

KINETIC SDES AND KINETIC FOKKER-PLANCK EQUATIONS WITH CRITICAL MORREY-TYPE DRIFTS, AND LOCAL BLOW-UPS

JEAN-FRANÇOIS JABIR, DAMIR KINZEBULATOV, AND CHENGJUN YUE

ABSTRACT. We develop a Sobolev theory for the kinetic Fokker-Planck equation subject to a singular Morrey-form $L^\infty + \mathcal{M}_{\mathcal{K}}^q$ force field which can produce local blow-ups due to strong attraction to a submanifold. Our approach is based on the fractional representation for the Duhamel series of the equation and on establishing kinetic Adams-type estimates. Accordingly, we construct weak solutions of Brownian-driven kinetic SDEs in a critical Morrey set-up, obtain kinetic Krylov-type estimates and a strongly continuous Feller propagator, and establish uniqueness criteria for a large class of weak solutions.

CONTENTS

1. Introduction	1
2. Kinetic geometry and kinetic Bessel potentials	7
3. Main results	10
4. Counterexample to weak existence for the kinetic SDE (blow-up)	17
5. Main Rellich-type estimate from kinetic Morrey class via Adams estimates	20
6. Proof of Theorem 3.3	25
7. Proof of Theorem 3.6	26
8. Proof of Theorem 3.9	27
9. Removal of global integrability condition using weights (Proof of Theorem 3.10)	38
Appendix A. Fractional powers of linear operators	44
Appendix B. Proof of Proposition 2.2	46
Appendix C. Proof of Proposition 2.3	48
Appendix D. Many-particle and Lebesgue class drifts	50
References	51

1. INTRODUCTION

1.1. Brief overview of the paper. Kinetic equations describe the statistical evolution of inertial particles subject to environmental noise. These equations arise naturally in Vlasov systems in statistical mechanics and the study of Langevin-type dynamics [44]. Notably, related SDEs generically correspond to stochastic Newtonian dynamics

$$\ddot{X}_t = F_t(X_t, \dot{X}_t) + \sqrt{2}\dot{W}_t$$

Key words and phrases. Kinetic Fokker-Planck equation; Morrey-type singular drift; Adams estimates; Hypo-elliptic PDEs; Singular kinetic SDEs; Feller propagator.

The research of J.-F. J. is supported by the Russian Science Foundation project (project No. 24-11-00123). The research of D.K. is supported by NSERC Discovery grant (RGPIN-2024-04236). The research of C. Y. is supported by the CRM-Laval postdoctoral fellowship.

which arise in models of molecular motion, electromagnetic phenomena, turbulence, social or economic multi-agent systems as well as in deep learning methods (we may refer to [13], [56], [48], [18] and references therein).

The goal of this paper is to develop a Sobolev theory for the forward equation

$$\lambda \eta(t, x, v) + \mathcal{K} \eta(t, x, v) - \operatorname{div}_v(b \eta)(t, x, v) = f(t, x, v), \quad (t, x, v) \text{ in } \mathbb{R}^{1+2d}, \quad d \geq 2, \quad (1.1)$$

with the kinetic forward Kolmogorov operator

$$\mathcal{K} := \partial_t + v \cdot \nabla_x - \Delta_v.$$

Here the diffusion acts only in velocity; however, transport couples velocity and position, and the commutator identity

$$[\partial_{v_j}, \partial_t + v \cdot \nabla_x] = \partial_{x_j}$$

shows how velocity regularization is transmitted to the position variables. The drift $b : \mathbb{R}^{1+2d} \rightarrow \mathbb{R}^d$ is Borel measurable and may have critical singularities in time, position, and velocity, i.e. can introduce singular attractive forces and irregular alignment interactions. We then apply our Sobolev theory to develop weak solution and Feller theory for the kinetic stochastic system

$$\begin{cases} dX_t = V_t dt, \\ dV_t = -b(t, X_t, V_t) dt + \sqrt{2} dW_t, \end{cases} \quad (X_0, V_0) = (x_0, v_0) \in \mathbb{R}^{2d}. \quad (1.2)$$

Recall that the kinetic dilation related to $\mathcal{K}$ is $(t, x, v) \mapsto (r^2 t, r^3 x, rv)$, and so the dimension of space-time is

$$Q := 2 + 3d + d = 4d + 2. \quad (1.3)$$

The corresponding definition of the kinetic cylinder¹ is

$$\mathcal{C}_r(t, x, v) := \{(s, y, w) : -r^2 < s - t < 0, |w - v| < r, |y - x - (s - t)v| < r^3\}, \quad (1.4)$$

which determines (see also Section 2) the Morrey class $g \in \mathcal{M}_{\mathcal{K}}^q$, $q > 1$, with

$$g \in L_{loc}^q(\mathbb{R}^{1+2d}), \quad \|g\|_{\mathcal{M}_{\mathcal{K}}^q} := \sup_{(t,x,v) \in \mathbb{R}^{1+2d}, r>0} r \left(\frac{1}{|\mathcal{C}_r(t, x, v)|} \int_{\mathcal{C}_r(t, x, v)} |g(\xi)|^q d\xi \right)^{1/q} < \infty. \quad (1.5)$$

This class is critical in the sense that it is invariant under the kinetic scaling $g_r(t, x, v) = r g(r^2 t, r^3 x, rv)$.

Our main hypothesis on the drift component b in (1.1) and (1.2) consists of the class

$$b = b_s + b_b, \quad b_s \in \mathcal{M}_{\mathcal{K}}^q, \quad b_b \in L^\infty(\mathbb{R}^{1+2d}, \mathbb{R}^d), \quad (\mathbf{H}_{\mathcal{K}})$$

with Morrey norm $\|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}$ required to be sufficiently small. The Morrey class $\mathcal{M}_{\mathcal{K}}^q$ becomes larger as q gets smaller, so it is reasonable to consider q close to 1 (thus, up to the strict threshold $q > 1$, we are dealing with the largest possible critical Morrey class). Fixing q determines also how small $\|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}$ must be; this smallness condition on the Morrey norm turns out to be necessary for (1.2): Proposition 4.1 gives a counterexample to weak existence if the Morrey norm of the drift is too large.

Our class $\mathcal{M}_{\mathcal{K}}^q + L^\infty$ of drifts is general in the sense that it does not presuppose any specific structure of the drift, such as being divergence-free or having gradient form. In particular, it is invariant under L^∞ multipliers. This provides us with flexible means for modifying and constructing interaction kernels in particle systems with attracting interactions, as we may briefly illustrate at once.

¹ $\mathcal{C}_r(t, x, v)$ has measure comparable with r^Q .

Examples. The kinetic Morrey class $\mathcal{M}_{\mathcal{K}}^q$ contains the following vector fields:

1. Let $e \in \mathbb{R}^d$ be fixed. If $1 < q < 2$, then, for $t_0 \in \mathbb{R}$,

$$b(t, x, v) = \kappa |t - t_0|^{-1/2} e \in \mathcal{M}_{\mathcal{K}}^q.$$

If $1 < q < d$, then

$$b(t, x, v) = \kappa \frac{v}{|v|^2} \in \mathcal{M}_{\mathcal{K}}^q,$$

If $1 < q < 3d$, then, for some reference point $x_0 \in \mathbb{R}^d$,

$$b(t, x, v) = \kappa |x - x_0|^{-1/3} e \in \mathcal{M}_{\mathcal{K}}^q,$$

the proof of the three above inclusions is in Proposition D.1. These previous examples can be combined i.e. we can add them up (so that e.g. the set of singularities constitutes a dense subset of $\mathbb{R}^{1+2d}$) or multiply by functions in L^∞ . Multiplying b by a constant $\kappa > 0$ scales its Morrey norm by κ .

2. If $|b|$ is in the Lebesgue class $L^{p_t}(\mathbb{R}_t, L^{p_v}(\mathbb{R}_v^d, L^{p_x}(\mathbb{R}_x^d)))$ (i.e. $\int (\int \|b(s, \cdot, v)\|_{L^{p_x}}^{p_v} dv)^{p_t/p_v} ds < \infty$) with $p_t, p_x, p_v \in (1, \infty]$ satisfying

$$\frac{2}{p_t} + \frac{d}{p_v} + \frac{3d}{p_x} = 1,$$

then it is in a Morrey class (Proposition D.2).

3. A motivation for allowing local blow-ups of the drift comes from interacting particle systems with singular attractive forces. Indeed, a system of N Langevin-type particles in the physical dimension d_0 can be viewed as a single kinetic system in dimension $d = d_0 N$ with b modeling internal forces. For example, the (time-homogeneous) pairwise interaction drifts of the form

$$b = (b^1, \dots, b^N) : \mathbb{R}^{d_0 N} \rightarrow \mathbb{R}^{d_0 N}, \quad b^i : \mathbb{R}^{d_0 N} \rightarrow \mathbb{R}^{d_0},$$

where

$$b^i(x) = \frac{1}{N} \sum_{j=1, j \neq i}^N \kappa_{ij} \frac{x^i - x^j}{|x^i - x^j|^{4/3}}, \quad x = (x^1, \dots, x^N) \in \mathbb{R}^{d_0 N}$$

or

$$b^i(v) = \frac{1}{N} \sum_{j=1, j \neq i}^N \gamma_{ij} \frac{v^i - v^j}{|v^i - v^j|^2}, \quad v = (v^1, \dots, v^N) \in \mathbb{R}^{d_0 N},$$

and where $1 \leq i \leq N$, $\kappa_{ij}, \gamma_{ij} \geq 0$ and both drifts are set to zero at the collisions, have kinetic degree -1 and therefore exhibit the critical behaviour considered in this paper. The pair-collision sets $\{x^i = x^j\}$ and the velocity-synchronization sets $\{v^i = v^j\}$ are then submanifolds of the phase space $\mathbb{R}^{2d}$. Other examples are the many-particle drifts $b^i(x, v) = \frac{1}{N} \sum_{j=1, j \neq i}^N K(x^i - x^j, v^i - v^j)$ with

$$K(\xi, \nu) = \kappa |\xi|^{-\frac{1}{3}} \frac{\nu}{1 + |\nu|^2}, \quad \xi, \nu \in \mathbb{R}^{d_0},$$

or more broadly the family of interaction kernels

$$b^i(x, v) = \frac{1}{N} \sum_{j=1, j \neq i}^N a_{ij}(x^i - x^j, v^i - v^j) \left[\kappa_{ij} \frac{x^i - x^j}{|x^i - x^j|^{4/3}} + \gamma_{ij} \frac{v^i - v^j}{|v^i - v^j|^2} \right],$$

where a_{ij} are bounded measurable. These can be viewed as extended singular versions of the characteristic interaction forces leading to swarming at large time (see [14] and references therein). The a_{ij} 's model non-negative communication rates (essentially acting for close-by particles) which, with a velocity drag force, introduce velocity alignment and agglomeration in compact clouds of particles at large time. This phenomenon is particularly known in the case $a_{i,j}(x^i - x^j, v^i - v^j) \frac{v^i - v^j}{|v^i - v^j|^2} = \frac{v^i - v^j}{(1 + |x^i - x^j|^2)^\kappa}$

in the deterministic and stochastic cases. Introducing a further concentrating force $\frac{v^i - v^j}{|v^i - v^j|^2}$ (assuming e.g. the communication rates are bounded) introduces a stronger form of velocity alignment, and the additional force $\frac{x^i - x^j}{|x^i - x^j|^{4/3}}$ a faster agglomeration of the dynamics. Although large-time behaviors are not discussed in our paper, such questions may be re-addressed in future works.

The inclusion of the above many-particle drifts in our Morrey class is proved in Proposition D.1. Let us also add that the possibility to select $q > 1$ close to 1 in our Morrey class hypothesis on the drift is important since it allows us to handle the above many-particle drifts in dimension $d_0 = 2$.

Our results thus provide a wellposedness and Feller theory for such stochastic particle systems when the Morrey norm of the attractive interaction is sufficiently small. Conversely, the nonexistence counterexample exhibited in Proposition 4.1 shows that a sufficiently strong inward Hardy-type position-velocity concentrating force can drive the process into the singular set and prevent its continuation as a weak solution. Let us finally add that our condition on the Morrey norm is d -dependent, and does not yield weak solution theory of particle systems with an N -independent condition on the strength of interaction; in the non-degenerate case, however, something can be said about such general critical interactions when $d_0 \geq 3$ [35].

1.2. Main contributions of the paper. Our first main result, Theorem 3.3, develops the Sobolev theory of (1.1) by constructing the Duhamel series for (1.1) in terms of the fractional powers of $\lambda + \mathcal{K}$ and b . To ensure convergence of the Duhamel series, we obtain Adams-type inequalities [1] on the kinetic homogeneous group: for $1 < p < \infty$,

$$\begin{aligned} \| |b|^{1/p} (\lambda + \mathcal{K})^{-1/(2p)} \|_{p \rightarrow p} &\lesssim \| b \|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p}, \\ \| |b|^{1/p} \nabla_v (\lambda + \mathcal{K})^{-1/2 - 1/(2p)} \|_{p \rightarrow p} &\lesssim \| b \|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p}. \end{aligned}$$

This yields the regularity range in (1.1):

$$\eta \in \mathbb{W}_{\mathcal{K}}^{1/p, p}, \quad f \in \mathbb{W}_{\mathcal{K}}^{-1 - 1/p', p}, \quad \frac{1}{p} + \frac{1}{p'} = 1,$$

(see (2.18) for the definition of $\mathbb{W}_{\mathcal{K}}^{\gamma, p}$) and uniqueness in the corresponding weak class. Since $Q = 4d + 2$, a point mass distribution in $\mathbb{R}^{1+2d}$ belongs to the right-hand-side space whenever

$$1 < p < \frac{Q-1}{Q-2} = \frac{4d+1}{4d},$$

and so, choosing appropriately the source term f we can consider (1.1) with a Dirac delta-function in the initial condition.

The second main result, Theorem 3.9, concerns the kinetic SDE (1.2). Our PDE theory provides kinetic Krylov estimates, existence of a weak solution, conditional weak uniqueness in the Krylov-type class and produces a strongly continuous backward Feller propagator.

1.3. Our approach. 1. The approach to the Sobolev theory of Kolmogorov operators via the fractional representation of the Duhamel/Neumann series appeared in the non-degenerate case in [33]. Therein the second-named author developed regularity theory of elliptic operator $-\Delta + b \cdot \nabla$ with drift b in the class of weakly form-bounded drifts that is, in fact, larger than the corresponding non-degenerate Morrey class with $q > 1$ close to 1. This approach can be viewed as a rather far-reaching extension of the idea behind the Birman-Schwinger principle for Schrödinger operators and is related to much earlier work of Semënov [10] on obtaining characterization of the regularity of the eigenfunctions by working in L^p .

The Sobolev theory of [33] and the ensuing Feller theory were used in [36] to prove for the associated non-degenerate SDEs weak existence and uniqueness among all weak solutions constructed

via regularization of the drift. A parabolic version of [33, 36] was developed in the non-degenerate setting for time-inhomogeneous Morrey drifts in [34] and extended to stable-driven SDEs in [37].

2. We prove conditional weak uniqueness for the kinetic stochastic system (1.2) in the class of weak solutions satisfying Krylov estimates (= reasonable occupation estimates). This result is not contained in [36] – indeed, it requires a parabolic argument – and has been proved earlier in the nondegenerate setting for distributional drifts [62], Lebesgue class drifts [58], and weak Lebesgue class drifts [55]. We also refer to Krylov [42] for recent results in the non-degenerate setting on weak existence and conditional uniqueness for SDEs with drift in the Morrey class with $q > c_d$, $c_d > \frac{d}{2}$, and a VMO diffusion matrix.

3. The maximal-function arguments that we use to prove kinetic Adams estimate build on works [5, 21, 23, 31] that deal with a more general setting of spaces of homogeneous type.

1.4. Literature on kinetic Kolmogorov equations and SDEs. The first result on the operator $\mathcal{K}$ can be traced back to Kolmogorov [38] who gave the explicit Gaussian transition density of integrated Brownian motion and exhibited smoothing despite the absence of diffusion in the position variables. Hörmander’s theorem [32] later placed this phenomenon in a general framework: the velocity diffusion fields and their first commutators with the transport field generate all spatial directions. The intrinsic geometry is therefore governed by noncommutative space–time translations and anisotropic dilations rather than by the Euclidean parabolic structure. Garofalo and Lanconelli [27] related the level sets of the fundamental solution to invariant Harnack inequalities, and Lanconelli and Polidoro [43] developed the associated Lie-group and dilation structure for a broad class of ultraparabolic Kolmogorov operators. These ideas led, among other results, to global lower and Gaussian upper bounds for fundamental solutions [51, 50]. We refer to [6] for a survey of this classical theory.

The regularity theory has developed along several lines. The intrinsic Sobolev–Morrey theory of Polidoro and Ragusa [52] was followed by variational kinetic Sobolev frameworks [4], kinetic maximal L^p estimates [47, 24], and sharp intrinsic Sobolev embeddings [49, 8]. Golse, Imbert, Mouhot and Vasseur extended De Giorgi–Nash–Moser methods to kinetic Fokker–Planck equations with bounded measurable principal coefficients, obtaining Hölder regularity and a Harnack inequality [28]. Fundamental solutions and Gaussian estimates for variable or rough principal coefficients have also been constructed, see, for example, [7]. These works principally concern roughness of the diffusion matrix, local regularity of already constructed weak solutions, or lower-order terms with substantially more integrability than those considered below.

For kinetic SDEs and their backward Kolmogorov equations, Hölder assumptions on the degenerate drift yield weak and strong well-posedness, with sharp or nearly sharp thresholds in several models [15, 16]. Mixed Sobolev assumptions were treated in [26, 60]. In particular, these works use the smoothing properties of the Kolmogorov equation to construct regular stochastic flows or to prove well-posedness of the associated martingale problem. Heat kernel and gradient estimates under low-regularity assumptions were obtained in [17]. Singular coefficients of distributional order have also been treated by paracontrolled methods in [30]. More recently, [19] established weak well-posedness for time-dependent distribution-valued drifts in weighted anisotropic Hölder spaces under a strict subcritical condition and the assumption that $\operatorname{div}_v b$ is bounded.

Uniqueness of probability solutions to Fokker–Planck–Kolmogorov equations with partially degenerate diffusion matrices and unbounded coefficients was studied, under suitable additional assumptions, by Bogachev, Röckner and Shaposhnikov [12].

Singular interaction kernels depending jointly on position and velocity also arise in the work of Hao, Jabir, Menozzi, Röckner and Zhang [29], who establish quantitative propagation of chaos for

moderately interacting particle approximations of kinetic McKean–Vlasov equations with interaction kernels in anisotropic Besov spaces.

For mixed-Lebesgue drifts in the strictly subcritical range, maximum principles and kinetic Krylov estimates leading to weak existence for even density-dependent kinetic SDEs were established in [61]. The scaling-critical Hardy fields considered below are outside those mixed-Lebesgue hypotheses.

The closest result to the present work is that of Ren and Zhang [54]. For kinetic SDEs with Hölder diffusion and an unbounded velocity drift in an intrinsic kinetic Kato class, they construct a weak solution and prove uniqueness within a class requiring the one-time marginal laws to be absolutely continuous for almost every time. They also obtain sharp two-sided heat kernel bounds under a strengthened Kato assumption. Their hypotheses include mixed Lebesgue drifts in the strictly subcritical range

$$\frac{2}{p_t} + \frac{d}{p_v} + \frac{3d}{p_x} < 1.$$

The present paper addresses a different scaling-critical endpoint. A drift of kinetic degree -1 , such as e.g. Lebesgue drift

$$\frac{2}{p_t} + \frac{d}{p_v} + \frac{3d}{p_x} = 1$$

or having even stronger singularity, such as $|v|^{-1}$, $|t - t_0|^{-1/2}$, or $|x - x_0|^{-1/3}$, does not belong to the Kato class, but it belongs to the critical kinetic Morrey class of this work. Our drifts are not required to be divergence-free or have gradient structure.

Let us add that, very recently, Chen and Ren [20] obtained PDE well-posedness for anisotropic mixed-Lebesgue drifts in a scaling-supercritical range by exploiting the cancellation $\operatorname{div}_v b = 0$. This result is complementary to ours since its cancellation mechanism is unavailable for attracting drifts, or more generally the drifts considered in the present paper.

Notations.

- We write

$$z = (x, v) \in \mathbb{R}^{2d}, \quad \zeta = (t, z) = (t, x, v) \in \mathbb{R}^{1+2d}, \quad \xi = (s, y, w) \in \mathbb{R}^{1+2d},$$

and use

$$dz = dx dv, \quad d\zeta = dt dx dv.$$

- The symbol $|\cdot|$ denotes the Euclidean norm when applied to a vector and the Lebesgue measure when applied to a measurable set. We write $\mathbf{1}_E$ for the indicator function of a measurable set E . When acting on a space–time function, $\mathbf{1}_{[s,T]}$ means $\mathbf{1}_{[s,T] \times \mathbb{R}^{2d}}$.
- L^p denotes the Lebesgue space

$$L^p(\mathbb{R}^{1+2d}, d\zeta), \quad \|f\|_p := \|f\|_{L^p(\mathbb{R}^{1+2d})}.$$

The same notation is used for finite-dimensional vector-valued functions, with $|f|$ denoting their Euclidean norm. We write L^p_{loc} for the space of measurable functions f such that

$$\|\mathbf{1}_K f\|_p < \infty \quad \text{for every compact } K \subset \mathbb{R}^{1+2d}.$$

The kinetic metric $d_*(\zeta, \xi)$ introduced below (see (2.4)) induces the usual Euclidean topology on $\mathbb{R}^{1+2d}$, so this definition is equivalent to requiring local L^p -integrability on every kinetic ball.

- For $m \geq 1$,

$$C_\infty(\mathbb{R}^m) := \left\{ f \in C_b(\mathbb{R}^m) : \lim_{|y| \rightarrow \infty} f(y) = 0 \right\},$$

endowed with the supremum norm.

- $C_c^\infty(\mathbb{R}^m)$ denotes the space of smooth compactly supported functions.

- $\mathcal{S}(\mathbb{R}^m)$ denotes the Schwartz space (i.e. the set of all $f \in C^\infty$ such that $\sup_{y \in \mathbb{R}^m} |y|^l |D^\alpha f(y)| < \infty$ for all $l \geq 1, \alpha \in \mathbb{N}^m$).
- For $1 \leq p \leq \infty$, the conjugate exponent is denoted by p' :

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

- If $T : L^p \rightarrow L^r$ is bounded, we write $\|T\|_{p \rightarrow r} := \sup_{\|f\|_p \leq 1} \|Tf\|_r$. In particular, $\|T\|_{p \rightarrow p}$ denotes the operator norm of T on L^p .
- For brevity, we often make use of the notation

$$\langle h \rangle := \int_{\mathbb{R}^{1+2d}} h d\zeta, \quad \langle f, g \rangle := \langle fg \rangle.$$

Throughout this paper, all functions and vector fields are real-valued.

- We write $A \lesssim B$ when $A \leq CB$ for some positive constant C . Constants denoted by c or C may change from line to line.
- For $f \in C([s, T], \mathbb{R}^m)$, we define the uniform modulus of continuity by

$$\mathfrak{w}_{s,T}(f, \delta) := \sup_{\substack{s \leq u, t \leq T \\ |t-u| \leq \delta}} |f(t) - f(u)|.$$

When the interval $[s, T]$ is fixed, we write simply $\mathfrak{w}(f, \delta)$.

- $[M, N]_t$ denotes the cross-variation bracket between two continuous local martingales and $[M]_t = [M, M]_t$ the quadratic variation of M .
- For locally finite Radon measures on a locally compact space, vague convergence $\rightarrow$ means convergence of integrals against every continuous compactly supported test function.
- We also use $\rightarrow$ for weak convergence in function spaces.
- Finally, we denote by $\Rightarrow$ weak convergence of probability measures i.e. for $\{\mu_n\}_n$ and μ probabilities on a metric space E ,

$$\mu_n \Rightarrow \mu \quad \Leftrightarrow \quad \int_E f d\mu_n \rightarrow \int_E f d\mu \quad \text{for every } f \in C_b(E).$$

2. KINETIC GEOMETRY AND KINETIC BESSEL POTENTIALS

2.1. The kinetic group. For $\zeta = (t, x, v)$ and $\xi = (s, y, w)$ in $\mathbb{R}^{1+2d}$, set the composition law

$$\zeta \circ \xi = (t + s, x + y + sv, v + w). \quad (2.1)$$

Then the inverse element is given by

$$\zeta^{-1} = (-t, -x + tv, -v).$$

Lebesgue measure is a left and right Haar measure, and the vector fields

$$Y := \partial_t + v \cdot \nabla_x, \quad X_j := \partial_{v_j}, \quad 1 \leq j \leq d,$$

are left invariant. The maps

$$\delta_r(t, x, v) := (r^2 t, r^3 x, rv), \quad r > 0,$$

are group automorphisms. Lebesgue measure scales by r^Q , with Q as in (1.3), and $\mathcal{K} = Y - \sum_j X_j^2$ is homogeneous of degree two.

Define the left-invariant, generally non-symmetric, distance

$$d(\zeta, \xi) := \max \left\{ |t - s|^{1/2}, |w - v|, |y - x - (s - t)v|^{1/3} \right\}. \quad (2.2)$$

It is immediate from the group law that $d(\zeta, \xi) = d(0, \zeta^{-1} \circ \xi)$. It also satisfies the triangle inequality. Indeed, let $\zeta = (t, x, v)$, $\theta = (r, z, u)$, $\xi = (s, y, w)$, $a = d(\zeta, \theta)$ and $c = d(\theta, \xi)$. Then

$$\begin{aligned} |y - x - (s - t)v| &\leq |z - x - (r - t)v| + |y - z - (s - r)u| + |(s - r)(u - v)| \\ &\leq a^3 + c^3 + ac^2 \leq (a + c)^3. \end{aligned}$$

The time and velocity components are straightforward. Moreover,

$$2^{-1/3}d(\xi, \zeta) \leq d(\zeta, \xi) \leq 2^{1/3}d(\xi, \zeta). \quad (2.3)$$

Consequently,

$$d_*(\zeta, \xi) := d(\zeta, \xi) + d(\xi, \zeta) \quad (2.4)$$

is a left-invariant metric equivalent to d . The corresponding balls $B_r(\zeta) := \{\xi : d_*(\zeta, \xi) < r\}$ satisfy

$$c_d r^Q \leq |B_r(\zeta)| \leq C_d r^Q. \quad (2.5)$$

Thus $(\mathbb{R}^{1+2d}, d_*, dt dx dv)$ is a doubling space of homogeneous type.

With this set-up, we can define the Morrey class $\mathcal{M}_{\mathcal{K}}^q$ equally as in Section 1.

Definition 2.1. Let $q > 1$. A Borel function g belongs to the critical kinetic Morrey class $\mathcal{M}_{\mathcal{K}}^q$ if $g \in L_{\text{loc}}^q$ and

$$\|g\|_{\mathcal{M}_{\mathcal{K}}^q} := \sup_{\zeta \in \mathbb{R}^{1+2d}, r > 0} r \left(\frac{1}{|B_r(\zeta)|} \int_{B_r(\zeta)} |g(\xi)|^q d\xi \right)^{1/q} < \infty.$$

This class where backward cylinders C_r defined in (1.4) are replaced by d_* -balls B_r , is equivalent to the one introduced in (1.5), with an equivalent norm. Indeed, recalling that

$$\mathcal{C}_r(t, x, v) := \{(s, y, w) : -r^2 < s - t < 0, |w - v| < r, |y - x - (s - t)v| < r^3\},$$

we have

$$\mathcal{C}_r(t_0, x_0, v_0) \subset B_{3r}(t_0, x_0, v_0)$$

and, conversely,

$$B_r(t_0, x_0, v_0) \subset \mathcal{C}_{\sqrt{2}r}(t_0 + r^2, x_0 + r^2 v_0, v_0).$$

2.2. The kinetic Gaussian kernel. Let $\{P_\tau\}_{\tau \geq 0}$ be the contraction semigroup on $L^p(\mathbb{R}^{1+2d})$, $1 \leq p \leq \infty$, generated by $-\mathcal{K}$. Explicitly,

$$(P_\tau f)(t, x, v) = \int_{\mathbb{R}^{2d}} p_\tau(x, v; y, w) f(t - \tau, y, w) dy dw, \quad (2.6)$$

where

$$C_\tau = \begin{pmatrix} \tau I_d & -\frac{\tau^2}{2} I_d \\ -\frac{\tau^2}{2} I_d & \frac{\tau^3}{3} I_d \end{pmatrix}, \quad \Xi_\tau = \begin{pmatrix} w - v \\ y - x + \tau v \end{pmatrix},$$

and

$$p_\tau(x, v; y, w) = \left(\frac{\sqrt{3}}{2\pi\tau^2} \right)^d \exp \left(-\frac{1}{4} \langle C_\tau^{-1} \Xi_\tau, \Xi_\tau \rangle \right). \quad (2.7)$$

The above corresponds to the transition density of

$$dX_r = -V_r dr, \quad dV_r = \sqrt{2} dW_r,$$

started from (x, v) . It follows directly from (2.7) that, for suitable positive constants $c_i = c_i(d)$,

$$\begin{aligned} p_\tau(x, v; y, w) &\leq c_1 \tau^{-2d} \exp \left[-c_2 \left(\frac{|w - v|^2}{\tau} + \frac{|y - x + \tau v|^2}{\tau^3} \right) \right], \\ |\nabla_v p_\tau(x, v; y, w)| &\leq c_3 \tau^{-1/2} \tilde{p}_\tau(x, v; y, w), \end{aligned} \quad (2.8)$$

where $\tilde{p}_\tau$ is a Gaussian of the same form as in the upper-bound with a smaller constant in the exponential. Analogous estimates hold for the adjoint semigroup.

For $\lambda > 0$ and $\mathcal{K}$ the forward Kolmogorov operator, define the operator $\mathcal{A}$ and its adjoint $\mathcal{A}^*$ as

$$\mathcal{A} := \lambda + \mathcal{K}, \quad \mathcal{A}^* = \lambda + \mathcal{K}^* = \lambda - \partial_t - v \cdot \nabla_x - \Delta_v.$$

(Although $\mathcal{A}$ depends on the choice of λ , for notation convenience, we make this dependency implicit in our writing.)

For $\gamma > 0$, define the negative fractional powers by

$$\mathcal{A}^{-\gamma/2} f = \frac{1}{\Gamma(\gamma/2)} \int_0^\infty \tau^{\gamma/2-1} e^{-\lambda\tau} \mathcal{P}_\tau f d\tau, \quad (2.9)$$

and define $(\mathcal{A}^*)^{-\gamma/2}$ and $\mathcal{A}^{\gamma/2} = (\mathcal{A}^{-\gamma/2})^{-1}$ analogously. In particular,

$$\mathcal{A}^{-1} f(t, x, v) = \int_{-\infty}^t e^{-\lambda(t-s)} \int_{\mathbb{R}^{2d}} p_{t-s}(x, v; y, w) f(s, y, w) dy dw ds.$$

Fundamentals on fractional powers of operators, the specific derivations of $\mathcal{A}^{\gamma/2}$ and other key properties used throughout are collected in Appendix A and below.

In particular, we write $\mathcal{A}^{-\gamma/2}(\cdot, \xi) := \mathcal{A}^{-\gamma/2} \delta_{\{\xi\}}$ and $(\mathcal{A}^*)^{-\gamma/2}(\cdot, \xi) := (\mathcal{A}^*)^{-\gamma/2} \delta_{\{\xi\}}$, where $\delta_{\{\xi\}}$ is the Dirac mass at $\xi \in \mathbb{R}^{1+2d}$.

Proposition 2.2 (Fractional kernel bounds). *Let $0 < \gamma < Q$, $1 < r < \infty$, and $\lambda > 0$. The following assertions hold, with constants independent of λ in the pointwise estimates.*

(i) *The integral kernels of $\mathcal{A}^{-\gamma/2}$ and $(\mathcal{A}^*)^{-\gamma/2}$ satisfy*

$$|\mathcal{A}^{-\gamma/2}(\zeta, \xi)| + |(\mathcal{A}^*)^{-\gamma/2}(\zeta, \xi)| \leq C_{d,\gamma} d_*(\zeta, \xi)^{\gamma-Q}. \quad (2.10)$$

(ii) *The differentiated kernels satisfy*

$$\begin{aligned} & |\nabla_v \mathcal{A}^{-(1+\gamma)/2}(\zeta, \xi)| \\ & + |\nabla_v (\mathcal{A}^*)^{-(1+\gamma)/2}(\zeta, \xi)| \leq C_{d,\gamma} d_*(\zeta, \xi)^{\gamma-Q}. \end{aligned} \quad (2.11)$$

(iii) *One has*

$$\|\mathcal{A}^{-\gamma/2}\|_{r \rightarrow r} + \|(\mathcal{A}^*)^{-\gamma/2}\|_{r \rightarrow r} \leq 2\lambda^{-\gamma/2}, \quad (2.12)$$

$$\|\nabla_v \mathcal{A}^{-(1+\gamma)/2}\|_{r \rightarrow r} + \|\nabla_v (\mathcal{A}^*)^{-(1+\gamma)/2}\|_{r \rightarrow r} \leq C_{d,r,\gamma} \lambda^{-\gamma/2}. \quad (2.13)$$

For the sake of completeness, the proof of Proposition 2.2, which essentially relies on suitable estimates on $\{\mathcal{P}_\tau\}_{\tau \geq 0}$, is included in Appendix B.

For later use, define the intrinsic Riesz potential

$$I_\gamma f(\zeta) := \int_{\mathbb{R}^{1+2d}} \frac{|f(\xi)|}{d_*(\zeta, \xi)^{Q-\gamma}} d\xi, \quad 0 < \gamma < Q. \quad (2.14)$$

Clearly, I_γ is positively homogeneous and subadditive. On nonnegative functions it coincides with the action of the linear operator having the same integral kernel. Thus Proposition 2.2 gives the pointwise bounds

$$|\mathcal{A}^{-\gamma/2} f| + |(\mathcal{A}^*)^{-\gamma/2} f| \leq C I_\gamma f, \quad (2.15)$$

$$|\nabla_v \mathcal{A}^{-(1+\gamma)/2} f| + |\nabla_v (\mathcal{A}^*)^{-(1+\gamma)/2} f| \leq C I_\gamma f. \quad (2.16)$$

Proposition 2.3 (Kinetic Sobolev embeddings). *Let $0 < \gamma < Q$.*

(i) *If $1 < r < Q/\gamma$ and $1/s = 1/r - \gamma/Q$, then*

$$\|\mathcal{A}^{-\gamma/2} f\|_s + \|(\mathcal{A}^*)^{-\gamma/2} f\|_s \leq C \|f\|_r. \quad (2.17)$$

(ii) If $1 < r < \infty$ and $\gamma > Q/r$, then $\mathcal{A}^{-\gamma/2}$ and $(\mathcal{A}^*)^{-\gamma/2}$ map L^r continuously into $C_\infty(\mathbb{R}^{1+2d})$.

We included the proof of Proposition 2.3 in Appendix C.

2.3. Kinetic Bessel-potential spaces. Fix $\lambda > 0$ and $1 < p < \infty$. For $\gamma > 0$, define

$$\mathbb{W}_{\mathcal{K}}^{\gamma,p} := \mathcal{A}^{-\gamma/2} L^p(\mathbb{R}^{1+2d}), \quad \|u\|_{\mathbb{W}_{\mathcal{K}}^{\gamma,p}} := \|\mathcal{A}^{\gamma/2} u\|_p.$$

For $\gamma < 0$, let $\mathbb{W}_{\mathcal{K}}^{\gamma,p}$ be the completion of L^p in the norm

$$\|f\|_{\mathbb{W}_{\mathcal{K}}^{\gamma,p}} := \|\mathcal{A}^{\gamma/2} f\|_p. \quad (2.18)$$

Define $\mathbb{W}_{\mathcal{K}^*}^{\gamma,p}$ using $\mathcal{A}^*$ in the same way. We have the duality relation

$$(\mathbb{W}_{\mathcal{K}}^{\gamma,p})^* = \mathbb{W}_{\mathcal{K}^*}^{-\gamma,p'}, \quad (2.19)$$

with the distributional pairing extending the usual L^2 pairing. These spaces refine the scale of kinetic Sobolev spaces; for further details see [47, 8].

The operator $\mathcal{A}$ is an isometric isomorphism

$$\mathcal{A} : \mathbb{W}_{\mathcal{K}}^{\gamma,p} \rightarrow \mathbb{W}_{\mathcal{K}}^{\gamma-2,p}.$$

In particular, if $f \in \mathbb{W}_{\mathcal{K}}^{-1-1/p',p} = \mathbb{W}_{\mathcal{K}}^{1/p-2,p}$, then $\mathcal{A}^{-(1-1/(2p))} f \in L^p$.

The following consequence of Proposition 2.3 will be needed in order to consider a delta-function in the initial condition for the forward equation.

Corollary 2.4 (Point evaluation). *If $s > Q/p'$, then*

$$\mathbb{W}_{\mathcal{K}^*}^{s,p'} \hookrightarrow C_\infty(\mathbb{R}^{1+2d}),$$

and therefore

$$\delta_{\{\zeta_0\}} \in \mathbb{W}_{\mathcal{K}}^{-s,p} \quad \text{for every } \zeta_0 \in \mathbb{R}^{1+2d}.$$

Proof. Choose γ with $Q/p' < \gamma < \min\{s, Q\}$. The factorization $(\mathcal{A}^*)^{-s/2} = (\mathcal{A}^*)^{-\gamma/2} (\mathcal{A}^*)^{-(s-\gamma)/2}$, Proposition 2.2(iii), and Proposition 2.3(ii) give the asserted embedding for every $s > Q/p'$, including $s \geq Q$. Owing to the duality $C_\infty^*(\mathbb{R}^{1+2d}) \hookrightarrow (\mathbb{W}_{\mathcal{K}^*}^{s,p'})^*$, we have $\delta_{\{\zeta_0\}} \in \mathbb{W}_{\mathcal{K}}^{-s,p}$. $\square$

3. MAIN RESULTS

3.1. Hypothesis, smallness condition and Duhamel blocks.

Hypothesis ($\mathcal{H}_{\mathcal{K}}$). There exist $q > 1$ and a decomposition of the drift into a singular and a bounded part, respectively:

$$b = b_s + b_b, \quad b_s \in \mathcal{M}_{\mathcal{K}}^q, \quad b_b \in L^\infty(\mathbb{R}^{1+2d}, \mathbb{R}^d). \quad (\mathcal{H}_{\mathcal{K}})$$

The norm $\|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}$ will be chosen sufficiently small later on; see $(\mathcal{H}_{s,p})$.

Define the power drift

$$b^{1/p'} := b|b|^{-1+1/p'} \quad \text{on } \{0 < |b| < \infty\}$$

and set it equal to zero elsewhere. For $1 < p < \infty$, define, on smooth test functions, the operators

$$H_p(b) := |b|^{1/p} \mathcal{A}^{-1/(2p)}, \quad (3.1)$$

$$G_p(b) := -\mathcal{A}^{-(1-1/(2p))} \operatorname{div}_v(b^{\frac{1}{p'}} \cdot), \quad (3.2)$$

which will serve as the building blocks for the Duhamel series for η , solution to (1.1).

As shown in Proposition 5.5, $H_p(b), G_p(b)$ are bounded operators on L^p and

$$\|H_p(b)\|_{p \rightarrow p} \leq C_{d,p,q} \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p} + C_{d,p} \lambda^{-1/(2p)} \|b_b\|_{\infty}^{1/p} =: h_{p,\lambda}(b) \quad (3.3)$$

$$\|G_p(b)\|_{p \rightarrow p} = \|G_p^*(b)\|_{p' \rightarrow p'} \leq C_{d,p',q} \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p'} + C_{d,p'} \lambda^{-1/(2p')} \|b_b\|_{\infty}^{1/p'} =: g_{p,\lambda}(b), \quad (3.4)$$

where the adjoint of $G_p(b)$ is

$$G_p^*(b) = b^{1/p'} \cdot \nabla_v (\mathcal{A}^*)^{-(1-1/(2p))}.$$

We will use the same notation for the extension of these operators by continuity to L^p .

(Let us emphasize that in the kinetic setting one cannot commute the gradient ∇_v and the free resolvent since $[\partial_{v_j}, \partial_t + v \cdot \nabla_x] = \partial_{x_j}$.)

For fixed $1 < p < \infty$, we can state our smallness condition on the drift b : there exists $\lambda > 0$ such that

$$g_{p,\lambda}(b) h_{p,\lambda}(b) < 1. \quad (\text{H}_{s,p})$$

In particular, by (3.3)-(3.4), for every fixed d, p, q , there is $c_{d,p,q} > 0$ such that $(\text{H}_{s,p})$ holds for all sufficiently large λ whenever

$$\|b_s\|_{\mathcal{M}_{\mathcal{K}}^q} < c_{d,p,q}.$$

3.2. Regularization of the drift.

Definition 3.1. A sequence of bounded smooth vector fields $\{b_n\}_n$ is an admissible (or p -admissible) regularization of b if

- (i) $b_n \rightarrow b$ almost everywhere and in L_{loc}^q where q is the exponent in the decomposition $(\text{H}_{\mathcal{K}})$;
- (ii) For all n , $b_n \in L^\infty + \mathcal{M}_{\mathcal{K}}^q$ where the singular and bounded parts satisfy, uniformly in n , the bounds in $(\text{H}_{s,p})$ with b_n in place of b , namely,

$$\sup_n g_{p,\lambda}(b_n) h_{p,\lambda}(b_n) < 1;$$

- (iii) As $n \rightarrow \infty$,

$$H_p(b_n) \rightarrow H_p(b), \quad G_p(b_n) \rightarrow G_p(b) \quad \text{strongly on } L^p. \quad (3.5)$$

Proposition 3.2 (Existence of admissible regularizations). *Let b satisfy $(\text{H}_{\mathcal{K}})$ and $(\text{H}_{s,p})$. Then there is an admissible regularization $\{b_n\}_n$ for b for which*

$$\|b_{n,s}\|_{\mathcal{M}_{\mathcal{K}}^q} \leq \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}, \quad \|b_{n,b}\|_{\infty} \leq \|b_b\|_{\infty}. \quad (3.6)$$

Proof. We separate truncation from mollification. This is useful because the operators in (3.1)-(3.2) depend nonlinearly on the total vector field b , rather than separately on the two terms in $(\text{H}_{\mathcal{K}})$.

Choose a sequence of cut-offs $\chi_n \in C_c^\infty(\mathbb{R}^{1+2d})$ such that $0 \leq \chi_n \leq 1$, $\chi_n = 1$ on $B_n(0)$, and $\chi_n \rightarrow 1$ pointwise. With the $\min\{\cdot, \cdot\}$ below interpreted as equal to 1 on $\{b = 0\}$, let us put for a.e. $\zeta \in \mathbb{R}^{1+2d}$,

$$a_n(\zeta) := \chi_n(\zeta) \min \left\{ 1, \frac{n}{|b(\zeta)|} \right\}, \quad c_n := a_n b, \quad c_{n,s} := a_n b_s, \quad c_{n,b} := a_n b_b. \quad (3.7)$$

Then $c_n = c_{n,s} + c_{n,b}$, $0 \leq a_n \leq 1$, and $a_n \rightarrow 1$ almost everywhere. Moreover,

$$|c_n| \leq n, \quad |c_{n,s}| \leq n + \|b_b\|_{\infty}, \quad |c_{n,b}| \leq \|b_b\|_{\infty}.$$

Thus both components are bounded and compactly supported. We also have, clearly,

$$\|c_{n,s}\|_{\mathcal{M}_{\mathcal{K}}^q} \leq \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}, \quad \|c_{n,b}\|_{\infty} \leq \|b_b\|_{\infty}. \quad (3.8)$$

Since $a_n \rightarrow 1$ and $0 \leq a_n \leq 1$, applying the dominated convergence theorem on compact sets also shows that

$$c_{n,s} \rightarrow b_s \quad \text{in } L_{\text{loc}}^q, \quad c_{n,b} \rightarrow b_b \quad \text{in } L_{\text{loc}}^q, \quad (3.9)$$

as well as a.e.

The reason for using the same scalar a_n in both components $c_{n,s}$ and $c_{n,b}$ is explained by observing the pair of exact identities: for all $f \in L^p$,

$$\begin{aligned} H_p(c_n)f &= a_n^{1/p} H_p(b)f, \\ G_p(c_n)f &= G_p(b)(a_n^{1/p'} f). \end{aligned} \quad (3.10)$$

Indeed, the first identity follows immediately from $|c_n|^{1/p} = a_n^{1/p} |b|^{1/p}$. For the second one, we have $c_n^{1/p'} = a_n^{1/p'} b^{1/p'}$ from which we note that, for $f \in L^p$ and $\varphi \in L^{p'}$, we have

$$\begin{aligned} \langle G_p(c_n)f, \varphi \rangle &= \int f c_n^{1/p'} \cdot \nabla_v(\mathcal{A}^*)^{-(1-1/(2p))} \varphi d\zeta \\ &= \int a_n^{1/p'} f b^{1/p'} \cdot \nabla_v(\mathcal{A}^*)^{-(1-1/(2p))} \varphi d\zeta \\ &= \langle G_p(b)(a_n^{1/p'} f), \varphi \rangle. \end{aligned}$$

The first identity in (3.10), dominated convergence in L^p , and the boundedness of $H_p(b)$ imply

$$H_p(c_n)f \rightarrow H_p(b)f \quad \text{in } L^p$$

for every $f \in L^p$. The second identity and $a_n^{1/p'} f \rightarrow f$ in L^p similarly imply

$$G_p(c_n)f \rightarrow G_p(b)f \quad \text{in } L^p.$$

After this truncation, we next apply a mollifier. Let $\rho \geq 0$ be smooth, supported in $B_1(0)$ and having integral equal to 1, and put $\rho_\varepsilon(\zeta) := \varepsilon^{-Q} \rho(\delta_{1/\varepsilon} \zeta) = \varepsilon^{-Q} \rho(t/\varepsilon^2, x/\varepsilon^3, v/\varepsilon)$. For the mollification, we will use the convolution along the composition law (2.1):

$$(\rho_\varepsilon * g)(\zeta) := \int \rho_\varepsilon(\eta) g(\eta^{-1} \circ \zeta) d\eta.$$

This convolution is a contraction in the critical Morrey norm $\|\cdot\|_{\mathcal{M}_{\mathcal{K}}^q}$. In fact, left invariance of the related distance d_* and of Lebesgue measure gives $\xi^{-1} \circ B_r(\zeta) = B_r(\xi^{-1} \circ \zeta)$, and hence

$$\|\rho_\varepsilon * g\|_{L^q(B_r(\zeta))} \leq \int \rho_\varepsilon(\xi) \|g(\xi^{-1} \circ \cdot)\|_{L^q(B_r(\zeta))} d\xi = \int \rho_\varepsilon(\xi) \|g\|_{L^q(B_r(\xi^{-1} \circ \zeta))} d\xi.$$

After multiplication by $r|B_r|^{-1/q}$ and taking the supremum over r and ζ , this yields

$$\|\rho_\varepsilon * g\|_{\mathcal{M}_{\mathcal{K}}^q} \leq \|g\|_{\mathcal{M}_{\mathcal{K}}^q} \quad \text{and} \quad \|\rho_\varepsilon * g\|_\infty \leq \|g\|_\infty. \quad (3.11)$$

For fixed n , write now $c_{n,\varepsilon} := \rho_\varepsilon * c_n$ for which we have the further decomposition $c_{n,\varepsilon} = \rho_\varepsilon * c_{n,b} + \rho_\varepsilon * c_{n,s} \in L^\infty + \mathcal{M}_{\mathcal{K}}^q$. The approximation of the identity gives $c_{n,\varepsilon} \rightarrow c_n$ in every finite local Lebesgue space and almost everywhere as $\varepsilon \downarrow 0$. For $0 < \varepsilon \leq 1$, all these functions are uniformly bounded and supported in a fixed compact set depending on n . It follows, first for any $f \in C_c^\infty$ (which, we may recall, is dense in L^p), that

$$H_p(c_{n,\varepsilon})f \rightarrow H_p(c_n)f \quad \text{in } L^p. \quad (3.12)$$

Indeed, the difference being $(|c_{n,\varepsilon}|^{1/p} - |c_n|^{1/p}) \mathcal{A}^{-1/(2p)} f$,

$$\|H_p(c_{n,\varepsilon})f - H_p(c_n)f\|_p \leq \| |c_{n,\varepsilon} - c_n|^{1/p} \mathcal{A}^{-1/(2p)} f \|_p$$

and the convergence to 0 follows next from (2.12) in Proposition 2.2(iii) and the dominated convergence theorem (recalling that $|c_n| \leq n$ and so $|c_{n,\varepsilon}| \leq n$).

To handle the second limit $G_p(c_{n,\varepsilon}) \rightarrow G_p(c_n)$ in L^p , let us introduce the operator

$$D_p := -\mathcal{A}^{-(1-1/(2p))} \operatorname{div}_v,$$

for which we have $G_p(c_n)(f) = D_p(c_n^{\frac{1}{p'}} f)$. Its adjoint is $D_p^* = \nabla_v(\mathcal{A}^*)^{-(1+1/p')/2}$. Then (2.13) in Proposition 2.2(iii) with $\gamma = 1/p'$ and the duality $\|D_p F\|_p = \sup_{\|G\|_{p'}=1} \langle F, D_p^* G \rangle$ give

$$\|D_p F\|_p \leq C_{d,p} \lambda^{-1/(2p')} \|F\|_p, \quad F \in L^p(\mathbb{R}^{1+2d}, \mathbb{R}^d).$$

Therefore, for $f \in C_c^\infty$, as $\varepsilon \rightarrow 0$,

$$\|G_p(c_{n,\varepsilon})f - G_p(c_n)f\|_p \leq C_{d,p} \lambda^{-1/(2p')} \|(c_{n,\varepsilon}^{\frac{1}{p'}} - c_n^{\frac{1}{p'}})f\|_p \rightarrow 0,$$

follows again by dominated convergence. Uniform bounds and density extend (3.12) and the corresponding convergence for G_p to every fixed $f \in L^p$.

It remains to choose one mollification scale in terms of n to eventually set c_{n,ε_n} as an admissible approximation. First note that we can choose $\varepsilon_n \downarrow 0$ rapidly such that

$$\|\rho_{\varepsilon_n} * c_{n,s} - c_{n,s}\|_{L^q(B_n)}^q + \|\rho_{\varepsilon_n} * c_{n,b} - c_{n,b}\|_{L^q(B_n)}^q < 2^{-n}. \quad (3.13)$$

Then let $\{f_j : j \geq 1\} \subset C_c^\infty(\mathbb{R}^{1+2d})$ be a countable dense subset of L^p . By the preceding operator convergences, we may choose $\varepsilon_n \downarrow 0$, with $0 < \varepsilon_n < 1/n$, such that for $1 \leq j \leq n$,

$$\begin{aligned} & \| (H_p(\rho_{\varepsilon_n} * c_n) - H_p(c_n)) f_j \|_p \\ & + \| (G_p(\rho_{\varepsilon_n} * c_n) - G_p(c_n)) f_j \|_p < 2^{-n}, \end{aligned} \quad (3.14)$$

and (3.13) holds simultaneously. Define

$$b_{n,s} := \rho_{\varepsilon_n} * c_{n,s}, \quad b_{n,b} := \rho_{\varepsilon_n} * c_{n,b}, \quad b_n := b_{n,s} + b_{n,b} = \rho_{\varepsilon_n} * c_n.$$

These vector fields are smooth, bounded, and compactly supported. Equations (3.9) and (3.13) give convergence of the two components in L_{loc}^q . (3.13), applied on each fixed ball, also gives $b_{n,s} - c_{n,s} \rightarrow 0$ and $b_{n,b} - c_{n,b} \rightarrow 0$ almost everywhere. Hence $b_n \rightarrow b$ almost everywhere.

Combining (3.10) with (3.14) proves (3.5) first on the dense set $\{f_j\}$ and then on all of L^p , using the uniform operator bounds. Finally, (3.8) and (3.11) give the stronger estimates

$$\|b_{n,s}\|_{\mathcal{M}_{\mathcal{K}}^q} \leq \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}, \quad \|b_{n,b}\|_\infty \leq \|b_b\|_\infty.$$

Thus (3.6) holds, and

$$\sup_n \mathfrak{g}_{p,\lambda}(b_n) \mathfrak{h}_{p,\lambda}(b_n) \leq \mathfrak{g}_{p,\lambda}(b) \mathfrak{h}_{p,\lambda}(b) < 1.$$

This proves admissibility. $\square$

3.3. Forward and backward Kolmogorov equations. We can now state our main results on the solution to Equation 1.1.

Theorem 3.3 (Forward PDE theory). *Let $d \geq 2$ and b satisfy (H $_{\mathcal{K}}$). Fix $1 < p < \infty$ and assume (H $_{s,p}$). Then:*

- (i) *For every $f \in \mathbb{W}_{\mathcal{K}}^{-1-1/p',p}$, the equation*

$$\mathcal{A}\eta - \operatorname{div}_v(b\eta) = f$$

has a unique weak solution $\eta \in \mathbb{W}_{\mathcal{K}}^{1/p,p}$ (see Definition 3.4), and, with $H_p(b)$ and $G_p(b)$ as in (3.1)-(3.2), the following Duhamel series representation holds:

$$\eta = \mathcal{A}^{-1/(2p)} (I + G_p(b)H_p(b))^{-1} \mathcal{A}^{-(1-1/(2p))} f. \quad (3.15)$$

Moreover,

$$\|\eta\|_{\mathbb{W}_{\mathcal{K}}^{1/p,p}} \leq \frac{1}{1 - \|G_p(b)\|_{p \rightarrow p} \|H_p(b)\|_{p \rightarrow p}} \|f\|_{\mathbb{W}_{\mathcal{K}}^{-1-1/p',p}}; \quad (3.16)$$

- (ii) Let $\{b_n\}_n$ be any admissible regularization of b as is introduced in Section 3.2, and let $f \in C_c^\infty(\mathbb{R}^{1+2d})$. Then η_n defined by (3.15) (with b_n in place of b) is a classical solution to the equation

$$\mathcal{A}\eta_n - \operatorname{div}_v(b_n\eta_n) = f,$$

and

$$\eta_n \rightarrow \eta \quad \text{in } \mathbb{W}_{\mathcal{K}}^{1/p,p} \text{ as } n \rightarrow \infty.$$

Thus, η does not depend on the choice of an admissible regularization $\{b_n\}$ of b .

- (iii) If

$$1 < p < \frac{Q-1}{Q-2} = \frac{4d+1}{4d}, \quad (3.17)$$

then, for every $\zeta_0 = (t_0, x_0, v_0) \in \mathbb{R}^{1+2d}$, the following Cauchy problem for the kinetic Fokker–Planck equation

$$\begin{cases} \lambda\eta + \partial_t\eta + v \cdot \nabla_x\eta - \Delta_v\eta - \operatorname{div}_v(b\eta) = 0, & t > t_0, \quad (x, v) \in \mathbb{R}^{2d}, \\ \eta|_{t=t_0} = \delta_{\{(x_0, v_0)\}}, \end{cases}$$

has a unique weak solution η whose extension by zero to $t < t_0$ belongs to $\mathbb{W}_{\mathcal{K}}^{1/p,p}$. This extension corresponds to choosing $f = \delta_{\{\zeta_0\}}$ in (i), and the convergence assertion in (ii) remains valid. The resulting solution gives the forward fundamental solution issued from ζ_0 .

The second assertion can be viewed as an approximation uniqueness result.

Definition 3.4. Let $f \in \mathbb{W}_{\mathcal{K}}^{1/p-2,p}$. A function $\eta \in \mathbb{W}_{\mathcal{K}}^{1/p,p}$ is a $\mathbb{W}_{\mathcal{K}}^{1/p,p}$ -weak solution of

$$\mathcal{A}\eta - \operatorname{div}_v(b\eta) = f \quad (3.18)$$

if

$$\langle \eta, \mathcal{A}^*\varphi \rangle + \int_{\mathbb{R}^{1+2d}} b\eta \cdot \nabla_v\varphi = \langle f, \varphi \rangle \quad \text{for every } \varphi \in \mathcal{S}(\mathbb{R}^{1+2d}). \quad (3.19)$$

We may highlight that the second term on the left hand side of (3.19) is finite. Indeed, write $h = \mathcal{A}^{1/(2p)}\eta \in L^p$. Then, for H_p as in (3.1), $|b|^{1/p}\eta = |b|^{1/p}\mathcal{A}^{-1/(2p)}(\mathcal{A}^{1/(2p)}\eta) = H_p(b)h \in L^p$. Since $b \in L_{\text{loc}}^1$, Hölder's inequality gives $b\eta \in L_{\text{loc}}^1$ and further

$$\int_{\mathbb{R}^{1+2d}} |b\eta \cdot \nabla_v\varphi| \leq \int_{\mathbb{R}^{1+2d}} |b||b|^{-1/p}|\nabla_v\varphi||H_p(b)h| \leq \left(\int_{\mathbb{R}^{1+2d}} \underbrace{|b|^{p'(1-1/p)}}_{=|b|} |\nabla_v\varphi|^{p'} \right)^{1/p'} \|H_p(b)h\|_p.$$

Using the definition of the Morrey norm and taking into account that the drift also has the bounded component, we conclude at most polynomial growth of $\int_{B_R(0)} |b|$ as $R \rightarrow \infty$. Therefore, $\int |b||\nabla_v\varphi|^{p'} < \infty$ for $\varphi \in \mathcal{S}$, and the same weighted Hölder inequality proves absolute convergence of the drift pairing in (3.19).

Lemma 3.5 (Factorization of the weak equation). *Let $(H_{s,p})$ hold. A function $\eta \in \mathbb{W}_{\mathcal{K}}^{1/p,p}$ is a weak solution of (3.18) if and only if $h = \mathcal{A}^{1/(2p)}\eta$ satisfies*

$$(I + G_p(b)H_p(b))h = \mathcal{A}^{-(1-1/(2p))}f \quad \text{in } L^p. \quad (3.20)$$

Proof. We have

$$\eta = \mathcal{A}^{-1/(2p)} h, \quad b\eta = b^{\frac{1}{p'}} H_p h.$$

Put $a = 1 - 1/(2p)$. The definition of $G_p(b)$ by duality, tested against Schwartz functions, gives

$$\mathcal{A}^{-a} \operatorname{div}_v(b\eta) = -G_p(b)H_p(b)h \in L^p.$$

The fractional powers preserve Schwartz space, so this is an identity of tempered distributions and it proves $\operatorname{div}_v(b\eta) \in \mathcal{A}^a L^p = \mathbb{W}_{\mathcal{K}}^{1/p-2,p}$, with

$$\|\operatorname{div}_v(b\eta)\|_{\mathbb{W}_{\mathcal{K}}^{1/p-2,p}} \leq \|G_p(b)\|_{p \rightarrow p} \|H_p(b)\|_{p \rightarrow p} \|\eta\|_{\mathbb{W}_{\mathcal{K}}^{1/p,p}}.$$

We can therefore apply $\mathcal{A}^{-(1-1/(2p))}$ to (3.18). Since $(1 - 1/(2p)) + 1/(2p) = 1$, the first term becomes h , while the second becomes $G_p(b)H_p(b)h$ by (3.2). This proves (3.20). All operations are continuous between the Bessel potential spaces introduced in Section 2.3, so this argument is reversible. $\square$

Define

$$R_p(b) = |b|^{1/p} \nabla_v \mathcal{A}^{-1/2-1/(2p)}, \quad Q_p(b) = \mathcal{A}^{-1/(2p')} (b^{\frac{1}{p'}}). \quad (3.21)$$

Definition 3.1, i.e. the notion of an admissible regularization of b , and Proposition 3.2, extend in a straightforward manner to the operators R_p, Q_p .

Theorem 3.6 (Backward PDE theory). *Assume $\|R_p(b)\|_{p \rightarrow p} \|Q_p(b)\|_{p \rightarrow p} < 1$. For every $f \in \mathbb{W}_{\mathcal{K}}^{-1/p',p}$, the equation*

$$\mathcal{A}u + b \cdot \nabla_v u = f$$

has a unique weak solution $u \in \mathbb{W}_{\mathcal{K}}^{1+1/p,p}$ characterized by

$$u = \mathcal{A}^{-1/2-1/(2p)} (I + Q_p R_p)^{-1} \mathcal{A}^{-1/(2p')} f. \quad (3.22)$$

If $p > Q - 1$, then $1 + 1/p > Q/p$, and the kinetic Sobolev embedding makes this solution continuous, possibly after a modification on a measure zero set.

The notion of weak solution for $(\mathcal{A} + b \cdot \nabla_v)u = f$ is analogous to the one in Definition 3.4. The last statement is the pointwise regularity that we will use to establish the Krylov estimates for the kinetic SDE.

3.4. Weak solutions of the kinetic SDE. We now apply the backward PDE theory to

$$\begin{aligned} X_t &= x + \int_s^t V_r dr, \\ V_t &= v - \int_s^t b(r, X_r, V_r) dr + \sqrt{2} (W_t - W_s), \quad s \leq t \leq T. \end{aligned} \quad (3.23)$$

Let $\Omega_{s,T} := C([s, T], \mathbb{R}^{2d})$, $Z_t = (X_t, V_t)$ be the related canonical process, and let $\mathcal{F}_t$ be the completed (with the usual conditions) version of $\sigma(Z_r : s \leq r \leq t)$.

Definition 3.7. Given $s \in [0, T]$, $z = (x, v)$, a weak solution of (3.23) is a probability measure on $\Omega_{s,T}$, such that the canonical process Z satisfies (3.23), for some standard $\mathcal{F}_t$ -Brownian motion $\{W_t - W_s\}_{s \leq t \leq T}$ and such that $\int_s^T |b(r, Z_r)| dr \in L^1(\Omega_{s,T})$.

The wellposedness of this weak solution is equivalently (e.g. [57], [25]) given by the martingale problem related to (3.23) which states as follows.

Definition 3.8. For $0 \leq s \leq T$ and $z = (x, v) \in \mathbb{R}^{2d}$, a probability measure $P_{s,z}$ on $(\Omega_{s,T}, \{\mathcal{F}_t\}_{t \in [s,T]})$ is said to be a solution of the martingale problem for b at the initial time-state (s, z) if under $\mathbb{P}_{s,z}$, $Z_s = z$ a.s., $\mathbb{E}_{s,z}[\int_s^T |b(r, Z_r)| dr] := \mathbb{E}^{\mathbb{P}_{s,z}}[\int_s^T |b(r, Z_r)| dr] < \infty$ and, for all $\varphi \in C_b^1([s, T]; \mathcal{C}_c^2(\mathbb{R}^{2d}))$ and for

$$L_r(\varphi)(r, x, v) := \partial_r \varphi(r, x, v) + v \cdot \nabla_x \varphi(r, x, v) - b(r, x, v) \cdot \nabla_v \varphi(r, x, v) + \Delta_v \varphi(r, x, v),$$

the process

$$M_t^\varphi = \varphi(t, Z_t) - \varphi(s, z) - \int_s^t L_r \varphi(r, Z_r) dr, \quad s \leq t \leq T,$$

is an $\mathcal{F}_t$ -martingale.

We now call an exponent p *backward admissible* when the adjoint counterparts of the two operators in (3.21) have product norm strictly smaller than one; see (8.1) and (8.2) in Section 8. This is again guaranteed by sufficiently small $\|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}$, after increasing λ when a bounded component is present.

We first state a special version of our SDE weak well-posedness result whose proof is shorter and demonstrates the main idea. Here the drift has no separate bounded component and satisfies

$$b = b_s \in \mathcal{M}_{\mathcal{K}}^q \cap L^{1+\varepsilon_0}(\mathbb{R}^{1+2d}, \mathbb{R}^d) \quad \text{for some } \varepsilon_0 > 0. \quad (H_{\mathcal{K}, \varepsilon_0})$$

Recall the notation $Q = 4d + 2$.

Theorem 3.9 (Weak SDE theory under global higher integrability of the drift). *Let $d \geq 2$, $T > 0$, and assume $(H_{\mathcal{K}, \varepsilon_0})$. Let p and p_1 be backward admissible and suppose*

$$p > Q - 1, \quad p_1 > p \frac{q}{q-1}. \quad (3.24)$$

Set $\varepsilon_* = \varepsilon_0/p$. Then, for every starting point (s, z) , the following hold.

- (i) *There exists a weak solution $\mathbb{P}_{s,z}$ of (3.23). It satisfies*

$$\mathbb{E}_{s,z} \int_s^T |b(r, Z_r)|^{1+\varepsilon_*} dr \leq C_T, \quad (3.25)$$

and hence, for all $s \leq u < t \leq T$,

$$\mathbb{E}_{s,z} \int_u^t |b(r, Z_r)| dr \leq C_T^{1/(1+\varepsilon_*)} |t - u|^{\varepsilon_*/(1+\varepsilon_*)}.$$

It also satisfies the unweighted and drift-weighted kinetic Krylov bounds

$$\begin{aligned} \mathbb{E}_{s,z} \int_s^T |h(r, Z_r)| dr &\leq C_T \|\mathbf{1}_{[s,T]} h\|_p, \\ \mathbb{E}_{s,z} \int_s^T |b(r, Z_r) h(r, Z_r)| dr &\leq C_T \|\mathbf{1}_{[s,T]} |b|^{1/p} h\|_p. \end{aligned}$$

- (ii) *The transition operators*

$$P^{s,t} g(z) = \mathbb{E}_{s,z} g(Z_t)$$

form a conservative, strongly continuous backward Feller propagator on $C_\infty(\mathbb{R}^{2d})$, and so $\{\mathbb{P}_{s,z}\}$ is a strong Markov family.

- (iii) *Weak uniqueness holds among solutions satisfying the two Krylov bounds*

$$\begin{aligned} \mathbb{E}_{s,z}^{\mathbb{Q}} \int_s^T |h(r, Z_r)| dr &\leq K_{\mathbb{Q}} \|\mathbf{1}_{[s,T]} h\|_p, \\ \mathbb{E}_{s,z}^{\mathbb{Q}} \int_s^T |b(r, Z_r) h(r, Z_r)| dr &\leq K_{\mathbb{Q}} \|\mathbf{1}_{[s,T]} |b|^{1/p} h\|_p, \end{aligned}$$

for $K_{\mathbb{Q}}$ a constant independent of h .

(iv) If the smallness condition also holds at some

$$1 < r < (Q - 1)/(Q - 2),$$

let $\eta_{s,z}^\lambda$ be the point-source solution given by Theorem 3.3 (taking $f = \delta_{\{(s,z)\}}$). Then the discounted space-time occupation measure $(u, A) \in [s, T] \times \mathcal{B}(\mathbb{R}^{2d}) \mapsto \int_s^u e^{-\lambda(t-s)} \mathbb{P}_{s,z}(Z_t \in A) dt$ satisfies

$$e^{-\lambda(t-s)} \mathbb{P}_{s,z}(Z_t \in d\zeta) dt = \eta_{s,z}^\lambda(t, \zeta) dt d\zeta.$$

In particular, for a.e. $t \in (s, T)$, the law of Z_t has density $e^{\lambda(t-s)} \eta_{s,z}^\lambda(t, \cdot)$.

In the next theorem we remove the global integrability condition $(H_{\mathcal{K}, \varepsilon_0})$ with only a few additional efforts (see Section 9). In particular, we will have to use some polynomial weights. The bounded part of the drift will be absorbed by the choice of sufficiently large λ in the PDE theory of the corresponding Kolmogorov backward equation. Introducing

$$\begin{aligned} \mathfrak{r}_{p,\lambda}^\rho(b) &:= C_{d,p,q,\kappa}^{R,s} \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p} + C_{d,p,\kappa}^{R,b} \lambda^{-1/(2p)} \|b_b\|_\infty^{1/p}, \\ \mathfrak{q}_{p,\lambda}^\rho(b) &:= C_{d,p,q,\kappa}^{Q,s} \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p'} + C_{d,p,\kappa}^{Q,b} \lambda^{-1/(2p')} \|b_b\|_\infty^{1/p'}, \end{aligned}$$

for some constants $C_{d,p,q,\kappa}^{R,s}$, $C_{d,p,\kappa}^{R,b}$, $C_{d,p,q,\kappa}^{Q,s}$ and $C_{d,p,\kappa}^{Q,b}$ (see Lemma 9.1), we have the following wellposedness result.

Theorem 3.10 (Weak SDE theory without global integrability). *Let b satisfy $(H_{\mathcal{K}})$. Suppose $p > Q - 1$ is backward admissible and satisfies the centered weighted product condition*

$$\mathfrak{q}_{p,\lambda}^\rho(b) \mathfrak{r}_{p,\lambda}^\rho(b) < 1,$$

and let p_1 be backward admissible with

$$p_1 > p \frac{q}{q-1}.$$

Choose

$$0 < \varepsilon_* < \min \left\{ 1, \frac{q-1}{p} \right\}.$$

Then all four conclusions (i) to (iv) of Theorem 3.9 hold, with this value of ε_* . In particular, weak existence and the tightness argument require no global L^a assumption on b . No smallness is imposed on b_b .

4. COUNTEREXAMPLE TO WEAK EXISTENCE FOR THE KINETIC SDE (BLOW-UP)

Weak existence for the SDE (3.23) may fail when the kinetic Morrey norm $\|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}$ of the drift is too large. The following position-velocity dependent modification of the classical Hardy drift counterexample, cf. [9, Example 7.4] (in a non-degenerate case), strongly concentrates the dynamics on the singularity set $\{v = x\}$.

Proposition 4.1 (An inward singularity along $\{v = x\}$). *Let $d \geq 2$ and $\kappa > 0$, and define*

$$b_\kappa(t, x, v) = \begin{cases} \kappa \frac{v-x}{|v-x|^2}, & v \neq x, \\ 0, & v = x. \end{cases}$$

Then $b_\kappa \in \mathcal{M}_{\mathcal{K}}^q$ for every $1 < q < d$, with

$$\|b_\kappa\|_{\mathcal{M}_{\mathcal{K}}^q} \leq C_{d,q} \kappa. \quad (4.1)$$

If $\kappa \geq d - 1$, then the SDE (3.23) with $b = b_\kappa$ and initial data $(s, x, v) = (0, 0, 0)$ has no weak solution on any interval $[0, T]$, $T > 0$.

The same nonexistence result holds for the compactly supported drift

$$b_{\kappa, \chi} = \chi b_\kappa, \quad \chi \in C_c^\infty(\mathbb{R}^{1+2d}), \quad 0 \leq \chi \leq 1,$$

provided that $\chi = 1$ on a neighborhood of $(0, 0, 0)$. This drift satisfies

$$b_{\kappa, \chi} \in \mathcal{M}_{\mathcal{H}}^q \cap L^{1+\varepsilon_0}(\mathbb{R}^{1+2d}, \mathbb{R}^d), \quad 1 < q < d, \quad 0 < \varepsilon_0 < d - 1. \quad (4.2)$$

Remark 4.2. The physical interpretation requires us to consider the drift defined in terms of $v - \omega x$ where ω has units of inverse time, rather than $v - x$, but we will skip this. The crux of the argument is that, in any putative weak solution, the singular drift b_κ forces X and V to coincide on a nontrivial time interval, which is incompatible with the finite total variation of X and the infinite total variation of V due to its Brownian component.

Proof. Morrey norm and global integrability. Write $\zeta = (t, x, v)$. By the definition of d_* , every $(s, y, w) \in B_r(\zeta)$ satisfies

$$|s - t| < r^2, \quad |w - v| < r, \quad |y - x - (s - t)v| < r^3.$$

For $q < d$, the elementary estimate

$$\int_{|w-v|<r} |w - y|^{-q} dw \leq C_{d,q} r^{d-q}$$

holds uniformly in y, v and $r > 0$. Indeed, if $|v - y| < 2r$, the integration domain is contained in $\{|w - y| < 3r\}$. If $|v - y| \geq 2r$, then $|w - y| \geq r$ throughout the integration domain.

We integrate first in w , using this estimate, and then in y and s . It follows that

$$\begin{aligned} \int_{B_r(\zeta)} |b_\kappa|^q &\leq \kappa^q \int_{t-r^2}^{t+r^2} \int_{|y-x-(s-t)v|<r^3} \int_{|w-v|<r} |w - y|^{-q} dw dy ds \\ &\leq C_{d,q} \kappa^q r^{2+3d} r^{d-q} = C_{d,q} \kappa^q r^{Q-q}. \end{aligned}$$

Together with (2.5), this proves (4.1). Since $|b_{\kappa, \chi}| \leq |b_\kappa|$, the same Morrey bound holds for $b_{\kappa, \chi}$. The change of variables $(x, v) \mapsto (x, v - x)$ has Jacobian one, and $|v - x|^{-a}$ is locally integrable for every $a < d$. Compact support therefore proves (4.2).

Localization of a hypothetical weak solution. Assume $\kappa \geq d - 1$. We prove nonexistence for $b_{\kappa, \chi}$; the proof for b_κ is identical with $\chi \equiv 1$. Suppose that a weak solution exists on $[0, T]$. Choose $h \in (0, T]$ and $a \in (0, 1]$ such that $\chi = 1$ whenever

$$0 \leq t \leq h, \quad |x| \leq a, \quad |v| \leq a.$$

Define the stopping time

$$\tau = h \wedge \inf\{t \in [0, T] : |X_t| \geq a \text{ or } |V_t| \geq a\},$$

with the convention $\inf \emptyset = \infty$. Continuity and the initial condition $(X_0, V_0) = (0, 0)$ imply $\tau > 0$ almost surely.

Set $Y_t = V_t - X_t$ and $R_t = |Y_t|$. Since $dX_t = V_t dt$, up to time τ we have

$$dY_t = -\kappa \frac{Y_t}{|Y_t|^2} \mathbf{1}_{\{Y_t \neq 0\}} dt - V_t dt + \sqrt{2} dW_t, \quad Y_0 = 0. \quad (4.3)$$

Thus the change of variables introduces the additional drift $-V_t$, which is bounded up to τ . The required integrability $\mathbb{E} \int_0^\tau |b_{\kappa, \chi}(r, X_r, V_r)| dr < \infty$ implies

$$A_t := \int_0^{t \wedge \tau} \frac{\mathbf{1}_{\{R_r > 0\}}}{R_r} dr < \infty \quad \text{almost surely,} \quad 0 \leq t \leq T.$$

All quotients with denominator R_r are assigned the value zero when $R_r = 0$.

Zero occupation time at the singularity. Each coordinate of $Y = V - X$ is a continuous semimartingale. Since X has finite variation, for each j -coordinate, $[Y^{(j)}]_t = [V^{(j)}]_t = 2t$. The occupation-times formula (see e.g. [53, Corollary 1, pp. 219]) for the first coordinate gives

$$2 \int_0^T \mathbf{1}_{\{Y_r^{(1)}=0\}} dr = \int_{\mathbb{R}} \mathbf{1}_{\{u=0\}} L_T^u(Y^{(1)}) du = 0.$$

Because $\{Y_r = 0\} \subset \{Y_r^{(1)} = 0\}$, we obtain, for any finite time-horizon T ,

$$\int_0^T \mathbf{1}_{\{R_r=0\}} dr = 0 \quad \text{almost surely.} \quad (4.4)$$

Arriving at a contradiction. We will now derive a contradiction with (4.4). To this end, for $m \geq 1$, put

$$\sigma_m := \tau \wedge \inf\{t \in [0, T] : A_t \geq m\},$$

and let us establish (through Itô's expansion) that $\int_0^{\sigma_m} \mathbf{1}_{\{R_r=0\}} dr > 0$ almost surely.

The process A is continuous, nondecreasing, and starts from zero. Hence $\sigma_m > 0$ is non-decreasing almost surely and $A_{\sigma_m} \leq m$.

The extra bounded drift $-V_t$ in (4.3) is handled by the increasing concave function $r \mapsto 1 - e^{-r}$. For $\varepsilon > 0$, define its smooth radial approximation as

$$\phi_\varepsilon(y) := (|y|^2 + \varepsilon^2)^{1/2}, \quad \psi_\varepsilon(y) := 1 - e^{-\phi_\varepsilon(y)}.$$

For $y \neq 0$ and $w \in \mathbb{R}^d$, direct differentiation yields

$$\begin{aligned} H_\varepsilon(y, w) &:= \left(\Delta_y - \left(\kappa \frac{y}{|y|^2} + w \right) \cdot \nabla_y \right) \psi_\varepsilon(y) \\ &= e^{-\phi_\varepsilon(y)} \left(\frac{d - \kappa}{\phi_\varepsilon(y)} - \frac{|y|^2}{\phi_\varepsilon(y)^3} - \frac{y \cdot w}{\phi_\varepsilon(y)} - \frac{|y|^2}{\phi_\varepsilon(y)^2} \right). \end{aligned}$$

In particular, for $y \neq 0$ and $|w| \leq 1$,

$$|H_\varepsilon(y, w)| \leq \frac{|d - \kappa| + 1}{|y|} + 2,$$

and

$$H_\varepsilon(y, w) \rightarrow e^{-|y|} \left(\frac{d - 1 - \kappa}{|y|} - \frac{y}{|y|} \cdot w - 1 \right).$$

At $y = 0$, the actual drift coefficient in Itô's formula for $\psi_\varepsilon(Y_t)$ equals $de^{-\varepsilon}/\varepsilon$: the singular drift is defined to be zero there and $\nabla \psi_\varepsilon(0) = 0$. Its contribution to the time integral vanishes by (4.4).

Apply Itô's formula to $\psi_\varepsilon(Y_{t \wedge \sigma_m})$. Since $|\nabla \psi_\varepsilon| \leq 1$, the stochastic integral has expectation zero and we have

$$\mathbb{E} \psi_\varepsilon(Y_{t \wedge \sigma_m}) = 1 - e^{-\varepsilon} + \mathbb{E} \int_0^{t \wedge \sigma_m} \mathbf{1}_{\{R_r > 0\}} H_\varepsilon(Y_r, V_r) dr.$$

Before time σ_m , we have $|V_r| \leq a \leq 1$, and so

$$\mathbb{E} \int_0^{t \wedge \sigma_m} \sup_\varepsilon \mathbf{1}_{\{R_r > 0\}} |H_\varepsilon(Y_r, V_r)| dr \leq \mathbb{E} \int_0^{t \wedge \sigma_m} \left((|d - \kappa| + 1) \frac{\mathbf{1}_{\{R_r > 0\}}}{R_r} + 2 \right) dr \leq (|d - \kappa| + 1)m + 2T.$$

Thus dominated convergence applies to the drift integral and since

$$\sup_y \left| \psi_\varepsilon(y) - (1 - e^{-|y|}) \right| \leq \varepsilon,$$

letting $\varepsilon \downarrow 0$ therefore gives

$$\begin{aligned} \mathbb{E}[1 - e^{-R_{t \wedge \sigma_m}}] &= \mathbb{E} \int_0^{t \wedge \sigma_m} \mathbf{1}_{\{R_r > 0\}} e^{-R_r} \left(\frac{d-1-\kappa}{R_r} - \frac{Y_r}{R_r} \cdot V_r - 1 \right) dr \\ &\leq 0. \end{aligned} \quad (4.5)$$

Indeed, $\kappa \geq d-1$ and $|V_r| \leq 1$ imply that the expression in parentheses is nonpositive. The term -1 , coming from the concavity of the radial test function, absorbs the additional drift $-V_r$.

The left-hand side of (4.5) is nonnegative. It follows that $R_{t \wedge \sigma_m} = 0$ almost surely for every $t \in [0, T]$. Taking a countable dense set of times and using continuity, we conclude that the stopped path is identically zero. Therefore,

$$\int_0^{\sigma_m} \mathbf{1}_{\{R_r=0\}} dr = \sigma_m > 0 \quad \text{almost surely,}$$

contrary to (4.4). This proves nonexistence for $\kappa \geq d-1$. $\square$

5. MAIN RELICH-TYPE ESTIMATE FROM KINETIC MORREY CLASS VIA ADAMS ESTIMATES

Recall that $Q = 4d + 2$ is the homogeneous dimension fixed in (1.3).

We will need the following maximal function and estimates. For $0 \leq \beta < Q$, let

$$\mathbf{M}_\beta h(\zeta) := \sup_{r>0} r^\beta \frac{1}{|B_r(\zeta)|} \int_{B_r(\zeta)} |h(\xi)| d\xi, \quad \mathbf{M} := \mathbf{M}_0. \quad (5.1)$$

Proposition 5.1. *On $(\mathbb{R}^{1+2d}, d_*, dt dx dv)$, the following assertions hold:*

- (i) *The operator $\mathbf{M}$ is of weak type $(1, 1)$ and is bounded on L^s for every $1 < s \leq \infty$.*
- (ii) *If $1 < q < \infty$ and $1 \leq a < q$, and $f \in \mathcal{M}_\kappa^q$, then*

$$\|(\mathbf{M}(|f|^a))^{\frac{1}{a}}\|_{\mathcal{M}_\kappa^q} \leq C_{d,q,a} \|f\|_{\mathcal{M}_\kappa^q}.$$

- (iii) *If $f \in L_{loc}^1$ and $\mathbf{M}f < \infty$ almost everywhere, then for $0 < \delta < 1$, $(\mathbf{M}f)^\delta \in A_1$, that is*

$$\mathbf{M}((\mathbf{M}f)^\delta) \leq C_{d,\delta} (\mathbf{M}f)^\delta \quad \text{a.e.}$$

Proof. (i) and (iii) are standard results from maximal-function theory on spaces of homogeneous type; see [22, pp.71] and [45, Lemma 3.4], respectively.

(ii) See [5, Corollary 4.2] for a related Morrey-space estimate. We give a direct proof for $\mathcal{M}_\kappa^q$. Fix $B = B_r(\zeta_0)$ and decompose

$$|f|^a = |f|^a \mathbf{1}_{2B} + |f|^a \mathbf{1}_{(2B)^c} =: F_1 + F_2.$$

Since $a < q$, the $L^{\frac{q}{a}}$ -boundedness of $\mathbf{M}$ gives

$$\frac{1}{|B|} \int_B (\mathbf{M}F_1)^{\frac{q}{a}} \leq C_1 \frac{1}{|B|} \int_{2B} |f|^q \leq C_2 \|f\|_{\mathcal{M}_\kappa^q}^q r^{-q}.$$

If $\zeta \in B$, every ball centered at ζ that meets $(2B)^c$ has radius at least cr . Then, by Hölder's inequality and the definition of $\mathcal{M}_\kappa^q$,

$$\mathbf{M}F_2(\zeta) \leq \sup_{\rho \geq cr} \left(\frac{1}{|B_\rho(\zeta)|} \int_{B_\rho(\zeta)} |f|^q \right)^{\frac{q}{a}} \leq C_3 \|f\|_{\mathcal{M}_\kappa^q}^q r^{-a},$$

whence

$$\frac{1}{|B|} \int_B (\mathbf{M}F_2)^{\frac{q}{a}} \leq C_4 \|f\|_{\mathcal{M}_\kappa^q}^q r^{-q}.$$

Combining the two estimates yields

$$r \left(\frac{1}{|B|} \int_B (\mathbf{M}(|f|^a))^{\frac{q}{a}} \right)^{\frac{1}{q}} \leq C_5 \|f\|_{\mathcal{M}_{\mathcal{K}}^q}.$$

Taking the supremum over B proves (ii). $\square$

Proposition 5.2 (Hedberg-type inequality). *If $0 < \alpha < \beta < Q$, then*

$$I_{\alpha} f \leq C_{\alpha, \beta, d} (\mathbf{M}_{\beta} f)^{\frac{\alpha}{\beta}} (\mathbf{M} f)^{1 - \frac{\alpha}{\beta}}. \quad (5.2)$$

Proof. Fix $\zeta \in \mathbb{R}^{1+2d}$ and $R > 0$. If either maximal function is infinite, the estimate is immediate. If either one vanishes, then $f = 0$ almost everywhere. We may therefore assume that

$$0 < \mathbf{M} f(\zeta), \mathbf{M}_{\beta} f(\zeta) < \infty.$$

Let us split the integral defining $I_{\alpha} f(\zeta)$ into the regions $d_*(\zeta, \xi) < R$ and $d_*(\zeta, \xi) \geq R$. For the near region, using a dyadic decomposition,

$$\begin{aligned} \int_{d_*(\zeta, \xi) < R} \frac{|f(\xi)|}{d_*(\zeta, \xi)^{Q-\alpha}} d\xi &= \sum_{j=0}^{\infty} \int_{2^{-j-1}R \leq d_*(\zeta, \xi) < 2^{-j}R} \frac{|f(\xi)|}{d_*(\zeta, \xi)^{Q-\alpha}} d\xi \\ &\leq \sum_{j=0}^{\infty} (2^{-j-1}R)^{\alpha-Q} \int_{B_{2^{-j}R}(\zeta)} |f(\xi)| d\xi \\ &\leq C_1 R^{\alpha} \mathbf{M} f(\zeta). \end{aligned}$$

For the far region, similarly,

$$\begin{aligned} \int_{d_*(\zeta, \xi) \geq R} \frac{|f(\xi)|}{d_*(\zeta, \xi)^{Q-\alpha}} d\xi &= \sum_{j=0}^{\infty} \int_{2^j R \leq d_*(\zeta, \xi) < 2^{j+1}R} \frac{|f(\xi)|}{d_*(\zeta, \xi)^{Q-\alpha}} d\xi \\ &\leq \sum_{j=0}^{\infty} (2^j R)^{\alpha-Q} \int_{B_{2^{j+1}R}(\zeta)} |f(\xi)| d\xi \\ &\leq C_2 \sum_{j=0}^{\infty} (2^j R)^{\alpha-\beta} (2^{j+1}R)^{\beta} \frac{1}{|B_{2^{j+1}R}(\zeta)|} \int_{B_{2^{j+1}R}(\zeta)} |f(\xi)| d\xi \\ &\leq C_2 R^{\alpha-\beta} \mathbf{M}_{\beta} f(\zeta), \end{aligned}$$

where the last series converges because $\alpha < \beta$. Hence

$$I_{\alpha} f(\zeta) \leq C_1 R^{\alpha} \mathbf{M} f(\zeta) + C_2 R^{\alpha-\beta} \mathbf{M}_{\beta} f(\zeta).$$

Taking $R = \left(\frac{\mathbf{M}_{\beta} f(\zeta)}{\mathbf{M} f(\zeta)} \right)^{\frac{1}{\beta}}$ gives (5.2). $\square$

Proposition 5.3 (Kinetic Adams inequality). *Let $1 < p < \infty$, $q > 1$, and $g \in \mathcal{M}_{\mathcal{K}}^q$. Then*

$$\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}} f|^p \leq C_{p,q,d} \|g\|_{\mathcal{M}_{\mathcal{K}}^q} \|f\|_p^p$$

or, equivalently,

$$\| |g|^{\frac{1}{p}} I_{\frac{1}{p}} \|_{p \rightarrow p} \leq C_{d,p,q} \|g\|_{\mathcal{M}_{\mathcal{K}}^q}^{\frac{1}{p}}. \quad (5.3)$$

Remark 5.4. As mentioned in the introduction, the maximal-function arguments that we use to extend the proof of the Adams estimate [1, Theorem 7.3] to the kinetic setting build on [5, 21, 23, 31] who consider a more general setup of spaces of homogeneous type.

It should also be noted that, for $p = 2$, the preceding estimate already appeared in [46, Appendix C], where the author used a fixed-point argument to establish the existence and uniqueness of solutions

to certain nonlinear equations. The proof of the estimate there, however, follows a different approach from that of Adams and ours. However, we need this estimate for the full range $1 < p < \infty$. Indeed, we need to be able to (a) take p large when handling the backward equation, in order to apply the Sobolev embedding theorem Proposition 2.3 to establish Krylov estimates and to construct the Feller propagator; (b) take p close to 1 when handling the forward equation, in order to have delta-function δ_x in the initial condition and thus prove uniqueness of the probability density evolving from a fixed point.

We refer to [2, 40] regarding the theory of Morrey spaces in the non-degenerate setting, and, in particular, in what concerns parabolic Morrey embeddings, to [3, 41].

Proof. Since the integral kernel of $I_{1/p}$ is positive, it suffices to consider $f \geq 0$.

We first recall the A_1 condition that will be used below. For a nonnegative locally integrable function w , write

$$[w]_{A_1} := \inf\{C > 0 : Mw \leq Cw \text{ a.e.}\},$$

where M is the intrinsic Hardy–Littlewood maximal operator defined in (5.1). Thus $w \in A_1$ precisely when $[w]_{A_1} < \infty$.

We separate the argument into two steps. We first prove the estimate under the auxiliary assumption that a suitable power of $|g|$ belongs to A_1 . The constant in this intermediate estimate is allowed to depend on its A_1 -characteristic. We then remove this assumption by replacing $|g|$ with a Coifman–Rochberg majorant whose A_1 -characteristic is bounded solely in terms of the fixed exponents and the doubling constant.

Step 1: Estimate under the A_1 assumption. Choose q_0 and $\sigma > 0$ such that

$$1 < q_0 < \min\{q, Q\}, \quad 1 + \sigma p < q_0, \quad \frac{1}{p} + \sigma p' < q_0, \quad \sigma p' < 1.$$

Such a choice is possible for every $p > 1$ and $q > 1$. Since $q_0 < q$, Hölder's inequality gives $\|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}} \leq \|g\|_{\mathcal{M}_{\mathcal{K}}^q}$. The case $g = 0$ is immediate, so assume $g \neq 0$. For the moment, suppose additionally that $|g|^{q_0} \in A_1$, and put $A_0 = [|g|^{q_0}]_{A_1}$. All constants in this step may depend on d, p, q_0, σ, A_0 . We claim that

$$\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}} f|^p \leq C_{d,p,q_0,\sigma,A_0} \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}} \|f\|_p^p. \quad (5.4)$$

Indeed, for $p \in (1, \infty)$, the symmetry and positivity of the kernel give

$$\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}} f|^p = \langle I_{\frac{1}{p}} (|g| (I_{\frac{1}{p}} f)^{p-1}), f \rangle \leq \|I_{\frac{1}{p}} (|g| (I_{\frac{1}{p}} f)^{p-1})\|_{p'} \|f\|_p. \quad (5.5)$$

Note that Hölder's inequality yields

$$\begin{aligned} I_{\frac{1}{p}} (|g| (I_{\frac{1}{p}} f)^{p-1})(\zeta) &= I_{\frac{1}{p}} (|g|^{\frac{1}{p}+\sigma} |g|^{\frac{1}{p'}-\sigma} |I_{\frac{1}{p}} f|^{p-1}) \\ &\leq (I_{\frac{1}{p}} (|g|^{1+\sigma p}))^{\frac{1}{p}} (I_{\frac{1}{p}} (|g|^{1-\sigma p'} (I_{\frac{1}{p}} f)^p))^{\frac{1}{p'}}. \end{aligned} \quad (5.6)$$

Applying the Hedberg-type inequality with $\alpha = \frac{1}{p}$ and $\beta = 1 + \sigma p$, we obtain

$$I_{\frac{1}{p}} (|g|^{1+\sigma p}) \leq C_1 (M_{1+\sigma p} (|g|^{1+\sigma p}))^{\frac{1}{p(1+\sigma p)}} (M(|g|^{1+\sigma p}))^{1-\frac{1}{p(1+\sigma p)}}.$$

Since $1 + \sigma p < q_0$, Hölder's inequality on each intrinsic ball gives

$$\begin{aligned} \mathbf{M}_{1+\sigma p}(|g|^{1+\sigma p})(\zeta) &= \sup_{r>0} r^{1+\sigma p} \frac{1}{|B_r(\zeta)|} \int_{B_r(\zeta)} |g|^{1+\sigma p} \\ &\leq \sup_{r>0} \left(r \left(\frac{1}{|B_r(\zeta)|} \int_{B_r(\zeta)} |g|^{q_0} \right)^{\frac{1}{q_0}} \right)^{1+\sigma p} \\ &\leq \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}}^{1+\sigma p}. \end{aligned}$$

For every $0 < \kappa < q_0$, $|g|^{q_0} \in A_1$ gives

$$\mathbf{M}(|g|^\kappa) \leq (\mathbf{M}(|g|^{q_0}))^{\frac{\kappa}{q_0}} \leq [|g|^{q_0}]_{A_1}^{\frac{\kappa}{q_0}} |g|^\kappa \quad \text{a.e.}$$

Hence,

$$I_{\frac{1}{p}}(|g|^{1+\sigma p}) \leq C_2 \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}}^{\frac{1}{p}} |g|^{1+\sigma p - \frac{1}{p}}. \quad (5.7)$$

Similarly, applying the Hedberg-type inequality again we have

$$I_{\frac{1}{p}}(|g|^{\frac{1}{p} + \sigma p'}) \leq C_3 (\mathbf{M}_{\frac{1}{p} + \sigma p'}(|g|^{\frac{1}{p} + \sigma p'}))^{\frac{1}{1+\sigma p p'}} (\mathbf{M}(|g|^{\frac{1}{p} + \sigma p'}))^{1 - \frac{1}{1+\sigma p p'}} \leq C_4 \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}}^{\frac{1}{p}} |g|^{\sigma p'}. \quad (5.8)$$

Combining (5.6)-(5.8), we obtain

$$\begin{aligned} \|I_{\frac{1}{p}}(|g| |I_{\frac{1}{p}} f|^{p-1})\|_{p'} &\leq \left(\int_{\mathbb{R}^{1+2d}} (I_{\frac{1}{p}}(|g|^{1+\sigma p}))^{\frac{p'}{p}} I_{\frac{1}{p}}(|g|^{1-\sigma p'} |I_{\frac{1}{p}} f|^p) \right)^{\frac{1}{p'}} \\ &\leq C_5 \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}}^{\frac{1}{p^2}} \left(\int_{\mathbb{R}^{1+2d}} |g|^{\frac{1}{p} + \sigma p'} I_{\frac{1}{p}}(|g|^{1-\sigma p'} |I_{\frac{1}{p}} f|^p) \right)^{\frac{1}{p'}} \\ &= C_5 \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}}^{\frac{1}{p^2}} \left(\int_{\mathbb{R}^{1+2d}} I_{\frac{1}{p}}(|g|^{\frac{1}{p} + \sigma p'}) |g|^{1-\sigma p'} |I_{\frac{1}{p}} f|^p \right)^{\frac{1}{p'}} \\ &\leq C_6 \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}}^{\frac{1}{p}} \left(\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}} f|^p \right)^{\frac{1}{p'}. \end{aligned}$$

Recalling (5.5), we conclude that

$$\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}} f|^p \leq C_6 \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}}^{\frac{1}{p}} \left(\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}} f|^p \right)^{\frac{1}{p'}} \|f\|_p.$$

If the weighted integral on the left-hand side is finite, then, with the zero case being trivial, we obtain (5.4).

It remains to remove this finiteness condition. To this end, we apply the preceding argument first to truncated potentials. Let $f \in C_c^\infty(\mathbb{R}^{1+2d})$, and define

$$I_{\frac{1}{p}}^{\varepsilon, R} f(\zeta) := \int_{\{\varepsilon < d_*(\zeta, \xi) < R\}} \frac{f(\xi)}{d_*(\zeta, \xi)^{Q - \frac{1}{p}}} d\xi, \quad 0 < \varepsilon < R < \infty.$$

The function $I_{\frac{1}{p}}^{\varepsilon, R} f$ is bounded and has bounded support. Since $g \in L_{\text{loc}}^1$, it follows that

$$\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}}^{\varepsilon, R} f|^p < \infty.$$

Replace every occurrence of the integral operator in (5.5) and (5.6) by the same truncated operator $I_{1/p}^{\varepsilon, R}$, and put $U = I_{1/p}^{\varepsilon, R} f$. Its kernel is still positive and symmetric. The two estimates (5.7)–(5.8) remain valid for the truncated operator, since its kernel is bounded by the full kernel. Repeating the duality calculation therefore gives, with $S_{\varepsilon, R} = \int |g| U^p < \infty$,

$$S_{\varepsilon, R} \leq C \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}}^{1/p} S_{\varepsilon, R}^{1/p'} \|f\|_p.$$

Division by $S_{\varepsilon,R}^{1/p'}$ when this quantity is nonzero gives, after renaming the constant, uniformly in ε and R ,

$$\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}}^{\varepsilon,R} f|^p \leq C_6 \|g\|_{\mathcal{M}_{\mathcal{K}}^{q_0}} \|f\|_p^p.$$

Since $f \geq 0$, the monotone convergence theorem gives

$$I_{\frac{1}{p}}^{\frac{1}{n},n} f \uparrow I_{\frac{1}{p}} f.$$

Another application of the monotone convergence theorem therefore implies that (5.4) holds for every $f \in C_c^\infty(\mathbb{R}^{1+2d})$.

The estimate extends to arbitrary nonnegative $f \in L^p$ by approximation with nonnegative smooth compactly supported functions, Fatou's lemma, and the positivity of the kernel. Applying this to $|f|$ gives (5.4) for every $f \in L^p$. With the convention (2.14), the resulting map is a bounded sublinear operator. Indeed,

$$|I_{1/p} f - I_{1/p} h| \leq I_{1/p}(f - h),$$

so the weighted estimate also gives its continuous extension by density.

Step 2: Construction of an A_1 majorant. Choose q_1 such that $q_0 < q_1 < q$ and define

$$\tilde{g} := (\mathbf{M}(|g|^{q_1}))^{\frac{1}{q_1}}.$$

Since $g \in \mathcal{M}_{\mathcal{K}}^q$ and $q_1 < q$, we have $|g|^{q_1} \in L_{\text{loc}}^1$. By the Lebesgue differentiation theorem [31, Theorem 1.8] applied to $|g|^{q_1}$, we have $|g| \leq \tilde{g}$ almost everywhere.

Moreover, Proposition 5.1(ii) implies

$$\|\tilde{g}\|_{\mathcal{M}_{\mathcal{K}}^q} = \|(\mathbf{M}(|g|^{q_1}))^{\frac{1}{q_1}}\|_{\mathcal{M}_{\mathcal{K}}^q} \leq C_7 \|g\|_{\mathcal{M}_{\mathcal{K}}^q}.$$

In particular, $\mathbf{M}(|g|^{q_1}) < \infty$ almost everywhere. Since $q_0 < q_1$, the Coifman–Rochberg theorem on spaces of homogeneous type [45, Lemma 3.4] implies

$$\tilde{g}^{q_0} = (\mathbf{M}(|g|^{q_1}))^{q_0/q_1} \in A_1, \quad \mathbf{M}(\tilde{g}^{q_0}) \leq C_8 \tilde{g}^{q_0}.$$

The A_1 characteristic in the last formula is bounded in terms of d, q_0, q_1 , independently of g . Thus $\tilde{g}$ satisfies the hypothesis of Step 1 with a uniform A_1 constant. We may now apply the estimate (5.4) with $\tilde{g}$ in place of g and obtain

$$\int_{\mathbb{R}^{1+2d}} |g| |I_{\frac{1}{p}} f|^p \leq \int_{\mathbb{R}^{1+2d}} |\tilde{g}| |I_{\frac{1}{p}} f|^p \leq C_6 \|\tilde{g}\|_{\mathcal{M}_{\mathcal{K}}^{q_0}} \|f\|_p^p \leq C_6 \|\tilde{g}\|_{\mathcal{M}_{\mathcal{K}}^q} \|f\|_p^p \leq C_9 \|g\|_{\mathcal{M}_{\mathcal{K}}^q} \|f\|_p^p.$$

This completes the proof. $\square$

Combining the fractional kernel estimates with the Adams inequality above yields the following Rellich-type estimates, which will be used in the subsequent PDE and SDE analysis.

Proposition 5.5 (Rellich estimates from kinetic Adams inequality). *Let $q > 1$, $1 < p < \infty$, and $\lambda > 0$. If $g \in \mathcal{M}_{\mathcal{K}}^q$, then*

$$\| |g|^{\frac{1}{p}} \mathcal{A}^{-\frac{1}{2p}} \|_{p \rightarrow p} + \| |g|^{\frac{1}{p}} (\mathcal{A}^*)^{-\frac{1}{2p}} \|_{p \rightarrow p} \leq C_{d,p,q} \|g\|_{\mathcal{M}_{\mathcal{K}}^q}^{\frac{1}{p}}, \quad (5.9)$$

$$\| |g|^{\frac{1}{p}} \nabla_v \mathcal{A}^{-\frac{1}{2} - \frac{1}{2p}} \|_{p \rightarrow p} + \| |g|^{\frac{1}{p}} \nabla_v (\mathcal{A}^*)^{-\frac{1}{2} - \frac{1}{2p}} \|_{p \rightarrow p} \leq C_{d,p,q} \|g\|_{\mathcal{M}_{\mathcal{K}}^q}^{\frac{1}{p}}. \quad (5.10)$$

If $g \in L^\infty$, the right-hand side of each estimate may be replaced by

$$C_{d,p} \lambda^{-\frac{1}{2p}} \|g\|_\infty^{\frac{1}{p}}.$$

Proof. Taking $\gamma = \frac{1}{p}$ in (2.15) and (2.16), and then applying the Adams inequality (5.3), yields (5.9) and (5.10) for $f \in C_c^\infty(\mathbb{R}^{1+2d})$. The estimates extend to all $f \in L^p$ by density. If $g \in L^\infty$, the corresponding bounds follow directly from (2.12) and (2.13), again with $\gamma = \frac{1}{p}$. $\square$

6. PROOF OF THEOREM 3.3

In addition to the representation for η in Theorem 3.3, we will use the equivalent identity

$$\eta = \mathcal{A}^{-1}f - \mathcal{A}^{-1/(2p)}G_p(b)(I + H_p(b)G_p(b))^{-1}H_p(b)\mathcal{A}^{-(1-1/(2p))}f. \quad (6.1)$$

For an admissible regularization $\{b_n\}_n$ of b we will prove a more general result: if $f_n \rightarrow f$ in $\mathbb{W}_{\mathcal{K}}^{-1-1/p',p}$, then

$$\eta_n \rightarrow \eta \quad \text{in } \mathbb{W}_{\mathcal{K}}^{1/p,p}. \quad (6.2)$$

In particular, the point-source equation is given, for $\zeta_0 \in \mathbb{R}^{1+2d}$, by

$$\mathcal{A}\eta - \operatorname{div}_v(b\eta) = \delta_{\{\zeta_0\}}.$$

Proof of Theorem 3.3. By (H_{s,p}), both $I + G_p(b)H_p(b)$ and $I + H_p(b)G_p(b)$ are invertible on L^p . Formula (3.15) is therefore well defined. Since $\mathcal{A}^{-(1-1/(2p))}$ maps $\mathbb{W}_{\mathcal{K}}^{1/p-2,p}$ isometrically into L^p and $\mathcal{A}^{-1/(2p)}$ maps L^p isometrically onto $\mathbb{W}_{\mathcal{K}}^{1/p,p}$ (see Appendix A), estimate (3.16) follows. The identity

$$(I + G_p(b)H_p(b))^{-1} = I - G_p(b)(I + H_p(b)G_p(b))^{-1}H_p(b)$$

gives (6.1).

By Lemma 3.5, (3.15) is a weak solution. Conversely, every weak solution produces an $h \in L^p$ satisfying (3.20). Invertibility gives a unique h , and hence a unique $\eta = \mathcal{A}^{-1/(2p)}h$. This proves existence, the estimate, and uniqueness in part (i).

For part (ii), admissibility gives $H_p(b_n) \rightarrow H_p(b)$ and $G_p(b_n) \rightarrow G_p(b)$ strongly, while their norms have a common product bound strictly below one. The resolvent identity for inverses, first on a dense set and then by uniform boundedness, yields

$$(I + G_p(b_n)H_p(b_n))^{-1}g \rightarrow (I + G_p(b)H_p(b))^{-1}g \quad \text{in } L^p$$

for every $g \in L^p$. Together with $f_n \rightarrow f$, this proves (6.2). For smooth b_n and f , the distributional solution η_n is classical. Indeed, the diffusion fields ∂_{v_j} and the transport field $\partial_t + v \cdot \nabla_x - b_n \cdot \nabla_v$, together with their commutators, span all space-time directions: the first commutators give ∂_{x_j} modulo velocity fields, and the transport field then gives ∂_t . The smooth zeroth-order coefficient $\lambda - \operatorname{div}_v b_n$ does not change this condition. Hörmander's hypoellipticity theorem [32] therefore applies. Specifically, for a smooth regularized drift b_n , put

$$X_j := \partial_{v_j}, \quad Y_n := \partial_t + v \cdot \nabla_x - b_n \cdot \nabla_v.$$

Then

$$\mathcal{A}\eta - \operatorname{div}_v(b_n\eta) = \lambda\eta + Y_n\eta - \sum_{j=1}^d X_j^2\eta - (\operatorname{div}_v b_n)\eta.$$

The zeroth-order term does not enter the bracket condition. Moreover,

$$[X_j, Y_n] = \partial_{x_j} - \sum_{k=1}^d (\partial_{v_j} b_n^k) X_k.$$

It follows that every ∂_{x_j} belongs pointwise to the span of the fields X_k and the commutators $[X_j, Y_n]$. Since

$$\partial_t = Y_n - \sum_{j=1}^d v_j \partial_{x_j} + \sum_{k=1}^d b_n^k X_k,$$

these fields and their first commutators span the entire tangent space $T_\zeta \mathbb{R}^{1+2d}$ at every ζ . The regularized operator therefore satisfies Hörmander's bracket condition. Hörmander's hypoellipticity theorem [32] consequently implies that if

$$\mathcal{A}u - \operatorname{div}_v(b_n u) = F$$

in the sense of distributions on an open set U and $F \in C^\infty(U)$, then $u \in C^\infty(U)$.

Finally, for part (iii), set $s = 1 + 1/p'$. Condition (3.17) is equivalent to $p' > Q - 1$, hence

$$s = 1 + \frac{1}{p'} > \frac{Q}{p'}.$$

By Corollary 2.4, $\delta_{\{\zeta_0\}} \in \mathbb{W}_{\mathcal{H}}^{-s,p} = \mathbb{W}_{\mathcal{H}}^{-1-1/p',p}$, and so part (i) applies. Note also that every negative power of $\mathcal{A}$ is causal, multiplication by b preserves support in time, and the Duhamel series in (3.15) is a sum of causal terms. The solution therefore vanishes for $t < t_0$. $\square$

Remark 6.1 (Operator-norm formulation). The Morrey hypothesis is a concrete sufficient condition for the two Rellich estimates. Theorem 3.3 remains true verbatim for every Borel drift for which the operators in (3.1)–(3.2) are bounded and $\|G_p\|_{p \rightarrow p} \|H_p\|_{p \rightarrow p} < 1$.

7. PROOF OF THEOREM 3.6

We consider here the backward PDE

$$\mathcal{A}u + b \cdot \nabla_v u = f. \quad (7.1)$$

Recall that

$$R_p(b) := |b|^{1/p} \nabla_v \mathcal{A}^{-1/2-1/(2p)}, \quad Q_p(b) := \mathcal{A}^{-1/(2p')} (b^{\frac{1}{p'}}).$$

The operator $R_p(b)$ maps scalar functions to vector fields and $Q_p(b)$ contracts vector fields with the polar factor $b^{\frac{1}{p'}}$. By Proposition 5.5, both operators are bounded on L^p , and their product/composition has norm less than one under the same smallness regime $(H_{s,p})$, with p and p' placed as in (3.3)–(3.4).

Proof of Theorem 3.6. Let us show that the weak solution to (7.1) implies necessarily the representation (3.22). Existence and uniqueness in $\mathbb{W}_{\mathcal{H}}^{1+1/p,p}$ then follow as in the proof of Theorem 3.3. Let $u \in \mathbb{W}_{\mathcal{H}}^{1+1/p,p}$ be a weak solution to (7.1). Write $u = \mathcal{A}^{-1/2-1/(2p)} h$ with $h \in L^p$ (this simply follows from the very definition of $\mathbb{W}_{\mathcal{H}}^{1+1/p,p}$). The gradient estimate (again, in Proposition 5.5) ensures $\nabla_v u \in L^p$ and $|b|^{1/p} \nabla_v u = R_p(b)h \in L^p(\mathbb{R}^{1+2d}, \mathbb{R}^d)$. Hölder's inequality on any compact set shows that $b \cdot \nabla_v u = |b|^{-1/p} \cdot |b|^{1/p} \nabla_v u$ is locally integrable. By duality,

$$\mathcal{A}^{-1/(2p')} (b \cdot \nabla_v u) = Q_p(b) R_p(b) h \in L^p,$$

so the first-order term $b \cdot \nabla_v u$ belongs to $\mathbb{W}_{\mathcal{H}}^{-1/p',p} = \mathcal{A}^{1/(2p')} L^p$. Since $1 - (1/2 + 1/(2p)) = 1/(2p')$, applying $\mathcal{A}^{-1/(2p')}$ to (7.1) gives

$$\mathcal{A}^{1-1/(2p')} u + \mathcal{A}^{-1/(2p')} (b \cdot \nabla_v u) = (I + Q_p(b) R_p(b)) h = \mathcal{A}^{-1/(2p')} f.$$

Inverting $I + Q_p(b) R_p(b)$, we obtain $h = (I + Q_p(b) R_p(b))^{-1} \mathcal{A}^{-1/(2p')} f$. By inverting $\mathcal{A}^{1-1/(2p')}$, we end the proof. $\square$

8. PROOF OF THEOREM 3.9

For most of this section we assume that the drift satisfies the global hypothesis $(H_{\mathcal{K}, \varepsilon_0})$. In particular, there is no separate bounded component. We recall that the role of global $L^{1+\varepsilon_0}$ integrability is only to shorten the proof of the higher-power estimate used in the tightness argument in the construction of a solution to (3.23) – we will remove it in Section 9.

To start, we will need the time-reversed variant of Theorem 3.6. We only record the corresponding notations: the same proof remains valid. For $1 < r < \infty$, set

$$\begin{aligned}\widehat{R}_r(b) &:= |b|^{1/r} \nabla_v (\mathcal{A}^*)^{-1/2-1/(2r)}, \\ \widehat{Q}_r(b) &:= (\mathcal{A}^*)^{-1/(2r')} (b^{1/r'}). \end{aligned} \tag{8.1}$$

The estimates of Proposition 5.5 apply to these adjoint operators. We call r *backward admissible* if

$$\|\widehat{Q}_r(b)\|_{r \rightarrow r} \|\widehat{R}_r(b)\|_{r \rightarrow r} < 1. \tag{8.2}$$

The notion of an admissible regularization of b , and Proposition 3.2, extend in a straightforward manner to the operators $\widehat{R}_r(b)$, $\widehat{Q}_r(b)$:

Proposition 8.1 (A variant of Proposition 3.2). *Let $b = b_s + b_b$ satisfy $(H_{\mathcal{K}})$, and let $E_f, E_b \subset (1, \infty)$ be finite sets of forward and backward exponents r , respectively given by*

$$g_{r,\lambda}(b) h_{r,\lambda}(b) < 1 \quad (r \in E_f), \quad \|\widehat{Q}_r(b)\|_{r \rightarrow r} \|\widehat{R}_r(b)\|_{r \rightarrow r} < 1 \quad (r \in E_b).$$

Then there exists a sequence

$$b_n = b_{n,s} + b_{n,b} \in C_c^\infty(\mathbb{R}^{1+2d}; \mathbb{R}^d)$$

such that

$$\begin{aligned} b_{n,s} &\rightarrow b_s \quad \text{in } L_{\text{loc}}^q, \quad \sup_n \|b_{n,s}\|_{\mathcal{M}_{\mathcal{K}}^q} < \infty, \\ b_{n,b} &\rightarrow b_b \quad \text{in } L_{\text{loc}}^a, \quad \sup_n \|b_{n,b}\|_\infty \leq \|b_b\|_\infty \quad (a \in [1, \infty)). \end{aligned} \tag{8.3}$$

In particular, $b_n \rightarrow b$ in L_{loc}^q . Moreover, for every $r \in E := E_f \cup E_b$,

$$\begin{aligned} H_r(b_n) &\rightarrow H_r(b), & G_r(b_n) &\rightarrow G_r(b), \\ \widehat{R}_r(b_n) &\rightarrow \widehat{R}_r(b), & \widehat{Q}_r(b_n) &\rightarrow \widehat{Q}_r(b), \end{aligned}$$

strongly on the corresponding L^r spaces. The required product bounds remain uniformly strict:

$$\begin{aligned} \sup_n g_{r,\lambda}(b_n) h_{r,\lambda}(b_n) &< 1 \quad \text{for all } r \in E_f, \\ \sup_n \|\widehat{Q}_r(b_n)\|_{r \rightarrow r} \|\widehat{R}_r(b_n)\|_{r \rightarrow r} &< 1 \quad \text{for all } r \in E_b. \end{aligned}$$

In fact, each individual backward operator norm does not increase under this regularization. If $b_b = 0$ and, in addition to $(H_{\mathcal{K}})$, $b \in L^a$ for some $1 < a < \infty$, the sequence may be chosen so that

$$\sup_n \|b_n\|_a \leq \|b\|_a.$$

The proof of the above follows closely the proof of Proposition 3.2, so, for the sake of brevity, we skip it.

Corollary 8.2. *Let $\{b_n\}_n$ be bounded smooth drifts such that $\widehat{R}_p(b_n)$ and $\widehat{Q}_p(b_n)$ converge strongly to the corresponding operators of b , and suppose*

$$\sup_n \|\widehat{Q}_p(b_n)\|_{p \rightarrow p} \|\widehat{R}_p(b_n)\|_{p \rightarrow p} < 1.$$

If $f_n \rightarrow f$ in $\mathbb{W}_{\mathcal{K}^*}^{-1/p',p}$, then the corresponding solutions of

$$\mathcal{A}^* u_n + b_n \cdot \nabla_v u_n = f_n$$

satisfy

$$u_n \rightarrow u \quad \text{in } \mathbb{W}_{\mathcal{K}^*}^{1+1/p,p}. \quad (8.4)$$

If $p > Q - 1$, the convergence is also uniform on $\mathbb{R}^{1+2d}$.

Proof. Put

$$S_n := \widehat{Q}_p(b_n) \widehat{R}_p(b_n), \quad S := \widehat{Q}_p(b) \widehat{R}_p(b).$$

Strong convergence of the individual operators and their uniform boundedness give $S_n h \rightarrow S h$ for every $h \in L^p$. If their common product bound is $\sup_n \|\widehat{Q}_p(b_n)\|_{p \rightarrow p} \|\widehat{R}_p(b_n)\|_{p \rightarrow p} =: \vartheta < 1$, then

$$\sup_n \|(I + S_n)^{-1}\|_{p \rightarrow p} \leq \frac{1}{1 - \vartheta}.$$

The identity

$$(I + S_n)^{-1} - (I + S)^{-1} = (I + S_n)^{-1} (S - S_n) (I + S)^{-1}$$

therefore shows that the inverses converge strongly on L^p . Using the adjoint version of (3.22), together with

$$\|f_n - f\|_{\mathbb{W}_{\mathcal{K}^*}^{-1/p',p}} = \|(\mathcal{A}^*)^{-1/(2p')}(f_n - f)\|_p,$$

proves (8.4). If $p > Q - 1$, then $1 + 1/p > Q/p$, and the uniform convergence follows from Proposition 2.3(ii). $\square$

8.1. Kinetic Krylov estimates. The next lemma is the bridge from the PDE theory to the SDE. The vector field k below is an arbitrary test drift.

Lemma 8.3 (Krylov estimate). *Let $p > Q - 1$ be backward admissible, and let $\{b_n\}_n$ be the approximating sequence from Proposition 8.1, with a uniform backward product bound strictly smaller than one. Let*

$$k \in \mathcal{M}_{\mathcal{K}}^q(\mathbb{R}^{1+2d}, \mathbb{R}^d) \cap L^1(\mathbb{R}^{1+2d}, \mathbb{R}^d). \quad (8.5)$$

For every bounded Borel function h such that $|k|^{1/p} h \in L^p$, the solution of

$$\mathcal{A}^* u + b_n \cdot \nabla_v u = |k h| \quad (8.6)$$

satisfies, with a constant independent of n ,

$$\|u\|_{\infty} \leq C \|k\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p'} \times \| |k|^{1/p} h \|_p. \quad (8.7)$$

If $Z^n (= (X^n, V^n))$ denotes the strong solution of (3.23) with b_n in place of b , then

$$\mathbb{E}_{s,z}^n \int_s^T |k(r, Z_r^n) h(r, Z_r^n)| dr \leq C_T \|k\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p'} \times \|\mathbf{1}_{[s,T]} |k|^{1/p} h\|_p. \quad (8.8)$$

One also has the unweighted estimate

$$\mathbb{E}_{s,z}^n \int_s^T |h(r, Z_r^n)| dr \leq C_T \|\mathbf{1}_{[s,T]} h\|_p. \quad (8.9)$$

All constants are uniform in the starting point (s, z) .

Proof. The adjoint version of Theorem 3.6 gives the (unique) representation

$$u = (\mathcal{A}^*)^{-1/2-1/(2p)}(I + \widehat{Q}_p(b_n)\widehat{R}_p(b_n))^{-1}(\mathcal{A}^*)^{-1/(2p')}F, \quad F := |kh|, \quad (8.10)$$

for any weak solution of (8.6). By duality and the Rellich estimates (5.9),

$$\begin{aligned} \|(\mathcal{A}^*)^{-1/(2p')}(|k|h)\|_p &= \|(\mathcal{A}^*)^{-1/(2p')}(|k|^{1/p'}|k|^{1/p}h)\|_p \\ &= \sup_{\|\varphi\|_{p'}=1} \langle |k|^{1/p'}(\mathcal{A}^*)^{-1/(2p')}\varphi, |k|^{1/p}h \rangle \\ &\leq C\|k\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p'}\| |k|^{1/p}h \|_p. \end{aligned}$$

Since $1 + 1/p > Q/p$, Proposition 2.3(ii) maps the range of $(\mathcal{A}^*)^{-1/2-1/(2p)}$ into C_∞ . This proves (8.7).

For $\{F_m\}_m$ a smooth compactly supported approximation of F such that $\|(\mathcal{A}^*)^{-1/(2p')}F_m\|_p \leq C\|k\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p'}\| |k|^{1/p}h \|_p$ (e.g. Proposition 3.2), Itô's formula applied to $e^{-\lambda(t-s)}u_m(t, \cdot)$ of (8.6) for F_m in place of $F = |kh|$ along Z_t^n gives

$$\mathbb{E}_{s,z} \int_s^T e^{-\lambda(r-s)} F_m(r, Z_r^n) dr = u_m(s, z) - \mathbb{E}_{s,z} \left[e^{-\lambda(T-s)} u_m(T, Z_T^n) \right] \leq 2\|u_m\|_\infty.$$

The resolvent parameter introduces $e^{-\lambda(r-s)}$ and removing it on $[s, T]$ costs at most $e^{\lambda T}$. Applying truncation and using Fatou's lemma we extend the estimate to the stated Borel right-hand sides.

For (8.9), we use (8.10) with $F = |h|$ and estimate $(\mathcal{A}^*)^{-1/(2p')}$ by (2.12). The same embedding and Itô argument then yield (8.9). $\square$

Lemma 8.4 (Higher-power Krylov estimate). *Let $p > Q - 1$ be backward admissible and assume $(H_{\mathcal{K}, \varepsilon_0})$. Select a regularization $\{b_n\}$ that additionally satisfies*

$$\sup_n \|b_n\|_{1+\varepsilon_0} \leq \|b\|_{1+\varepsilon_0}. \quad (8.11)$$

Then

$$\sup_n \sup_{0 \leq s < T} \sup_{z \in \mathbb{R}^{2d}} \mathbb{E}_{s,z}^n \int_s^T |b_n(r, Z_r^n)|^{1+\varepsilon_0/p} dr \leq C_T < \infty. \quad (8.12)$$

Proof. Apply (8.8) with $k = b_n$ and $h_R = (|b_n| \wedge R)^{\varepsilon_0/p}$. Then

$$\mathbb{E}_{s,z}^n \int_s^T |b_n(r, Z_r^n)| (|b_n(r, Z_r^n)| \wedge R)^{\varepsilon_0/p} dr \leq C_T \|b_n\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p'} \left(\int_s^T \int_{\mathbb{R}^{2d}} |b_n|^{1+\varepsilon_0} dx dv dr \right)^{1/p}.$$

The right-hand side is bounded uniformly in n , s , and z by (8.11). Letting $R \rightarrow \infty$ and using the monotone convergence theorem proves (8.12). The regularization in Proposition 8.1 can be chosen to satisfy (8.11) since truncation, cutoff and convolution do not increase the $L^{1+\varepsilon_0}$ norm. $\square$

8.2. Proof of Theorem 3.9 completed. We divide the proof into five steps. Steps 1 and 2 are dedicated to existence and distribution estimates parts of Theorem 3.9: In Step 1, we show tightness of $\{\mathbb{P}_{s,z}^n := \mathbb{P} \circ (Z^n)^{-1}\}_n$. In Step 2, we identify subsequential limits as martingale (and so weak) solutions to (3.23), and establish the stability (as $n \rightarrow \infty$) of the Krylov bounds obtained in Lemmas 8.4 and 8.3 and the estimate (3.25) on b . Steps 3 and 4 address respectively the conditional uniqueness and Markov-Feller properties of this solution, and eventually Step 5 identifies the estimates of the one-time marginal densities.

Step 1: Tightness. Choose the single sequence b_n from Proposition 8.1 at p, p_1 , and, when needed, r . Choose it also to satisfy (8.11). For every n , the kinetic SDE with drift b_n has a unique nonexplosive strong solution $Z^n = (X^n, V^n)$. Put

$$\alpha := \frac{\varepsilon_*}{1 + \varepsilon_*} \in (0, 1).$$

Apply Lemma 8.4. Then

$$D_n := \left(\int_s^T |b_n(r, Z_r^n)|^{1+\varepsilon_*} dr \right)^{1/(1+\varepsilon_*)} \quad (8.13)$$

satisfies

$$\sup_n \mathbb{E} D_n^{1+\varepsilon_*} = \sup_n \mathbb{E} \int_s^T |b_n(r, Z_r^n)|^{1+\varepsilon_*} dr \leq C_T.$$

Put $B_t^n := \int_s^t b_n(r, Z_r^n) dr$. Hölder's inequality gives, for $s \leq u < t \leq T$,

$$|B_t^n - B_u^n| \leq |t - u|^{\varepsilon_*/(1+\varepsilon_*)} D_n. \quad (8.14)$$

In particular,

$$\mathbb{P}\left(\max_{s \leq r \leq T} |B_r^n| > R\right) \leq \frac{C_T(T-s)^{\varepsilon_*}}{R^{1+\varepsilon_*}}, \quad \mathbb{P}(\mathfrak{w}(B^n, \delta) > \gamma) \leq \frac{C_T \delta^{\varepsilon_*}}{\gamma^{1+\varepsilon_*}}, \quad (8.15)$$

for $\delta \in (0, 1)$ and $\mathfrak{w}(B^n, \delta)$ the modulus of continuity

$$\mathfrak{w}(B^n, \delta) := \sup_{t, r \in [s, T], |t-r| \leq \delta} |B_t^n - B_r^n|.$$

The first bound in (8.15) together with standard Brownian (moment) estimates immediately gives

$$\lim_{R \rightarrow \infty} \sup_n \mathbb{P}\left(\sup_{s \leq r \leq T} (|X_r^n| + |V_r^n|) > R\right) = 0, \quad (8.16)$$

while the second bound, together with standard Brownian (increment) estimates yields, for any $\epsilon > 0$, and compact $K \subset \mathbb{R}^{2d}$,

$$\lim_{\delta \rightarrow 0} \limsup_n \sup_{\substack{s \in [0, T] \\ z \in K}} \mathbb{P}(\mathfrak{w}(Z^n, \delta) > \epsilon) = 0. \quad (8.17)$$

Prokhorov's theorem (see e.g. [11, Theorem 7.3, Chapter 2]) so applies, ensuring tightness of the measures $\{\mathbb{P}_{s,z}^n = \mathbb{P} \circ (Z^n)^{-1}\}_n$ on the path space $\Omega_{s,T}$.

Step 2: Occupation measures, limit to a weak solution and Krylov bounds. 1. Following Step 1, and without loss of generality, let $\mathbb{P}_{s,z}^n$ denote a converging (sub-)sequence of the laws of Z^n with $\mathbb{P}_{s,z}^n \Rightarrow \mathbb{P}_{s,z}$ on $\Omega_{s,T}$. Define the approximating occupation measures

$$\mu_n(D) := \mathbb{E}_{s,z}^n \int_s^T \mathbf{1}_D(r, Z_r) dr, \quad D, \text{ a Borel set of } \mathbb{R}^{1+2d},$$

(we recall that Z is the canonical process of $\Omega_{s,T}$). The unweighted PDE estimate (8.9) and L^p duality give

$$\mu_n(dr, dz) = \rho_n(r, z) dr dz \quad \sup_n \|\rho_n\|_{p'} < \infty. \quad (8.18)$$

The weak convergence $\mathbb{P}_{s,z}^n \Rightarrow \mathbb{P}_{s,z}$ on $\Omega_{s,T}$ implies vague convergence of μ_n to the occupation measure $\mu(D) = \mathbb{E}_{s,z} \int_s^T \mathbf{1}_D(r, Z_r) dr$, while the compactness in $L^{p'}$ and weak lower semicontinuity of the $L^{p'}$ -norm therefore give, up to extracting a further converging subsequence of $\mathbb{P}_{s,z}^n$,

$$\rho_n \rightharpoonup \rho \text{ in } L^{p'}, \quad \|\rho\|_{p'} \leq \liminf_{n \rightarrow \infty} \|\rho_n\|_{p'} \leq C_T < \infty. \quad (8.19)$$

As such $d\mu = \rho dr d\zeta$ and

$$\begin{aligned} \mathbb{E}_{s,z} \int_s^T |h(r, Z_r)| dr &= \int_{\mathbb{R}^{1+2d}} \mathbf{1}_{[s,T]}(r) |h(r, \zeta)| \mu(dr, d\zeta) = \int_{\mathbb{R}^{1+2d}} \mathbf{1}_{[s,T]}(r) |h(r, \zeta)| \rho(r, \zeta) dr d\zeta \\ &\leq \|\mathbf{1}_{[s,T]} h\|_p \|\rho\|_{p'} \leq C_T \|\mathbf{1}_{[s,T]} h\|_p. \end{aligned} \quad (8.20)$$

This proves the first claimed Krylov bound.

2. Let k satisfy (8.5) and set $d\nu_k = |k| dr d\zeta$. The weighted PDE estimate (8.8) gives

$$\sup_n \|\rho_n\|_{L^{p'}(\nu_k)} \leq C_T \|k\|_{\mathcal{M}_{\mathcal{X}}^q}^{1/p'}. \quad (8.21)$$

From this, we can check that, for all bounded and compactly supported Borel function φ ,

$$\lim_{n \rightarrow \infty} \int \varphi \rho_n d\nu_k = \int \varphi \rho d\nu_k. \quad (8.22)$$

For every $R > 0$, the function $\varphi(|k| \wedge R)$ belongs to L^p with respect to Lebesgue measure. Hence (8.19) gives

$$\int \varphi \rho_n (|k| \wedge R) dr d\zeta \rightarrow \int \varphi \rho (|k| \wedge R) dr d\zeta.$$

Moreover, by Hölder's inequality with respect to $d\nu_k = |k| dr d\zeta$,

$$\begin{aligned} &\left| \int \varphi \rho_n (|k| - |k| \wedge R) dr d\zeta \right| \\ &\leq \|\rho_n\|_{L^{p'}(\nu_k)} \left(\int_{\{|k| > R\}} |\varphi|^p |k| dr d\zeta \right)^{1/p}. \end{aligned}$$

The right-hand side tends to zero as $R \rightarrow \infty$, uniformly in n , by (8.21) and the assumption $k \in L^1$. Thus (8.22) follows.

Now, arguing as in the previous step, i.e. using weak lower semicontinuity of the weighted $L^{p'}$ norm and the a priori weighted Krylov bound, we conclude

$$\mathbb{E}_{s,z} \int_s^T |k(r, Z_r) h(r, Z_r)| dr \leq C_T \|k\|_{\mathcal{M}_{\mathcal{X}}^q}^{1/p'} \|\mathbf{1}_{[s,T]} |k|^{1/p} h\|_p. \quad (8.23)$$

3. The higher-power estimate on $b_n(r, Z_r^n)$ in Lemma 8.4 also survives the passage to the limit. Put $r_* = 1 + \varepsilon_*$. Fix a cutoff $\chi \in C_c^\infty(\mathbb{R}^{1+2d})$ with $0 \leq \chi \leq 1$. We first observe that

$$\mathbb{E}_{s,z}^n \int_s^T (\chi |b_n - b|)(r, Z_r) dr = \int \chi |b_n - b| d\mu_n \rightarrow 0.$$

Indeed, set $k_n = \chi(b_n - b)$. These vector fields belong to $\mathcal{M}_{\mathcal{X}}^q \cap L^1$, their Morrey norms are uniformly bounded, and $\|k_n\|_1 \rightarrow 0$ by (8.3). Applying (8.8) with $k = k_n$ and $h = 1$ gives

$$\int \chi |b_n - b| d\mu_n \leq C_T \|k_n\|_{\mathcal{M}_{\mathcal{X}}^q}^{1/p'} \|k_n\|_1^{1/p} \rightarrow 0.$$

For $M > 0$, the truncated power satisfies the elementary inequality

$$|(|a| \wedge M)^{r_*} - (|c| \wedge M)^{r_*}| \leq r_* M^{r_*-1} |a - c|, \quad a, c \in \mathbb{R}^d.$$

This follows from the mean value theorem on $[0, M]$ and $||a| - |c|| \leq |a - c|$. Hence, for fixed M ,

$$\left| \int \chi (|b_n| \wedge M)^{r_*} d\mu_n - \int \chi (|b| \wedge M)^{r_*} d\mu_n \right| \leq r_* M^{r_*-1} \int \chi |b_n - b| d\mu_n \rightarrow 0.$$

On the other hand, $\chi(|b| \wedge M)^{r^*}$ is a bounded, compactly supported Borel function and hence belongs to L^p with respect to Lebesgue measure. The weak convergence $\rho_n \rightharpoonup \rho$ in $L^{p'}$ therefore gives

$$\int \chi(|b| \wedge M)^{r^*} d\mu_n \rightarrow \int \chi(|b| \wedge M)^{r^*} d\mu.$$

Combining these two convergences, we obtain

$$\int \chi(|b_n| \wedge M)^{r^*} d\mu_n \rightarrow \int \chi(|b| \wedge M)^{r^*} d\mu. \quad (8.24)$$

Since $0 \leq \chi \leq 1$, it follows that

$$\int \chi(|b| \wedge M)^{r^*} d\mu \leq \liminf_{n \rightarrow \infty} \int |b_n|^{r^*} d\mu_n.$$

First let $M \rightarrow \infty$. Then choose cutoffs $0 \leq \chi_j \leq 1$ with $\chi_j \uparrow 1$ on $\mathbb{R}^{1+2d}$ and let $j \rightarrow \infty$. Two applications of the monotone convergence theorem, followed by (8.12), yield

$$\begin{aligned} \mathbb{E}_{s,z} \int_s^T |b(r, Z_r)|^{r^*} dr &= \int |b|^{r^*} d\mu \\ &\leq \liminf_{n \rightarrow \infty} \int |b_n|^{r^*} d\mu_n \leq C_T. \end{aligned} \quad (8.25)$$

This proves the first estimate in (3.25). The second follows from Hölder's inequality:

$$\mathbb{E}_{s,z} \int_u^t |b(r, Z_r)| dr \leq (t-u)^{1-1/r^*} \left(\mathbb{E}_{s,z} \int_u^t |b(r, Z_r)|^{r^*} dr \right)^{1/r^*} \leq C_T^{1/r^*} (t-u)^{1-1/r^*}. \quad (8.26)$$

4. We next show that the probability measure $\mathbb{P}_{s,z}$, on $(\Omega_{s,T}, \{\mathcal{F}_t\}_{s \leq t \leq T})$, constructed above solves the martingale problem for the drift b and initial condition z at time s in the sense of Definition 3.8. For $\varphi \in C^1([s, T]; C_c^2(\mathbb{R}^{2d}))$, and

$$L_r^n \varphi(r, x, v) := \partial_r \varphi(r, x, v) + v \cdot \nabla_x \varphi(r, x, v) - b_n(r, x, v) \cdot \nabla_v \varphi(r, x, v) + \Delta_v \varphi(r, x, v),$$

under $\mathbb{P}_{s,z}^n$, the process $\{M_t^{n,\varphi} := \varphi(t, Z_t) - \varphi(s, z) - \int_s^t L_r^n \varphi(r, Z_r) dr\}_{t \in [s, T]}$ is a martingale and so, for all $s \leq t_1 \leq t_2 \leq \dots \leq t_\ell \leq u < t \leq T$ and $G_u^\ell := G(Z_{t_1}, \dots, Z_{t_\ell})$ for $G : \mathbb{R}^{2d\ell} \rightarrow \mathbb{R}$ continuous bounded

$$\mathbb{E}_{s,z}^n \left[G_u^\ell (M_t^n - M_u^n) \right] = 0.$$

(Owing to a monotone-class argument, the "test function" G_u^ℓ extends the identity from cylinder functions $G(Z_{t_1}, \dots, Z_{t_\ell})$ to every bounded $\mathcal{F}_u$ -measurable random variable.) As $n \rightarrow \infty$ and since $\mathbb{P}_{s,z}^n \Rightarrow \mathbb{P}_{s,z}$,

$$\begin{aligned} \mathbb{E}_{s,z}^n \left[G_u^\ell \left(\varphi(t, Z_t) - \varphi(u, Z_u) - \int_u^t \partial_r \varphi(r, Z_r) + V_r \cdot \nabla_x \varphi(r, Z_r) + \Delta_v \varphi(r, Z_r) dr \right) \right] \\ \rightarrow \mathbb{E}_{s,z} \left[G_u^\ell \left(\varphi(t, Z_t) - \varphi(u, Z_u) - \int_u^t \partial_r \varphi(r, Z_r) + V_r \cdot \nabla_x \varphi(r, Z_r) + \Delta_v \varphi(r, Z_r) dr \right) \right], \end{aligned}$$

and it remains to prove that

$$\mathbb{E}_{s,z}^n \left[G_u^\ell \int_u^t b_n(r, Z_r) \cdot \nabla_v \varphi(r, Z_r) dr \right] \rightarrow \mathbb{E}_{s,z} \left[G_u^\ell \int_u^t b(r, Z_r) \cdot \nabla_v \varphi(r, Z_r) dr \right]. \quad (8.27)$$

To this end, let us introduce a two-parameter approximation $\{b_{n,m}\}_{n,m}$ of b such that $b_m := \lim_{n \rightarrow \infty} b_{n,m}$ and $b_n := \lim_{m \rightarrow \infty} b_{n,m}$ still approximate b in the sense of Proposition 3.2 (Such

approximation can be constructed by a slight modification of the truncation – mollification construction in the proof of Proposition 3.2, e.g. duplicating the cut-off χ_n into $\chi_n \chi_m$ and modifying the calibrated mollifier ρ_{ϵ_n} into $\rho_{\epsilon_{n+m}}$ with the bound $2^{-n} + 2^{-m}$ in (3.13).) With this, it follows that

$$\begin{aligned} & \left| \mathbb{E}_{s,z}^n \left[G_u^\ell \int_u^t b_n(r, Z_r) \cdot \nabla_v \varphi(r, Z_r) dr \right] - \mathbb{E}_{s,z} \left[G_u^\ell \int_u^t b(r, Z_r) \cdot \nabla_v \varphi(r, Z_r) dr \right] \right| \\ & \leq \left| \mathbb{E}_{s,z}^n \left[G_u^\ell \int_u^t b_{n,m}(r, Z_r) \cdot \nabla_v \varphi(r, Z_r) dr \right] - \mathbb{E}_{s,z} \left[G_u^\ell \int_u^t b_m(r, Z_r) \cdot \nabla_v \varphi(r, Z_r) dr \right] \right| \\ & \quad + \left| \mathbb{E}_{s,z}^n \left[G_u^\ell \int_u^t (b_{n,m}(r, Z_r) - b_n(r, Z_r)) \cdot \nabla_v \varphi(r, Z_r) dr \right] \right| \\ & \quad + \left| \mathbb{E}_{s,z} \left[G_u^\ell \int_u^t (b_m(r, Z_r) - b(r, Z_r)) \cdot \nabla_v \varphi(r, Z_r) dr \right] \right|. \end{aligned}$$

Invoking again $\mathbb{P}_{s,z}^n \Rightarrow \mathbb{P}_{s,z}$, the first upper-bound vanishes as $n \rightarrow \infty$, for all fixed m . For the rest, choose $\chi \in C_c^\infty(\mathbb{R}^{1+2d})$ equal to one on $\text{supp } \nabla_v \varphi$ and put $k_{n,m} := \chi(b_{n,m} - b_n)$. Then $k_{n,m} \in \mathcal{M}_{\mathcal{X}}^q \cap L^1$, its Morrey norms are uniformly bounded in n, m , and $\lim_m \sup_n \|k_{n,m}\|_1 = 0$ by (8.3). Hence the estimate (8.8) for Z^n gives

$$\mathbb{E}_{s,z}^n \int_s^T |b_{n,m} - b_n|(r, Z_r) |\nabla_v \varphi(r, Z_r)| dr \leq C_{T, \|\nabla_v \varphi\|_\infty} \|k_{n,m}\|_{\mathcal{M}_{\mathcal{X}}^q}^{1/p'} \|k_{n,m}\|_1^{1/p} \rightarrow 0. \quad (8.28)$$

In the same way, taking $k_m = \chi(b_m - b)$, the estimate (8.23) for Z yields

$$\mathbb{E}_{s,z} \int_s^T |b_m - b|(r, Z_r) |\nabla_v \varphi(r, Z_r)| dr \leq C_{T, \|\nabla_v \varphi\|_\infty} \|k_m\|_{\mathcal{M}_{\mathcal{X}}^q}^{1/p'} \|k_m\|_1^{1/p} \rightarrow 0.$$

Subsequently, (8.27) follows. The estimates in 2. and 3. naturally still hold and this proves (i) in Theorem 3.9.

Step 3: Conditional uniqueness. For any $(s_0, z_0) \in [0, T] \times \mathbb{R}^{2d}$, let us introduce the class $\mathcal{Q}_{s_0, z_0}$ consisting of all probability measures $\mathbb{Q}_{s_0, z_0}$ on $\Omega_{s_0, T}$ that solve the martingale problem (in the sense of Definition 3.8) for b with initial time-state (s_0, z_0) , and lying in the uniqueness class set in (iii) i.e. satisfying, for every Borel bounded function h on $[s_0, T] \times \mathbb{R}^{2d}$,

$$\mathbb{E}_{s_0, z_0}^{\mathbb{Q}} \int_{s_0}^T |h(r, Z_r)| dr \leq K_{\mathbb{Q}} \|\mathbf{1}_{[s_0, T]} h\|_p, \quad (8.29)$$

$$\mathbb{E}_{s_0, z_0}^{\mathbb{Q}} \int_{s_0}^T |b(r, Z_r) h(r, Z_r)| dr \leq K_{\mathbb{Q}} \|\mathbf{1}_{[s_0, T]} |b|^{1/p} h\|_p, \quad (8.30)$$

with $K_{\mathbb{Q}}$ independent of h . Then weak uniqueness of (3.23) holds provided that, for all $(s_0, z_0) \in [0, T] \times \mathbb{R}^{2d}$, $\mathbb{Q}_{s_0, z_0} \in \mathcal{Q}_{s_0, z_0}$, $t \in [s_0, T]$ and for all Borel function $F = F(r, z)$ with $z \mapsto F(r, z)$ belonging to a determining class (or dense in a determining class) for $\mathcal{P}(\mathbb{R}^{2d})$,

$$u(s_0, z_0) := \mathbb{E}_{s_0, z_0}^{\mathbb{Q}} \left[\int_{s_0}^t F(r, Z_r) dr \right] \quad (8.31)$$

is independent of the choice of $\mathbb{Q}_{s_0, z_0}$. This characterization is classically enough to ensure uniqueness of the solution to the martingale problem of b at (s_0, z_0) (we refer to e.g. [57, Theorem 6.2.3 and Corollaries 6.2.4 and 6.2.6] and [25, Chapter 4.4]), and to the uniqueness in law of (3.23).

Fix $F \in C_c^\infty([s_0, T] \times \mathbb{R}^{2d})$, and, applying Theorem 3.6, let ω_n be the solution to the backward PDE $\mathcal{A}^*(\omega_n) + b_n \cdot \nabla_v \omega_n = e^{\lambda(t-T)} F$. Backward causality gives $\omega_n(r, z) = 0$ for all $r \geq T$, and so $u_n(r, z) = e^{\lambda(T-r)} \omega_n(r, z)$ solves the backward PDE:

$$\begin{cases} (-\partial_t - v \cdot \nabla_x - \Delta_v + b_n \cdot \nabla_v) u_n = F, & s_0 < t < T, \\ u_n(T, \cdot) = 0. \end{cases} \quad (8.32)$$

Let us next apply Corollary 8.2 to $\omega_n = e^{\lambda(t-T)}u_n$ at both the exponents p and p_1 given in (3.24). The estimate at p and the kinetic Sobolev embedding give uniform convergence on $[s_0, T] \times \mathbb{R}^{2d}$. The estimate at p_1 , followed by the velocity-gradient bound for the fractional resolvent, gives

$$u_n \rightarrow u_F \quad \text{uniformly,} \quad \sup_n \|\nabla_v u_n\|_{L^{p_1}([s_0, T] \times \mathbb{R}^{2d})} \leq C_T \|F\|_{p_1}. \quad (8.33)$$

In particular, u_F is bounded and continuous, vanishes at time T , and is only determined by (8.32).

For each fixed n , the function u_n is smooth. Indeed, the whole-space resolvent construction above has smooth coefficients and a smooth source, so the same hypoellipticity argument used in the proof of Theorem 3.3 applies to the adjoint equation. The extension of F by zero is smooth because its support lies strictly between s_0 and T . Thus u_n is smooth on a neighborhood of $[s_0, T] \times \mathbb{R}^{2d}$ and is bounded by the PDE estimates. In particular, its derivatives are bounded on every compact space-time cylinder, with bounds that may depend on n .

Consider now $\mathbb{Q}_{s_0, z_0} \in \mathcal{Q}_{s_0, z_0}$ for which, since u_n is smooth,

$$u_n(t, Z_t) - u_n(s_0, z_0) - \int_{s_0}^t (\partial_r + V_r \cdot \nabla_x + \Delta_v - b \cdot \nabla_v) u_n(r, Z_r) dr, \quad s_0 \leq t \leq T,$$

is a martingale. As such, define the bounded stopping time

$$\tau_R := T \wedge \inf\{r \geq s_0 : |Z_r| \geq R\}.$$

Owing to (8.32), it follows that

$$u_n(s_0, z_0) = \mathbb{E}^{\mathbb{Q}_{s_0, z_0}} \left[u_n(\tau_R, Z_{\tau_R}) + \int_{s_0}^{\tau_R} F(r, Z_r) dr \right] - \mathbb{E}^{\mathbb{Q}_{s_0, z_0}} \left[\int_{s_0}^{\tau_R} ((b_n - b) \cdot \nabla_v u_n)(r, Z_r) dr \right]. \quad (8.34)$$

Let us control the last “error” term. Recall from (3.7) and Theorem 8.1 that the approximating sequence $\{b_n\}_n$ used in Step 1 has the form $b_n = \rho_{\varepsilon_n} * (a_n b)$, with a_n the b -dependent truncation given by (3.7). Thus

$$b_n - b = (b_n - a_n b) + (a_n - 1)b,$$

and we can estimate these two contributions separately and show both tend to 0 as $n \rightarrow \infty$.

As a preliminary, for p and p_1 as above, put

$$a := (p^{-1} - p_1^{-1})^{-1} \Leftrightarrow \frac{1}{a} = \frac{1}{p} - \frac{1}{p_1}. \quad (8.35)$$

Since $p_1 > p$, a is finite. Moreover, $\frac{a}{p} = \frac{p_1}{p_1 - p} < q$ by (3.24). In the original diagonal choice of the mollification scales ε_n (see (3.13) in the proof of Proposition 3.2), we have imposed the additional requirement

$$\|\rho_{\varepsilon_n} * (a_n b) - a_n b\|_a \leq 2^{-n}. \quad (8.36)$$

This is possible because, for each fixed n , the vector field $a_n b$ is bounded and compactly supported, and hence belongs to L^a . In turn, by the standard properties of the mollifiers,

$$\|\rho_\varepsilon * (a_n b) - a_n b\|_a \rightarrow 0 \quad \text{as } \varepsilon \downarrow 0.$$

Thus, (8.36) can be imposed together with all the earlier requirements in the same diagonal construction. The functions u_n are defined using this fixed sequence.

We first estimate the truncation contribution $(a_n - 1)b$. Fix R and write

$$K_R = [s_0, T] \times \{|z| \leq R\}, \quad E_n = K_R \cap \{|b| > n\}.$$

For all sufficiently large n , the spatial-temporal cutoff χ_n in (3.7) equals one on K_R . Therefore, on this cylinder,

$$a_n = \min \left\{ 1, \frac{n}{|b|} \right\}, \quad |(a_n - 1)b| = (|b| - n)_+ \leq |b| \mathbf{1}_{E_n},$$

with the usual convention that $a_n = 1$ on $\{b = 0\}$. Using this inequality, the weighted Krylov estimate (8.30) ensures that

$$\begin{aligned} & \left| \mathbb{E}^{\mathbb{Q}_{s_0, z_0}} \left[\int_{s_0}^{\tau_R} (a_n - 1) b(r, Z_r) \cdot \nabla_v u_n(r, Z_r) dr \right] \right| \\ & \leq \mathbb{E}^{\mathbb{Q}_{s_0, z_0}} \int_{s_0}^T \mathbf{1}_{E_n}(r, Z_r) |b(r, Z_r)| |\nabla_v u_n(r, Z_r)| dr \\ & \leq K_{\mathbb{Q}} \|\mathbf{1}_{E_n} |b|^{1/p} \nabla_v u_n\|_p. \end{aligned}$$

Hölder's inequality with the expression (8.35) of a gives

$$\|\mathbf{1}_{E_n} |b|^{1/p} \nabla_v u_n\|_p \leq \left(\int_{E_n} |b|^{a/p} dr d\zeta \right)^{1/a} \|\nabla_v u_n\|_{L^{p_1}(K_R)} \rightarrow 0.$$

Indeed, the second factor is uniformly bounded by (8.33), while

$$\int_{E_n} |b|^{a/p} dr d\zeta \leq n^{a/p-q} \int_{K_R} |b|^q dr d\zeta \rightarrow 0,$$

because $b \in L_{\text{loc}}^q$ and $a/p < q$. This proves that the truncation contribution tends to zero for every fixed R .

For the mollification contribution $b_n - a_n b$, apply the unweighted Krylov estimate (8.29), followed by Hölder's inequality:

$$\begin{aligned} \left| \mathbb{E}^{\mathbb{Q}_{s_0, z_0}} \left[\int_{s_0}^{\tau_R} (b_n - a_n b)(r, Z_r) \cdot \nabla_v u_n(r, Z_r) dr \right] \right| & \leq K_{\mathbb{Q}} \|\mathbf{1}_{K_R} |b_n - a_n b| |\nabla_v u_n|\|_p \\ & \leq K_{\mathbb{Q}} \|b_n - a_n b\|_a \|\nabla_v u_n\|_{L^{p_1}(K_R)} \\ & \leq 2^{-n} K_{\mathbb{Q}} C_T \|F\|_{p_1} \rightarrow 0, \end{aligned}$$

using, for the last inequality, (8.36) and (8.33).

Combining the two bounds proves, for every fixed R ,

$$\lim_n \left| \mathbb{E}^{\mathbb{Q}_{s_0, z_0}} \left[\int_{s_0}^{\tau_R} (b_n - b)(r, Z_r) \cdot \nabla_v u_n(r, Z_r) dr \right] \right| = 0. \quad (8.37)$$

Therefore, letting $n \rightarrow \infty$ in the stopped identity (8.34), using the uniform convergence in (8.33), gives

$$u_F(s_0, z_0) - \mathbb{E}^{\mathbb{Q}_{s_0, z_0}} [u_F(\tau_R, Z_{\tau_R})] = \mathbb{E}^{\mathbb{Q}_{s_0, z_0}} \left[\int_{s_0}^{\tau_R} F(r, Z_r) dr \right].$$

For every continuous path on $[s_0, T]$, $\tau_R \rightarrow T$ as $R \rightarrow \infty$, and the boundedness of u_F and F , together with $u_F(T, \cdot) = 0$ yields the claim (8.31).

Step 4: The Feller and Markov properties. We write, for all bounded Borel function g ,

$$P^{s,t} g(z) := \mathbb{E}_{s,z} g(Z_t), \quad 0 \leq s \leq t \leq T.$$

The Markov property of the weak solution Z of (3.23) is a direct consequence of Steps 1. and 2. yielding the wellposedness of the martingale problem for b at initial time-state (s, z) (see [57, Theorem 10.1.1]).

We can highlight that, if $(s_n, z_n) \rightarrow (s, z)$, the estimates in Step 1 are uniform for initial points in compact sets and for starting times in $[0, T]$. Every limit point is therefore identified by Step 2 as a weak solution started from (s, z) and satisfies the same Krylov estimates. Conditional uniqueness gives $\mathbb{P}_{s_n, z_n} \Rightarrow \mathbb{P}_{s, z}$, so $P^{s,t} g \in C_b$ for $g \in C_{\infty}$.

Since each $\mathbb{P}_{s,z}$ is a probability measure on $\Omega_{s,T}$, its transition kernel is conservative:

$$P^{s,t} \mathbf{1}(z) = \mathbb{P}_{s,z}(Z_t \in \mathbb{R}^{2d}) = 1, \quad 0 \leq s \leq t \leq T.$$

We now prove strong continuity of $P^{s,t}$ in $C_\infty(\mathbb{R}^{2d})$. Fix $g \in C_\infty(\mathbb{R}^{2d})$ and write $h = t - s$. Set $\Phi_h(x, v) = (x + hv, v)$. The integral form of the SDE (3.23) gives

$$Z_t = \Phi_h z + R_{s,t},$$

where the velocity and position components of $R_{s,t}$ are

$$\begin{aligned} R_{s,t}^v &= - \int_s^t b(r, Z_r) dr + \sqrt{2}(W_t - W_s), \\ R_{s,t}^x &= - \int_s^t (t-r)b(r, Z_r) dr + \sqrt{2} \int_s^t (W_r - W_s) dr. \end{aligned}$$

From our PDE theory, uniformly in (s, z) ,

$$\mathbb{E}_{s,z} \int_s^t |b(r, Z_r)| dr \leq C_T^{1/(1+\varepsilon_*)} |t-s|^{\varepsilon_*/(1+\varepsilon_*)} \rightarrow 0 \quad \text{as } t-s \downarrow 0.$$

This yields

$$\sup_z \mathbb{E}_{s,z} |R_{s,t}| \leq C_T \left(h^{\varepsilon_*/(1+\varepsilon_*)} + h^{1/2} \right), \quad 0 \leq s < t \leq T.$$

Every function in $C_\infty(\mathbb{R}^{2d})$ is uniformly continuous. Let

$$\omega_g(\delta) := \sup\{|g(y) - g(y')| : |y - y'| \leq \delta\}$$

be the modulus of continuity of g . Splitting according to whether $|R_{s,t}| \leq \delta$ or not, and applying Markov's inequality, we obtain

$$\begin{aligned} |P^{s,t}g(z) - g(\Phi_h z)| &= |\mathbb{E}_{s,z} g(\Phi_h z + R_{s,t}) - g(\Phi_h z)| \\ &\leq \omega_g(\delta) + 2\|g\|_\infty \mathbb{P}_{s,z}(|R_{s,t}| > \delta) \\ &\leq \omega_g(\delta) + \frac{2\|g\|_\infty}{\delta} \mathbb{E}_{s,z} |R_{s,t}|. \end{aligned}$$

Taking the supremum over s and z , then letting $h \downarrow 0$ and afterwards $\delta \downarrow 0$, proves

$$\sup_{0 \leq s \leq T-h} \|P^{s,s+h}g - g \circ \Phi_h\|_\infty \rightarrow 0.$$

It remains to show that $\|g \circ \Phi_h - g\|_\infty \rightarrow 0$. Here uniform continuity alone is insufficient, since $|\Phi_h z - z| = h|v|$ is unbounded as v varies. Fix $\epsilon > 0$ and choose R so that $|g(y)| < \epsilon$ whenever $|y| > R$. For $0 \leq h \leq 1$, the inverse of Φ_h has operator norm at most two, so $|\Phi_h z| \geq |z|/2$. Thus, when $|z| > 2R$,

$$|g(\Phi_h z) - g(z)| \leq 2\epsilon.$$

When $|z| \leq 2R$, we instead use $|\Phi_h z - z| \leq 2Rh$ to obtain

$$|g(\Phi_h z) - g(z)| \leq \omega_g(2Rh).$$

These two estimates prove the required convergence. Combining it with the preceding estimate gives

$$\lim_{h \downarrow 0} \sup_{0 \leq s \leq T-h} \|P^{s,s+h}g - g\|_\infty = 0.$$

Let us verify joint strong continuity in the two time variables. Suppose $(s_n, t_n) \rightarrow (s, t)$ with $s_n \leq t_n$ and $s \leq t$. The weak continuity of the solution laws established above implies that, whenever $z_n \rightarrow z$,

$$P^{s_n, t_n}g(z_n) \rightarrow P^{s, t}g(z).$$

To interpret this statement on a common path space, extend each path by a constant before its starting time so that it is still in $C([0, T]; \mathbb{R}^{2d})$. The assertion then follows because $(\omega, r) \mapsto \omega(r)$ is continuous for the uniform topology on paths. This sequential statement implies

$$\sup_{|z| \leq R} |P^{s_n, t_n} g(z) - P^{s, t} g(z)| \rightarrow 0 \quad \text{for every } R < \infty.$$

Indeed, otherwise one could choose a sequence of points in the closed ball witnessing failure of convergence and then pass to a convergent subsequence, contradicting the preceding statement and continuity of $P^{s, t} g$.

To pass from compact sets to the whole space, we use the uniform version of the vanishing-at-infinity estimate. Since

$$|\Phi_{t-s} z| \geq \frac{|z|}{1+T}, \quad \sup_{0 \leq s \leq t \leq T} \sup_z \mathbb{E}_{s, z} |R_{s, t}| \leq C_T,$$

we have, for $|z| > (1+T)R$,

$$|P^{s, t} g(z)| \leq \sup_{|y| > R} |g(y)| + \frac{C_T \|g\|_\infty}{|z|/(1+T) - R}.$$

Here the second term bounds the probability that Z_t lies in the ball of radius R : on this event, $|R_{s, t}| \geq |\Phi_{t-s} z| - R$. First choosing R large and then $|z|$ large shows that $P^{s, t} g(z) \rightarrow 0$ uniformly over $0 \leq s \leq t \leq T$. Together with convergence on compact sets, this proves

$$\|P^{s_n, t_n} g - P^{s, t} g\|_\infty \rightarrow 0,$$

which is joint strong continuity.

Thus the transition operators form a strongly continuous Feller propagator on $C_\infty(\mathbb{R}^{2d})$.

Remark 8.5. Another argument that could be applied above consists of adjoining time to the state space and using the standard result that weak continuity of semigroups implies strong continuity.

Step 5: The probability density of Z_t . Fix

$$1 < r < \frac{Q-1}{Q-2} \tag{8.38}$$

from Theorem 3.3(iii) and let $\eta_{s, z}^\lambda$ denote the solution of the point-source PDE:

$$\mathcal{A} \eta - \operatorname{div}_v(b \eta) = \delta_{\{(s, z)\}}.$$

Our goal is to prove

$$e^{-\lambda(t-s)} \mathbb{P}_{s, z}(Z_t \in d\zeta) dt = \eta_{s, z}^\lambda(t, \zeta) dt d\zeta. \tag{8.39}$$

For smooth b_n , the discounted space-time transition density from (s, z) is the point-source solution of

$$\mathcal{A} \eta_n - \operatorname{div}_v(b_n \eta_n) = \delta_{\{(s, z)\}}.$$

At the exponent r in (8.38), Theorem 3.3(ii)–(iii) gives $\eta_n \rightarrow \eta_{s, z}^\lambda$ in $\mathbb{W}_{\mathcal{X}}^{1/r, r}$. On the other hand, weak convergence $\mathbb{P}_{s, z}^n \Rightarrow \mathbb{P}_{s, z}$ of the path laws following Step 1, identifies the distributional limit of $\eta_n dt d\zeta$ with $e^{-\lambda(t-s)} \mathbb{P}_{s, z}(Z_t \in d\zeta) dt$. The limits agree, which proves (8.39) and yields (iv). $\square$

Remark 8.6. We could give an alternative proof of Theorem 3.9 that reverses the order of the preceding proof: first one constructs a conservative Feller propagator directly from the backward PDE, and only then identifies the corresponding probability measures, forming a strong Markov family parameterized by the initial time and the initial point, as weak solutions of (3.23). Our proof in [37] has this structure. This construction does not use tightness of the laws of the smooth SDEs

and does not require conditional uniqueness in the Krylov class. The same applies to the proof of Theorem 3.10 that we give next.

9. REMOVAL OF GLOBAL INTEGRABILITY CONDITION USING WEIGHTS (PROOF OF THEOREM 3.10)

We now remove the global $L^{1+\varepsilon_0}$ assumption from Theorem 3.9 and restore the bounded component of the drift permitted by the PDE theory. Specifically, throughout this section we replace the global hypothesis $(H_{\mathcal{K},\varepsilon_0})$ by the original condition $(H_{\mathcal{K}})$:

$$b = b_s + b_b, \quad b_s \in \mathcal{M}_{\mathcal{K}}^q(\mathbb{R}^{1+2d}, \mathbb{R}^d), \quad b_b \in L^\infty(\mathbb{R}^{1+2d}, \mathbb{R}^d), \quad q > 1, \quad (9.1)$$

where no global Lebesgue $L^{1+\varepsilon}$ -integrability is imposed on b_s . The Morrey norm of b_s is small, whereas the size of b_b is arbitrary and is absorbed by taking λ large. All backward and forward smallness conditions on the denominator of the Duhamel series used below are assumed strict.

The localization used to handle $(H_{\mathcal{K}})$ will be carried out using a vanishing at infinity weight, centered at the point ζ_0 from which the diffusion $Z_t = (X_t, V_t)$ in (3.23) is started. Namely, define

$$\kappa > Q, \quad \rho_{\zeta_0}(\zeta) := (1 + d_*(\zeta, \zeta_0))^{-\kappa}, \quad \zeta_0, \zeta \in \mathbb{R}^{1+2d},$$

and write

$$\|f\|_{L_{\rho_{\zeta_0}}^p} := \left(\int_{\mathbb{R}^{1+2d}} |f(\zeta)|^p \rho_{\zeta_0}(\zeta) d\zeta \right)^{1/p}.$$

The centering at ζ_0 is essential: $\rho_{\zeta_0}(\zeta_0) = 1$, so a weighted estimate followed by evaluation at ζ_0 is uniform in the starting point. Note also that the triangle inequality for d_* gives

$$\frac{\rho_{\zeta_0}(\zeta)}{\rho_{\zeta_0}(\xi)} \leq C_{d,\kappa} (1 + d_*(\zeta, \xi))^\kappa, \quad (9.2)$$

uniformly in ζ_0 . Moreover, by (2.5) and $\kappa > Q$,

$$\sup_{\zeta_0} \int \rho_{\zeta_0}(\zeta) d\zeta < \infty. \quad (9.3)$$

Lemma 9.1 (Centered weighted backward estimate). *Let $p > Q - 1$, $q > 1$, and let $b = b_s + b_b$ satisfy (9.1). For $\lambda \geq 1$, define*

$$\mathfrak{r}_{p,\lambda}^\rho(b) := C_{d,p,q,\kappa}^{R,s} \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p} + C_{d,p,\kappa}^{R,b} \lambda^{-1/(2p)} \|b_b\|_\infty^{1/p}, \quad (9.4)$$

$$\mathfrak{q}_{p,\lambda}^\rho(b) := C_{d,p,q,\kappa}^{Q,s} \|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}^{1/p'} + C_{d,p,\kappa}^{Q,b} \lambda^{-1/(2p')} \|b_b\|_\infty^{1/p'}. \quad (9.5)$$

The operators in (8.1) extend to $L_{\rho_{\zeta_0}}^p$, uniformly in the center ζ_0 , and satisfy

$$\|\widehat{R}_p(b)\|_{L_{\rho_{\zeta_0}}^p \rightarrow L_{\rho_{\zeta_0}}^p} \leq \mathfrak{r}_{p,\lambda}^\rho(b), \quad (9.6)$$

$$\|\widehat{Q}_p(b)\|_{L_{\rho_{\zeta_0}}^p \rightarrow L_{\rho_{\zeta_0}}^p} \leq \mathfrak{q}_{p,\lambda}^\rho(b). \quad (9.7)$$

Moreover,

$$|(\mathcal{A}^*)^{-1/2-1/(2p)} f(\zeta_0)| \leq C_{d,p,\lambda,\kappa} \|f\|_{L_{\rho_{\zeta_0}}^p}. \quad (9.8)$$

Therefore, if

$$\Theta_{p,\lambda}^\rho(b) := \mathfrak{q}_{p,\lambda}^\rho(b) \mathfrak{r}_{p,\lambda}^\rho(b) < 1, \quad (9.9)$$

then $I + \widehat{Q}_p(b) \widehat{R}_p(b)$ is invertible on $L_{\rho_{\zeta_0}}^p$, with inverse norm bounded uniformly in ζ_0 .

Proof. Put

$$M := \|b_s\|_{\mathcal{M}_{\mathcal{X}}^q}, \quad B := \|b_b\|_{\infty}.$$

All constants in the estimates below are independent of the center ζ_0 .

Step 1: Off-diagonal kernel bounds. Let K_γ^* be the kernel of $(\mathcal{A}^*)^{-\gamma/2}$ (see (B.1) in Appendix B) and let D_γ^* be the kernel of $\nabla_v(\mathcal{A}^*)^{-(1+\gamma)/2}$. The explicit formula used in the proof of Proposition 2.2 (see again Appendix B) shows, on the time half-space on which either kernel is nonzero, that

$$|K_\gamma^*(\zeta, \xi)| + |D_\gamma^*(\zeta, \xi)| \leq C_{d,\gamma} \tau^{(\gamma-Q)/2} e^{-\lambda\tau} e^{-c_d \Phi_\tau(\zeta, \xi)}. \quad (9.10)$$

Here τ is the absolute time difference and Φ_τ is the sum of the two normalized quadratic expressions appearing in (2.8). Notice that the powers of τ are the same for the two kernels: differentiation contributes $\tau^{-1/2}$, while the fractional power in D_γ^* is higher by one.

Set $r = d_*(\zeta, \xi)$. By the equivalence of d and d_* , either $\tau \geq cr^2$, or the Gaussian exponent satisfies $\Phi_\tau(\zeta, \xi) \geq cr^2/\tau$. In the first case the factor $e^{-\lambda\tau}$ is bounded by e^{-cr^2} because $\lambda \geq 1$. In the second case,

$$e^{-\lambda\tau - cr^2/\tau} \leq e^{-c'r}$$

by the arithmetic–geometric mean inequality. It follows from (9.10) that, for every $L > 0$,

$$|K_\gamma^*(\zeta, \xi)| + |D_\gamma^*(\zeta, \xi)| \leq C_{d,\gamma,L} r^{\gamma-Q} (1+r)^{-L}, \quad r \geq 1. \quad (9.11)$$

Together with the near-diagonal estimates of Proposition 2.2, this implies that, for every $\theta \geq 0$ and $0 < \gamma < Q$,

$$(1 + d_*(\zeta, \xi))^\theta (|K_\gamma^*(\zeta, \xi)| + |D_\gamma^*(\zeta, \xi)|) \leq C_{d,\gamma,\theta} d_*(\zeta, \xi)^{\gamma-Q}. \quad (9.12)$$

Indeed, the polynomial factor is bounded near the diagonal and is absorbed by the arbitrary power $(1+r)^{-L}$ in (9.11) away from it. The same statements hold for the non-adjoint kernels. We will use (9.12) with $\gamma = 1/p$ and $1/p'$; the off-diagonal form (9.11) with $\gamma = 1 + 1/p$ will be used for point evaluation.

The constants in (9.11) and (9.12) can be chosen uniformly for $\lambda \geq 1$: for these bounds one uses $e^{-\lambda\tau} \leq e^{-\tau}$ throughout. This fixes the singular-part constants independently of λ . We will also retain the λ -dependence needed for the bounded component. Upon writing

$$\mathcal{E}_\gamma^*(\zeta, \xi) := |K_\gamma^*(\zeta, \xi)| + |D_\gamma^*(\zeta, \xi)|,$$

Gaussian integration and the kinetic scaling give, for every $\theta \geq 0$ and $0 < \gamma < Q$,

$$\begin{aligned} & \sup_\zeta \int (1 + d_*(\zeta, \xi))^\theta \mathcal{E}_\gamma^*(\zeta, \xi) d\xi \\ & + \sup_\xi \int (1 + d_*(\zeta, \xi))^\theta \mathcal{E}_\gamma^*(\zeta, \xi) d\zeta \leq C_{d,\gamma,\theta} (\lambda^{-\gamma/2} + \lambda^{-(\gamma+\theta)/2}) \leq C_{d,\gamma,\theta} \lambda^{-\gamma/2}. \end{aligned} \quad (9.13)$$

Indeed, after integration of the normalized Gaussian in the space variables, the terms without and with the distance moment are bounded by constant multiples of

$$\int_0^\infty \tau^{\gamma/2-1} e^{-\lambda\tau} d\tau \quad \text{and} \quad \int_0^\infty \tau^{(\gamma+\theta)/2-1} e^{-\lambda\tau} d\tau,$$

respectively. The same computation applies with the variables reversed.

Step 2: Conjugation by the weight. Put

$$\omega_{\zeta_0} := \rho_{\zeta_0}^{1/p}, \quad (U_{\zeta_0} f)(\zeta) := \omega_{\zeta_0}(\zeta) f(\zeta).$$

Then

$$U_{\zeta_0} : L_{\rho_{\zeta_0}}^p \rightarrow L^p$$

is an isometric isomorphism. Thus the weighted operator norm of an operator T is exactly the ordinary L^p -operator norm of $U_{\zeta_0} T U_{\zeta_0}^{-1}$. From (9.2),

$$\frac{\omega_{\zeta_0}(\zeta)}{\omega_{\zeta_0}(\xi)} \leq C_{d,\kappa,p} (1 + d_*(\zeta, \xi))^{\kappa/p}, \quad (9.14)$$

uniformly in ζ_0 .

We first treat $\widehat{R}_p(b)$. Since

$$|b|^{1/p} \leq |b_s|^{1/p} + B^{1/p},$$

its conjugated kernel is bounded by the sum of a singular part and a bounded part. For the singular part, (9.14) and (9.12), with $\theta = \kappa/p$ and $\gamma = 1/p$, give

$$|U_{\zeta_0} \widehat{R}_p(b) U_{\zeta_0}^{-1} h| \leq C |b_s|^{1/p} I_{1/p}(|h|) + B^{1/p} \mathcal{D}_{p,\lambda}^\rho(|h|),$$

where $\mathcal{D}_{p,\lambda}^\rho$ is the positive integral operator with kernel

$$C(1 + d_*(\zeta, \xi))^{\kappa/p} |D_{1/p}^*(\zeta, \xi)|.$$

The Adams estimate (5.3) controls the first term, while (9.13) and the Schur test give

$$\|\mathcal{D}_{p,\lambda}^\rho\|_{p \rightarrow p} \leq C_{d,p,\kappa} \lambda^{-1/(2p)}.$$

Hence

$$\|U_{\zeta_0} \widehat{R}_p(b) U_{\zeta_0}^{-1}\|_{p \rightarrow p} \leq C M^{1/p} + C \lambda^{-1/(2p)} B^{1/p} = \mathfrak{r}_{p,\lambda}^\rho(b),$$

after fixing the constants in (9.4). This is (9.6).

For $\widehat{Q}_p(b)$, use

$$|b|^{1/p'} \leq |b_s|^{1/p'} + B^{1/p'}.$$

The singular part of the conjugated kernel satisfies

$$|U_{\zeta_0} \widehat{Q}_p(b) U_{\zeta_0}^{-1} h| \leq C I_{1/p'}(|b_s|^{1/p'} |h|) + B^{1/p'} \mathcal{K}_{p,\lambda}^\rho(|h|),$$

where $\mathcal{K}_{p,\lambda}^\rho$ has kernel

$$C(1 + d_*(\zeta, \xi))^{\kappa/p} |K_{1/p'}^*(\zeta, \xi)|.$$

Since $I_{1/p'}$ has a symmetric kernel, duality and (5.3), used with exponent p' , yield

$$\begin{aligned} \|I_{1/p'}(|b_s|^{1/p'} |h|)\|_p &= \sup_{\|\varphi\|_{p'} \leq 1} \int |b_s|^{1/p'} |h| I_{1/p'}(|\varphi|) \\ &\leq \|h\|_p \sup_{\|\varphi\|_{p'} \leq 1} \| |b_s|^{1/p'} I_{1/p'}(|\varphi|) \|_{p'} \\ &\leq C M^{1/p'} \|h\|_p. \end{aligned}$$

The Schur test and (9.13), now with $\gamma = 1/p'$, give

$$\|\mathcal{K}_{p,\lambda}^\rho\|_{p \rightarrow p} \leq C_{d,p,\kappa} \lambda^{-1/(2p')}.$$

It follows that

$$\|U_{\zeta_0} \widehat{Q}_p(b) U_{\zeta_0}^{-1}\|_{p \rightarrow p} \leq C M^{1/p'} + C \lambda^{-1/(2p')} B^{1/p'} = \mathfrak{q}_{p,\lambda}^\rho(b),$$

which proves (9.7). Initially the calculation is made on compactly supported functions. Since C_c^∞ is dense in $L_{\rho_{\zeta_0}}^p$, the two conjugated estimates give the asserted extensions.

Step 3: Evaluation at the center and inversion. By (2.10), the kernel of $(\mathcal{A}^*)^{-1/2-1/(2p)}$ is bounded near the diagonal by $C d_*(\zeta, \xi)^{1+1/p-Q}$. Its p' -th power is locally integrable if and only if

$$\left(1 + \frac{1}{p} - Q\right) p' + Q > 0, \quad \text{equivalently} \quad 1 + \frac{1}{p} > \frac{Q}{p}. \quad (9.15)$$

Condition (9.15) is precisely $p > Q - 1$.

Let $G^*(\zeta_0, \xi)$ be the kernel of $(\mathcal{A}^*)^{-1/2-1/(2p)}$. Hölder's inequality with the weight gives

$$|(\mathcal{A}^*)^{-1/2-1/(2p)} f(\zeta_0)| \leq \|f\|_{L_{\rho_{\zeta_0}}^p} \left(\int |G^*(\zeta_0, \xi)|^{p'} \rho_{\zeta_0}(\xi)^{-p'/p} d\xi \right)^{1/p'}.$$

On $B_1(\zeta_0)$ the last integral is finite by (9.15). On its complement, (9.11) and $\rho_{\zeta_0}^{-p'/p} \leq (1 + d_*(\zeta_0, \cdot))^{\kappa p'/p}$ make it finite after taking L sufficiently large. This proves (9.8).

Finally, $L_{\rho_{\zeta_0}}^p$ is a Banach space, and (9.6)–(9.7) imply

$$\|\widehat{Q}_p(b) \widehat{R}_p(b)\|_{L_{\rho_{\zeta_0}}^p \rightarrow L_{\rho_{\zeta_0}}^p} \leq q_{p,\lambda}^\rho(b) \mathfrak{r}_{p,\lambda}^\rho(b) = \Theta_{p,\lambda}^\rho(b) < 1.$$

Therefore, the series

$$(I + \widehat{Q}_p(b) \widehat{R}_p(b))^{-1} = \sum_{n=0}^{\infty} (-\widehat{Q}_p(b) \widehat{R}_p(b))^n$$

converges in operator norm on $L_{\rho_{\zeta_0}}^p$, uniformly in the center. This completes the proof. $\square$

The weighted space does not impose an additional assumption on the drift.

In the next lemma the choice $\kappa > Q$ in the definition of the weight is used in an essential manner.

Lemma 9.2 (Centered weighted moments of the full drift). *Let $b = b_s + b_b$ satisfy (9.1), $0 < a \leq q$, and let $\kappa > Q$. Then*

$$\sup_{\zeta_0 \in \mathbb{R}^{1+2d}} \int_{\mathbb{R}^{1+2d}} |b(\zeta)|^a \rho_{\zeta_0}(\zeta) d\zeta \leq C_{d,q,a,\kappa} \left(\|b_s\|_{\mathcal{M}_{\mathcal{X}}^q}^a + \|b_b\|_{\infty}^a \right).$$

Proof. Put $M = \|b_s\|_{\mathcal{M}_{\mathcal{X}}^q}$ and $B = \|b_b\|_{\infty}$. From the definition of the Morrey norm and (2.5),

$$\int_{B_R(\zeta_0)} |b_s|^q \leq C_d M^q R^{Q-q}.$$

Hence Hölder's inequality gives, for every $R > 0$,

$$\int_{B_R(\zeta_0)} |b_s|^a \leq |B_R(\zeta_0)|^{1-a/q} \left(\int_{B_R(\zeta_0)} |b_s|^q \right)^{a/q} \leq C_{d,q,a} M^a R^{Q-a}. \quad (9.16)$$

Now set $A_0 = B_1(\zeta_0)$ and

$$A_{k+1} = B_{2^{k+1}}(\zeta_0) \setminus B_{2^k}(\zeta_0), \quad k \geq 0.$$

On A_{k+1} one has $\rho_{\zeta_0} \leq 2^{-k\kappa}$, and therefore (9.16) with $R = 2^{k+1}$ yields

$$\int_{A_{k+1}} |b_s|^a \rho_{\zeta_0} \leq C M^a 2^{-k\kappa} 2^{k(Q-a)} = C M^a 2^{-k(\kappa-Q+a)}.$$

The contribution of A_0 is bounded by (9.16) with $R = 1$, and the displayed geometric series converges because $\kappa > Q$. Thus

$$\sup_{\zeta_0} \int |b_s|^a \rho_{\zeta_0} \leq C M^a.$$

Moreover, (9.3) gives

$$\sup_{\zeta_0} \int |b_b|^a \rho_{\zeta_0} \leq C B^a.$$

The result follows from $|b_s + b_b|^a \leq C_a(|b_s|^a + |b_b|^a)$. $\square$

We call p *SDE admissible* if it is backward admissible in the sense of (8.2) and also satisfies $\Theta_{p,\lambda}^\rho(b) < 1$ in (9.9). By (9.4) and (9.5), this condition holds after increasing λ , provided that the Morrey norm of b_s satisfies the corresponding strict smallness condition. No smallness is imposed on $\|b_b\|_\infty$.

Lemma 9.3 (Higher-power Krylov estimate). *Let $p > Q - 1$ be SDE admissible and choose*

$$0 < \varepsilon < \min \left\{ 1, \frac{q-1}{p} \right\}.$$

Let $b_n = b_{n,s} + b_{n,b}$ be the regularization from Proposition 8.1, chosen with a uniform strict bound in (9.9). Then

$$\sup_n \sup_{0 \leq s < T} \sup_{z \in \mathbb{R}^{2d}} \mathbb{E}_{s,z}^n \int_s^T |b_n(r, Z_r^n)|^{1+\varepsilon} dr \leq C_T < \infty.$$

Proof. Fix n, s, T, z and put $\zeta_0 = (s, z)$. Set $F_n = \mathbf{1}_{[s,T]} |b_n|^{1+\varepsilon}$ and define the vector fields

$$\ell_n := \mathbf{1}_{[s,T]} \operatorname{sgn}(b_n) |b_n|^{1/p+\varepsilon}, \quad \operatorname{sgn}(b_n) := \frac{b_n}{|b_n|} \mathbf{1}_{\{b_n \neq 0\}},$$

so that $F_n = b_n^{1/p'} \cdot \ell_n$. Since $1 + p\varepsilon < q$, Lemma 9.2, with $a = 1 + p\varepsilon$, gives

$$\|\ell_n\|_{L_{\rho_{\zeta_0}}^p}^p = \int_{\{s \leq t \leq T\}} |b_n|^{1+p\varepsilon} \rho_{\zeta_0} \leq C_{d,p,q,\varepsilon,\kappa} \left(\|b_{n,s}\|_{\mathcal{M}_{\mathcal{K}}^q}^{1+p\varepsilon} + \|b_{n,b}\|_\infty^{1+p\varepsilon} \right) \leq C. \quad (9.17)$$

The centered weighted backward factorization is therefore well-defined and gives

$$u_n = (\mathcal{A}^*)^{-1/2-1/(2p)} (I + \widehat{Q}_p(b_n) \widehat{R}_p(b_n))^{-1} \widehat{Q}_p(b_n) \ell_n.$$

Because b_n and ℓ_n are compactly supported, the same Duhamel partial sums occur in the global L^p factorization from Theorem 3.6. They converge both in global L^p and in $L_{\rho_{\zeta_0}}^p$. Since both convergences imply local L^p convergence, the two limits coincide. Thus u_n is the usual smooth resolvent solution. By (9.7), (9.8), (9.9), and (9.17),

$$|u_n(s, z)| \leq C q_{p,\lambda}^\rho(b_n) \|\ell_n\|_{L_{\rho_{\zeta_0}}^p} \leq C. \quad (9.18)$$

The adjoint resolvent is backward causal, so u_n vanishes for $t > T$. Apply Itô's formula to $u_n(t, Z_t^n)$, first with a smooth approximation of F_n . The resolvent parameter gives the “discount” factor $e^{-\lambda(r-s)}$. Positivity of F_n , (9.18), and removal of the discount on $[s, T]$ yield

$$\mathbb{E}_{s,z}^n \int_s^T |b_n(r, Z_r^n)|^{1+\varepsilon} dr \leq e^{\lambda T} |u_n(s, z)| \leq C_T.$$

The constant is uniform in n, s, z . Applying Fatou's lemma ends the proof. $\square$

We are now in a position to demonstrate Theorem 3.10.

Proof of Theorem 3.10. Choose, simultaneously at p, p_1 satisfying $p > Q - 1$, $p_1 > pq/(q-1)$, and, when necessary, r , the single sequence

$$b_n = b_{n,s} + b_{n,b}$$

from Proposition 8.1. Both approximating components use the same cutoff index and the same convolution scale. In view of (8.3) and (9.4)–(9.5), the sequence may be chosen so that

$$\sup_n \Theta_{p,\lambda}^\rho(b_n) < 1.$$

Here one first fixes the strict smallness margin for $\|b_s\|_{\mathcal{M}_{\mathcal{K}}^q}$ and then takes λ large enough to absorb the common L^∞ bound of $b_{n,b}$.

Starting from the diffusion $Z^n = (X^n, V^n)$ with drift $b_n = b_{n,s} + b_{n,b}$, we can extend each main step in the preceding proof of Theorem 3.9.

Tightness. Apply Lemma 9.3 with $\varepsilon = \varepsilon_*$ and its conclusion is exactly the control (8.13). Subsequently, the pathwise Hölder estimate (8.14), the tail bounds (8.15), and the control (8.17) in Step 1 of Theorem 3.9 apply verbatim. Thus, again up to a subsequence $\mathbb{P}_{s,z}^n := \mathbb{P} \circ (Z^n)^{-1}$ are tight and converge to a limit $\mathbb{P}_{s,z}$.

Occupation measures, passage to a weak solution and Krylov bounds. The proof of the bounds (8.20) and (8.23) proceeds as in parts 1 and 2 in Step 2 of the proof of Theorem 3.9. Regarding the control $\mathbb{E}_{s,z} \int_s^T |b(r, Z_r)|^{1+\varepsilon_*} dr$, we verify explicitly how the two components b_s, b_b enter the arguments of part 3 of Step 2. Let $\chi \in C_c^\infty(\mathbb{R}^{1+2d})$, $0 \leq \chi \leq 1$, be one on the space-time support of the derivative of a martingale test function, and set

$$k_{n,s} := \chi(b_{n,s} - b_s), \quad h_{n,b} := \chi(b_{n,b} - b_b).$$

Then

$$\sup_n \|k_{n,s}\|_{\mathcal{M}_{\mathcal{X}}^q} < \infty, \quad \|k_{n,s}\|_1 \rightarrow 0, \quad \|h_{n,b}\|_p \rightarrow 0.$$

The first two assertions follow from the Morrey bound and local L^q -convergence; the last follows from the local L^p -convergence of the bounded components. Applying respectively the weighted and unweighted Krylov estimates (8.8) and (8.9) gives

$$\begin{aligned} \mathbb{E}_{s,z}^n \int_s^T |b_n - b|(r, Z_r) |\nabla_v \varphi(Z_r)| dr \\ \leq C_{T,\varphi} \left(\|k_{n,s}\|_{\mathcal{M}_{\mathcal{X}}^q}^{1/p'} \|k_{n,s}\|_1^{1/p} + \|h_{n,b}\|_p \right) \rightarrow 0. \end{aligned}$$

Thus the varying coefficient causes no new difficulty.

For the fixed coefficient, repeat the weighted occupation-density argument of Step 2 with $k = \chi b_s$. The bounded component is simpler: the signed occupation densities are bounded in $L^{p'}$ by (8.9), while $\chi b_b \cdot \nabla_v \varphi \in L^p$; hence ordinary L^p - $L^{p'}$ duality passes this term to the limit. The same decomposition proves, for every $\psi \in C_c(\mathbb{R}^{1+2d}, \mathbb{R}^d)$,

$$\int b_n \cdot \psi d\mu_n \rightarrow \int b \cdot \psi d\mu.$$

Consequently the argument (8.24), (8.25) applies to the full drift and gives

$$\mathbb{E}_{s,z} \int_s^T |b(r, Z_r)|^{1+\varepsilon_*} dr \leq C_T.$$

Owing to this estimate, (8.28) and all the arguments of Step 2.4 hold and ensure that $\mathbb{P}_{s,z}$ is a solution to the martingale problem for b .

Conditional uniqueness. Only the error term (8.37) in Step 3 of Theorem 3.9 needs comments. Let

$$\frac{1}{a} = \frac{1}{p} - \frac{1}{p_1};$$

then $a/p < q$. On the compact cylinder determined by the stopping time, write

$$\tilde{b}_{n,s} := a_n b_s, \quad b_{n,s} = \rho_{\varepsilon_n} * \tilde{b}_{n,s}.$$

Here a_n is the common scalar cutoff of (3.7). On the fixed cylinder, $\chi_n = 1$ for all sufficiently large n , and

$$|b_s - \tilde{b}_{n,s}| \leq |b_s| \mathbf{1}_{\{|b|>n\}} \leq (|b| + \|b_b\|_\infty) \mathbf{1}_{\{|b|>n\}}.$$

Set $E_n = \{|b| > n\}$ on this cylinder. These sets decrease to a null set. The two conditional Krylov bounds in Theorem 3.9(iii) give

$$\begin{aligned} & \mathbb{E}_{s,z}^{\mathbb{Q}} \int_s^T |b_s - \tilde{b}_{n,s}|(r, Z_r) |\nabla_v u_n(r, Z_r)| dr \\ & \leq C_{\mathbb{Q}} \left(\| |b|^{1/p} \mathbf{1}_{E_n} \nabla_v u_n \|_p + \|b_b\|_{\infty} \| \mathbf{1}_{E_n} \nabla_v u_n \|_p \right) \\ & \leq C_{\mathbb{Q}} \| \nabla_v u_n \|_{p_1} \left(\| |b|^{1/p} \mathbf{1}_{E_n} \|_a + \|b_b\|_{\infty} \| \mathbf{1}_{E_n} \|_a \right) \rightarrow 0. \end{aligned} \quad (9.19)$$

All norms in this display are restricted to the fixed cylinder. The last limit follows from $a/p < q$, $b = b_s + b_b \in L_{\text{loc}}^{a/p}$, and absolute continuity of the integral.

The single scale ε_n in Proposition 8.1 can simultaneously be chosen so that

$$\| \tilde{b}_{n,s} - \rho_{\varepsilon_n} * \tilde{b}_{n,s} \|_a \leq 2^{-n}$$

on the n th compact cylinder, without changing any factor convergence. The remaining singular error is therefore controlled by the unweighted conditional Krylov estimate and the uniform L^{p_1} -bound on $\nabla_v u_n$. For the bounded component the same estimate gives, on every fixed cylinder,

$$\begin{aligned} & \mathbb{E}_{s,z}^{\mathbb{Q}} \int_s^T |b_{n,b} - b_b|(r, Z_r) |\nabla_v u_n(r, Z_r)| dr \\ & \leq C_{\mathbb{Q}} \|b_{n,b} - b_b\|_a \| \nabla_v u_n \|_{p_1} \rightarrow 0 \end{aligned} \quad (9.20)$$

by (8.3). It follows from (9.19) and (9.20) that the error term tends to zero. The same estimates hold after multiplication by any bounded $\mathcal{F}_{t_0}$ -measurable factor G , with the right-hand sides multiplied by $\|G\|_{\infty}$. Consequently they prove the representation (8.31), and by extension uniqueness of the martingale problem for the full drift $b = b_s + b_b$.

The Feller and Markov properties. Finally, the Feller and strong Markov arguments use only the uniform higher-power estimate for the full drift, and the density identification is Theorem 3.3(ii)–(iii) for the decomposition $b = b_s + b_b$. Steps 4 and 5 of Theorem 3.9 therefore apply without further changes and complete the proof. $\square$

APPENDIX A. FRACTIONAL POWERS OF LINEAR OPERATORS

We recall the construction of the fractional powers of operators that we use throughout the paper.

We follow Komatsu [39, Sections 7 and 11]. Specifically, see Propositions 11.1 and 11.4 there, and Theorem 11.8 for the definition of positive fractional powers via Bochner-Phillips subordination. (Regarding the definition via subordination, see also the exposition in Yosida [59, Chapter IX, §11].)

1. For $\lambda > 0$ and $s \in \mathbb{R}$, recall that the Fourier transform definition of the fractional power of the (shifted) Laplacian on $L^2 = L^2(\mathbb{R}^d)$ is

$$\begin{aligned} \mathcal{F}[(\lambda - \Delta)^s f](\xi) &= (\lambda + |\xi|^2)^s \widehat{f}(\xi), \\ D((\lambda - \Delta)^s) &= \{f \in L^2 : (\lambda + |\xi|^2)^s \widehat{f} \in L^2\}, \end{aligned} \quad (\text{A.1})$$

where $\mathcal{F}$ is the unitary Fourier transform normalized by $\mathcal{F}(-\Delta f)(\xi) = |\xi|^2 \widehat{f}(\xi)$.

2. The semigroup formulation below also works in L^p , and agrees with the Fourier transform definition.

Let $1 < p < \infty$, and let $-B$ generate a strongly continuous semigroup $T(t)$ on $L^p = L^p(E, \mu)$, where (E, μ) is a fixed σ -finite measure space, such that $\|T(t)\|_{p \rightarrow p} \leq M e^{-\omega t}$ for some $M \geq 1$ and

$\omega > 0$. For every $\beta > 0$, the negative (fractional) power $B^{-\beta}$ of B is the bounded operator

$$B^{-\beta} f := \frac{1}{\Gamma(\beta)} \int_0^\infty t^{\beta-1} T(t) f dt, \quad \|B^{-\beta}\|_{p \rightarrow p} \leq M \omega^{-\beta}, \quad (\text{A.2})$$

where the integral converges in L^p . Using the identity

$$\int_0^1 u^{\beta-1} (1-u)^{\delta-1} du = \frac{\Gamma(\beta)\Gamma(\delta)}{\Gamma(\beta+\delta)}, \quad \beta, \delta > 0, \quad (\text{A.3})$$

the semigroup property $T(t)T(s) = T(t+s)$, Fubini's theorem and the change of variables $r = t+s$, $u = t/(t+s)$, we obtain $B^{-\beta} B^{-\delta} = B^{-(\beta+\delta)}$.

Remark A.1 (Injectivity and dense range). The Laplace-resolvent formula identifies B^{-1} with the inverse of B . To see that $B^{-\beta}$ is injective, choose an integer $m > \beta$. Then $B^{-\beta} f = 0$ implies $B^{-m} f = B^{-(m-\beta)} B^{-\beta} f = 0$, hence $f = 0$, since $B^{-m} = (B^{-1})^m$ is injective. Moreover, $\text{Ran}(B^{-\beta})$ contains $\text{Ran}(B^{-m}) = D(B^m)$, which is dense by the standard resolvent approximation.

We therefore define the positive power as the closed operator $B^\beta = (B^{-\beta})^{-1}$, with $D(B^\beta) = \text{Ran}(B^{-\beta})$. Thus $B^\beta B^{-\beta} = I$ on L^p and $B^{-\beta} B^\beta = I$ on $D(B^\beta)$. For $0 < \alpha < 1$, the scalar identity

$$r^\alpha = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (1-e^{-tr}) \frac{dt}{t^{1+\alpha}}, \quad r \geq 0, \quad (\text{A.4})$$

motivates the corresponding semigroup representation: B^α is equivalently the closure of

$$B^\alpha f = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (f - T(t)f) \frac{dt}{t^{1+\alpha}}, \quad f \in D(B). \quad (\text{A.5})$$

In particular, $D(B)$ is a core for B^α . The positive powers satisfy $B^\beta B^\delta = B^{\beta+\delta}$ on the natural composition domain.

Böchner subordination further shows that $-B^\alpha$, $0 < \alpha < 1$, generates a strongly continuous semigroup and represents it as an average of T . Namely, let $\eta_{\alpha,t}$ be the probability measures on $[0, \infty)$ characterized by $\int_{[0,\infty)} e^{-\rho r} \eta_{\alpha,t}(dr) = e^{-t\rho^\alpha}$, $\rho \geq 0$. Bochner subordination gives

$$e^{-tB^\alpha} f = \int_{[0,\infty)} T(r) f \eta_{\alpha,t}(dr), \quad \|e^{-tB^\alpha}\|_{p \rightarrow p} \leq M e^{-t\omega^\alpha}. \quad (\text{A.6})$$

The representation also preserves positivity whenever T is positive.

If B^* is the Banach adjoint on $L^{p'}$, $p' = p/(p-1)$, its powers satisfy $(B^{-\beta})^* = (B^*)^{-\beta}$ and $(B^\beta)^* = (B^*)^\beta$. Semigroups T_p and T_q on $L^p(\Omega, \mu)$ and $L^q(\Omega, \mu)$ are called consistent if $T_p(t)f = T_q(t)f$ almost everywhere for every $t \geq 0$ and $f \in L^p \cap L^q$. For such semigroups satisfying the preceding bounds, the negative powers agree on $L^p \cap L^q$, and the positive powers agree on their common domains.

3. We now specify the previous construction to the kinetic operator. Recall the operators introduced in Section 2:

$$\mathcal{H} = \partial_t + v \cdot \nabla_x - \Delta_v, \quad \mathcal{A} = \lambda + \mathcal{H}, \quad \lambda > 0, \quad \mathcal{A}^* = \lambda - \partial_t - v \cdot \nabla_x - \Delta_v. \quad (\text{A.7})$$

The realization of $-\mathcal{H}$ on $L^p(\mathbb{R}^{1+2d})$ is the generator of the strongly continuous contraction semigroup

$$(\mathbf{P}_\tau f)(t, x, v) = \int_{\mathbb{R}^{2d}} p_\tau(x, v; y, w) f(t - \tau, y, w) dy dw, \quad (\text{A.8})$$

where p_τ is the kernel in (2.7); thus (A.8) is (2.6). Consequently $e^{-\tau \mathcal{A}} = e^{-\lambda \tau} \mathbf{P}_\tau$, and the preceding statements apply with $M = 1$ and $\omega = \lambda$. We suppress the L^p subscript on the realizations of $\mathcal{A}$ and $\mathcal{A}^*$; in adjoint identities they act on the dual exponents p and p' . For $\gamma > 0$,

$$\mathcal{A}^{-\gamma/2} f = \frac{1}{\Gamma(\gamma/2)} \int_0^\infty \tau^{\gamma/2-1} e^{-\lambda \tau} \mathbf{P}_\tau f d\tau, \quad \|\mathcal{A}^{-\gamma/2}\|_{p \rightarrow p} \leq \lambda^{-\gamma/2}. \quad (\text{A.9})$$

This recovers (2.9) and the corresponding bound in (2.12). The same formula and bound hold for $\mathcal{A}^*$ with $\mathbf{P}_\tau^*$. In particular, $\langle \mathcal{A}^{-\gamma/2} f, g \rangle = \langle f, (\mathcal{A}^*)^{-\gamma/2} g \rangle$ for $f \in L^p$ and $g \in L^{p'}$. For positive power, as previously discussed, $\mathcal{A}^{\gamma/2} = (\mathcal{A}^{-\gamma/2})^{-1}$, $\gamma > 0$.

The following lemma is needed to establish the compatibility with test functions used in the weak formulation (3.19) and in Lemma 3.5.

Lemma A.2 (Compatibility with test functions). *Let $\lambda > 0$. For every $s \in \mathbb{R}$, the operators $\mathcal{A}^s$ and $(\mathcal{A}^*)^s$ restrict to continuous automorphisms of $\mathcal{S}(\mathbb{R}^{1+2d})$, with inverses $\mathcal{A}^{-s}$ and $(\mathcal{A}^*)^{-s}$, respectively. Their extensions to $\mathcal{S}'$ by transposition satisfy*

$$\langle \mathcal{A}^s u, \varphi \rangle = \langle u, (\mathcal{A}^*)^s \varphi \rangle, \quad \mathcal{A}^s \mathcal{A}^r = \mathcal{A}^{s+r} \quad \text{on } \mathcal{S}', \quad r, s \in \mathbb{R}, \quad (\text{A.10})$$

and the analogous identities hold with $\mathcal{A}$ replaced by $\mathcal{A}^*$.

Proof. Both $\mathbf{P}_\tau$ and $\mathbf{P}_\tau^*$ act continuously on $\mathcal{S}(\mathbb{R}^{1+2d})$, and the Gaussian formula gives, for every Schwartz seminorm q , a seminorm $\tilde{q}$ and constants C, N such that

$$q(\mathbf{P}_\tau \varphi) + q(\mathbf{P}_\tau^* \varphi) \leq C(1 + \tau)^N \tilde{q}(\varphi), \quad \tau \geq 0. \quad (\text{A.11})$$

Indeed, differentiation produces polynomial factors in τ , while the polynomial weights are controlled by the corresponding Gaussian moments. The factor $e^{-\lambda\tau}$ therefore makes (A.9) convergent in $\mathcal{S}$. For $s > 0$, choosing an integer $m > s$ and using $\mathcal{A}^s = \mathcal{A}^m \mathcal{A}^{-(m-s)}$ shows that positive powers also preserve $\mathcal{S}$. Together with the inverse identities, this proves the automorphism assertions on $\mathcal{S}$. Transposition gives the extensions to $\mathcal{S}'$. Then the power identities follow from those on $\mathcal{S}$ for the adjoint powers. $\square$

For the Bessel-potential spaces of Section 2.3, this gives, for $\gamma > 0$, $\mathbb{W}_{\mathcal{H}}^{\gamma,p} = D(\mathcal{A}^{\gamma/2}) = \mathcal{A}^{-\gamma/2} L^p$ with norm $\|\mathcal{A}^{\gamma/2} u\|_p$. With $\mathbb{W}_{\mathcal{H}}^{0,p} = L^p$ and the negative-order spaces defined by completion in the norms $\|\mathcal{A}^{\gamma/2} \cdot\|_p$, one obtains, for every $\gamma \in \mathbb{R}$,

$$\mathcal{A} : \mathbb{W}_{\mathcal{H}}^{\gamma,p} \longrightarrow \mathbb{W}_{\mathcal{H}}^{\gamma-2,p} \quad \text{isometrically onto} \quad (\mathbb{W}_{\mathcal{H}}^{\gamma,p})^* = \mathbb{W}_{\mathcal{H}^*}^{-\gamma,p'}. \quad (\text{A.12})$$

The duality is the distributional pairing in (2.19). These identities justify the fractional factorizations in Lemma 3.5 and Section 3.

APPENDIX B. PROOF OF PROPOSITION 2.2

Step 1: The kernel of the fractional resolvent. Let $\zeta = (t, x, v)$ and $\xi = (s, y, w)$. Formula (2.6) shows that the space-time kernel of $\mathcal{A}^{-\gamma/2}$ vanishes for $t \leq s$. If $t > s$ and $\tau := t - s$, the integration variable in (2.9) is forced to equal $t - s$ by the time translation in (2.6). Consequently

$$\mathcal{A}^{-\gamma/2} f(\zeta) = \int_{\mathbb{R}^{1+2d}} K_{\gamma,\lambda}(\zeta, \xi) f(\xi) d\xi,$$

where the integral kernel is

$$K_{\gamma,\lambda}(\zeta, \xi) = \frac{\mathbf{1}_{\{t>s\}}}{\Gamma(\gamma/2)} \tau^{\gamma/2-1} e^{-\lambda\tau} p_\tau(x, v; y, w). \quad (\text{B.1})$$

For $t > s$, the definition (2.2) reads

$$R := d(\zeta, \xi) = \max \left\{ \sqrt{\tau}, |w - v|, |y - x + \tau v|^{1/3} \right\}. \quad (\text{B.2})$$

In particular, $R/\sqrt{\tau} \geq 1$. From (2.8) and $Q = 4d + 2$ we obtain

$$\begin{aligned} |K_{\gamma,\lambda}(\zeta, \xi)| &\leq C\tau^{\gamma/2-1-2d} \exp \left[-c \left(\frac{|w-v|^2}{\tau} + \frac{|y-x+\tau v|^2}{\tau^3} \right) \right] \\ &= C\tau^{(\gamma-Q)/2} \exp \left[-c \left(\frac{|w-v|^2}{\tau} + \frac{|y-x+\tau v|^2}{\tau^3} \right) \right]. \end{aligned} \quad (\text{B.3})$$

Here and below we have used $e^{-\lambda\tau} \leq 1$, which is why the pointwise constants do not depend on λ .

We now compare the Gaussian exponent with the intrinsic distance. If the maximum in (B.2) is attained by the velocity term, then $|w-v|^2/\tau = R^2/\tau$. If it is attained by the position term, then

$$\frac{|y-x+\tau v|^2}{\tau^3} = \left(\frac{R}{\sqrt{\tau}} \right)^6 \geq \frac{R^2}{\tau}.$$

If it is attained by $\sqrt{\tau}$, then $R^2/\tau = 1$. After changing the constants, all three cases are therefore summarized by

$$\exp \left[-c \left(\frac{|w-v|^2}{\tau} + \frac{|y-x+\tau v|^2}{\tau^3} \right) \right] \leq C \exp \left(-c' \frac{R^2}{\tau} \right). \quad (\text{B.4})$$

Set $u = R/\sqrt{\tau} \geq 1$. Combining (B.3) and (B.4) gives

$$|K_{\gamma,\lambda}(\zeta, \xi)| \leq CR^{\gamma-Q} u^{Q-\gamma} e^{-c'u^2} \leq C_{d,\gamma} R^{\gamma-Q}, \quad (\text{B.5})$$

because $Q - \gamma > 0$ and $\sup_{u \geq 1} u^{Q-\gamma} e^{-c'u^2} < \infty$. Finally, $d \leq d_* \leq (1 + 2^{1/3})d$ by (2.3)–(2.4). Since $\gamma - Q < 0$, replacing $R = d(\zeta, \xi)$ by the equivalent metric d_* changes only the constant. This proves the estimate for $\mathcal{A}^{-\gamma/2}$ in (2.10).

The kernel of $(\mathcal{A}^*)^{-\gamma/2}$ is $K_{\gamma,\lambda}^*(\zeta, \xi) = K_{\gamma,\lambda}(\xi, \zeta)$. Since d_* is symmetric, the same bound proves the adjoint part of (2.10).

Step 2: Differentiated kernels. Differentiating the explicit quadratic form in (2.7) shows that

$$|\nabla_v p_\tau(x, v; y, w)| + |\nabla_w p_\tau(x, v; y, w)| \leq C\tau^{-1/2} \bar{p}_\tau(x, v; y, w), \quad (\text{B.6})$$

where $\bar{p}_\tau$ satisfies the same estimate as (2.8), with a smaller positive constant in the exponential. To see the power $\tau^{-1/2}$ directly, differentiation of the exponent produces terms bounded by

$$C \left(\frac{|w-v|}{\tau} + \frac{|y-x+\tau v|}{\tau^2} \right) p_\tau.$$

Each normalized linear factor can be absorbed into a slightly wider Gaussian by the elementary inequality $ae^{-ca^2} \leq C_e e^{-ca^2/2}$. The derivative with respect to w is handled in exactly the same way. This proves (B.6).

For test functions, differentiation under the integral is justified by (B.6). The kernel of $\nabla_v \mathcal{A}^{-(1+\gamma)/2}$ is therefore bounded by

$$C\tau^{(1+\gamma)/2-1}\tau^{-1/2}\tau^{-2d} \exp \left[-c \left(\frac{|w-v|^2}{\tau} + \frac{|y-x+\tau v|^2}{\tau^3} \right) \right].$$

The power of τ is again $(1+\gamma)/2 - 1 - 1/2 - 2d = (\gamma - Q)/2$. Thus the argument in (B.4)–(B.5) applies without change and proves the first term in (2.11). The adjoint kernel uses the w -derivative in (B.6); symmetry of d_* then gives the second term.

Step 3: L^r bounds. Both P_τ and its adjoint are contractions on L^r . Hence

$$\|\mathcal{A}^{-\gamma/2} f\|_r \leq \frac{\|f\|_r}{\Gamma(\gamma/2)} \int_0^\infty \tau^{\gamma/2-1} e^{-\lambda\tau} d\tau = \lambda^{-\gamma/2} \|f\|_r.$$

The identical calculation for $\mathcal{A}^*$ proves (2.12).

The broadened Gaussian in (B.6) has uniformly bounded mass. Therefore,

$$\|\nabla_v \mathbf{P}_\tau f\|_r + \|\nabla_v \mathbf{P}_\tau^* f\|_r \leq C_{d,r} \tau^{-1/2} \|f\|_r. \quad (\text{B.7})$$

Initially this is proved for smooth compactly supported functions and extended to L^r using a standard density argument. Using (B.7) in the fractional power formula gives

$$\begin{aligned} \|\nabla_v \mathcal{A}^{-(1+\gamma)/2} f\|_r &\leq \frac{C_{d,r}}{\Gamma((1+\gamma)/2)} \int_0^\infty \tau^{(1+\gamma)/2-1} \tau^{-1/2} e^{-\lambda\tau} d\tau \|f\|_r \\ &= C_{d,r,\gamma} \lambda^{-\gamma/2} \|f\|_r. \end{aligned}$$

The same calculation for $\mathcal{A}^*$ completes the proof of (2.13). $\square$

APPENDIX C. PROOF OF PROPOSITION 2.3

We give the proof for $\mathcal{A}$ and then indicate why it also applies to $\mathcal{A}^*$. We will only need the doubling property, the volume estimate (2.5) and the kernel estimates of Proposition 2.2.

Step 1: Hardy–Littlewood–Sobolev estimate. Let $\mathbf{M}$ be the centered Hardy–Littlewood maximal operator associated with the metric d_* and Lebesgue measure. Since the space is doubling, $\mathbf{M}$ is bounded on L^r for every $1 < r < \infty$, see, for example, [31, Chapter 2]. We recall the standard argument that converts this maximal estimate into the fractional integration estimate needed here.

Fix $\zeta \in \mathbb{R}^{1+2d}$ and $R > 0$. Decompose the integral defining $I_\gamma f(\zeta)$ into the regions $d_*(\zeta, \xi) < R$ and $d_*(\zeta, \xi) \geq R$. A decomposition of the first region into the annuli

$$2^{-k-1}R \leq d_*(\zeta, \xi) < 2^{-k}R, \quad k \geq 0,$$

together with (2.5), gives

$$\begin{aligned} \int_{\{d_*(\zeta, \xi) < R\}} \frac{|f(\xi)|}{d_*(\zeta, \xi)^{Q-\gamma}} d\xi &\leq C \sum_{k=0}^\infty (2^{-k}R)^{\gamma-Q} \int_{B_{2^{-k}R}(\zeta)} |f| d\xi \\ &\leq CR^\gamma \mathbf{M}f(\zeta). \end{aligned} \quad (\text{C.1})$$

For the outer region, Hölder's inequality gives

$$\int_{\{d_*(\zeta, \xi) \geq R\}} \frac{|f(\xi)|}{d_*(\zeta, \xi)^{Q-\gamma}} d\xi \leq \|f\|_r \left(\int_{\{d_*(\zeta, \xi) \geq R\}} d_*(\zeta, \xi)^{(\gamma-Q)r'} d\xi \right)^{1/r'}.$$

Decomposing into the annuli $2^k R \leq d_*(\zeta, \xi) < 2^{k+1}R$ and again using (2.5), we find

$$\int_{\{d_*(\zeta, \xi) \geq R\}} \frac{|f(\xi)|}{d_*(\zeta, \xi)^{Q-\gamma}} d\xi \leq C \|f\|_r R^{\gamma-Q/r}. \quad (\text{C.2})$$

The geometric series used here converges precisely when $\gamma < Q/r$.

At points where $\mathbf{M}f(\zeta) > 0$, choose

$$R = \left(\frac{\|f\|_r}{\mathbf{M}f(\zeta)} \right)^{r/Q}.$$

The cases $\mathbf{M}f(\zeta) = 0$ and $f = 0$ are immediate. Equations (C.1)–(C.2) yield the pointwise Hedberg inequality

$$I_\gamma f(\zeta) \leq C \|f\|_r^{\gamma r/Q} (\mathbf{M}f(\zeta))^{1-\gamma r/Q}. \quad (\text{C.3})$$

The relation $1/s = 1/r - \gamma/Q$ is equivalent to

$$s \left(1 - \frac{\gamma r}{Q} \right) = r.$$

Taking the L^s norm in (C.3) and using the L^r boundedness of $\mathbf{M}$ therefore gives

$$\|I_\gamma f\|_s \leq C_{d,\gamma,r} \|f\|_r.$$

The pointwise estimates (2.15) now imply (2.17) for both $\mathcal{A}$ and $\mathcal{A}^*$.

Step 2: Uniform $L^{r'}$ bound for the kernel. Assume now that $\gamma > Q/r$. We first verify explicitly the kernel integrability used for point evaluation. From (B.1), for fixed $\zeta = (t, x, v)$,

$$\int_{\mathbb{R}^{1+2d}} |K_{\gamma,\lambda}(\zeta, \xi)|^{r'} d\xi = C_{\gamma,r} \int_0^\infty \tau^{(\gamma/2-1)r'} e^{-\lambda r' \tau} \left(\int_{\mathbb{R}^{2d}} p_\tau(x, v; y, w)^{r'} dy dw \right) d\tau.$$

The change of variables

$$a = \frac{w - v}{\sqrt{\tau}}, \quad c = \frac{y - x + \tau v}{\tau^{3/2}}$$

has Jacobian τ^{2d} . The explicit Gaussian formula (2.7) consequently gives

$$\int_{\mathbb{R}^{2d}} p_\tau(x, v; y, w)^{r'} dy dw = C_{d,r} \tau^{-2d(r'-1)}.$$

It follows that

$$\int |K_{\gamma,\lambda}(\zeta, \xi)|^{r'} d\xi = C \int_0^\infty \tau^{((\gamma-Q)r'+Q)/2-1} e^{-\lambda r' \tau} d\tau.$$

The integral is finite near zero exactly when

$$(\gamma - Q)r' + Q > 0,$$

which is equivalent to $\gamma > Q/r$; it is finite at infinity because $\lambda > 0$. The Gamma-integral also shows that

$$\sup_{\zeta \in \mathbb{R}^{1+2d}} \|K_{\gamma,\lambda}(\zeta, \cdot)\|_{r'} \leq C_{d,\gamma,r} \lambda^{-(\gamma-Q/r)/2}.$$

The adjoint kernel is obtained by interchanging ζ and ξ , and Lebesgue measure is both left and right invariant. Thus the same estimate holds for its kernel sections. Hölder's inequality now gives the L^r -to- L^∞ bound

$$\|\mathcal{A}^{-\gamma/2} f\|_\infty + \|(\mathcal{A}^*)^{-\gamma/2} f\|_\infty \leq C_{d,\gamma,r} \lambda^{-(\gamma-Q/r)/2} \|f\|_r. \quad (\text{C.4})$$

Step 3: Continuity. The operator $\mathcal{A}^{-\gamma/2}$ is left invariant, so its kernel sections are left translates of one fixed function in $L^{r'}$. Left translations are strongly continuous on $L^{r'}$ for $1 \leq r' < \infty$: this follows first for continuous compactly supported functions and then for all $L^{r'}$ functions by density and invariance of Haar measure. Hence, if $\zeta_n \rightarrow \zeta$,

$$\|K_{\gamma,\lambda}(\zeta_n, \cdot) - K_{\gamma,\lambda}(\zeta, \cdot)\|_{r'} \rightarrow 0.$$

By Hölder's inequality,

$$|\mathcal{A}^{-\gamma/2} f(\zeta_n) - \mathcal{A}^{-\gamma/2} f(\zeta)| \leq \|K_{\gamma,\lambda}(\zeta_n, \cdot) - K_{\gamma,\lambda}(\zeta, \cdot)\|_{r'} \|f\|_r,$$

which proves continuity. The same reasoning applies to $\mathcal{A}^*$. More explicitly, if $K(\zeta, \xi) = k(\zeta^{-1} \circ \xi)$, then $K^*(\zeta, \xi) = k^\vee(\zeta^{-1} \circ \xi)$ with $k^\vee(\eta) = k(\eta^{-1})$. Inversion preserves Haar measure on this unimodular group, so $k^\vee \in L^{r'}$ and its left translates are strongly continuous as well.

Step 4: Vanishing at infinity. Let first $f \in C_c^\infty$ and let $E = \text{supp } f$. Since d_* is a proper metric compatible with the Euclidean topology, the triangle inequality implies

$$\inf_{\xi \in E} d_*(\zeta, \xi) \rightarrow \infty \quad \text{as } \zeta \rightarrow \infty.$$

Because $\gamma < Q$, (2.10) gives

$$|\mathcal{A}^{-\gamma/2} f(\zeta)| \leq C \|f\|_1 \left(\inf_{\xi \in E} d_*(\zeta, \xi) \right)^{\gamma-Q} \rightarrow 0.$$

The same argument applies to $\mathcal{A}^*$. For a general $f \in L^r$, choose $f_n \in C_c^\infty$ with $f_n \rightarrow f$ in L^r . Estimate (C.4) shows that the corresponding potentials converge uniformly. As a uniform limit of functions in C_∞ , each limiting potential belongs to C_∞ . This proves part (ii). $\square$

APPENDIX D. MANY-PARTICLE AND LEBESGUE CLASS DRIFTS

Proposition D.1 (Morrey norm for pairwise interaction kernels). *Let $N \geq 2$, $d_0 \geq 2$, and $d = Nd_0$. Write $x = (x^1, \dots, x^N)$ and $v = (v^1, \dots, v^N)$, with $x^i, v^i \in \mathbb{R}^{d_0}$ and suppose that $b(t, x, v) = b(x, v)$ with*

$$b_i(x, v) = \sum_{j \neq i} K_{ij}(x^i - x^j, v^i - v^j), \quad b = (b_1, \dots, b_N),$$

where the measurable kernels satisfy, almost everywhere,

$$|K_{ij}(\xi, \nu)| \leq \kappa_{ij}|\xi|^{-1/3} + \gamma_{ij}|\nu|^{-1}, \quad \kappa_{ij}, \gamma_{ij} \geq 0.$$

Then, for every $1 < q < d_0$,

$$\|b\|_{\mathcal{M}_{\mathcal{K}}^q} \leq C_{N, d_0, q} \sum_{i \neq j} (\kappa_{ij} + \gamma_{ij}).$$

If all γ_{ij} vanish, the same conclusion holds for $1 < q < 3d_0$.

Proof. We first recall the elementary estimate

$$\sup_{a \in \mathbb{R}^{d_0}} \int_{|y-a| < R} |y|^{-\beta} dy \leq C_{d_0, \beta} R^{d_0 - \beta}, \quad 0 < \beta < d_0.$$

Indeed, on the set $\{|y| < R\}$ one integrates in spherical coordinates. On its complement the integrand is at most $R^{-\beta}$, while the integration domain has volume at most CR^{d_0} .

Consider a kinetic cylinder $\mathcal{C}_r := \mathcal{C}_r(t_0, x_0, v_0)$ as in (1.4). Its time interval has length r^2 , and each time slice is contained in the product of the sets

$$|x^k - x_0^k - (t - t_0)v_0^k| < r^3, \quad |v^k - v_0^k| < r, \quad k = 1, \dots, N.$$

Fix $i \neq j$. Integrating first in x^i , with all other variables fixed, and applying the preceding estimate with $\beta = q/3$ gives

$$\begin{aligned} \int_{\mathcal{C}_r} |x^i - x^j|^{-q/3} dt dx dv &\leq Cr^2 r^d r^{3(d-d_0)} r^{3d_0-q} \\ &= Cr^{4d+2-q}, \quad q < 3d_0. \end{aligned}$$

Similarly, integrating first in v^i gives

$$\begin{aligned} \int_{\mathcal{C}_r} |v^i - v^j|^{-q} dt dx dv &\leq Cr^2 r^{3d} r^{d-d_0} r^{d_0-q} \\ &= Cr^{4d+2-q}, \quad q < d_0. \end{aligned}$$

Both estimates are uniform in the cylinder center (t_0, x_0, v_0) : the elementary integral estimate is uniform under translations, including the time-dependent spatial shifts $x_0 + (t - t_0)v_0$.

Since $|b| \leq \sum_{i \neq j} |K_{ij}|$, we have

$$\|b\|_{L^q(\mathcal{C}_r)} \leq C \sum_{i \neq j} (\kappa_{ij} + \gamma_{ij}) r^{(4d+2)/q-1}.$$

Dividing by $|\mathcal{C}_r|^{1/q}$, multiplying by r , and taking the supremum proves the assertion since $|\mathcal{C}_r| = cr^{4d+2}$ and the definitions of the Morrey norm using kinetic cylinders and metric balls are equivalent. $\square$

Proposition D.2. *Let $p_t, p_v, p_x \in (1, \infty]$ satisfy*

$$\frac{2}{p_t} + \frac{d}{p_v} + \frac{3d}{p_x} = 1.$$

If

$$|b| \in L_t^{p_t} L_v^{p_v} L_x^{p_x} := L^{p_t}(\mathbb{R}_t, L^{p_v}(\mathbb{R}_v^d, L^{p_x}(\mathbb{R}_x^d))),$$

then, for every $1 < q \leq \min\{p_t, p_v, p_x\}$,

$$\|b\|_{\mathcal{M}_K^q} \leq C \|b\|_{L_t^{p_t} L_v^{p_v} L_x^{p_x}}.$$

The same assertion holds with the integration order of the x and v norms reversed.

Proof. Let C_r be as in (1.4). Applying Hölder's inequality successively in x , v , and t , we obtain, for $q \in (1, \min\{p_t, p_v, p_x\}]$,

$$\|b\|_{L^q(C_r)} \leq C r^{3d(1/q-1/p_x)} r^{d(1/q-1/p_v)} r^{2(1/q-1/p_t)} \|b\|_{L_t^{p_t} L_v^{p_v} L_x^{p_x}}.$$

By the assumed relation between the exponents, the total power of r is given by

$$\frac{4d+2}{q} - \frac{3d}{p_x} - \frac{d}{p_v} - \frac{2}{p_t} = \frac{4d+2}{q} - 1.$$

Since $|C_r| = cr^{4d+2}$, it follows that

$$r \left(\frac{1}{|C_r|} \int_{C_r} |b|^q \right)^{1/q} \leq C \|b\|_{L_t^{p_t} L_v^{p_v} L_x^{p_x}}.$$

Taking the supremum over all centers and radii proves the claim.

One can finally note that, Hölder's inequality (applied to v , x and t) also yields

$$r \left(\frac{1}{|C_r|} \int_{C_r} |b|^q \right)^{1/q} \leq C \|b\|_{L_t^{p_t} L_x^{p_x} L_v^{p_v}},$$

and conclude the estimate $\|b\|_{\mathcal{M}_K^q} \leq C \|b\|_{L_t^{p_t} L_x^{p_x} L_v^{p_v}}$. $\square$

AI USAGE

The authors used OpenAI ChatGPT 5.6 Sol and 6 Astra for light editorial assistance (namely, grammar and misprints checks, fixing typos, provide style suggestions), and also for light assistance in carrying out additional literature surveys on pages 5 and 6. The motivations, key ideas and all results and proof arguments are the sole outcomes of the authors' works.

REFERENCES

- [1] D. R. Adams, Weighted nonlinear potential theory, Trans. Amer. Math. Soc. **297** (1986), 73–94.
- [2] D. R. Adams and J. Xiao, Morrey spaces in harmonic analysis, Ark. Mat. **50** (2012), no. 2, 201–230.
- [3] D. R. Adams and J. Xiao, Elliptic-to-parabolic Morrey spaces-potentials-capacities with applications to certain evolution PDE, J. Lond. Math. Soc. (2) **111** (2025), no. 4, Paper No. e70131, 38 pp.
- [4] D. Albritton, S. Armstrong, J.-C. Mourrat and M. Novack, Variational methods for the kinetic Fokker–Planck equation, Anal. PDE **17** (2024), no. 6, 1953–2010.
- [5] H. Arai and T. Mizuhara, Morrey spaces on spaces of homogeneous type and estimates for $\square_b$ and the Cauchy–Szegő projection, Math. Nachr. **185** (1997), 5–20.
- [6] F. Anceschi and S. Polidoro, A survey on the classical theory for Kolmogorov equation, Le Matematiche **75** (2020), no. 1, 221–258.
- [7] P. Auscher, C. Imbert and L. Niebel, Fundamental solutions to Kolmogorov–Fokker–Planck equations with rough coefficients: existence, uniqueness, upper estimates, SIAM J. Math. Anal. **57** (2025), no. 2, 2114–2137.
- [8] P. Auscher and L. Niebel, Kinetic Sobolev spaces, arXiv:2603.17491 (2026).

- [9] L. Beck, F. Flandoli, M. Gubinelli and M. Maurelli, Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness, Electron. J. Probab. **24** (2019), Paper No. 136, 72 pp.
- [10] A. G. Belyi and Yu. A. Semenov, On the L^p -theory of Schrödinger semigroups. II. Sibirsk. Mat. Zh., **31** (1990), 16–26; English transl. in Siberian Math. J., **31** (1990), 540–549.
- [11] P. Billingsley, Convergence of probability measures (Second Edition), John Wiley and Sons, 1999.
- [12] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, Uniqueness problems for degenerate Fokker–Planck–Kolmogorov equations, J. Math. Sci. (N.Y.) **207** (2015), no. 2, 147–165.
- [13] M. Bossy, Some stochastic particle methods for nonlinear parabolic PDEs, ESAIM Proc., **15** (2005), 18–57.
- [14] P. Cattiaux, F. Delebecque, and L. Pédèches, Stochastic Cucker–Smale models: old and new, Ann. Appl. Probab. **28** (2018), no. 5, 3239–3286.
- [15] P.-E. Chaudru de Raynal, Strong existence and uniqueness for degenerate SDE with Hölder drift, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 259–286.
- [16] P.-E. Chaudru de Raynal, Weak regularization by stochastic drift: result and counter example, Discrete Contin. Dyn. Syst. **38** (2018), no. 3, 1269–1291.
- [17] P.-E. Chaudru de Raynal, S. Menozzi, A. Pesce and X. Zhang, Heat kernel and gradient estimates for kinetic SDEs with low regularity coefficients, Bull. Sci. Math. **183** (2023), Paper No. 103229, 56 pp.
- [18] X. Cheng, N. S. Chatterji, P. L. Bartlett and M. I. Jordan, Underdamped Langevin MCMC: A non-asymptotic analysis, Proceedings of Machine Learning Research, **75** (2018), 300–323.
- [19] Z. Chen, Z. Hao and X. Zhang, Kinetic SDEs with subcritical distributional drifts, Adv. Nonlinear Stud. (2026), published online.
- [20] Z. Chen and C. Ren, Kinetic Fokker–Planck equations with drifts in a supercritical range, arXiv:2609.02385 (2026).
- [21] R. R. Coifman and R. Rochberg, Another characterization of BMO, Proc. Amer. Math. Soc. **79** (1980), 249–254.
- [22] R. Coifman and G. Weiss, Analyse harmonique non-commutative sur certains espaces homogènes. Lecture Notes in Math. Vol. 242. Springer-Verlag, Berlin-New York, 1971.
- [23] D. Deng and Y. Han, Harmonic analysis on spaces of homogeneous type. With a preface by Yves Meyer. Lecture Notes in Math., 1966, Springer-Verlag, Berlin, 2009. xii+154 pp.
- [24] H. Dong and T. Yastrzhembskiy, Global L^p estimates for kinetic Kolmogorov–Fokker–Planck equations in nondivergence form, Arch. Ration. Mech. Anal. **245** (2022), no. 1, 501–564.
- [25] S. N. Ethier and T. G. Kurtz, Markov processes – Characterization and convergence, John Wiley and Sons Publications, 2005.
- [26] E. Fedrizzi, F. Flandoli, E. Priola and J. Vovelle, Regularity of stochastic kinetic equations, Electron. J. Probab. **22** (2017), Paper No. 48, 42 pp.
- [27] N. Garofalo and E. Lanconelli, Level sets of the fundamental solution and Harnack inequality for degenerate equations of Kolmogorov type, Trans. Amer. Math. Soc. **321** (1990), no. 2, 775–792.
- [28] F. Golse, C. Imbert, C. Mouhot and A. F. Vasseur, Harnack inequality for kinetic Fokker–Planck equations with rough coefficients and application to the Landau equation, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **19** (2019), no. 1, 253–295.
- [29] Z. Hao, J.-F. Jabir, S. Menozzi, M. Röckner and X. Zhang, Propagation of chaos for moderately interacting particle systems related to singular kinetic McKean–Vlasov SDEs, arXiv:2405.09195.
- [30] Z. Hao, X. Zhang, R. Zhu and X. Zhu, Singular kinetic equations and applications, Ann. Probab. **52** (2024), no. 2, 576–657.
- [31] J. Heinonen, Lectures on Analysis on Metric Spaces, Universitext, Springer-Verlag, New York, 2001.
- [32] L. Hörmander, Hypoelliptic second order differential equations, Acta Math. **119** (1967), 147–171.
- [33] D. Kinzebulatov, A new approach to the L^p -theory of $-\Delta + b \cdot \nabla$, and its applications to Feller processes with general drifts, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **17** (2017), no. 2, 685–711.
- [34] D. Kinzebulatov, Parabolic equations and SDEs with time-inhomogeneous Morrey drift, NoDEA Nonlinear Differential Equations Appl. **32** (2025), no. 6, Paper No. 117, 30 pp.
- [35] D. Kinzebulatov, On particle systems and critical strengths of general singular interactions, Ann. Inst. H. Poincaré (B) Probab. Statist. **62** (2026), no. 2, 1191–1251.
- [36] D. Kinzebulatov and Yu. A. Semënov, Brownian motion with general drift, Stochastic Process. Appl. **130** (2020), no. 5, 2737–2750.
- [37] D. Kinzebulatov and C. Yue, Non-local SDEs, critical drifts and local blow-ups, arXiv:2606.02786.

- [38] A. N. Kolmogoroff, Zufällige Bewegungen (zur Theorie der Brownschen Bewegung), Ann. of Math. (2) **35** (1934), no. 1, 116–117.
- [39] H. Komatsu, Fractional powers of operators, Pacific J. Math. **19** (1966), no. 2, 285–346.
- [40] N. V. Krylov, Essentials of real analysis and Morrey–Sobolev spaces for second order elliptic and parabolic PDEs with singular first order coefficients, J. Math. Sci. **294** (2025), 802–899.
- [41] N. V. Krylov, On parabolic Adams’s, the Chiarenza-Frasca theorems, and some other results related to parabolic Morrey spaces, Math. Eng., **5** (2023) no. 2, 1–20.
- [42] N. V. Krylov, Once again on weak solutions of time inhomogeneous Itô’s equations with VMO diffusion and Morrey drift, Electron. J. Probab., **29**, paper 95 (2024), 1–19.
- [43] E. Lanconelli and S. Polidoro, On a class of hypoelliptic evolution operators, Rend. Sem. Mat. Univ. Politec. Torino **52** (1994), no. 1, 29–63.
- [44] P. Langevin, Sur la théorie du mouvement brownien, C. R. Acad. Sci. (Paris) **146** (1908), 530–533.
- [45] M.-Y. Lee, J. Li, and L. A. Ward, On weak-star convergence in product Hardy spaces on spaces of homogeneous type, Studia Math. **235** (2016), 251–267.
- [46] P. G. Lemarié-Rieusset, Sobolev multipliers, maximal functions and parabolic equations with a quadratic nonlinearity, J. Funct. Anal. **274** (2018), 659–694.
- [47] L. Niebel and R. Zacher, Kinetic maximal L^p -regularity with temporal weights and application to quasilinear kinetic diffusion equations, J. Differential Equations **307** (2022), 29–82.
- [48] L. Pareschi and G. Toscani, Interacting Multiagent Systems: Kinetic equations and Monte Carlo methods, 2013, Oxford University Press.
- [49] A. Pascucci and A. Pesce, Sobolev embeddings for kinetic Fokker–Planck equations, J. Funct. Anal. **286** (2024), Paper No. 110344.
- [50] A. Pascucci and S. Polidoro, A Gaussian upper bound for the fundamental solutions of a class of ultraparabolic equations, J. Math. Anal. Appl. **282** (2003), no. 1, 396–409.
- [51] S. Polidoro, A global lower bound for the fundamental solution of Kolmogorov–Fokker–Planck equations, Arch. Rational Mech. Anal. **137** (1997), no. 4, 321–340.
- [52] S. Polidoro and M. A. Ragusa, Sobolev–Morrey spaces related to an ultraparabolic equation, Manuscripta Math. **96** (1998), 371–392.
- [53] P. Protter, Stochastic Integration and Differential Equations, Second edition, Springer-Verlag, Berlin-New York, 2005.
- [54] C. Ren and X. Zhang, Heat kernel estimates for kinetic SDEs with drifts being unbounded and in Kato’s class, Bernoulli **31** (2025), no. 2, 1402–1427.
- [55] M. Röckner and G. Zhao, SDEs with critical time dependent drifts: weak solutions, Bernoulli **29** (2023), no. 1, 757–784.
- [56] S. Serfaty, Lectures on Coulomb and Riesz Gases, 2026, AMS, Colloquium Publications, Volume 70.
- [57] D. W. Stroock and S. R. S. Varadhan, Multidimensional diffusion processes (Reprint of 1997 edition), Springer-Verlag Berlin-Heidelberg, 2006.
- [58] P. Xia, L. Xie, X. Zhang and G. Zhao, $L^q(L^p)$ -theory of stochastic differential equations, Stochastic Process. Appl. **130** (2020), no. 8, 5188–5211.
- [59] K. Yosida, Functional Analysis, sixth ed., Grundlehren der mathematischen Wissenschaften, vol. 123, Springer-Verlag, Berlin–Heidelberg–New York, 1980; reprinted in Classics in Mathematics, 1995.
- [60] X. Zhang, Stochastic Hamiltonian flows with singular coefficients, Sci. China Math. **61** (2018), no. 8, 1353–1384.
- [61] X. Zhang, Second order McKean–Vlasov SDEs and kinetic Fokker–Planck–Kolmogorov equations, Discrete Contin. Dyn. Syst. Ser. S **18** (2025), no. 10, 2973–3026.
- [62] X. Zhang and G. Zhao, Heat kernel and ergodicity of SDEs with distributional drifts, preprint, arXiv:1710.10537.

HSE UNIVERSITY, DEPARTMENT OF STATISTICS AND DATA ANALYSIS AND LSA, MOSCOW, RUSSIA.
Email address: `jjabir@hse.ru`

UNIVERSITÉ LAVAL, DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, QUÉBEC, QC, CANADA
Email address: `damir.kinzebulatov@mat.ulaval.ca`

UNIVERSITÉ LAVAL, DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, QUÉBEC, QC, CANADA
Email address: `chengjun.yue.1@ulaval.ca`