u_n - (u_n)_{B_{2R}})^2 \mathbf{1}_{B_{2R}} \rangle + K_2 \langle f^2 \mathbf{1}_{B_{2R}} \rangle, \quad 0 < R < \frac{1}{2}.$$

By the Sobolev-Poincaré inequality,

$$\left( \frac{1}{|B_{2R}|} \langle (u_n - (u_n)_{B_{2R}})^2 \mathbf{1}_{B_{2R}} \rangle \right)^{\frac{1}{2}} \leq C |B_R|^{\frac{1}{d}} \left( \frac{1}{|B_{2R}|} \langle |\nabla u_n|^{\frac{2d}{d+2}} \mathbf{1}_{B_{2R}} \rangle \right)^{\frac{d+2}{2d}},$$

i.e.

$$\langle (u_n - (u_n)_{B_{2R}})^2 \mathbf{1}_{B_{2R}} \rangle \leq C^2 |B_R|^{\frac{2}{d}+1} \left( \frac{1}{|B_{2R}|} \langle |\nabla u_n|^{\frac{2d}{d+2}} \mathbf{1}_{B_{2R}} \rangle \right)^{\frac{d+2}{d}},$$

so the condition of the Gehring lemma is verified with  $g = |\nabla u_n|^{\frac{2d}{d+2}}$ ,  $g^q = |\nabla u_n|^2$  (so  $q = \frac{d+2}{d}$ ) and  $h = c|f|^{\frac{2d}{d+2}}$ . Hence there exists  $s > \frac{d+2}{d}$  such that

$$\left( \frac{1}{|B_R|} \langle |\nabla u_n|^{s \frac{2d}{d+2}} \mathbf{1}_{B_R} \rangle \right)^{\frac{1}{s}} \leq C_1 \left( \frac{1}{|B_{2R}|} \langle |\nabla u_n|^2 \mathbf{1}_{B_{2R}} \rangle \right)^{\frac{d}{d+2}} + C_2 \left( \frac{1}{|B_{2R}|} \langle |f|^{s \frac{2d}{d+2}} \mathbf{1}_{B_{2R}} \rangle \right)^{\frac{1}{s}},$$

where all constants are independent of  $n$ .



Now, passing in both sides of the previous inequality to the cubes (inscribed in  $B_R$  in the left-hand side and circumscribed over  $B_{2R}$  in the right-hand side), then considering an equally spaced grid in  $\mathbb{R}^d$  so that the smaller cubes centered at the nodes of the grid cover  $\mathbb{R}^d$ , applying the previous estimate on each cube, and then summing up, we obtain the global estimate

$$\|\nabla u_n\|_{s\frac{2d}{d+2}}^2 \leq C_3 \|\nabla u_n\|_2^2 + C_4 \|f\|_{s\frac{2d}{d+2}}^2.$$

Step 2. Let us show that  $\sup_n \|\nabla u_n\|_2^2 < \infty$ . To this end, we multiply  $(\mu - \Delta + b_n \cdot \nabla)u_n = f$  by  $u_n$  and integrate:  $\mu \|u_n\|_2^2 + \|\nabla u_n\|_2^2 + \langle b_n \cdot \nabla u_n, u_n \rangle = \langle f, u_n \rangle$ , where, after integrating by parts,  $\langle b_n \cdot \nabla u_n, u_n \rangle = -\frac{1}{2} \langle \operatorname{div} b_n, u_n^2 \rangle \geq -\frac{1}{2} \langle (\operatorname{div} b_n)_+, u_n^2 \rangle$ . Hence, by our form-boundedness assumption on  $(\operatorname{div} b_n)_+$ ,

$$\left(\mu - \frac{c_{\delta_+}}{2}\right) \|u_n\|_2^2 + \left(1 - \frac{\delta_+}{2}\right) \|\nabla u_n\|_2^2 \leq \langle f, u_n \rangle. \quad (9.1)$$

So, applying the quadratic inequality in the right-hand side, we arrive at  $\left(\mu - \frac{c_{\delta_+}}{2} - \frac{1}{2}\right) \|u_n\|_2^2 + \left(1 - \frac{\delta_+}{2}\right) \|\nabla u_n\|_2^2 \leq \frac{1}{2} \|f\|_2^2$ . Since  $\delta_+ < 2$ ,  $\sup_n \|\nabla u_n\|_2^2 < \infty$  for  $\mu \geq \mu_0 := \frac{c_{\delta_+}}{2} + \frac{1}{2}$ .

It follows from Steps 1 and 2 that  $\sup_n \|\nabla u_n\|_{s\frac{2d}{d+2}}^2 < \infty$ .

Step 3. Put  $h := u_n - u_m$ . Then

$$\mu \|h\|_2^2 + \|\nabla h\|_2^2 + \langle b_n \cdot \nabla h, h \rangle + \langle (b_n - b_m) \cdot \nabla u_m, h \rangle = 0.$$

So,

$$\left(\mu - \frac{c_{\delta_+}}{2} - \frac{1}{2}\right) \|u_n\|_2^2 + \left(1 - \frac{\delta_+}{2}\right) \|\nabla u_n\|_2^2 \leq |\langle (b_n - b_m) \cdot \nabla u_m, h \rangle|. \quad (9.2)$$

In turn, the right-hand side

$$|\langle (b_n - b_m) \cdot \nabla u_m, h \rangle| \leq \|b_n - b_m\|_{2-\varkappa} \|\nabla u_m\|_{s\frac{2d}{d+2}} \|f\|_\infty$$

where  $0 < \varkappa < 1$  is defined by

$$2 - \varkappa := \left(s \frac{2d}{d+2}\right)' = \frac{s \frac{2d}{d+2}}{s \frac{2d}{d+2} - 1}$$

(recall that  $s \frac{2d}{d+2} > 2$ ). Since  $\sup_m \|\nabla u_m\|_{s\frac{2d}{d+2}} < \infty$  and, by our assumption,  $\{b_n\}$  converge in  $L^{2-\varkappa}$ , we obtain that the RHS of (9.2) converges to zero as  $n, m \rightarrow \infty$ , so  $\{u_n\}$  is a Cauchy sequence in  $L^2$ . This yields the uniqueness result (since we can always combine two different approximations of  $b$  obtaining again a Cauchy sequence of the approximating solutions).  $\square$

## 10. PROOF OF THEOREMS 1 AND 2

This follows right away, in view of Lemmas 1, 2, from Theorem 5(i), (ii) where we consider the general SDE in  $\mathbb{R}^{Nd}$  with  $Y = (X_1, \dots, X_N)$ ,  $B = (B_1, \dots, B_N)$ ,  $y = (x_1, \dots, x_N)$  and drift  $b : \mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$  defined by (2.13).  $\square$

## 11. PROOF OF THEOREM 3

(i) Since the sum of two form-bounded vector fields is again form-bounded, we only need to improve Lemma 2 for  $K(y) = \sqrt{\kappa} \frac{d-2}{2} |y|^{-2} y$  and then simply repeat the proof of Theorem 2. In Lemma 2 we have three estimates (2.14), (2.15) (2.16) for  $b = (b_1, \dots, b_N) : \mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$ , where now

$$b_i(x) := \sqrt{\kappa} \frac{d-2}{2} \frac{1}{N} \sum_{j=1, j \neq i}^N \frac{x_i - x_j}{|x_i - x_j|^2}, \quad x = (x_1, \dots, x_N) \in \mathbb{R}^{Nd}, \quad 1 \leq i \leq N. \quad (11.1)$$

We do not need to change (2.14) and (2.16) since the actual values of the form-bounds there are not important for the sake of repeating the proof of Theorem 2, only their finiteness matters. The form-bound  $\delta_+$  in (2.15), however, plays a crucial role. Let us estimate it using the multi-particle Hardy's inequality (2.18):

$$\begin{aligned} (\operatorname{div} b)_+ &= \operatorname{div} b = \sum_{i=1}^N \frac{1}{N} \sum_{j=1, j \neq i}^N \operatorname{div} K(x_i - x_j) \\ &= \sqrt{\kappa} \frac{(d-2)^2}{N} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^2}. \end{aligned}$$

Applying (2.18), we obtain that  $(\operatorname{div} b)_+^{\frac{1}{2}} \in \mathbf{F}_{\delta_+}$  with  $\delta_+ = \sqrt{\kappa}$ . Armed with this result, i.e. a replacement of Lemma 2, we repeat the proof of Theorem 2 (i.e. we apply Theorem 5 where we still have  $\delta_+ < 4$ ).

(ii) The assertions of Theorem 4 are also valid for such interaction kernels  $K_{ij}$ . In view of (2.22), Dunford-Pettis' theorem yields that  $e^{-t\Lambda_p}$  is a semigroup of integral operators. Their integral kernel  $e^{-t\Lambda}(x, z)$  does not depend on  $p$  and is defined to be the heat kernel of particle system (2.1).

Now, we apply Theorem 11 from Appendix A. There  $\Omega := \mathbb{R}^d$  and  $\mu$  is the Lebesgue measure on  $\mathbb{R}^d$ . The semigroup  $e^{-t\Lambda}$  is from Theorem 4(i) with  $K(x) = -\sqrt{\kappa} \frac{d-2}{2} |x|^{-2} x$ . The weights  $\{\varphi_s\}_{s>0}$  are defined by

$$\varphi_s(x) := \prod_{1 \leq i < j \leq N} \eta(s^{-\frac{1}{2}} |x_i - x_j|), \quad s > 0,$$

where  $\eta$  is defined in Theorem 3(ii). It is easily seen that these weights  $\varphi_s$  satisfy conditions  $(S_2)$  and  $(S_3)$  of Theorem 11. In turn, condition  $(S_1)$  with  $j = \frac{d}{d-2}$  and  $r > 2(2 - \frac{N-1}{N} \sqrt{\kappa})^{-1}$  was verified in Theorem 4(i) under hypothesis (2.21), see (2.22). Let us verify the ‘‘desingularizing  $L^1 \rightarrow L^1$  bound’’  $(S_4)$  for  $0 < s \leq t$ :

Step 1. Set

$$\eta_s(r) := \eta(s^{-\frac{1}{2}} r), \quad r > 0$$

and put

$$\varphi_s^\varepsilon(x) \equiv \varphi^\varepsilon(x) := \prod_{1 \leq i < j \leq N} \eta_s(|x_i - x_j|_\varepsilon), \quad |x_i - x_j|_\varepsilon := \sqrt{|x_i - x_j|^2 + \varepsilon}, \quad \varepsilon > 0.$$

Define

$$\psi^\varepsilon(x) := \prod_{1 \leq i < j \leq N} (s^{-\frac{1}{2}} |x_i - x_j|_\varepsilon)^{-\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N}}.$$

and put

$$b_\varepsilon := -\frac{\nabla_x \psi^\varepsilon}{\psi^\varepsilon} \quad (\text{clearly, independent of } s).$$

This is a vector field  $\mathbb{R}^{Nd} \rightarrow \mathbb{R}^{Nd}$  such that

$$b_\varepsilon \cdot \nabla_x = \sqrt{\kappa} \frac{d-2}{2} \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{x_i - x_j}{|x_i - x_j|_\varepsilon^2} \cdot \nabla_{x_i}.$$

Without loss of generality, we discuss the (minus) first component  $\mathbb{R}^{Nd} \rightarrow \mathbb{R}^d$  of  $b_\varepsilon$ :

$$\begin{aligned} \frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} &= \frac{1}{\psi^\varepsilon} \sum_{2 \leq k \leq N} \prod_{1 \leq i < j \leq N, i \neq k} |x_i - x_j|_\varepsilon^{-\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N}} \nabla_{x_1} \left( |x_1 - x_k|_\varepsilon^{-\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N}} \right) \\ &= \sum_{2 \leq k \leq N} \frac{\nabla_{x_1} |x_1 - x_k|_\varepsilon^{-\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N}}}{|x_1 - x_k|_\varepsilon^{-\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N}}} \\ &= -\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N} \sum_{2 \leq k \leq N} \frac{x_1 - x_k}{|x_1 - x_k|_\varepsilon^2}. \end{aligned}$$

In the same way,

$$\frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} = \sum_{2 \leq k \leq N} \frac{\nabla_{x_1} \eta_s(|x_1 - x_k|_\varepsilon)}{\eta_s(|x_1 - x_k|_\varepsilon)}.$$

We now compare these quantities (this will be needed at the next step):

(a) If  $|x_1 - x_k|_\varepsilon \leq \sqrt{s}$  for all  $2 \leq k \leq N$ , then, by the definition of  $\eta$ ,

$$-\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} = 0.$$

Therefore,

$$\frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \cdot \left( -\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \right) = 0, \quad \operatorname{div}_{x_1} \left( -\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \right) = 0.$$

(b) If there exists one  $k_0$  such that  $|x_1 - x_{k_0}|_\varepsilon \geq 2\sqrt{s}$ , but for the other  $k \neq k_0$   $|x_1 - x_k|_\varepsilon \leq \sqrt{s}$ , then, since  $x_1 \mapsto \eta_s(|x_1 - x_{k_0}|_\varepsilon)$  is constant and so  $\nabla_{x_1} \varphi^\varepsilon = 0$ , we have

$$\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} - \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} = -\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N} \frac{x_1 - x_{k_0}}{|x_1 - x_{k_0}|_\varepsilon^2}.$$

Hence

$$\frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \cdot \left( -\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \right) = 0, \quad \left| \operatorname{div}_{x_1} \left( -\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \right) \right| \leq \sqrt{\kappa} \frac{(d-2)^2}{2} \frac{1}{N} 4s^{-1}.$$

(c) More generally, if there exist  $2 \leq M \leq N-1$  indices  $k_0$  such that  $|x_1 - x_{k_0}|_\varepsilon \geq 2\sqrt{s}$ , but for the other  $k \neq k_0$   $|x_1 - x_k|_\varepsilon \leq \sqrt{s}$ , then we have

$$\frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \cdot \left( -\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \right) = 0, \quad \left| \operatorname{div}_{x_1} \left( -\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \right) \right| \leq \sqrt{\kappa} \frac{(d-2)^2}{2} \frac{M}{N} 4s^{-1}.$$

Over the annuli  $\sqrt{s} < |x_1 - x_k|_\varepsilon < 2\sqrt{s}$  we make a change of variable to finally obtain, for all possible values of  $|x_1 - x_k|_\varepsilon$ ,  $2 \leq k \leq N$ ,

$$\left| \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \cdot \left( -\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \right) \right| \leq c_1 \frac{N-1}{N} s^{-1}, \quad \left| \operatorname{div}_{x_1} \left( -\frac{\nabla_{x_1} \psi^\varepsilon}{\psi^\varepsilon} + \frac{\nabla_{x_1} \varphi^\varepsilon}{\varphi^\varepsilon} \right) \right| \leq c_2 \frac{N-1}{N} s^{-1}$$

for constants  $c_1$  and  $c_2$  independent of  $\varepsilon$  and  $s$ .

The same holds for the other components of  $b_\varepsilon = -\frac{\nabla_x \psi_\varepsilon}{\psi_\varepsilon}$ . Thus,

$$\left| \frac{\nabla_x \varphi^\varepsilon}{\varphi^\varepsilon} \cdot \left( b_\varepsilon + \frac{\nabla_x \varphi^\varepsilon}{\varphi^\varepsilon} \right) \right| \leq c_1 \frac{N-1}{\sqrt{N}} s^{-1}, \quad \left| \operatorname{div} \left( b_\varepsilon + \frac{\nabla_x \varphi^\varepsilon}{\varphi^\varepsilon} \right) \right| \leq c_2 \frac{N-1}{\sqrt{N}} s^{-1}. \quad (11.2)$$

Step 2. Define the approximating operators  $\Lambda_\varepsilon := -\Delta_x + b_\varepsilon \cdot \nabla_x$  having domain  $\mathcal{W}^{2,1} = (1 - \Delta)^{-1} L^1$ . Since  $\varphi^\varepsilon, (\varphi^\varepsilon)^{-1}$  are bounded and continuous, one sees right away that  $\varphi^\varepsilon e^{-t\Lambda_\varepsilon} (\varphi^\varepsilon)^{-1}$  is a strongly continuous semigroup in  $L^1$  whose generator coincides with  $-\varphi_\varepsilon \Lambda_\varepsilon (\varphi_\varepsilon)^{-1}$  having domain  $\mathcal{W}^{2,1}$ . This generator can be computed explicitly:

$$\varphi^\varepsilon \Lambda_\varepsilon (\varphi^\varepsilon)^{-1} = -\Delta + \nabla \cdot \left( b_\varepsilon + 2 \frac{\nabla \varphi^\varepsilon}{\varphi^\varepsilon} \right) + W_\varepsilon, \quad (11.3)$$

$$W_\varepsilon := -\frac{\nabla \varphi^\varepsilon}{\varphi^\varepsilon} \cdot \left( b_\varepsilon + \frac{\nabla \varphi^\varepsilon}{\varphi^\varepsilon} \right) - \operatorname{div} \left( b_\varepsilon + \frac{\nabla \varphi^\varepsilon}{\varphi^\varepsilon} \right).$$

By (11.2), potential  $W_\varepsilon$  is (uniformly in  $\varepsilon$ ) bounded:  $|W_\varepsilon| \leq \frac{N-1}{\sqrt{N}} \frac{c}{s}$  for a constant  $c$  independent of  $\varepsilon$ . Employing formula (11.3) and using the general fact that  $e^{t(\Delta - \nabla \cdot f)}$  is an  $L^1$  contraction, we obtain

$$\|\varphi^\varepsilon e^{-t\Lambda_\varepsilon} (\varphi^\varepsilon)^{-1} h\|_1 \leq e^{c \frac{N-1}{\sqrt{N}} \frac{t}{s}} \|h\|_1, \quad h \in L^1. \quad (11.4)$$

It remains to pass to the limit  $\varepsilon \downarrow 0$  in (11.4). This is done at the next step.

Step 3. Define  $b = -\frac{\nabla_x \psi}{\psi}$ , where  $\psi(x) = \prod_{1 \leq i < j \leq N} |x_i - x_j|^{-\sqrt{\kappa} \frac{d-2}{2} \frac{1}{N}}$  is a Lyapunov function of the formal adjoint of  $\Lambda$  (see Remark 2). Then

$$b \cdot \nabla_x = \sqrt{\kappa} \frac{d-2}{2} \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{x_i - x_j}{|x_i - x_j|^2} \cdot \nabla_{x_i}$$

It is seen using e.g. the Monotone convergence theorem that  $b_\varepsilon \rightarrow b$  in  $[L^2_{\text{loc}}]^{Nd}$ . Moreover, the vector fields  $b_\varepsilon$  do not increase the form-bound  $\delta = \kappa \left( \frac{N-1}{N} \right)^2 (< 4)$  of  $b$ . Therefore, by Theorem 5(iii),

$$e^{-t\Lambda_\varepsilon} \rightarrow e^{-t\Lambda} \quad \text{in } L^r(\mathbb{R}^{Nd}), \quad (11.5)$$

where  $r > \frac{2}{2 - \frac{N-1}{N} \sqrt{\kappa}}$ .

Now, from (11.4) we have

$$\|\varphi^\varepsilon e^{-t\Lambda_\varepsilon} g\|_1 \leq e^{c \frac{N-1}{\sqrt{N}} \frac{t}{s}} \|\varphi^\varepsilon g\|_1, \quad g \in \varphi L^1 \cap L^\infty.$$

In view of (11.5) and since  $\varphi^\varepsilon \rightarrow \varphi$  a.e., we can use Fatou's lemma to obtain  $\|\varphi e^{-t\Lambda} g\|_1 \leq e^{c \frac{t}{s}} \|\varphi g\|_1$ , which yields condition  $(S_4)$  of Theorem 11 (recall that by our assumption  $s \geq t$ ).

Thus, Theorem 11 applies and gives assertion (ii) of Theorem 3.  $\square$

## 12. PROOF OF THEOREM 4

(i) follows from Theorem 5(iii) and Lemmas 1, 2.

(ii) follows from the uniqueness result in [26], see also [24], and Lemma 1.

(iii) follows upon applying appropriate (straightforward) modification of Lemma 1.

(iv) follows from the result in [25] upon applying Lemma 1.

## 13. PROOF OF THEOREM 10

We will need the following result on the regularization of the vector field  $b$  in Theorem 10.

**Lemma 11.** *Assume that  $b \in [W_{\text{loc}}^{1,1}(\mathbb{R}^d)]^d$  has symmetric Jacobian  $Db = (\nabla_k b_i)_{k,i=1}^d$  and the negative part  $B_-$  of matrix*

$$B(b) := Db - \frac{\text{div } b}{q} I, \quad \text{for some } q > (d-2) \vee 2,$$

has normalized eigenvectors  $e_j$  and eigenvalues  $\lambda_j \geq 0$  satisfying  $\sqrt{\lambda_j} e_j \in \mathbf{F}_{\nu_j}$ . Set  $\nu := \sum_{j=1}^d \nu_j$ . Set  $b_\varepsilon := E_\varepsilon b$ . The following are true:

1.

$$B(b_\varepsilon) + E_\varepsilon B_- \geq 0,$$

2.

$$\langle B_- h, h \rangle \leq \nu \langle |\nabla| h|^2 \rangle + c_\nu \langle |h|^2 \rangle, \quad (13.1)$$

and

$$\langle (E_\varepsilon B_-) h, h \rangle \leq \nu \langle |\nabla| h|^2 \rangle + c_\nu \langle |h|^2 \rangle, \quad \varepsilon > 0,$$

for all  $h \in [C_c^\infty(\mathbb{R}^d)]^d$ , with  $c_\nu := \sum_{j=1}^d c_{\nu_j}$ .

*Proof.* 1. We have, by definition,  $B(b) = B_+ - B_-$ , and  $B(b_\varepsilon) = E_\varepsilon B_+ - E_\varepsilon B_-$ . Clearly,  $E_\varepsilon B_+ \geq 0$ , which yields the required.

2. We have  $B_- = \sum_{j=1}^d \lambda_j e_j e_j^\top$ . Put for brevity  $\lambda = \lambda_j$  and  $e = e_j$ . Denote the components of  $e$  by  $e^k$ ,  $k = 1, \dots, d$ . Then

$$\langle h \lambda (e e^\top), h \rangle = \sum_{k,i=1}^d \langle h_k \sqrt{\lambda} e^k \sqrt{\lambda} e^i h_i \rangle = \langle \lambda (h \cdot e)^2 \rangle \leq \langle \lambda |h|^2 |e|^2 \rangle.$$

Therefore,

$$\begin{aligned} \langle B_- h, h \rangle &\leq \sum_{j=1}^d \langle \lambda_j |h|^2 |e_j|^2 \rangle \\ &\quad (\text{we use } \sqrt{\lambda_j} e_j \in \mathbf{F}_{\nu_j}) \\ &\leq \sum_{j=1}^d \nu_j \langle |\nabla| h|^2 \rangle + \sum_{j=1}^d c_{\nu_j} \langle |h|^2 \rangle, \end{aligned}$$

which gives us the first inequality in assertion 2.

Let us prove the second inequality in assertion 2. Writing again  $\lambda = \lambda_j$  and  $e = e_j$  and denoting the  $k$ -th component of  $e$  by  $e^k$ , we have

$$\begin{aligned} \langle h E_\varepsilon (\lambda e e^\top), h \rangle &= \sum_{k,i=1}^d \langle h_k E_\varepsilon (\sqrt{\lambda} e^k \sqrt{\lambda} e^i) h_i \rangle = \sum_{k,i=1}^d \langle \sqrt{\lambda} e^k \sqrt{\lambda} e^i E_\varepsilon (h_k h_i) \rangle \\ &\leq \sum_{k,i=1}^d \langle \sqrt{E_\varepsilon |h_k|^2} \sqrt{\lambda} |e^k| \sqrt{\lambda} |e^i| \sqrt{E_\varepsilon |h_i|^2} \rangle \\ &\leq \langle \lambda |e|^2, |h_\varepsilon|^2 \rangle, \end{aligned}$$

where  $h_\varepsilon$  denotes the vector field with  $k$ -th component  $\sqrt{E_\varepsilon|h_k|^2}$ . Hence, using the previous estimate, we obtain

$$\begin{aligned} \langle (E_\varepsilon B_-)h, h \rangle &= \sum_{j=1}^d \langle h E_\varepsilon(\lambda_j e_j e_j^\top), h \rangle \leq \sum_{j=1}^d \langle \lambda_j |e_j|^2, |h_\varepsilon|^2 \rangle \\ &\quad (\text{use } \sqrt{\lambda_j} e_j \in \mathbf{F}_{\nu_j}) \\ &\leq \nu \langle |\nabla |h_\varepsilon||^2 \rangle + c_\nu \langle |h_\varepsilon|^2 \rangle \\ &\quad (\text{note that } |h_\varepsilon| = \sqrt{E_\varepsilon|h|^2} \text{ and apply (3.1)}) \\ &\leq \langle |\nabla |h||^2 \rangle + c_\nu \langle |h|^2 \rangle, \end{aligned}$$

as needed.  $\square$

*Proof of Theorem 10 in the case drift  $b$  satisfies condition  $(\mathbb{B}_2)$ .* We start with the proof of assertion (ii). Put

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

Multiplying equation  $(\mu - \Delta + b \cdot \nabla)u = f$  by the test function

$$\phi := - \sum_{i=1}^d \nabla_i (w_i |w|^{q-2}) = -\nabla \cdot (w |w|^{q-2})$$

and integrating by parts, we obtain

$$\mu \langle |w|^q \rangle + I_q + (q-2)J_q + \langle b \cdot w, \phi \rangle = \langle f, \phi \rangle, \quad (13.2)$$

where

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

Step 1. Regarding term  $\langle b \cdot w, \phi \rangle$  in (13.2), we have

$$\begin{aligned} \langle b \cdot w, \phi \rangle &= \langle \tilde{B}w, w |w|^{q-2} \rangle + \langle b \cdot \nabla |w|, |w|^{q-1} \rangle \quad \tilde{B} := (\nabla_k b_i)_{k,i=1}^d \\ &= \langle \tilde{B}w, w |w|^{q-2} \rangle - \frac{1}{q} \langle \operatorname{div} b, |w|^q \rangle \\ &\geq -\langle B_- w, w |w|^{q-2} \rangle. \end{aligned}$$

Hence, applying (13.1), we arrive at

$$\begin{aligned} \langle b \cdot w, \phi \rangle &\geq -\nu \langle |\nabla |w|^{\frac{q}{2}}|^2 \rangle - c \langle |w|^q \rangle \\ &= -\nu \frac{q^2}{4} J_q - c_\nu \langle |w|^q \rangle, \end{aligned}$$

so (13.2) yields

$$(\mu - c_\nu) \langle |w|^q \rangle + I_q + \left( q - 2 - \frac{q^2}{4} \nu \right) J_q \leq \langle f, \phi \rangle. \quad (13.3)$$

Step 2. Let us estimate  $\langle f, \phi \rangle$  in the previous inequality. To this end, we evaluate  $\phi$ :

$$\langle f, \phi \rangle = -\langle f, |w|^{q-2} \Delta u \rangle - (q-2) \langle f, |w|^{q-3} w \cdot \nabla |w| \rangle. \quad (13.4)$$

(a) We estimate

$$|\langle f, |w|^{q-2} \Delta u \rangle| \leq \varepsilon_0 \langle |w|^{q-2} |\Delta u|^2 \rangle + \frac{1}{4\varepsilon_0} \langle f^2, |w|^{q-2} \rangle, \quad (13.5)$$

where  $\varepsilon_0 > 0$  will be chosen sufficiently small.

Let us deal with the first term in the RHS of (13.5). Representing  $|\Delta u|^2 = |\nabla \cdot w|^2$  and integrating by parts twice, we obtain

$$\begin{aligned} \langle |w|^{q-2} |\Delta u|^2 \rangle &= -\langle \nabla |w|^{q-2} \cdot w, \Delta u \rangle + \sum_{i=1}^d \langle w_i \nabla |w|^{q-2}, \nabla w_i \rangle + I_q \\ &\leq (q-2) \left[ \frac{1}{4\kappa} \langle |w|^{q-2} |\Delta v|^2 \rangle + \kappa J_q \right] + (q-2) \left( \frac{1}{2} I_q + \frac{1}{2} J_q \right) + I_q. \end{aligned}$$

So, for any fixed  $\kappa > \frac{q-2}{4}$ ,

$$\left( 1 - \frac{q-2}{4\kappa} \right) \langle |w|^{q-2} |\Delta v|^2 \rangle \leq I_q + (q-2) \left( \kappa J_q + \frac{1}{2} I_q + \frac{1}{2} J_q \right). \quad (13.6)$$

Let us handle the second term in the RHS of (13.5):

$$\begin{aligned} \langle f^2, |w|^{q-2} \rangle &\leq \|f\|_{\frac{qd}{d+q-2}}^2 \|w\|_{\frac{qd}{d-2}}^{q-2} \\ &\leq c_S \|f\|_{\frac{qd}{d+q-2}}^2 \|\nabla |w|^{\frac{q}{2}}\|_2^{2\frac{(q-2)}{q}} = C \|f\|_{\frac{qd}{d+q-2}}^2 J_q^{\frac{q-2}{q}}, \quad C = \frac{4c}{q^2} \\ &\leq \frac{q-2}{q} C \varepsilon^{\frac{q}{q-2}} J_q + \frac{2}{q} C \varepsilon^{-\frac{q}{2}} \|f\|_{\frac{qd}{d+q-2}}^q. \end{aligned}$$

(b) We estimate

$$\begin{aligned} (q-2) |\langle -f, |w|^{q-3} w \cdot \nabla |w| \rangle| &\leq (q-2) J_q^{\frac{1}{2}} \langle f^2, |w|^{q-2} \rangle^{\frac{1}{2}} \\ &\leq (q-2) (\varepsilon_1 J_q + 4\varepsilon_1^{-1} \langle f^2, |w|^{q-2} \rangle), \end{aligned}$$

where we estimate the very last term in the same way as above.

Substituting the above estimates in (13.4), we obtain

$$|\langle f, \phi \rangle| \leq c\varepsilon_0 I_q + \frac{c_1(\varepsilon, \varepsilon_1)}{\varepsilon_0} J_q + \frac{c_2(\varepsilon, \varepsilon_1)}{\varepsilon_0} \|f\|_{\frac{qd}{d+q-2}}^q, \quad (13.7)$$

where  $c_1(\varepsilon, \varepsilon_1) > 0$  can be made as small as needed by first selecting  $\varepsilon_1$  sufficiently small, and then selecting  $\varepsilon$  even smaller.

Step 3. Now, we return to (13.3). By the pointwise inequality  $|\nabla |w|^2| \leq \sum_{i=1}^d |\nabla w_i|^2$ , we have

$$J_q \leq I_q.$$

The latter, and (13.7) with  $\varepsilon_0$  chosen sufficiently small, yield

$$(\mu - c_\nu) \langle |w|^q \rangle + \left( q-1 - \frac{q^2}{4} \nu - c(\varepsilon_0, \varepsilon, \varepsilon_1) \right) J_q \leq C(\varepsilon_0, \varepsilon, \varepsilon_1) \|f\|_{\frac{qd}{d+q-2}}^q, \quad (13.8)$$

where constant  $c(\varepsilon_0, \varepsilon, \varepsilon_1)$  can be made as small as needed by first selecting  $\varepsilon_1$  sufficiently small, and then selecting  $\varepsilon$  even smaller. Take  $\mu_0 := c_\nu$ . The required gradient estimate now follows from (13.8).

*Proof of assertion (i).* Steps 1 and 3 do not change. Step 2 now consists of estimating  $\langle |g|f, \phi \rangle$ , which we represent as

$$\langle |g|f, \phi \rangle = -\langle |g|f, |w|^{q-2}\Delta u \rangle - (q-2)\langle |g|f, |w|^{q-3}w \cdot \nabla |w| \rangle.$$

(a') We have

$$|\langle |g|f, |w|^{q-2}\Delta u \rangle| \leq \varepsilon_0 \langle |w|^{q-2}|\Delta u|^2 \rangle + \frac{1}{4\varepsilon_0} \langle |g|^2 f^2, |w|^{q-2} \rangle, \quad \varepsilon_0 > 0,$$

where  $\langle |w|^{q-2}|\Delta u|^2 \rangle$  is estimates in the same way as in (a) above, and

$$\begin{aligned} \langle |g|^2 f^2, |w|^{q-2} \rangle &= \langle |g|^{2-\frac{4}{q}} |w|^{q-2}, |g|^{\frac{4}{q}} f^2 \rangle \\ &\leq \frac{q-2}{q} \varepsilon^{\frac{q}{q-2}} \langle |g|^2 |w|^q \rangle + \frac{2}{q} \varepsilon^{-\frac{q}{2}} \langle \rho |g|^2 f^q \rangle \\ &\text{(we are using } g \in \mathbf{F}_{\delta_1} \text{)} \\ &\leq \frac{q-2}{q} \varepsilon^{\frac{q}{q-2}} \left[ \delta_1 \frac{q^2}{4} J_q + c_{\delta_1} \langle |w|^q \rangle \right] + \frac{2}{q} \varepsilon^{-\frac{q}{2}} \langle |g|^2 f^q \rangle. \end{aligned}$$

(b') We estimate

$$\begin{aligned} (q-2) |\langle |g|f, |w|^{q-3}w \cdot \nabla |w| \rangle| &\leq (q-2) J_q^{\frac{1}{2}} \langle |g|^2 f^2, |w|^{q-2} \rangle^{\frac{1}{2}} \\ &\leq (q-2) (\varepsilon_1 J_q + 4\varepsilon_1^{-1} \langle |g|^2 f^2, |w|^{q-2} \rangle), \end{aligned}$$

where we bound  $\langle |g|^2 f^2, |w|^{q-2} \rangle$  as in (a').

Now, arguing as above, we arrive at

$$(\mu - c_\nu - c_0(\varepsilon_0, \varepsilon_1, \varepsilon)) \langle |w|^q \rangle + (q-1 - \frac{q^2}{4}\nu - c(\varepsilon_0, \varepsilon, \varepsilon_1)) J_q \leq C(\varepsilon_0, \varepsilon, \varepsilon_1) \langle |g|^2 f^q \rangle,$$

where constant  $c(\varepsilon_0, \varepsilon, \varepsilon_1)$  can be made as small as needed by selecting  $\varepsilon_1$  sufficiently small and then selecting  $\varepsilon$  even smaller. So, taking  $\mu_0 := c_\mu + c_0(\varepsilon_0, \varepsilon_1, \varepsilon)$ , we obtain the required gradient estimate.

*Proof of Theorem 10 in the case drift  $b$  satisfies condition  $(\mathbb{B}_1)$ .* One needs to estimate term  $\langle b \cdot w, \phi \rangle$  in (13.2) differently. Indeed,  $b$  is no longer differentiable and hence one cannot integrate by parts. Instead, arguing as in [36], we evaluate the test function  $\phi$  as

$$\langle b \cdot w, \phi \rangle = -\langle b \cdot w, |w|^{q-2}\Delta u \rangle - (q-2)\langle b \cdot w, |w|^{q-3}w \cdot \nabla |w| \rangle,$$

and then re-uses the elliptic equation to express  $\Delta u$  in terms of  $\mu u$ ,  $b \cdot w$  and  $f$  (or  $|g|f$ ). Then we repeat the argument from [36] up to the estimates on  $|\langle f, \phi \rangle|$  (assertion (ii)) and  $|\langle |g|f, \phi \rangle|$  (assertion (i)), which we take from Step 2 above.

## 14. PROOF OF THEOREM 7

Let  $b_n$  be constructed as in Lemma 3, i.e.

$$b_n = E_{\varepsilon_n} b, \quad \varepsilon_n \downarrow 0,$$

so that  $b_n$  are bounded, smooth, converge to  $b$  locally in  $L^2$  and, crucially, do not increase neither form-bound  $\delta$  of  $b$  nor constant  $c_\delta$ .



A comment regarding the case when  $b$  satisfies condition  $(\mathbb{B}_2)$ . Below we use gradient bounds from Theorem 10 for vector fields  $b_n$ . The proof of these gradient bounds depends on a somewhat less restrictive condition than  $(\mathbb{B}_2)$ , i.e.  $b \in \mathbf{F}_\delta$ ,  $\delta < \infty$ , and

$$\langle (E_{\varepsilon_n} B_-)h, h \rangle \leq \nu \langle |\nabla|h|^2 \rangle + c_\nu \langle |h|^2 \rangle, \quad (14.1)$$

where  $B_-$  is the negative part of matrix  $(\nabla_k b^i)_{k,i=1}^d - \frac{\operatorname{div} b}{q} I$ . (Indeed, if  $B_+$  denotes the positive part of the last matrix, we have

$$(\nabla_k b_n^i)_{k,i=1}^d - \frac{\operatorname{div} b_n}{q} I = E_{\varepsilon_n} B_+ - E_{\varepsilon_n} B_-, \quad E_{\varepsilon_n} B_\pm \geq 0,$$

and can repeat the proof of Theorem 10 for  $b_n$  and  $E_{\varepsilon_n} B_-$ .) By Lemma 11, inequality (14.1) does hold with constants  $\nu = \sum_{j=1}^d \nu_j$  and  $c_\nu = \sum_{j=1}^d c_{\nu_j}$  that are, obviously, independent of  $\{\varepsilon_n\}$ , and so the constants in the gradient bounds in Theorem 10 for  $b_n$  do not depend on  $n$ .

*Proof of assertion (i).* Let  $\{\mathbb{P}_x\}_{x \in \mathbb{R}^d}$  be the strong Markov family of martingale solutions to (2.30) constructed in Theorem 5. Fix some  $y$ . Our goal is prove the following estimate: there exists generic  $q > (d-2) \vee 2$  and  $C$  such that, for all  $\mathbf{g} \in \mathbf{F}_{\delta_1}$ ,  $\delta_1 < \infty$ , and all  $\lambda$  greater than some generic  $\lambda_0$ ,

$$\mathbb{E}_y \int_0^\infty e^{-\lambda s} |\mathbf{g}f|(\omega_s) ds \leq C \|\mathbf{g}|f|^{\frac{q}{2}}\|_2^{\frac{2}{q}} \quad (14.2)$$

for all  $f \in C_c$ . Let  $\mathbf{g}_m$  the bounded smooth regularization of  $\mathbf{g}$  constructed according to Lemma 3. Using the gradient estimate of Theorem 10(i), after applying the Sobolev embedding theorem twice, we obtain

$$\mathbb{E}_{\mathbb{P}_y^n} \int_0^\infty e^{-\lambda s} |\mathbf{g}_m f|(\omega_s) ds \leq C \|\mathbf{g}_m |f|^{\frac{q}{2}}\|_2^{\frac{2}{q}}, \quad n, m = 1, 2, \dots,$$

where  $\mathbb{P}_x^n$  is the martingale solution of the regularized SDE

$$Y(t) = y - \int_0^t b_n(Y(s)) ds + \sqrt{2} B(t), \quad t \geq 0$$

and, by the construction of  $\mathbb{P}_x$  in the proof of Theorem 5,  $\mathbb{P}_x^n \rightarrow \mathbb{P}_x$  weakly (we pass to a subsequence of  $\{b_n\}$  if necessary). Thus, we have

$$\mathbb{E}_{\mathbb{P}_y} \int_0^\infty e^{-\lambda s} |\mathbf{g}_m f|(\omega_s) ds \leq C \|\mathbf{g}_m |f|^{\frac{q}{2}}\|_2^{\frac{2}{q}}, \quad m = 1, 2, \dots$$

Fatou's lemma applied in  $m$  now yields (14.2) and thus ends the proof of (i).

*Proof of assertion (i').* Let  $\{\mathbb{P}_x^1\}_{x \in \mathbb{R}^d}$ ,  $\{\mathbb{P}_x^2\}_{x \in \mathbb{R}^d}$  be two Markov families of martingale solutions to SDE

$$Y(t) = y - \int_0^t b(Y(s)) ds + \sqrt{2} B(t), \quad t \geq 0.$$

Fix some  $y$ . By our assumption, there exists  $q > (d-2) \vee 2$  such that, for all  $\mathbf{g} \in \mathbf{F}_{\delta_1}$ ,  $\delta_1 < \infty$ , and all  $\lambda$  greater than some generic  $\lambda_0$ ,

$$\mathbb{E}_y^i \int_0^\infty e^{-\lambda s} |\mathbf{g}f|(\omega_s) ds \leq C \|\mathbf{g}|f|^{\frac{q}{2}}\|_2^{\frac{2}{q}} \quad (14.3)$$

for all  $f \in C_c$ . Let  $v_n$  be the classical solution to equation

$$(\lambda - \Delta + b_n \cdot \nabla) v_n = -F,$$

where  $F \in C_c(\mathbb{R}^d)$ . We will need weight  $\rho(x) = (1 + k|x|^2)^{-\beta}$ ,  $k > 0$ , where constant  $\beta$  is fixed greater than  $\frac{d}{2}$  so that  $\rho \in L^1(\mathbb{R}^d)$ . By Itô's formula applied to  $e^{-\lambda t} \rho v_n$ , we have

$$\begin{aligned} \mathbb{E}_y^i[e^{-\lambda t} \rho v_n(X_t)] &= \rho(y) v_n(y) + \mathbb{E}_y^i \int_0^t \rho e^{-\lambda s} F(\omega_s) ds \\ &\quad + \mathbb{E}_y^i \int_0^t [e^{-\lambda s} \rho(b - b_n) \cdot \nabla v_n](\omega_s) ds + S_n, \end{aligned} \quad (14.4)$$

where  $S_n$  is the remainder term given by

$$S_n := \mathbb{E}_y^i \int_0^t e^{-\lambda s} [-(\Delta \rho) v_n - 2 \nabla \rho \cdot \nabla v_n + b_n \cdot (\nabla \rho) v_n](\omega_s) ds.$$

**Proposition 7.** *For every  $k > 0$ ,  $\mathbb{E}_y^i \int_0^t e^{-\lambda s} [\rho(b - b_n) \cdot \nabla v_n](\omega_s) ds$  as  $n \uparrow \infty$  uniformly in  $t > 0$ .*

*Proof.* We have

$$\begin{aligned} |\mathbb{E}_y^i \int_0^t [e^{-\lambda s} \rho(b - b_n) \cdot \nabla v_n](\omega_s) ds| &\leq |\mathbb{E}_x^i \int_0^\infty [e^{-\lambda s} \rho(b - b_n) \cdot \nabla v_n](\omega_s) ds| \\ &\quad (\text{we apply (14.3) with } \mathbf{g} := \rho(b - b_n) \in \mathbf{F}_{2\delta}) \\ &\leq K \|\rho(b - b_n) |\nabla v_n|^{\frac{q}{2}}\|_2^{\frac{2}{q}}. \end{aligned}$$

In turn, for a  $0 < \theta < 1$ , we have

$$\|\rho(b - b_n) |\nabla v_n|^{\frac{q}{2}}\|_2 \leq \|\rho(b - b_n)\|_2^\theta \|\rho(b - b_n) |\nabla v_n|^{\frac{q}{2(1-\theta)}}\|_2^{1-\theta}. \quad (14.5)$$

Regarding the second multiple in the RHS of (14.5): we assume that  $\theta$  is chosen to be sufficiently close to 0 so that  $\frac{q}{1-\theta} > (d-2) \vee 2$ . Then, by  $b - b_n \in \mathbf{F}_{2\delta}$ ,

$$\begin{aligned} \|\rho(b - b_n) |\nabla v_n|^{\frac{q}{2(1-\theta)}}\|_2^2 &\leq \|(b - b_n) |\nabla v_n|^{\frac{q}{2(1-\theta)}}\|_2^2 \\ &\leq 2\delta \|\nabla |\nabla v_n|^{\frac{q}{2(1-\theta)}}\|_2^2 + 2c_\delta \|\nabla v_n\|_2^{\frac{q}{2(1-\theta)}}\|_2^2. \end{aligned}$$

Hence, by the gradient estimate of Theorem 10(i),  $\sup_n \|\rho(b - b_n) |\nabla v_n|^{\frac{q}{2(1-\theta)}}\|_2^2 < \infty$ .

The first multiple in the RHS of (14.5):

$$\begin{aligned} \|\rho(b - b_n)\|_2^2 &\leq \langle \mathbf{1}_{B_R(0)}(b - b_n) \rangle + \langle (1 - \mathbf{1}_{B_R(0)}) \rho, \rho(b - b_n)^2 \rangle \\ &\leq \langle \mathbf{1}_{B_R(0)}(b - b_n) \rangle + (1 + kR^2)^{-\beta} \langle \rho(b - b_n)^2 \rangle. \end{aligned}$$

Since  $b_n \rightarrow b$  in  $L_{\text{loc}}^2$ , the first integral can be made as small as needed (uniformly in  $R$ ) by selecting  $n$  sufficiently large. In the second integral  $\sup_n \langle \rho(b - b_n)^2 \rangle < \infty$ , since, by  $b - b_n \in \mathbf{F}_{2\delta}$ ,

$$\langle \rho(b - b_n)^2 \rangle \leq 2\delta \langle (\nabla \sqrt{\rho})^2 \rangle + 2c_\delta \langle \rho \rangle,$$

so it remains to apply  $|\nabla \rho| \leq \beta \sqrt{k} \rho$ . At the same time,  $(1 + kr^2)^{-\beta}$  can be made as small as needed by selecting  $r$  sufficiently large. This completes the proof.  $\square$

**Proposition 8.**  *$S_n \rightarrow 0$  as  $k \downarrow 0$  uniformly in  $n$  and  $t$ .*

*Proof.* Using  $|\nabla \rho| \leq \beta \sqrt{k} \rho$ ,  $|\Delta \rho| \leq \beta^2 k$ , we have

$$|S_n| \leq \sqrt{k} C \mathbb{E}_x^i \int_0^t [\rho |v_n| + 2\rho |\nabla v_n| + \rho |b_n| |v_n|](\omega_s) ds.$$

Now we can argue as in the proof of the previous proposition, using additionally  $\|v_n\|_\infty \leq \lambda^{-1} \|F\|_\infty$ , to show that  $\sup_n \mathbb{E}_x^i \int_0^t [\rho |v_n| + 2\rho |\nabla v_n| + \rho |b_n| |v_n|](\omega_s) ds < \infty$ . In fact, in this case

the proof is easier since none of the terms contains simultaneously  $b_n$  and  $\nabla v_n$ . Selecting  $k$  sufficiently small, we can make  $S_n$  as small as needed.  $\square$

We now complete the proof of assertion (i'). Let us note that, for every  $k > 0$ ,

$$\mathbb{E}_y^i[e^{-\lambda t} \rho v_n(\omega_t)] \rightarrow 0 \quad \text{as } t \rightarrow \infty \text{ uniformly in } n.$$

Indeed,  $\|v_n\|_\infty \leq \lambda^{-1} \|F\|_\infty$ , so  $|\mathbb{E}_y^i[e^{-\lambda t} \rho v_n(\omega_t)]| \leq \lambda^{-1} e^{-\lambda t}$ , which yields the required. Combining this result with Propositions 7 and 8, and taking into account that, by Theorem 8(iv),  $\{v_n\}$  converge uniformly as  $n \rightarrow \infty$  to a continuous function  $v$ , we obtain from (14.4) upon taking  $n \rightarrow \infty$  and then taking  $k \downarrow 0$ :

$$0 = v(y) + \mathbb{E}_y^i \int_0^\infty e^{-\lambda s} F(\omega_s) ds, \quad i = 1, 2.$$

Taking into account the continuity of  $F$  and  $\omega$ , and invoking the uniqueness of Laplace transform, we obtain that  $\mathbb{E}_y^1 F(\omega_t) = \mathbb{E}_y^2 F(\omega_t)$  for all  $F \in C_c$ ,  $t > 0$ . We deduce from here that the one-dimensional distributions of  $\mathbb{P}_y^1$  and  $\mathbb{P}_y^2$  coincide. Since we are dealing with Markov families of probability measures, we conclude that  $\mathbb{P}_y^1 = \mathbb{P}_y^2$  for every  $y \in \mathbb{R}^d$ .

*Proof of assertion (ii).* The proof follows closely the proof of (i), but uses the gradient estimate of Theorem 10(i) for  $q > (d-2) \vee 2$  chosen closely to  $(d-2) \vee 2$ . In fact, this proof is easier since we no longer need to take care of extra form-bounded vector fields  $\mathbf{g}$  as in (i).

*Proof of assertion (ii').* We modify the previous proof of (i'). By our assumption,

$$\mathbb{E}_x^i \int_0^\infty e^{-\lambda s} |f|(\omega_s) ds \leq C \|f\|_{\frac{d}{2-\varepsilon} \wedge \frac{2}{1-\varepsilon}}, \quad \forall f \in C_c, \quad \lambda > \lambda_0. \quad (14.6)$$

The analogue of Proposition 7 is proved as follows. Clearly, hypothesis

$$(1 + |x|^{-2})^{-\beta} |b|^{\frac{d}{2-\varepsilon_1} \vee \frac{2}{1-\varepsilon_1}} \in L^1, \quad \varepsilon_1 \in ]\varepsilon, 1[$$

implies that, for any  $k > 0$ ,  $\rho |b|^{\frac{d}{2-\varepsilon_1} \vee \frac{2}{1-\varepsilon_1}} \in L^1$ . We have

$$\begin{aligned} |\mathbb{E}_x^i \int_0^t [e^{-\lambda s} \rho(b - b_n) \cdot \nabla v_n](\omega_s) ds| &\leq |\mathbb{E}_x^i \int_0^\infty [e^{-\lambda s} \rho(b - b_n) \cdot \nabla v_n](\omega_s) ds| \\ &\quad (\text{we apply (14.6) using Fatou's lemma}) \\ &\leq K \|\rho(b - b_n) \cdot \nabla v_n\|_r \quad r := \frac{d}{2-\varepsilon} \wedge \frac{2}{1-\varepsilon} \\ &\leq K \|\rho(b - b_n)\|_{s'} \|\nabla v_n\|_s, \quad \frac{1}{s'} + \frac{1}{s} = \frac{1}{r}, \end{aligned} \quad (14.7)$$

where  $s' = \frac{d}{2-\varepsilon_1} \vee \frac{2}{1-\varepsilon_1}$  and  $s = \frac{q_* d}{d-2}$ , where  $q_*$  was defined in assertion (ii'') of Theorem 7 that we are proving. Theorem 10(ii), which applies by our assumptions on  $\delta$ ,  $\nu$  and  $q_*$  in the end of assertion (ii''), and the Sobolev embedding theorem, yield

$$\sup_n \|\nabla v_n\|_{\frac{q_* d}{d-2}} < \infty.$$

Therefore, the second multiple in the RHS of (14.7) is uniformly (in  $n$ ) bounded.

In turn, for every fixed  $k$ , the first multiple in the RHS of (14.7) tends to zero as  $n \rightarrow \infty$ . Indeed, since  $0 < \rho \leq 1$ , we have

$$\sup_n \|\rho^{s'} b_n^{s'}\|_1 \leq \sup_n \|\rho b_n^{s'}\|_1 < \infty,$$

where the finiteness is seen, after integrating by parts, from  $E_{\varepsilon_n}\rho \leq C\rho$  with constant  $C$  independent of  $n$  (here we simply use the fact that the Friedrichs mollifier is a convolution with a function having compact support) and our hypothesis  $\|\rho|b|^{s'}\|_1 < \infty$ . Now, we represent

$$\begin{aligned} \|\rho(b - b_n)\|_{s'} &= \|\mathbf{1}_{B_R(0)}(b - b_n)\|_{s'} + \|(1 - \mathbf{1}_{B_R(0)})\rho(b - b_n)\|_{s'} \\ &\leq \|\mathbf{1}_{B_R(0)}(b - b_n)\|_{s'} + (1 + kR^2)^{-\beta(s'-1)}(\langle \rho b^{s'} \rangle + \langle \rho b_n^{s'} \rangle). \end{aligned}$$

The second term can be made as small as needed by selecting  $R$  sufficiently large (uniformly in  $n$ ). Then, for  $R$  thus fixed, the first term can be made as small as needed by selecting  $n$  sufficiently large, since  $b_n \rightarrow b$  in  $L^1_{\text{loc}}$  by the properties of Friedrichs mollifier.

Arguing as above, we prove  $\sup_n \mathbb{E}_x^i \int_0^t [\rho|v_n| + 2\rho|\nabla v_n| + \rho|b_n||v_n|](\omega_s) ds < \infty$ , and hence have the analogue of Proposition 8.

The rest of the proof of (ii') repeats the proof of (i').

#### APPENDIX A. A DESINGULARIZATION THEOREM FROM [33]

Let  $X$  be a locally compact topological space, and  $\mu$  a  $\sigma$ -finite Borel measure on  $X$ . In what follows,  $L^r = L^r(X, \mu)$  ( $1 \leq r \leq \infty$ ). Let  $j > 1$ , put  $j' := \frac{j}{j-1}$ .

Let  $\Lambda$  be the generator of a strongly continuous semigroup  $e^{-t\Lambda}$  on  $L^r$  for some  $r > 1$ , such that for some constants  $c, j > 1$ , for all  $t > 0$ ,

$$\|e^{-t\Lambda}\|_{r \rightarrow \infty} \leq ct^{-\frac{j'}{r}}. \quad (S_1)$$

We consider a family of weights  $\varphi = \{\varphi_s\}_{s>0}$  in  $X$  such that

$$0 \leq \varphi_s, \frac{1}{\varphi_s} \in L^1_{\text{loc}}(X, \mu) \quad \text{for all } s > 0, \quad (S_2)$$

$$\inf_{s>0, x \in X} \varphi_s(x) \geq c_0 > 0. \quad (S_3)$$

**Theorem 11.** *Assume that conditions (S<sub>1</sub>) - (S<sub>3</sub>) hold and there exists constant  $c_1$ , independent of  $s$ , such that, for all  $0 < t \leq s$ ,*

$$\|\varphi_s e^{-t\Lambda} \varphi_s^{-1} f\|_1 \leq c_1 \|f\|_1, \quad f \in L^1 \cap L^\infty. \quad (S_4)$$

*Then, for each  $t > 0$ ,  $e^{-t\Lambda}$  is an integral operator, and there is a constant  $C = C(j, c_1, c_0)$  such that, up to change of  $e^{-t\Lambda}(x, y)$  on a measure zero set, the weighted Nash initial estimate*

$$|e^{-t\Lambda}(x, y)| \leq Ct^{-j'} \varphi_t(y) \quad (A.1)$$

*is valid for  $\mu$  a.e.  $x, y \in X$ .*

**REMARK 12.** The first desingularization result of this type appeared in the context of studying Schrödinger operator with the inverse-square potential and is due to [42]. There the authors introduced a weighted variant of Nash's method with the "desingularizing weight" that either explodes or vanishes at the origin, depending on the sign of the potential. That said, the non-symmetric situation considered in Theorem 11 is quite different from the setting of [42] since it cannot be recast, even formally, as following Nash's proof of the on-diagonal bound in a weighted space, unless one wants to impose rather restrictive assumptions on the adjoint operator that would rule out strong singularities of the drift.

For the sake of keeping the paper self-contained, we reproduce here the proof of Theorem 11 from [33].

*Proof of Theorem 11.* 1. We will use a weighted variant of the Coulhon-Raynaud extrapolation theorem. Put

$$0 \leq \psi \in L^1 + L^\infty, \quad \|f\|_{p, \sqrt{\psi}} := \langle |f|^p \psi \rangle^{1/p}.$$

Let  $U^{t, \theta}$  be a two-parameter family of operators

$$U^{t, \theta} f = U^{t, \tau} U^{\tau, \theta} f, \quad f \in L^1 \cap L^\infty, \quad 0 \leq \theta < \tau < t \leq \infty.$$

If for some  $1 \leq p < q < r \leq \infty$ ,  $\nu > 0$

$$\begin{aligned} \|U^{t, \theta} f\|_p &\leq M_1 \|f\|_{p, \sqrt{\psi}}, \\ \|U^{t, \theta} f\|_r &\leq M_2 (t - \theta)^{-\nu} \|f\|_q \end{aligned}$$

for all  $(t, \theta)$  and  $f \in L^1 \cap L^\infty$ , then

$$\|U^{t, \theta} f\|_r \leq M (t - \theta)^{-\nu/(1-\beta)} \|f\|_{p, \sqrt{\psi}}, \quad (\text{A.2})$$

where  $\beta = \frac{r}{q} \frac{q-p}{r-p}$  and  $M = 2^{\nu/(1-\beta)^2} M_1 M_2^{1/(1-\beta)}$ . Here is the proof of (A.2) for reader's convenience. Put  $t_\theta := \frac{t+\theta}{2}$ . We have

$$\begin{aligned} \|U^{t, \theta} f\|_r &\leq M_2 (t - t_\theta)^{-\nu} \|U^{t_\theta, \theta} f\|_q \\ &\leq M_2 (t - t_\theta)^{-\nu} \|U^{t_\theta, \theta} f\|_r^\beta \|U^{t_\theta, \theta} f\|_p^{1-\beta} \\ &\leq M_2 M_1^{1-\beta} (t - t_\theta)^{-\nu} \|U^{t_\theta, \theta} f\|_r^\beta \|f\|_{p, \sqrt{\psi}}^{1-\beta}, \end{aligned}$$

and hence

$$(t - \theta)^{\nu/(1-\beta)} \|U^{t, \theta} f\|_r / \|f\|_{p, \sqrt{\psi}} \leq M_2 M_1^{1-\beta} 2^{\nu/(1-\beta)} [(t - \theta)^{\nu/(1-\beta)} \|U^{t_\theta, \theta} f\|_r / \|f\|_{p, \sqrt{\psi}}]^\beta.$$

Setting  $R_{2T} := \sup_{t-\theta \in ]0, T]} [(t - \theta)^{\nu/(1-\beta)} \|U^{t, \theta} f\|_r / \|f\|_{p, \sqrt{\psi}}]$ , we obtain from the last inequality that  $R_{2T} \leq M^{1-\beta} (R_T)^\beta$ . But  $R_T \leq R_{2T}$ , and so  $R_{2T} \leq M$ . This gives us (A.2).

2. We are in position to complete the proof of Theorem 11. By  $(S_4)$  and  $(S_3)$ ,

$$\begin{aligned} \|e^{-t\Lambda} h\|_1 &\leq c_0^{-1} \|\varphi_s e^{-t\Lambda} \varphi_s^{-1} \varphi_s h\|_1 \\ &\leq c_0^{-1} c_1 \|h\|_{1, \sqrt{\varphi_s}}, \quad h \in L_{\text{com}}^\infty. \end{aligned}$$

The latter,  $(S_1)$  and the Coulhon-Raynaud extrapolation theorem with  $\psi := \varphi_s$  yield

$$\|e^{-t\Lambda} f\|_\infty \leq M t^{-j'} \|\varphi_s f\|_1, \quad 0 < t \leq s, \quad f \in L_c^\infty.$$

Note that  $(S_1)$  verifies the assumptions of the Dunford-Pettis theorem, which yields that  $e^{-t\Lambda}$  is an integral operator. Therefore, taking  $s = t$  in the previous estimate, we obtain (A.1).  $\square$

## REFERENCES

- [1] D. Adams, Weighted nonlinear potential theory, *Trans. Amer. Math. Soc.*, **297** (1986), 73-94.
- [2] R. Adams and J. Fournier, Sobolev Spaces, Second Edition, *Elsevier*, 2003.
- [3] S. Albeverio, Yu. Kondratiev and M. Röckner, Strong Feller properties for distorted Brownian motion and applications to finite particle systems with singular interactions, “Finite and Infinite Dimensional Analysis in Honor of Leonard Gross” (H. H. Kuo et al., eds.), *Contemporary Mathematics*, **317**, Amer. Math. Soc., 2003.
- [4] D. Albrighon and H. Dong, Regularity properties of passive scalars with rough divergence-free drifts, *Arch. Rat. Mech. Anal.*, **247**, article no. 75 (2023) (arXiv:2107.12511).
- [5] R. Bass, Diffusions and Elliptic Operators, *Springer*, 1997.
- [6] R. Bass, Probabilistic Techniques in Analysis, *Springer-Verlag*, 1995.
- [7] 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).
- [8] 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).
- [9] 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.
- [10] 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.
- [11] 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.
- [12] C. Fefferman, The uncertainty principle, *Bull. Amer. Math. Soc.*, **9** (1983), 129-206.
- [13] 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.
- [14] N. Fournier and T. Tardy, A simple proof of non-explosion for measure solutions of the Keller-Segel equation, arXiv:2202.03508.
- [15] E. Giusti, Direct Methods in the Calculus of Variations, *World Scientific*, 2003.
- [16] D. Guzu, Many-particle Hardy type inequalities, *J. Math. Sci.* **239** (2019), 292-298.
- [17] Z. Hao, M. Röcker and X. Zhang, Strong convergence of propagation of chaos for McKean-Vlasov SDE with singular interactions, arXiv:2204.07952.
- [18] Z. Hao and X. Zhang, SDEs with supercritical distributional drifts, arXiv:2312.11145.
- [19] T. Hoffmann-Ostenhof and A. Laptev, Hardy inequalities with homogeneous weights, *J. Funct. Anal.*, **268** (2015), 3278-3289.
- [20] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, A. Laptev and J. Tidblom, Many-particle Hardy inequalities, *J. Lond. Math. Soc.*, **77** (2008), 99-114.
- [21] D. Kinzebulatov, Feller evolution families and parabolic equations with form-bounded vector fields, *Osaka J. Math.*, **54** (2017), 499-516 (arXiv:1407.4861).
- [22] D. Kinzebulatov, Laplacian with singular drift in a critical borderline case, arXiv:2309.04436.
- [23] D. Kinzebulatov, Form-boundedness and SDEs with singular drift, arXiv:2305.00146.
- [24] D. Kinzebulatov, Parabolic equations and SDEs with time-inhomogeneous Morrey drift, arXiv:2301.13805.
- [25] D. Kinzebulatov and K.R. Madou, Strong solutions of SDEs with singular (form-bounded) drift via Roeckner-Zhao approach, arXiv:2306.04825.
- [26] D. Kinzebulatov and K.R. Madou, Stochastic equations with time-dependent singular drift, *J. Differential Equations*, **337** (2022), 255-293 (arXiv:2105.07312).
- [27] D. Kinzebulatov and Yu.A. Semënov, Brownian motion with general drift, *Stoch. Proc. Appl.*, **130** (2020), 2737-2750 (arXiv:1710.06729).
- [28] D. Kinzebulatov and Yu. A. Semënov, Sharp solvability for singular SDEs, *Electr. J. Probab.*, **28** (2023), article no. 69, 1-15. (arXiv:2110.11232).
- [29] D. Kinzebulatov and Yu. A. Semënov, On the theory of the Kolmogorov operator in the spaces  $L^p$  and  $C_\infty$ , *Ann. Sc. Norm. Sup. Pisa (5)* **21** (2020), 1573-1647 (arXiv:1709.08598).
- [30] D. Kinzebulatov and Yu. A. Semënov, Heat kernel bounds for parabolic equations with singular (form-bounded) vector fields, *Math. Ann.*, **384** (2022), 1883-1929.
- [31] D. Kinzebulatov and Yu. A. Semënov, Regularity for parabolic equations with singular non-zero divergence vector fields, *J. Differential Equations*, to appear.

- [32] D. Kinzebulatov and Yu. A. Semënov, Fractional Kolmogorov operator and desingularizing weights, *Publ. Res. Inst. Math. Sci. Kyoto*, to appear (arXiv:2005.11199).
- [33] D. Kinzebulatov, Yu. A. Semënov and K. Szczypkowski, Heat kernel of fractional Laplacian with Hardy drift via desingularizing weights, *J. London Math. Soc.*, 104 (2021), 1861-1900 (arXiv:1904.07368).
- [34] D. Kinzebulatov and R. Vafadar, On divergence-free (form-bounded type) drifts, *Discrete Contin. Dyn. Syst. Ser. S.*, to appear (arXiv:2209.04537).
- [35] 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.
- [36] 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.
- [37] N. V. Krylov, On diffusion processes with drift in  $L_d$ , *Probab. Theory Related Fields* **179** (2021), no. 1-2, 165-199 (arXiv:2001.04950).
- [38] N. V. Krylov, On strong solutions of Itô's equations with  $A \in W^{1,d}$  and  $B \in L^d$ , *Ann. Probab.* **49** (2021), no. 6, 3142-3167 (arXiv:2007.06040).
- [39] N. V. Krylov, On strong solutions of Itô's equations with  $D\sigma$  and  $b$  in Morrey classes containing  $L^d$ , *Ann. Probab.* **51** (2023), no. 5, 1729-1751 (arXiv:2111.13795).
- [40] N. V. Krylov and M. Röckner, Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131 (2005), 154-196.
- [41] G. Metafune, L. Negro and C. Spina, Sharp kernel estimates for elliptic operators with second-order discontinuous coefficients, *J. Evol. Equ.* **18** (2018), 467-514.
- [42] P. D. Milman and Yu. A. Semënov, Global heat kernel bounds via desingularizing weights, *J. Funct. Anal.*, **212** (2004), p. 373-398.
- [43] M. Röckner and G. Zhao, SDEs with critical time dependent drifts: weak solutions, *Bernoulli*, **29** (2023), 757-784 (arXiv:2012.04161).
- [44] M. Röckner and G. Zhao, SDEs with critical time dependent drifts: strong solutions, arXiv:2103.05803.
- [45] Yu. A. Semënov, Regularity theorems for parabolic equations, *J. Funct. Anal.*, **231** (2006), 375-417.
- [46] D. W. Stroock, S. Varadhan, *Multidimensional Diffusion Processes*, Springer, 2007.
- [47] Y. Tardy, Weak convergence of the empirical measure for the Keller-Segel model in both sub-critical and critical cases, arXiv:2205.04968.
- [48] M. Tomašević, Propagation of chaos for stochastic particle systems with singular mean-field interaction of  $L^q$ - $L^p$  type, HAL-03086253 (2020).
- [49] Q. S. Zhang, A strong regularity result for parabolic equations, *Comm. Math. Phys.* **244** (2004), 245-260.
- [50] Q. S. Zhang, Local estimates of two linear parabolic equations with singular coefficients, *Pacific J. Math.* **223** (2006), 367-396.
- [51] X. Zhang and G. Zhao, Stochastic Lagrangian paths for Leray solutions of 3D Navier-Stokes equations, *Comm. Math. Phys.*, **381**(2) (2021), 491-525.
- [52] X. Zhang and G. Zhao, Singular Brownian diffusion processes, *Communications in Mathematics and Statistics* (2018), 1-49.
- [53] G. Zhao, Stochastic Lagrangian flows for SDEs with rough coefficients, arXiv:1911.05562.