# FELLER GENERATORS WITH SINGULAR DRIFTS IN THE CRITICAL RANGE

D. KINZEBULATOV AND YU. A. SEMËNOV

Abstract. We consider diffusion operator $-\Delta + b\cdot\nabla$ in $\mathbb{R}^d$, $d\geq 3$, with drift $b$ in a large class of locally unbounded vector fields that can have critical-order singularities. Covering the entire range of admissible magnitudes of singularities of $b$, we construct a strongly continuous Feller semigroup on the space of continuous functions vanishing at infinity, thus completing a number of results on well-posedness of SDEs with singular drifts. Our approach uses De Giorgi's method ran in $L^p$ for $p$ sufficiently large, hence the gain in the assumptions on singular drift.

For the critical borderline value of the magnitude of singularities of $b$, we construct a strongly continuous semigroup in a "critical" Orlicz space on $\mathbb{R}^d$ whose topology is stronger than the topology of $L^p$ for any $2\leq p<\infty$ but is slightly weaker than that of $L^\infty$.

## 1. Introduction

**1.** The paper concerns with the following question: what are the minimal assumptions on a locally unbounded vector field $b:\mathbb{R}^d\to\mathbb{R}^d$, $d\geq 3$, such that operator $-\Delta+b\cdot\nabla$ generates a strongly continuous Feller semigroup? We deal with the drift singularities that substantially affect the behaviour of the heat kernel of $-\Delta+b\cdot\nabla$. For instance, the heat kernel can vanish or blow up at some points in space. However, the Feller semigroup structure ensures that the corresponding strong Markov process exists and has a number of important properties that make it of practical interest (e.g. properties related to continuity, existence of invariant measure, solvability of a martingale problem [7, 10]). It is almost impossible to survey the literature on Feller generators. We only mention some results related to the diffusion operators with irregular drifts, including drifts having strong growth at infinity [28, 29], generators of distorted Brownian motion [2, 4, 5], general locally unbounded drifts $b$ [3, 14, 26, 27]. See also [6, 30, 31].

The question of what local singularities of drift $b$ are admissible has two dimensions: the order of singularities (for example, for the model singular drift $b(x)=\sqrt{\delta}\frac{d-2}{2}|x|^{-\alpha}x$ the order of singularities is determined by $\alpha-1>0$) and their magnitude (i.e. factor $\delta$ in the previous formula if $\alpha$ is chosen to be critical, which, as can be seen by rescaling the equation, is $\alpha=2$). The following is a large class of vector fields that can have critical-order singularities:

**Definition 1.** A Borel measurable vector field $b:\mathbb{R}^d\to\mathbb{R}^d$ is said to be form-bounded if

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

for some constants $\delta$ and $c(\delta)$ (here and in what follows, $\|\cdot\|_p:=\|\cdot\|_{L^p}$, $W^{1,2}$ is the Sobolev space of functions with square integrable derivatives). Condition (1) is written as $b\in\mathbf{F}_\delta$.

2020 *Mathematics Subject Classification.* 60G53, 60H10 (primary), 47D07 (secondary).

*Key words and phrases.* Singular drifts, Feller semigroups, form-bounded vector fields, Orlicz spaces, stochastic differential equations.

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

The form-boundedness with form-bound $\delta<1$ is a classical condition on $|b|$: it provides coercivity of the corresponding to $-\Delta+b\cdot\nabla$ quadratic form in $L^2$.

Constant $\delta$ measures the magnitude of singularities of the vector field $b$. If $\delta>4$, then there are various counterexamples to the regularity theory of $-\Delta+b\cdot\nabla$ and to the theory of the corresponding diffusion process. We explain below that the critical threshold value of $\delta$ is 4. The present paper concerns with the value of $\delta$ going up to (and including) $\delta=4$.

**2.** There is a plethora of results devoted to verifying inclusion $b\in\mathbf{F}_\delta$ [1, 8, 9, 11]. Here are some examples of sub-classes of $\mathbf{F}_\delta$ that appear in the literature on PDEs and stochastic differential equations (SDEs). For example, class $\mathbf{F}_\delta$ contains vector fields $b$ from $[L^d+L^\infty]^d$ (with $\delta$ that can be chosen arbitrarily small), weak $L^d$ class, which includes

$$b(x)=\pm\sqrt{\delta}\frac{d-2}{2}|x|^{-2}x\in\mathbf{F}_\delta\quad\text{(but not in any }\mathbf{F}_{\delta'}\text{ with }\delta'<\delta), \tag{2}$$

and, more generally, the scaling-invariant Morrey class

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

where $B_r(x)$ is the ball of radius $r$ centered at $x$, and $\varepsilon>0$ is fixed arbitrarily small, so $\delta=C\|b\|_{M_{2+\varepsilon}}$ for appropriate constant $C=C(\varepsilon)$ [11]. Some other examples can be found, in particular, in [17, 25].

**3.** It was proved in [24], using De Giorgi's iterations in $L^p$, $p>\frac{2}{2-\sqrt{\delta}}$, and a compactness argument, that if $b\in\mathbf{F}_\delta$ with $\delta<4$, then the corresponding to $-\Delta+b\cdot\nabla$ SDE

$$X_t=x-\sqrt{\delta}\frac{d-2}{2}\int_0^t b(X_s)ds+\sqrt{2}B_t, \tag{3}$$

where $B_t$ is the $d$-dimensional Brownian motion, has a martingale solution for every initial point $x\in\mathbb{R}^d$ (see Theorem 2.2 below). This is important in light of the following counterexample: if we take a particular form-bounded singular vector field $b(x)=\sqrt{\delta}\frac{d-2}{2}\mathbf{1}_{x\neq 0}|x|^{-2}x$ introducing strong attraction of $X_t$ to the origin, then, whenever

$$\delta>4\biggl(\frac{d}{d-2}\biggr)^2,$$

the corresponding SDE does not have a weak solution departing from $x=0$. Thus, the constraint $\delta<4$ in [24] is sharp at least asymptotically (i.e. in high dimensions). It should also be added that if $\delta>4$, then for every initial point $x\neq 0$ the corresponding solution of (3) (which, one can prove, still exists locally in time) arrives to the origin with positive probability.

We explain where does the condition $p>\frac{2}{2-\sqrt{\delta}}$ come from in the end of this introduction. Let us add that it was known for some time that $-\Delta+b\cdot\nabla$, $b\in\mathbf{F}_\delta$, $\delta<4$, generates a strongly continuous semigroup in $L^p$, $p>\frac{2}{2-\sqrt{\delta}}$ [26]. Although this semigroup is an $L^\infty$ contraction and $p$ can be taken arbitrarily large, this result on its own does not provide a path to constructing strongly continuous Feller semigroup.

There already exist various methods for constructing Feller semigroup for $-\Delta+b\cdot\nabla$ with $b\in\mathbf{F}_\delta$ with some small $\delta$. The first paper where such construction was carried out for $\delta<1\wedge(\frac{2}{d-2})^2$, using Moser-type iterations, was [26]. [14] gave a different approach via fractional resolvent representations in $L^q$, $q>d-2$, to constructing the Feller generator, reaching the same condition on $\delta$ as in [26], and also providing additional information about regularity of the Feller semigroup, cf. Theorem 2.1($v$). All these results require $\delta \ll 1$. The reasons for this is that the argument in [26] uses rather strong gradient bound on solutions of the corresponding elliptic equations, while the construction in [14] automatically provides such gradient bounds, so the Feller semigroup arises as a by-product of this construction (a more detailed discussion can be found in survey [17]).

The question of what happens with operator $-\Delta + b \cdot \nabla$ and the corresponding parabolic equation in the critical case $\delta = 4$ was addressed in [16]. It turned out one still has a strongly continuous Markov semigroup but in Orlicz space with gauge function $\cosh -1$, moreover, the corresponding elliptic equation has a unique weak solution, and a variant of energy inequality holds. The local topology of this Orlicz space is stronger than the local topology of $L^p$ with any finite $p$, but is weaker than the topology of $L^\infty$ ([16] dealt with the dynamics on the torus $\mathbb{R}^d/\mathbb{Z}^d$ or, rather on a compact Riemannian manifold). The result of [16] was summarized there as follows: strengthening the topology of the space,where the semigroup of $-\Delta + b \cdot \nabla$ is considered, allows to relax the assumptions on $\delta$[1]. In the same vein, the Feller semigroup for $-\Delta + b \cdot \nabla$, which is acting in a space with an even stronger local topology (i.e. space $C_\infty$ of continuous functions vanishing at infinity with the sup-norm), should be defined for all values of $\delta$ going up to 4. Below we show that this is indeed the case for all $\delta < 4$.

Our main results in this paper, stated briefly, are as follows.

**Theorem.** *Let $b \in \mathbf{F}_\delta$. The following are true:*

*Theorem 2.1: If $\delta < 4$, then the constructed in* [24] *probability measures $\{\mathbb{P}_x\}_{x \in \mathbb{R}^d}$ solving the martingale problem for* (3) *in fact determine a Feller semigroup. Its generator is an appropriate realization of formal operator $-\Delta + b \cdot \nabla$ in $C_\infty$. This Feller semigroup is unique among Feller semigroups that can be constructed via an approximation of $b$ by bounded smooth vector fields that do not increase form-bound $\delta$ and constant $c(\delta)$.*

*Theorem 3.1: If $\delta \leq 4$ and $b$ satisfies some additional constraints on its behaviour outside of a large ball (e.g. bounded), then there is an analogous semigroup theory of $-\Delta + b \cdot \nabla$ but in the Orlicz space with gauge function* $\cosh -1$ *on $\mathbb{R}^d$.*

The proof of Theorem 2.1 uses, in particular, some regularity results for non-homogeneous elliptic equations obtained in [19] by means of De Giorgi's method ran in $L^p$, and some convergence theorems obtained in [26]. This allows to verify conditions of the Trotter approximation theorem in $C_\infty$.

Theorem 3.1 is proved directly, by verifying Cauchy's criterion for solutions of the approximating parabolic equations. Let us add that in [16] the volume of the torus enters the estimates, so simply blowing it up, in order to work on $\mathbb{R}^d$, is not an option. We address this in the present paper (Theorem 3.1) by working carefully with appropriate weights.

Theorem 3.1 admits more or less direct extension to time-inhomogeneous form-bounded vector fields. On the other hand, the proof of Theorem 2.1 so far uses in an essential manner (via Trotter's approximation theorem) the fact that we are working with elliptic equations, so it is limited to time-homogeneous $b = b(x)$.

[1] Retrospectively, condition $p > \frac{2}{2-\sqrt{\delta}}$ could be interpreted as saying the same thing, but, since semigroup $e^{t(\Delta - b \cdot \nabla)}$ in $L^p$ is automatically strongly continuous in all $L^q$, $p < q < \infty$ by the interpolation with the $L^\infty$ contraction estimate, the link between the strength of topology and the value of $\delta$ was somewhat less transparent in the $L^p$ setting.

The literature on the regularity theory of diffusion operator $-\Delta + b \cdot \nabla$ and on the corresponding SDE also deals with larger classes of singular vector fields $b$, i.e. those that contain $\mathbf{F}_\delta$, such as the class of weakly form-bounded vector fields [15, 22] or (basically the largest possible scaling-invariant time-inhomogeneous) Morrey class [18]. However, in the cited papers it is essential that the form-bound $\delta$ is smaller than a dimension-dependent constant $\ll 1$, and it is not yet clear what is the critical value of $\delta$ for these classes of vector fields. There is also the Kato class of vector fields that contains drifts having strong hypersurface singularities, see e.g. [3], but, on the other hand, the Kato class does not even contain $|b| \in L^d$ and itself is contained in the class of weakly form-bounded vector fields.

**4.** As was mentioned above, if $b \in \mathbf{F}_\delta$, $\delta < 4$, then one can construct a quasi contraction strongly continuous Markov semigroup $e^{-t\Lambda_p}$ in $L^p$, $\Lambda_p \supset -\Delta + b \cdot \nabla$, $p \in ]\frac{2}{2-\sqrt{\delta}}, \infty[$. We proved in [25] that the last statement remains valid for all $p$ in a larger interval

$$I_c := [\frac{2}{2-\sqrt{\delta}}, \infty[ \quad \text{(“interval of quasi contractive solvability”)},$$

moreover, the corresponding semigroup inherits many important properties of the heat semigroup $e^{t\Delta}$ such as $L^p \to L^q$ bounds and holomorphy. The interval of quasi contractive solvability $I_c$ can be further extended to the interval of quasi bounded solvability

$$I_m := ]\frac{2}{2-\frac{d-2}{d}\sqrt{\delta}}, \infty[,$$

i.e. for all $p \in I_m$ one still has a strongly continuous semigroup $e^{-t\Lambda_p}$, $\Lambda_p \supset -\Delta + b \cdot \nabla$, but now it satisfies a weaker bound

$$\|e^{-t\Lambda_p}\|_p \leq M_{p,\delta} e^{\lambda_{p,\delta} t} \|f\|_p \quad \text{for some } M_{p,\delta} > 1.$$

The interval of quasi bounded solvability $I_m$ is sharp. See [25]. We note that if $\delta \uparrow 4$, then, while the interval of quasi contractive solvability $I_c$ tends to the empty set, the interval of quasi bounded solvability $I_m$ tends to a non-empty interval $]\frac{d}{2}, \infty[$. That said, as $\delta \uparrow 4$, one has $M_{p,\delta} \uparrow \infty$, so this result still does not allow to include $\delta = 4$.

Where does the condition $\delta < 4$, $p \in I_c$, come from can be seen from the following elementary calculation. Let $b \in \mathbf{F}_\delta$ be additionally bounded and smooth. Consider Cauchy problem $(\partial_t - \Delta + b \cdot \nabla)u = 0$, $u(0) = f \in C_c^\infty$. Without loss of generality, $f \geq 0$, and so $u \geq 0$. Set $v = e^{-\lambda t}u$, $\lambda > 0$. Multiply equation $(\lambda + \partial_t - \Delta + b \cdot \nabla)v = 0$ by $v^{p-1}$ and integrate by parts:

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

($\langle \cdot \rangle$ denotes the integration over $\mathbb{R}^d$, $\langle \cdot, \cdot \rangle$ is the inner product in $L^2$ over reals).

Applying quadratic inequality in the last term, we arrive at

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

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

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

In order to keep the dispersion term non-negative, one needs $\frac{4(p-1)}{p} - 2\sqrt{\delta} \geq 0$, i.e. $\delta < 4$ and $p \in I_c$, which then yields $\|u\|_p \leq e^{\frac{c(\delta)t}{p\sqrt{\delta}}}\|f\|_p$.

**Notations.** $B_r(x)$ denotes the open ball of radius $r$ centered at $x \in \mathbb{R}^d$, $B_r := B_r(0)$.

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

The space of $d$-dimensional vectors with entries in $X$ is denoted by $[X]^d$.

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

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

Put $L^p = L^p(\mathbb{R}^d)$, $W^{1,p} = W^{1,p}(\mathbb{R}^d)$. Set $\|\cdot\|_p := \|\cdot\|_{L^p}$ and $\|\cdot\|_{p\to q} := \|\cdot\|_{L^p \to L^q}$.

Put

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

(all functions considered in this paper are real-valued).

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

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

Set

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

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

$$E_\varepsilon h := \gamma_\varepsilon * h.$$

## 2. Feller semigroup in regime $\delta < 4$

For a given $b \in \mathbf{F}_\delta$, define $b_n := E_{\varepsilon_n} b$ ($\varepsilon_n \downarrow 0$), where $E_\varepsilon$ is the Friedrichs mollifier. Then $b_n$ are bounded, smooth, converge to $b$ component-wise locally in $L^2_{\rm loc}$, and do not increase the form-bound $\delta$ and constant $c(\delta)$ of $b$, i.e.

$$\|b_n\varphi\|_2^2 \leq \delta\|\nabla\varphi\|_2^2 + c(\delta)\|\varphi\|_2^2 \quad \forall\, \varphi \in W^{1,2}$$

(see e.g. [23] for the proof). By the classical theory, for every $n \geq 1$, Cauchy problem

$$(\partial_t + \Lambda_n)u_n = 0, \quad u_n(0) = f \in C_\infty,$$

$$\text{where } \Lambda_n := -\Delta + b_n \cdot \nabla, \quad D(\Lambda_n) := (1-\Delta)^{-1}C_\infty,$$

has unique solution $u_n(t,x) =: e^{-t\Lambda_n}f(x)$, and $e^{-t\Lambda_n}$ is a strongly continuous Feller semigroup on $C_\infty$.

Put $\rho_x(y) := \rho(y-x)$, $\rho(y) = (1+\kappa|y|^2)^{-\frac{d}{2}-1}$, $y \in \mathbb{R}^d$.

**Theorem 2.1** (1st Main Result)**.** *Let $b \in \mathbf{F}_\delta$ with $\delta < 4$. Then*

(*i*) *The limit*

$$s\text{-}C_\infty\text{-}\lim_n e^{-t\Lambda_n} \quad (\textit{loc. uniformly in } t \geq 0)$$

*exists and is a strongly continuous Feller semigroup on $C_\infty$, say, $e^{-t\Lambda}$. Its generator $\Lambda$ is an appropriate operator realization of the formal operator $-\Delta + b\cdot\nabla$ in $C_\infty$ (in general, no longer an algebraic sum of $-\Delta$ and $b\cdot\nabla$, see remark after the theorem regarding domain $D(\Lambda)$).*

*($ii$) Feller semigroup $e^{-t\Lambda}$ is unique in the sense of approximations, i.e. does not depend on the choice of a bounded smooth approximation $b_n$ of $b$ in (i), as long as $b_n$ converge to $b$ in $[L^2_{\rm loc}]^d$ and do not increase the form-bound $\delta$ of $b$ and constant $c(\delta)$.*

*($iii$) Strong Feller property for resolvent:*

$$\|(\mu+\Lambda)^{-1}f\|_{C_\infty} \leq K \sup_{x\in\frac{1}{2}\mathbb{Z}^d}\bigg((\mu-\mu_1)^{-\frac{1}{p\theta}}\langle|f|^{p\theta}\rho_x\rangle^{\frac{1}{p\theta}} + \mu^{-\frac{\beta}{p}}\langle|f|^{p\theta'}\rho_x\rangle^{\frac{1}{p\theta'}}\bigg), \quad f\in L^{p\theta}\cap L^{p\theta'} \qquad ([19])$$

*for fixed $1<\theta<\frac{d}{d-2}$ and $p\geq 2$ such that $p>\frac{2}{2-\sqrt{\delta}}$, for all $\mu$ strictly greater than certain $\mu_1$. In particular, taking into account that $\langle\rho_x\rangle<\infty$, we have, appealing to the Dominated convergence theorem,*

$$\|(\mu+\Lambda)^{-1}f\|_{C_\infty}\leq C\|f\|_\infty, \quad f\in L^\infty.$$

*($iv$) For all $\frac{2}{2-\sqrt{\delta}}\leq p\leq q<\infty$,*

$$\|e^{-t\Lambda_p}\|_{p\to q}\leq C_{\delta,d}e^{\omega_p t}t^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}, \quad \omega_p=\frac{c(\delta)}{2(p-1)}. \qquad ([25, 30])$$

*($v$) If additionally $\delta<\frac{4}{(d-2)^2}\wedge 1$, then the resolvent $u=(\mu+\Lambda)^{-1}f$ satisfies, for every $q\in[2,\frac{2}{\sqrt{\delta}}[$,*

$$\|\nabla u\|_q\leq K_1(\mu-\mu_0)^{-\frac{1}{2}}\|f\|_q, \quad \|\nabla|\nabla u|^{\frac{q}{2}}\|_2\leq K_2(\mu-\mu_0)^{-\frac{1}{2}+\frac{1}{q}}\|f\|_q, \qquad ([26])$$

*and*

$$\|(\mu-\Delta)^{\frac{1}{2}+\frac{1}{s}}u\|_q\leq K\|(\mu-\Delta)^{-\frac{1}{2}+\frac{1}{r}}f\|_q, \quad \textit{for all } 2\leq r<q<s \qquad ([14])$$

*for all $\mu$ greater than some generic $\mu_0$. In particular, we can select $q>d-2$ (and, in the second assertion, $s$ close to $q$) so that, by the Sobolev embedding theorem, the elements on the domain $D(\Lambda)$ are Hölder continuous.*

Remarks. 1. A crucial feature of assertions ($i$)-($iii$) is that they cover the entire range $0<\delta<4$ of magnitudes of singularities of $b$.

2. Assertions ($iv$), ($v$) are included for the sake of completeness. Assertion ($v$) demonstrates that as $\delta$ becomes smaller the information that we have about the Feller generator $\Lambda$ becomes more detailed.

3. The Feller semigroup $e^{-t\Lambda}$ from Theorem 2.1 determines probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ on the canonical space of càdlàg trajectories $\omega_t$, i.e.

$$e^{-t\Lambda}f(x)=\mathbb{E}_{\mathbb{P}_x}f(\omega_t), \quad f\in C_\infty.$$

By a classical result, the process

$$t\mapsto u(\omega_t)-u(x)+\int_0^t\Lambda u(\omega_s)ds, \quad u\in D(\Lambda), \quad \omega \text{ is càdlàg},$$

is a $\mathbb{P}_x$-martingale. That said, there is no description of the domain $D(\Lambda)$ of generator $\Lambda$ even if $|b|\in L^\infty$ with compact support; one can be certain that $C_c^\infty\not\subset D(\Lambda)$. So, for the continuous martingale characterization of $\mathbb{P}_x$, we have the following results.

**Theorem 2.2** ([19, 24])**.** *Let $b \in \mathbf{F}_\delta$ with $\delta < 4$.*

*1)* [24] *For every $x \in \mathbb{R}^d$ there exists a martingale solution of SDE* (3)*, i.e. a probability measure $\mathbb{P}_x$ on the canonical space of continuous trajectories $\big(C([0,1],\mathbb{R}^d), \mathcal{B}_t = \sigma\{\omega_s \mid 0 \leq s \leq t\}\big)$, such that $\mathbb{P}_x[\omega_0 = x] = 1$,*

$$\mathbb{E}_x \int_0^t |b(\omega_s)| < \infty, \quad 0 < t \leq 1 \qquad (\mathbb{E}_x := \mathbb{E}_{\mathbb{P}_x})$$

*and, for every $\varphi \in C_2^2$, the process*

$$M_t^\varphi := \varphi(\omega_t) - \varphi(\omega_0) + \int_0^t (-\Delta\varphi + b \cdot \nabla\varphi)(s, \omega_s) ds$$

*is a continuous martingale, so*

$$\mathbb{E}_x[M_{t_1}^\varphi \mid \mathcal{B}_{t_0}] = M_{t_0}^\varphi$$

*for all $0 \leq t_0 < t_1 \leq 1$ $\mathbb{P}_x$-a.s.*

*2)* [19] *The probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ are unique in the sense of approximation (Theorem 2.1(ii)) and constitute a strong Markov family.*

The probability measures from Theorems 2.1 and 2.2 are obtained via the same approximation of $b$ by $b_n$ and thus coincide. Moreover, as follows from Theorem 2.1, the probability measures $\{\mathbb{P}_x\}_{x\in\mathbb{R}^d}$ from Theorem 2.2 determine a strongly continuous Feller semigroup on $C_\infty$ by formula $e^{-t\Lambda}f(x) := \mathbb{E}_{\mathbb{P}_x} f(x)$.

Together with the conditional weak uniqueness results of [18, 20] for SDE (3) and the strong well-posedness result of [21] (via the approach of Röckner-Zhao), we consider Theorem 2.1 as tentatively completing the description of the diffusion process with form-bounded drift $b$ for $\delta < 4$.

## 3. Semigroup in Orlicz space in the critical regime $\delta = 4$

Here we treat the borderline case $\delta = 4$ which forces us to consider the problem in a suitable Orlicz space. Namely, put

$$\Phi(t) = \cosh t - 1, \quad \cosh t := \frac{e^t + e^{-t}}{2}, \quad t \in \mathbb{R}.$$

Clearly, this function is convex, $\Phi(t) = \Phi(|t|)$, $\Phi(t)/t \to 0$ as $t \to 0$, $\Phi(t)/t \to \infty$ as $t \to \infty$, and $\Phi(t) = 0$ if and only if $t = 0$. So the space $\mathcal{L}_\Phi = \mathcal{L}_\Phi(\mathbb{R}^d)$ of real-valued $\mathcal{L}^d$ measurable functions on $\mathbb{R}^d$ endowed with the gauge norm

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

is a Banach space (recall that $\langle \cdot \rangle$ denotes integration over $\mathbb{R}^d$).

Note that

$$\Phi(t) = \int_0^t \sinh \tau d\tau, \quad \Phi(t) = \sum_{m=1}^\infty \frac{t^{2m}}{(2m)!} \quad \text{and} \quad \Big\langle \Phi\Big(\frac{f}{\|f\|_\Phi}\Big) \Big\rangle \leq 1.$$

In particular,

$$\|f\|_{2m} \leq \big((2m)!\big)^{\frac{1}{2m}} \|f\|_\Phi, \; m = 1, 2, \dots, \tag{4}$$

so

$$f \in \mathcal{L}_\Phi \Rightarrow f \in L^p \text{ and } \lim_n \|f_n - f\|_\Phi = 0 \Rightarrow \lim_n \|f_n - f\|_p = 0$$

for each $p \in [2, \infty[$ and $f_n \in \mathcal{L}_\Phi$.

**Definition 2.** Let $L_\Phi$ denote the closure of $C_c^\infty$ with respect to gauge norm $\|\cdot\|_\Phi$. This is our Orlicz space.

It follows from (4) that locally the topology in $L_\Phi$ is weaker than the topology in $L^\infty$. On the other hand, the functions in $L_\Phi$ must vanish at infinity sufficiently rapidly, i.e. in particular, no slower than functions in $L^2$. We also note that $\mathcal{S} \subset L_\Phi$.

**Theorem 3.1** (2nd Main Result)**.** *Assume that* $b \in \mathbf{F}_4$, *i.e.*

$$\|b\varphi\|_2^2 \leq 4\|\nabla\varphi\|_2^2 + c(4)\|\varphi\|_2^2 \quad \forall\, \varphi \in W^{1,2} \tag{5}$$

*and that* $b$ *has compact support:*

$$\operatorname{sprt} b \subset B_{R_1} \quad \textit{for some } R_1 < \infty.$$

*Let* $\{b_n\}_{n\geq 1}$ *be any sequence of* $C^\infty$ *smooth vector fields that satisfy* (5) *with the same constants as* $b$ *and are such that*

$$\lim_{n\to\infty} \|b - b_n\|_2 = 0 \qquad \textit{and } \cup_{n \geq n_0} \operatorname{sprt} b_n \subset B_R \textit{ for some } R\textit{, where } R_1 < R < \infty \textit{ and } n_0 \gg 1$$

*(e.g. one can take* $b_n := E_{\varepsilon_n} b$, $\varepsilon_n \downarrow 0$, *where* $E_\varepsilon$ *is the Friedrichs mollifier. Then* $\operatorname{sprt}\ b_n \subset B_{R_1 + \frac{1}{n}}$*). Let* $u_n = u_n(t,x)$ *denote the classical solution to Cauchy problem*

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

*Put*

$$T_n^t f := u_n(t), \quad t \geq 0.$$

*The following are true:*

*(i) For every* $n \geq 1$,

$$T_n^t f \in L_\Phi \textit{ and } \|T_n^t f\|_\Phi \leq e^{\omega t}\|f\|_\Phi, \quad f \in C_c^\infty(\mathbb{R}^d),$$

*where constant* $\omega \geq 0$ *depends only on* $d$, $c(4)$, $R$.

*The operators* $\{T_n^t\}_{t\geq 0}$ *extend by continuity to a positivity preserving quasi contraction strongly continuous semigroup in* $L_\Phi$, *say,* $e^{-t\Lambda_n}$. *Its generator* $\Lambda_n$ *in an appropriate operator realization of* $-\Delta + b_n \cdot \nabla$ *in* $L_\Phi$.

*(ii) The limit*

$$s\text{-}L_\Phi\text{-}\lim_n e^{-t\Lambda_n} \quad (\textit{loc. uniformly in } t \geq 0)$$

*exists and determines a positivity preserving quasi contraction strongly continuous semigroup in* $L_\Phi$, *say,* $e^{-t\Lambda}$. *For every* $g \in L_\Phi$, $u := e^{-t\Lambda} g$ *satisfies* $u \in L^2_{\rm loc}([0,\infty[, W^{1,2})$ *and is a weak solution to parabolic equation* $(\partial_t - \Delta + b \cdot \nabla)u = 0$ *in the sense that*

$$-\langle u, \partial_t \psi\rangle + \langle \nabla u, \nabla \psi\rangle + \langle b \cdot \nabla u, \psi\rangle = 0 \quad \textit{for all } \psi \in C_c^1(]0,\infty[\times\mathbb{R}^d).$$

*(iii) The semigroup* $T^t$ *is unique in the sense of approximations, i.e. does not depend on the choice of a regularization* $b_n$ *of* $b$ *as long as* $b_n$ *satisfies the above assumptions.*

3.1. **Some extensions of Theorem 3.1.** We can remove the assumption of the compact support of $b$, but we still need some assumptions on the rate of decay of $b$ at infinity. That is, assume that vector field $b:\mathbb{R}^d\to\mathbb{R}^d$ can be represented as the sum $b=b^{(1)}+b^{(2)}$ where

$$b^{(1)}\in\mathbf{F}_4,\quad b^{(2)}\in[L^\infty\cap L^2]^d$$

are such that

$$\operatorname{sprt} b^{(1)}\subset B_R \ \text{ and } \ \operatorname{sprt} b^{(1)}\cap\operatorname{sprt} b^{(2)}=\varnothing.$$

Set $b^{(1)}_n=\mathbf{1}_{|b^{(1)}|\leq n}b^{(1)}$ and put $b_n:=b^{(1)}_n+b^{(2)}$.

**Theorem 3.2.** *Let $b$ be as above. Then the assertions of Theorem 3.1 remain valid with the following modification: for every $n\geq 1$,*

$$\|T^t_nf\|_\Phi\leq e^{(\lambda+G)t}\|f\|_\Phi,\quad t\geq 0,$$

*where $\lambda=2^{-1}c_5+2^{-1}\|b^{(2)}\|^2_\infty$ and $G=c_5\langle\mathbf{1}_{B_{aR_1}}\rangle+\|b^{(2)}\|_2^2$, $c_5=c(4)+4(d-1)R_1^{-2}$, $a=1+\theta^{-1}$ (constants $c_5$, $a$ and $\theta$ are from the proof of Theorem 3.1).*

The previous theorem applies to vector field

$$b(x)=(d-2)\mathbf{1}_{B_R}(x)|x|^{-2}x+C\mathbf{1}_{B_R^c}(x)|x|^{-\alpha-1}x,\quad \alpha>\frac{d}{2},$$

where $R>0$, $C<\infty$. That said, a model example of a vector field $b\in\mathbf{F}_\delta$ having critical-order singularity at the origin and critical decay at infinity is

$$b(x)=\frac{\sqrt{\delta}}{2}(d-2)|x|^{-x}x \tag{6}$$

(note the sign in front of $\sqrt{\delta}$). As the previous example shows, Theorem 3.2 allows us to take $\delta=4$, but it still imposes a stronger requirement, in comparison with (6), on the rate of decay of $b$ outside of a ball of large radius. The next theorem and the remark after address that.

**Theorem 3.3.** *Let $|b|\in L^2$, $\operatorname{div} b\in L^1_{\rm loc}$. Set $V=0\vee\operatorname{div} b$ and assume that $V=V_1+V_2$,*

$$V_2\in L^\infty,\quad \|V_1^{\frac{1}{2}}\varphi\|_2^2\leq 4\|\nabla\varphi\|_2^2+c(4)\|\varphi\|_2^2 \text{ for all } \varphi\in W^{1,2}, \text{ and } \operatorname{sprt} V_1\subset B_{R_1}.$$

*Then the assertions of Theorem 3.1 remain valid with the following modification: for every $n\geq 1$,*

$$\|T^t_nf\|_\Phi\leq e^{(\lambda+\|V_2\|_\infty+G)t}\|f\|_\Phi,\quad t\geq 0,$$

*where $\lambda=c(4)+4(d-1)R^{-2}$, $G=2\lambda\langle\mathbf{1}_{B_{aR_1}}\rangle$, $a=3$.*

Furthermore, one can remove condition $|b|\in L^2$ in Theorem 3.3 by considering $\tilde{b}=b+\mathsf{f}$, where $b$ satisfies the assumptions of Theorem 3.3, and $|\mathsf{f}|\in L^\infty$, $\operatorname{div}\mathsf{f}\in L^1_{\rm loc}$, $V_3:=0\vee\operatorname{div}\mathsf{f}\in L^\infty$. See Remark 2 after the proof of Theorem 3.3 for details. This allows to include model drift (6), i.e. take

$$\tilde{b}(x)=(d-2)|x|^{-2}x.$$

(Set $\tilde{b}_n=(d-2)E_n(\mathbf{1}_{B_1}|x|^{-2}x)+(d-2)\mathbf{1}_{B_1^c}|x|^{-2}x$.)

Remark 1. One can combine drifts considered in the previous theorems, e.g. one can consider drift $b+\mathsf{f}$ with $b$ from Theorem 3.2 and $\mathsf{f}$ from Theorem 3.3, such that

$$b^{(1)}\in\mathbf{F}_{\delta_1},\quad V_1^{\frac{1}{2}}\in\mathbf{F}_{\delta_2},\quad \delta_1+\delta_2=4.$$

The main disadvantage of the previous results is that the singularities of $b$ are contained in a bounded set. In the next theorem we improve these results as follows.

**Theorem 3.4.** *Let $\{x_m\} \subset \mathbb{R}^d$, $\{R_m\} \subset \mathbb{R}_+$ be such that $\lim_m |x_m| = \infty$ and $B(x_m, R_m) \cap B(x_k, R_k) = \emptyset$ for all $m \neq k$. Let $b(x) = \sum_{m=1}^\infty b^{(m)}(x)$ be such that*

$$\text{sprt } b^{(m)} \subset B(x_m, R_m), \ \|b^{(m)}\varphi\|_2^2 \leq \delta_m \|\nabla \varphi\|_2^2 + c(\delta_m)\|\varphi\|_2^2 \quad \varphi \in W^{1,2},$$

$$\sum_{m=1}^\infty \delta_m = 4, \quad \sum_{m=m_0}^\infty (R_m^{-2} + R_m^d)\delta_m < \infty, \ \text{and} \ \sum_{m=m_0}^\infty (1 + R_m^d) c(\delta_m) < \infty \text{ for some } m_0 >> 1.$$

*Then all assertions of Theorem 3.1 remain valid.*

## 4. Proof of Theorem 2.1

Assertion $(i)$ will follow from the Trotter approximation theorem, which, applied to semigroups $\{e^{-t\Lambda_n}\}_{n \geq 1}$ in $C_\infty$, can be formulated as follows:

**Theorem 4.1** (see [13, IX.2.5]). *Assume that exists $\mu_0 > 0$ independent of $n$ such that*

1) *$\sup_n \|(\mu + \Lambda_n)^{-1} f\|_\infty \leq \mu^{-1} \|f\|_\infty$, $\mu \geq \mu_0$;*

2) *there exists $s$-$C_\infty$-$\lim_n (\mu + \Lambda_n)^{-1}$ for some $\mu \geq \mu_0$;*

3) *$\mu(\mu + \Lambda_n)^{-1} \to 1$ in $C_\infty$ as $\mu \uparrow \infty$ uniformly in $n$.*

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

$$e^{-t\Lambda_n} \to e^{-t\Lambda} \quad \text{strongly in } C_\infty$$

*locally uniformly in $t \geq 0$.*

Condition 1) follows from the classical theory, that is, from the fact that $e^{-t\Lambda_n}$ are $L^\infty$ contractions.

Condition 2) is verified as follows. In view of 1), it suffices to verify the existence of the limit on $f$ in a countable dense subset of $C_c^\infty$. Set $u_n := (\mu + \Lambda_n)^{-1} f$. Fix $R > 0$ sufficiently large so that, by Corollary A.1, $\sup_{\mathbb{R}^d \setminus B_R(0)} |u|$ is sufficiently small uniformly in $n$. (To this end, we note that $\langle |f|^{p\theta} \rho_x \rangle$, $\langle |f|^{p\theta'} \rho_x \rangle$ in Corollary A.1 are small if $x \in \mathbb{R}^d \setminus B_R(0)$ for $R$ sufficiently large, i.e. $x$ is far away from the support of $f$.) Next, applying Theorem A.1 and the Arzelà-Ascoli theorem on $\bar{B}_R(0)$, we obtain that there is a subsequence $n_k$ such that $\{u_{n_k}\}$ converges uniformly on $\bar{B}_R(0)$. Taking into account the previous observation regarding smallness of $|u_n|$ on $\mathbb{R}^d \setminus B_R(0)$, we use the diagonal argument to construct a subsequence $u_{n_\ell}$ such that the limit $C_\infty$-$\lim_\ell (\mu + \Lambda_{n_\ell})^{-1} f$ exists. Finally, using the existence of the limit $s$-$L^p$-$\lim_n (\mu + \Lambda_n)^{-1} f$, $p > \frac{2}{2 - \sqrt{\delta}}$, see [26], we obtain that the subsequential limit $C_\infty$-$\lim_\ell (\mu + \Lambda_{n_\ell})^{-1} f$ does not depend on the choice of $n_\ell$. This gives us condition 2).

Let us verify condition 3). Once again, in view of 1), it suffices to verify 3) on a dense subset of $C_\infty$, e.g. all $g \in C_c^\infty$. We invoke the resolvent identity:

$$\begin{aligned} \mu(\mu + \Lambda_n)^{-1} g - \mu(\mu - \Delta)^{-1} g &= \mu(\mu + \Lambda_n)^{-1} b_n \cdot \nabla (\mu - \Delta)^{-1} g \\ &= (\mu + \Lambda_n)^{-1} b_n \cdot \mu(\mu - \Delta)^{-1} \nabla g. \end{aligned}$$

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

$$\|(\mu + \Lambda_n)^{-1} b_n \cdot \mu(\mu - \Delta)^{-1} \nabla g\|_\infty \leq \|(\mu + \Lambda_n)^{-1} |b_n| \mu(\mu - \Delta)^{-1} |\nabla g|\|_\infty \to 0 \tag{7}$$

as $\mu \to \infty$ uniformly in $n$. This is proved in [26, Lemma 4] under additional hypothesis $|b| \in L^2 + L^\infty$, but the proof there can be modified to excludes this hypothesis, see [25, Lemma 4.16]. (Alternatively, one can prove (7) using Theorem A.2 below after taking supremum in $x \in \frac{1}{2}\mathbb{Z}^d$ in (12) and using the fact that $f = |\mu(\mu-\Delta)^{-1}g|$ is bounded on $\mathbb{R}^d$ uniformly in $\mu$.)

## 5. Proofs of Theorems 3.1-3.4

5.1. **Proof of Theorem 3.1.** $(i)$, $(ii)$ We have

$$\langle \partial_t v + \lambda v - \Delta v + b_n \cdot \nabla v, e^v - e^{-v}\rangle = 0 \quad \text{ where } v = e^{-\lambda t} u_n.$$

Let us introduce the weight function $\zeta_r(x) := \eta\big(\frac{|x|}{r}\big)$, where

$$\eta(t) := \begin{cases} 1 & \text{if } \ t \leq 1 \\ (1-\theta(t-1)))^{\frac{1}{\theta}} & \text{if } \ 1 < t < 1+\theta^{-1}, \quad 0<\theta<\frac{1}{2} \\ 0 & \text{if } \ 1+\theta^{-1} \leq t, \end{cases}$$

Put $\mathcal{C}(r, ar) = \{y \in \mathbb{R}^d \mid r \leq |y| \leq ar\}$, $a = 1+\theta^{-1}$. It is easy to check that

$$|\nabla \zeta_r| \leq r^{-1}\mathbf{1}_{\mathcal{C}(r,ar)} \text{ and } -\Delta\zeta_r \leq (d-1)r^{-2}\mathbf{1}_{\mathcal{C}(r,ar)}.$$

1. A direct calculation yields (clearly, $|\nabla \zeta_M| \leq M^{-1}$, $-\Delta v \in L^1$, $|\nabla v| \in L^2 \cap L^1$, $v \in L^\infty$):

$$\begin{aligned} \langle -\Delta v, e^v - e^{-v}\rangle &= \lim_{M\to\infty} \langle -\Delta v, \zeta_M(e^v - e^{-v})\rangle \\ &= \lim_{M\to\infty} \big(\langle |\nabla v|^2, \zeta_M(e^v + e^{-v})\rangle + \langle \nabla v, (e^v - e^{-v})\nabla\zeta_M\rangle\big) \\ &= \langle |\nabla v|^2, (e^v + e^{-v})\rangle = \langle |\nabla v|^2, (e^{\frac{v}{2}} - e^{-\frac{v}{2}})^2 + 2\rangle \\ &= 4\|\nabla(e^{\frac{v}{2}} + e^{-\frac{v}{2}})\|_2^2 + 2\|\nabla v\|_2^2. \end{aligned}$$

Therefore,

$$\lambda\langle v(e^v - e^{-v})\rangle + \partial_t\langle e^v + e^{-v} - 2\rangle + 2\|\nabla v\|_2^2 + 4\|\nabla(e^{\frac{v}{2}} + e^{-\frac{v}{2}})\|_2^2 + 2\langle b_n(e^{\frac{v}{2}} + e^{-\frac{v}{2}}), \nabla(e^{\frac{v}{2}} + e^{-\frac{v}{2}})\rangle = 0,$$

so

$$\lambda\langle v\sinh v\rangle + \partial_t\langle \cosh v - 1\rangle + \|\nabla v\|_2^2 + 8\|\nabla \cosh\frac{v}{2}\|_2^2 \leq 4\|b_n \cosh\frac{v}{2}\|_2 \|\nabla\cosh\frac{v}{2}\|_2.$$

Using our assumption on $b_n$, we write

$$\|b_n \cosh\frac{v}{2}\|_2^2 = \|b_n\big(\zeta_R\cosh\frac{v}{2}\big)\|_2^2 \leq 4\|\nabla\big(\zeta_R\cosh\frac{v}{2}\big)\|_2^2 + c(4)\|\zeta_R\cosh\frac{v}{2}\|_2^2$$

with $R$ such that $\operatorname{sprt} b_n \subset B_R$ (for this, we increase $R$ slightly, or simply redenote $R + \frac{1}{n}$ from the assumption on $b_n$ by $R$), where, setting $w := \cosh\frac{v}{2}$, we have

$$\begin{aligned} \|\nabla\big(\zeta_R\cosh\frac{v}{2}\big)\|^2 \equiv \|\nabla(\zeta_R w)\|_2^2 &= \|\zeta_R\nabla w\|_2^2 + \|w\nabla\zeta_R\|_2^2 + \langle \zeta_R\nabla\zeta_R, \nabla w^2\rangle \\ &= \|\zeta_R\nabla w\|_2^2 - \langle \zeta_R\Delta\zeta_R, w^2\rangle \\ &\leq \|\zeta_R\nabla w\|_2^2 + (d-1)R^{-2}\langle \zeta_R w^2\rangle. \end{aligned}$$

Therefore,

$$\begin{aligned} 4\|b_n\cosh\frac{v}{2}\|_2\|\nabla\cosh\frac{v}{2}\|_2 &\leq \|b_n\cosh\frac{v}{2}\|_2^2 + 4\|\nabla\cosh\frac{v}{2}\|_2^2 \\ &\leq 8\|\nabla\cosh\frac{v}{2}\|_2^2 + c_5\langle\zeta_R w^2\rangle, \quad c_5 = c(4) + 4(d-1)R^{-2} \end{aligned}$$

and so

$$\lambda\langle v\sinh v\rangle+\partial_t\langle\cosh v-1\rangle+\|\nabla v\|_2^2\leq c_5\|\mathbf{1}_{B_{aR}}\cosh\frac{v}{2}\|_2^2.$$

Since $v\sinh v\geq\cosh v-1=2(\cosh^2\frac{v}{2}-1)$,

$$(\lambda-2^{-1}c_5)\langle v\sinh v\rangle+\partial_t\langle\cosh v-1\rangle+\|\nabla v\|_2^2\leq c_5\|\mathbf{1}_{B_{aR}}\|_1,$$

or setting $\lambda=2^{-1}c_5$ and changing $v$ to $\frac{v}{c}$, $c>0$,

$$\langle\cosh\frac{v(t)}{c}-1\rangle+\int_0^t\|\nabla\frac{v(s)}{c}\|_2^2ds\leq\langle\cosh\frac{f}{c}-1\rangle+tc_5\|\mathbf{1}_{B_{aR}}\|_1. \tag{$\star$}$$

From $(\star)$ we obtain that $\int_0^t\|\nabla v(s)\|_2^2ds\leq c_1^2(\langle\cosh\frac{f}{c_1}-1\rangle+tc_5\|\mathbf{1}_{B_{aR}}\|_1)$ with $c_1=\|f\|_\Phi$. Therefore,

$$\int_0^t\|\nabla v(s)\|_2^2ds\leq(1+tc_5\|\mathbf{1}_{B_{aR}}\|_1)\|f\|_\Phi^2. \tag{$\star_1$}$$

From $(\star)$ we obtain also the inequality

$$\langle\cosh\frac{v(\frac{t}{m})}{c}-1\rangle\leq\langle\cosh\frac{f}{c}-1\rangle+c_5\|\mathbf{1}_{B_{aR}}\|_1\frac{t}{m},\quad m=1,2,\dots \tag{$\star_2$}$$

2. Set $c=\frac{\|f\|_\Phi}{1-\gamma_m}$, $m\geq m_0$, where $\gamma_m=c_5\|\mathbf{1}_{B_{aR}}\|_1\frac{t}{m}$ and $\gamma_{m_0}<1$. Then, due to

$$\langle\cosh\frac{f}{(1-\gamma_m)c}-1\rangle\leq 1 \text{ and inequality } \cosh\frac{f}{c}-1\leq(1-\gamma_m)\big(\cosh\frac{f}{(1-\gamma_m)c}-1\big),$$

we obtain from $(\star_2)$ that

$$\Big\langle\cosh\frac{v(\frac{t}{m})}{c}-1\Big\rangle\leq 1-\gamma_m+\gamma_m=1,\quad\text{i.e. }\big\|v(\frac{t}{m})\big\|_\Phi\leq c=\frac{1}{1-\gamma_m}\|f\|_\Phi.$$

Therefore, setting $G=c_5\|\mathbf{1}_{B_{aR}}\|_1$, and using semigroup property of $v(t)$, we arrive at

$$\|v(t)\|_\Phi\leq(1-G\frac{t}{m})^{-m}\|f\|_\Phi$$

and so

$$\|v(t)\|_\Phi\leq e^{Gt}\|f\|_\Phi.$$

Thus, setting $T_n^tf:=u_n(t)$,

$$\|T_n^tf\|_\Phi\leq e^{(2^{-1}c_5+G)t}\|f\|_\Phi. \tag{8}$$

Thus, every $T_n^t$ admits extension by continuity from $C_c^\infty$ to $L_\Phi$, which we denote again by $T_n^t$. We have $\lim_{t\downarrow 0}\|T_n^tf-f\|_\Phi=0$ for all $f\in C_c^\infty$. Since $n$ is finite, the latter is evident from the classical theory, which allows to pass to the limit in $n$ under the gague norm of $T_n^tf-f$. Now, combined with (8), this yields

$$s\text{-}L_\Phi\text{-}\lim_{t\downarrow 0}T_n^t=1,\quad n\geq 1,$$

i.e. semigroups $T_n^t$ are strongly continuous. (So, we can write $T_n^t=e^{-t\Lambda_n}$, where generator $\Lambda_n$ should be considered as appropriate operator realization of $-\Delta+b\cdot\nabla$ in $L_\Phi$. For the sake of uniformity, however, we will continue to use notation $T_n^t$ throughout the rest of the proof.)

3. Next, we claim that $\{T_n^t f\}$ is a Cauchy sequence in $L^\infty([0,T],\mathcal{L}_\Phi)$ and in $L^2([0,T],W^{1,2}(\mathbb{R}^d))$. Indeed, set $h=\frac{v_n-v_k}{c}$, $c>0$. Then

$$\lambda h+\partial_t h-\Delta h+b_n\cdot\nabla h=c^{-1}(b_k-b_n)\cdot\nabla v_k,\quad h(0)=0,$$

so

$$\sup_{0\leq s\leq t}\langle\cosh h(s)-1\rangle+\int_0^t\|\nabla h(s)\|_2^2ds\leq c_5\langle\mathbf{1}_{B_{aR}}\rangle t+c^{-1}e^{2c^{-1}\|f\|_\infty}\int_0^t\langle|b_k-b_n||\nabla v_k(s)|\rangle ds. \tag{$\star\star$}$$

We estimate, using $(\star_1)$,

$$\begin{aligned}\int_0^t\langle|b_k-b_n||\nabla v_k(s)|\rangle ds&\leq\left(\int_0^t\|b_k-b_n\|_2^2ds\right)^{\frac12}\left(\int_0^t\|\nabla v_k(s)\|_2^2ds\right)^{\frac12}\\&\leq\sqrt{t}\|b_k-b_n\|_2(1+tG)^{\frac12}\|f\|_\Phi.\end{aligned}$$

Thus, for every fixed $c>0$, $t>0$,

$$\lim_{n,k\to\infty}\sup_{0\leq s\leq t}\langle\cosh h(s)-1\rangle+\lim_{n,k\to\infty}\int_0^t\|\nabla h(s)\|_2^2ds\leq c_5\langle\mathbf{1}_{B_{aR}}\rangle t.$$

In particular, $\lim_{n,k\to\infty}\int_0^t\|\nabla(v_n(s)-v_k(s))\|_2^2ds\leq c^2c_5\langle\mathbf{1}_{B_{aR}}\rangle t$ for any $c>0$, i.e.

$$\lim_{n,k\to\infty}\int_0^t\|\nabla v_n(s)-\nabla v_k(s)\|_2^2ds=0.$$

Now fix $t_0$ by $c_5\langle\mathbf{1}_{B_{aR}}\rangle t_0\leq 1$, then

$$\lim_{n,k\to\infty}\sup_{0\leq s\leq t_0}\langle\cosh h(s)-1\rangle\leq 1\text{ for any }c>0.$$

The latter means that $\lim_{n,k\to\infty}\sup_{0\leq s\leq t_0}\|v_n(s)-v_k(s)\|_\Phi=0$. The claim is established.

4. Set $T^tf:=L_\Phi\text{-}\lim_n T_n^tf$, $f\in C_c^\infty$. Then, clearly, by (8)

$$\|T^tf\|_\Phi\leq e^{(2^{-1}c_5+G)t}\|f\|_\Phi.$$

We extend $T^t$ by continuity from $C_c^\infty$ to $L_\Phi$. Then, clearly, $T^{t+s}=T^tT^s$,

$$s\text{-}L_\Phi\text{-}\lim_{t\downarrow 0}T^t=1.$$

This is the sought semigroup $e^{-t\Lambda_\Phi}:=T^t$. Moreover, in view of $(\star_1)$, we have

$$T^tg\in L^2([0,T],W^{1,2}(\mathbb{R}^d))\quad g\in L_\Phi.$$

The weak solution characterization of $u=T^tg$ now follows right away from the convergence results established above. The proof of $(i)$, $(ii)$ is completed.

$(iii)$ This uniqueness result follows right away from the construction of the semigroup $T^t$ by verifying Cauchy's criterion. □

5.2. **Proof of Theorem 3.2.** Set $v := e^{-t(\lambda+\Lambda(b_n))}f$. We have

$$\lambda\langle v\sinh v\rangle + \partial_t\langle \cosh v - 1\rangle + \|\nabla v\|_2^2 + 8\|\nabla \cosh\frac{v}{2}\|_2^2 = -4\langle b_n\cosh\frac{v}{2}, \nabla\cosh\frac{v}{2}\rangle.$$

Put $w := \cosh\frac{v}{2}$. Let us first establish the estimate

$$4|\langle b_n w, \nabla w\rangle| \leq 8\|\nabla w\|_2^2 + c_5\langle \zeta_{R_1} w^2\rangle + \|b^{(2)}w\|_2^2. \tag{9}$$

Writing $b_n^{(1)} = (b_{n,1}^{(1)}, b_{n,2}^{(1)}, \dots, b_{n,d}^{(1)})$, $b^{(2)} = (b_1^{(2)}, b_2^{(2)}, \dots, b_d^{(2)})$ and using the assumptions and inequality $4|\alpha\beta| \leq |\alpha|^2 + 4\|\beta|^2$, we have

$$\langle b_{n,i}^{(1)} w, \nabla_i w\rangle = \langle b_{n,i}^{(1)}\zeta_{R_1} w, \mathbf{1}_{\mathrm{sprt}\, b_{n,i}^{(1)}}\nabla_i w\rangle, \quad \langle b_i^{(2)} w, \nabla_i w\rangle = \langle b_i^{(2)} w, \mathbf{1}_{\mathrm{sprt}\, b_i^{(2)}}\nabla_i w\rangle,$$

so

$$\begin{aligned}4|\langle b_n w, \nabla w\rangle| &\leq \|b_n^{(1)}\zeta_{R_1} w\|_2^2 + \|b^{(2)}w\|_2^2 + 4\sum_{i=1}^d \langle (\mathbf{1}_{\mathrm{sprt}\, b_{n,i}^{(1)}} + \mathbf{1}_{\mathrm{sprt}\, b_i^{(2)}})|\nabla_i w|^2\rangle \\ &\leq \|b_n^{(1)}\zeta_{R_1} w\|_2^2 + \|b^{(2)}w\|_2^2 + 4\|\nabla w\|_2^2 \\ &\leq 4\|\nabla(\zeta_{R_1} w)\|_2^2 + c_4\|\zeta_{R_1} w\|_2^2 + \|b^{(2)}w\|_2^2 + 4\|\nabla w\|_2^2.\end{aligned}$$

Recalling that $\|\nabla(\zeta_{R_1} w)\|_2^2 \leq \|\nabla w\|_2^2 + (d-1)R_1^{-2}\langle \zeta_{R_1} w^2\rangle$, we arrive at (9).

The proof of the crucial bounds

$$\int_0^t \|e^{-\lambda s}\nabla T_n^s f\|_2^2 ds \leq (1+tG)\|f\|_\Phi^2, \quad \|T_n^t f\|_\Phi \leq e^{(\lambda+G)t}\|f\|_\Phi$$

follows the proof of Theorem 3.1, i.e. using (9) we obtain inequality

$$\lambda\langle v\sinh v\rangle + \partial_t\langle \cosh v - 1\rangle + \|\nabla v\|_2^2 \leq (c_4 + (d-1)2R_1^{-2})\langle \mathbf{1}_{B_{aR_1}}\cosh^2\frac{v}{2}\rangle + \langle |b^{(2)}|^2\cosh^2\frac{v}{2}\rangle,$$

and hence inequality $\partial_t\langle \cosh v - 1\rangle + \|\nabla v\|_2^2 \leq G$. Integrating the latter over $[0,t]$, we have

$$\langle \cosh v(t) - 1\rangle + \int_0^t \|\nabla v(s)\|_2^2 ds \leq \langle \cosh f - 1\rangle + Gt.$$

The rest of the proof essentially repeats the proof of Theorem 3.1. □

5.3. **Proof of Theorem 3.3.** We start with identity

$$\lambda'\langle v\sinh v\rangle + \partial_t\langle \cosh v - 1\rangle + \|\nabla v\|_2^2 + 8\|\nabla\cosh\frac{v}{2}\|_2^2 = -\langle b_n\cdot\nabla(\cosh v - 1)\rangle, \quad v = e^{-t(\lambda'+\Lambda(b_n))}f.$$

Let us estimate $-\langle b_n\cdot\nabla(\cosh v - 1)\rangle$ from above. Define $\hat\eta(t)$ to be 1 if $t \leq 1$, $2-t$ if $1 < t < 2$ and 0 if $t \geq 2$. Set $\eta_R(x) = \hat\eta(\frac{|x|}{R})$. Then

$$\begin{aligned}-\langle b_n\cdot\nabla(\cosh v - 1)\rangle &= -\lim_{R\to\infty}\langle \eta_R b_n\cdot\nabla(\cosh v - 1)\rangle \\ &= \lim_R\langle \eta_R \operatorname{div} b_n, \cosh v - 1\rangle + \lim_R\langle \nabla\eta_R, b_n(\cosh v - 1)\rangle \\ &\leq \langle E_n V_1, \cosh v - 1\rangle + \langle E_n V_2, \cosh v - 1\rangle;\end{aligned}$$

$$\begin{aligned}\langle E_n V_1, \cosh v - 1\rangle =& 2\langle V_1, E_n(\cosh^2\frac{v}{2} - 1)\rangle = -2\langle V_1\rangle + 2\big\|V_1^{\frac12}(\zeta_{R_1}\sqrt{E_n\cosh^2\frac{v}{2}})\big\|_2^2 \\ &\leq 8\big\|\nabla\big(\zeta_{R_1}\sqrt{E_n\cosh^2\frac{v}{2}}\big)\big\|_2^2 + 2c_4\big\langle \zeta_{R_1}E_n\cosh^2\frac{v}{2}\big\rangle\end{aligned}$$

and setting $w = \cosh \frac{v}{2}$

$$\begin{aligned}
\big\|\nabla(\zeta_{R_1}\sqrt{E_n w^2})\big\|_2^2 &= \big\|\zeta_{R_1}\nabla\sqrt{E_n w^2}\big\|_2^2 - \big\langle \zeta_{R_1}\Delta\zeta_{R_1}, E_n w^2\big\rangle \\
&\leq \|\zeta_{R_1}\nabla\sqrt{E_n w^2}\|_2^2 + (d-1)R_1^{-2}\langle E_n\zeta_{R_1}, w^2\rangle \\
\Big(\text{we are using } \big|\nabla\sqrt{E_n w^2}\big| &= \frac{\big|E_n(w\nabla w)\big|}{\sqrt{E_n w^2}} \leq \sqrt{E_n|\nabla w|^2}\Big) \\
&\leq \big\|\nabla w\big\|_2^2 + (d-1)R_1^{-2}\big\langle E_n\zeta_{R_1}, w^2\big\rangle.
\end{aligned}$$

Thus, $-\langle b_n\cdot\nabla(\cosh v - 1)\rangle \leq 8\|\nabla \cosh\frac{v}{2}\|_2^2 + [2c_4 + (d-1)8R_1^{-2}]\langle E_n\zeta_{R_1}, \cosh^2\frac{v}{2}\rangle + \langle E_n V_2, \cosh v - 1\rangle$ and the inequality

$$\lambda'\langle v\sinh v\rangle + \partial_t\langle \cosh v - 1\rangle + \|\nabla v\|_2^2 \leq [2c_4 + (d-1)8R_1^{-2}]\langle E_n\zeta_{R_1}, \cosh^2\frac{v}{2}\rangle + \langle E_n V_2, \cosh v - 1\rangle$$

is derived and yields (with $\lambda' = \lambda + \|V_2\|_\infty$)

$$\partial_t\langle \cosh v - 1\rangle + \|\nabla v\|_2^2 \leq G.$$

The rest of the proof is practically identical to the proof of Theorem 3.1. □

Remark 2. As we noted earlier, one can remove condition $|b| \in L^2$ in Theorem 3.3 by considering $\tilde{b} = b + \mathsf{f}$, where $b$ satisfies the assumptions of Theorem 3.3, and $|\mathsf{f}| \in L^\infty$, $\operatorname{div}\mathsf{f} \in L^1_{\rm loc}$, $V_3 := 0 \vee \operatorname{div}\mathsf{f} \in L^\infty$. Indeed, set $\tilde{b}_n = b_n + \mathsf{f}$ and let $v = e^{-t(\lambda' + \Lambda(\tilde{b}_n))}f$, where $\lambda' = \lambda + \|V_2\|_\infty + \|V_3\|_\infty$. Then

$$\partial_t\langle \cosh v - 1\rangle + \|\nabla v\|_2^2 \leq G,$$

$$\int_0^t \langle |\tilde{b}_k - \tilde{b}_n||\nabla v_k(s)|\rangle ds \to 0 \text{ as } k, n \to \infty$$

due to $|\tilde{b}_k - \tilde{b}_n| = |b_k - b_n|$.

5.4. **Proof of Theorem 3.4.** Clearly, we are left to estimate $4\|b_n w\|_2\|\nabla w\|_2$, where $b_n = b\mathbf{1}_{|b|\leq n}$, $w = \cosh\frac{v}{2}$, and $v = e^{-t\lambda}u_n$, as follows. Set $\varrho_{R_m}(x) = \zeta_{R_m}(x - x_m)$. We have

$$\|b_n w\|_2^2 = \sum_{m=1}^\infty \|b_n^{(m)}\varrho_{R_m} w\|_2^2 \leq \sum_{m=1}^\infty \delta_m\|\nabla(\varrho_{R_m} w)\|_2^2 + \sum_{m=1}^\infty c(\delta_m)\|\varrho_{R_m} w\|_2^2,$$

$$\begin{aligned}
\|\nabla(\varrho_{R_m} w)\|_2^2 &= \|\varrho_{R_m}\nabla w\|_2^2 + \|w\nabla\varrho_{R_m}\|_2^2 + \langle \varrho_{R_m}\nabla\varrho_{R_m}, \nabla w^2\rangle \\
&= \|\varrho_{R_m}\nabla w\|_2^2 - \langle \varrho_{R_m}\Delta\varrho_{R_m}, w^2\rangle \\
&\leq \|\nabla w\|_2^2 + (d-1)R_m^{-2}\langle \varrho_{R_m} w^2\rangle,
\end{aligned}$$

$$\sum_{m=1}^\infty \delta_m\|\nabla(\varrho_{R_m} w)\|_2^2 \leq 4\|\nabla w\|_2^2 + (d-1)\sum_{m=1}^\infty \delta_m R_m^{-2}\langle \varrho_{R_m} w^2\rangle,$$

$$4\|b_n w\|_2\|\nabla w\|_2 \leq 8\|\nabla w\|_2^2 + \sum_{m=1}^\infty C_m\langle \varrho_{R_m} w^2\rangle, \quad C_m = c(\delta_m) + 4(d-1)\delta_m R_m^{-2}.$$

Finally,

$$\sum_{m=1}^{\infty} C_m \langle \varrho_{R_m} w^2 \rangle \leq \sum_{m=1}^{\infty} C_m \langle \mathbf{1}_{B(x_m, aR_m)} w^2 \rangle$$
$$\leq \langle w^2 - 1 \rangle \sum_{m=1}^{\infty} C_m + \omega_d a^d \sum_{m=1}^{\infty} C_m R^d.$$

□

## Appendix A. Hölder continuity of solutions and embedding theorems

Throughout this section, $b \in \mathbf{F}_\delta$, $\delta < 4$. We use notations introduced in the previous sections: $b_n = E_{\varepsilon_n} b$, $\varepsilon_n \downarrow 0$, and

$$\Lambda_n := -\Delta + b_n \cdot \nabla.$$

**Theorem A.1** ([19, Theorem 5]). *The classical solution $u = u_n$ to non-homogeneous equation*

$$\big(\mu + \Lambda_n\big) u = f, \quad f \in C_c^\infty, \quad \mu > 0, \tag{10}$$

*is Hölder continuous in every ball $B_1(x)$ with constants that do not depend on $n$ (i.e. boundedness or smoothness of $b_n$) or $x \in \mathbb{R}^d$.*

**Theorem A.2** (special case of [19, Theorem 6]). *Let $u = u_n$ denote the classical solution to non-homogeneous equation*

$$(\mu + \Lambda_n) u = |b_n| f, \quad f \in C \cap L^1. \tag{11}$$

*Then for fixed $1 < \theta < \frac{d}{d-2}$ and $p > \frac{2}{2-\sqrt{\delta}}$, $p \geq 2$, there exist constants $\mu_1 > 0$, $\kappa$, $C$ and $\beta \in ]0,1[$ independent of $n$ such that, for every $x \in \mathbb{R}^d$,*

$$\sup_{B_{\frac{1}{2}}(x)} |u| \leq C\bigg( (\mu - \mu_1)^{-\frac{1}{p\theta}} \langle \big(\mathbf{1}_{|b_n|>1} + |b_n|^{p\theta} \mathbf{1}_{|b_n| \leq 1}\big) |f|^{p\theta} \rho_x \rangle^{\frac{1}{p\theta}}$$
$$+ \mu^{-\frac{\beta}{p}} \langle \big(\mathbf{1}_{|b_n|>1} + |b_n|^{p\theta'} \mathbf{1}_{|b_n| \leq 1}\big) |f|^{p\theta'} \rho_x \rangle^{\frac{1}{p\theta'}} \bigg) \tag{12}$$

*for all $\mu > \mu_1$, where $\rho_x(y) := \rho(y - x)$, $\rho(y) = (1 + \kappa |y|^2)^{-\frac{d}{2} - 1}$, $y \in \mathbb{R}^d$.*

*It follows that*

$$\|u\|_\infty \leq C \sup_{x \in \frac{1}{2}\mathbb{Z}^d} \bigg( (\mu - \mu_0)^{-\frac{1}{p\theta}} \Big\langle \big(\mathbf{1}_{|b_n|>1} + |b_n|^{p\theta} \mathbf{1}_{|b_n| \leq 1}\big) |f|^{p\theta} \rho_x \Big\rangle^{\frac{1}{p\theta}}$$
$$+ \mu^{-\frac{\beta}{p}} \Big\langle \big(\mathbf{1}_{|b_n|>1} + |b_n|^{p\theta'} \mathbf{1}_{|b_n| \leq 1}\big) |f|^{p\theta'} \rho_x \Big\rangle^{\frac{1}{p\theta'}} \bigg).$$

To make the paper self-contained, below we reproduce more or less verbatim the proofs of [19, Theorem 5] and [19, Theorem 6].

A.1. **Proof of Theorem A.1.** Fix throughout this proof $p > \frac{2}{2-\sqrt{\delta}}$, $p \geq 2$. Set

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

Fix $R_0 \leq 1$.

**Proposition A.1** ([19, Prop. 1], Remark 12). *For all $0 < r < R \leq R_0$,*

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

*for generic constants $K_1$, $K_2$ (in particular, independent of $k$ or $r$, $R$).*

**Lemma A.1** ([12, Lemma 7.1]). *If $\{z_m\}_{m=0}^\infty \subset \mathbb{R}_+$ is a sequence of positive real numbers such that*

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

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

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

*Then $\lim_m z_m = 0$.*

**Lemma A.2** ([12, Lemma 7.3]). *Let $\varphi(t)$ be a positive function, and assume that there exists a constant $q$ and a number $0 < \tau < 1$ such that for every $0 < R < R_0$*

$$\varphi(\tau R) \leq \tau^\delta \varphi(R) + BR^\beta$$

*with $0 < \beta < \delta$, and*

$$\varphi(t) \leq q\varphi(\tau^k R)$$

*for every $t$ in the interval $(\tau^{k+1}R, \tau^k R)$. Then, for every $0 < \rho < R < R_0$, we have*

$$\varphi(\rho) \leq C\left(\left(\frac{\rho}{R}\right)^\beta \varphi(R) + B\rho^\beta\right)$$

*with constant $C$ that depends only on $q$, $\tau$, $\delta$ and $\beta$.*

The proof follows De Giorgi's method as it is presented in [12, Ch. 7] with appropriate modifications to account for our somewhat different definition of $L^p$ De Giorgi's classes, i.e. functions satisfying the inequality in Proposition A.1.

**Proposition A.2** ([19, Prop. 2]). *For all $0 < r < R \leq R_0$,*

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

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

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

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

Hence

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

On the other hand, Proposition A.1 yields:

$$
\|(\nabla v^{\frac{p}{2}})\mathbf{1}_{B_{\frac{r+\rho}{2}}}\|_2^2 \leq \frac{K_1}{(\rho-r)^2}\|v^{\frac{p}{2}}\mathbf{1}_{B_\rho}\|_2^2+K_2\|f-\mu u\|_\infty^p\,\bigl|B_\rho\cap\{u>k\}\bigr|, \tag{13}
$$

so

$$
\begin{aligned}
\|v^{\frac{p}{2}}\mathbf{1}_{B_r}\|_2^2 &\leq C|B_r\cap\{u>k\}|^{\frac{2}{d}}\biggl(\frac{1}{(\rho-r)^2}\|v^{\frac{p}{2}}\mathbf{1}_{B_\rho}\|_2^2+\|f-\mu u\|_\infty^p\,\bigl|B_\rho\cap\{u>k\}\bigr|\biggr) \\
&\leq \frac{C|B_\rho\cap\{u>k\}|^{\frac{2}{d}}}{(\rho-r)^2}\|v^{\frac{p}{2}}\mathbf{1}_{B_\rho}\|_2^2+C\|f-\mu u\|_\infty^p|B_\rho\cap\{u>k\}|^{1+\frac{2}{d}}.
\end{aligned} \tag{14}
$$

Now, returning from notation $v$ to $(u-k)_+$, we note that if $h<k$, then $\|(u-k)^{\frac{p}{2}}\mathbf{1}_{B_\rho\cap\{u>k\}}\|_2 \leq \|(u-h)^{\frac{p}{2}}\mathbf{1}_{B_\rho\cap\{u>h\}}\|_2$ and $\|(u-h)^{\frac{p}{2}}\mathbf{1}_{B_\rho\cap\{u>h\}}\|_2^2 \geq (k-h)^p|B_\rho\cap\{u>k\}|$. Therefore, we obtain from (14)

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

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

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

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

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

we have

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

hence (using $\xi\geq R^{\frac{2}{p}}$)

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

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

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

where, recall, $z_0 = \langle u^p \mathbf{1}_{B_R \cap \{u>0\}} \rangle |B_R \cap \{u > 0\}|^\alpha$. The latter amounts to requiring

$$\xi \geq C_1 R^{-\frac{2}{p\alpha}} z_0^{\frac{1}{p}}.$$

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

Set

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

**Proposition A.3** ([19, Prop. 3])**.** *Fix $k_0$ by*

$$2k_0 = M(2R) + m(2R) := \sup_{B_{2R}} u + \inf_{B_{2R}} u.$$

*Assume that $|B_R \cap \{u > k_0\}| \leq \gamma |B_R|$ for some $\gamma < 1$. If*

$$\operatorname{osc}(u, 2R) \geq 2^{n+1} C R^{\frac{2}{p}}, \tag{15}$$

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

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

*Proof.* 1. For $h \in ]k_0, k[$, set $w := (u-h)^{\frac{p}{2}}$ if $h < u < k$, set $w := (k-h)^{\frac{p}{2}}$ if $u \geq k$, and $w := 0$ if $u \leq h$. Note that $w = 0$ in $B_R \setminus (B_R \cap \{u > k_0\})$. The measure of this set is greater than $\gamma|B_R|$, so the Sobolev embedding theorem applies and yields

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

where

$$\Delta := B_R \cap \{u > h\} \setminus (B_R \cap \{u > k\}).$$

Now, it follows from Proposition A.1 that

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

For $h \leq k_n$ we have $M(2R) - h \geq M(2R) - k_n \geq C R^{\frac{2}{p}}$, where we have used (15). Therefore, summarizing what was written above, we have

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

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

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

so

$$|B_R \cap \{u > k_n\}|^{\frac{2(d-1)}{d}} \leq |B_R \cap \{u > k_i\}|^{\frac{2(d-1)}{d}} \leq C |\Delta_i| R^{d-2},$$

where $\Delta_i := B_R \cap \{u > k_i\} \setminus (B_R \cap \{u > k_{i-1}\})$. Summing up in $i$ from 1 to $n$, we obtain

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

and the claimed inequality follows. □

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

$$\begin{aligned}\sup_{B_{\frac{R}{2}}}(u - k_n) &\leq C_1\Big(\frac{1}{|B_R|}\langle (u-k_n)^p \mathbf{1}_{B_R \cap \{u > k_n\}}\rangle\Big)^{\frac{1}{p}} \left(\frac{|B_R \cap \{u > k_n\}|}{|B_R|}\right)^{\frac{\alpha}{p}} + C_2 R^{\frac{2}{p}} \\ &\leq C_1 \sup_{B_R}(u - k_n)\left(\frac{|B_R \cap \{u > k_n\}|}{|B_R|}\right)^{\frac{1+\alpha}{p}} + C_2 R^{\frac{2}{p}}\end{aligned}$$

We now apply Proposition A.3 (with, say, $C = 1$). Fix $n$ by

$$cn^{-\frac{d}{2(d-1)}} \leq \left(\frac{1}{2C_1}\right)^{\frac{p}{1+\alpha}}.$$

Then, if $\mathrm{osc}\,(u, 2R) \geq 2^{n+1}R^{\frac{2}{p}}$, we obtain from Proposition A.3

$$M\left(\frac{R}{2}\right) - k_n \leq \frac{1}{2}(M(2R) - k_n) + C_2 R^{\frac{2}{p}},$$

so,

$$M\left(\frac{R}{2}\right) \leq M(2R) - \frac{1}{2^{n+1}}\mathrm{osc}\,(u, 2R) + \frac{1}{2}\frac{1}{2^{n+1}}\mathrm{osc}\,(u, 2R) + C_2 R^{\frac{2}{p}},$$

$$M\left(\frac{R}{2}\right) - m\left(\frac{R}{2}\right) \leq M(2R) - m(2R) - \frac{1}{2}\frac{1}{2^{n+1}}\mathrm{osc}\,(u, 2R) + C_2 R^{\frac{2}{p}}.$$

Hence, since $\mathrm{osc}\,(u, 2R) = M(2R) - m(2R)$,

$$\mathrm{osc}\left(u, \frac{R}{2}\right) \leq \left(1 - \frac{1}{2^{n+2}}\right)\mathrm{osc}\,(u, 2R) + C_2 R^{\frac{2}{p}}. \tag{16}$$

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

$$\mathrm{osc}\left(u, \frac{R}{2}\right) \leq \left(1 - \frac{1}{2^{n+2}}\right)\mathrm{osc}\,(u, 2R) + \frac{1}{2}R^{\frac{2}{p}}. \tag{17}$$

This yields the sought Hölder continuity of $u$ via Lemma A.2 with $\tau = \frac{1}{4}$, $\delta = \log_\tau(1 - 2^{-n-2})$ and $0 < \beta < \frac{2}{p} \wedge \delta$. (Note that the second inequality in the conditions of Lemma A.2 holds if $q = 1$ and $\varphi$ is non-decreasing, which is our case.) □

A.2. **Proof of Theorem A.2.** Recall that $v := (u - k)_+$, where $u = u_n$ solves

$$(\mu + \Lambda_n)u = |b_n|f, \quad f \in C \cap L^1.$$

It suffices to carry out the proof for the case $f \geq 0$. We will need

**Proposition A.4** ([19, Prop. 4])**.** *Fix $R_0 \leq 1$ and $p > \frac{2}{2-\sqrt{\delta}}$, $p \geq 2$. Then, for all $0 < r < R \leq R_0$ and every $k \geq 0$,*

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

*for constants $K_1$, $K_2$ independent of $r$, $R$, $k$ and $n$.*

Recall that we have fixed $1 < \theta < \frac{d}{d-2}$.

**Proposition A.5** ([19, Prop. 5])**.** *There exists constants $K$ and $\beta \in ]0,1[$ such that, for all $\mu \geq 1$,*

$$\sup_{B_{\frac{1}{2}}} u_+ \leq K\bigg(\langle u_+^{p\theta}\mathbf{1}_{B_1}\rangle^{\frac{1}{p\theta}} + \mu^{-\frac{\beta}{p}}\big\langle(\mathbf{1}_{|b_n|>1} + |b_n|^p\mathbf{1}_{|b_n|\leq 1})^{\theta'}|f|^{p\theta'}\mathbf{1}_{B_1}\big\rangle^{\frac{1}{p\theta'}}\bigg). \tag{18}$$

*Proof.* Proposition A.4 yields

$$\mu\|v^{\frac{p}{2}}\mathbf{1}_{B_r}\|_2^2 + \|v^{\frac{p}{2}}\|_{W^{1,2}(B_r)}^2 \leq \tilde{K}_1(R-r)^{-2}\|v\|_{L^p(B_R)}^p + K_2\|\Theta^{\frac{1}{p}}f\mathbf{1}_{u>k}\|_{L^p(B_R)}^p, \qquad v = (u-k)_+, \;\; k \geq 0,$$

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

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

Next, we estimate the left-hand side from below using interpolation inequality:

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

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

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

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

Applying Hölder's inequality in the RHS, we obtain

$$\mu^\beta\|v\|_{L^{p\theta_0}(B_r)}^p \leq \tilde{C}_1(R-r)^{-2}|B_R|^{\frac{\theta-1}{\theta}}\|v\|_{L^{p\theta}(B_R)}^p + \tilde{C}_2\|\Theta^{\frac{1}{p}}f\mathbf{1}_{u>k}\|_{L^p(B_R)}^p. \tag{19}$$

Set

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

so $B^m \equiv B_{R_m}$ is a decreasing sequence of balls converging to the ball of radius $\frac{1}{2}$. By (19),

$$\begin{aligned}\mu^\beta\|v\|_{L^{p\theta_0}(B^{m+1})}^p &\leq \hat{C}_1 2^{2m}\|v\|_{L^{p\theta}(B^m)}^p + \tilde{C}_2\|\Theta^{\frac{1}{p}}f\mathbf{1}_{u>k}\|_{L^p(B^m)}^p \\ &\leq \hat{C}_1 2^{2m}\|v\|_{L^{p\theta}(B^m)}^p + \tilde{C}_2 H|B^m \cap \{v>0\}|^{\frac{1}{\theta}},\end{aligned} \tag{20}$$

where

$$H := \langle\Theta^{\theta'}|f|^{p\theta'}\mathbf{1}_{B^0}\rangle^{\frac{1}{\theta'}} \quad (B^0 = B_1, \text{ i.e. ball of radius } 1)$$

On the other hand, by Hölder's inequality,

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

Applying (20) in the first multiple in the RHS, we obtain

$$\|v\|^{p\theta}_{L^{p\theta}(B^{m+1})} \leq \tilde{C}\mu^{-\beta\theta}\bigg(2^{2\theta m}\|v\|^{p\theta}_{L^{p\theta}(B^m)} + H^\theta|B^m \cap \{v>0\}|\bigg)\bigg(|B^m \cap \{v>0\}|\bigg)^{1-\frac{\theta}{\theta_0}}.$$

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

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

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

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

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

$$\begin{aligned}&\frac{1}{\xi^{p\theta}}\|v_{m+1}\|^{p\theta}_{L^{p\theta}(B^{m+1})} \\ &\leq \tilde{C}\bigg(\frac{2^{p\theta m}}{\xi^{p\theta}}\|v_{m+1}\|^{p\theta}_{L^{p\theta}(B^m)} + |B^m \cap \{u>k_{m+1}\}|\bigg)\bigg(|B^m \cap \{u>k_{m+1}\}|\bigg)^{1-\frac{\theta}{\theta_0}}.\end{aligned} \tag{21}$$

Now,

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

so using in (21) $\|v_{m+1}\|_{L^{p\theta}(B^m)} \leq \|v_m\|_{L^{p\theta}(B^m)}$ and applying the previous inequality, we obtain

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

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

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

and $z_0 = \frac{1}{\xi^{p\theta}}\langle u_+^{p\theta}\mathbf{1}_{B^0}\rangle \leq \tilde{C}^{-\frac{1}{\alpha}}\gamma^{-\frac{1}{\alpha^2}}$ (recall: $B^0 := B_{R_0} \equiv B_1$) provided that we fix $c$ by

$$\xi^{p\theta} := \tilde{C}^{\frac{1}{\alpha}}\gamma^{\frac{1}{\alpha^2}}\langle u_+^{p\theta}\mathbf{1}_{B^0}\rangle + \mu^{-\beta\theta}H^\theta.$$

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

To end the proof of Theorem A.2, we need to estimate $\langle u_+^{p\theta}\mathbf{1}_{B_1}\rangle^{1/p\theta}$ in the RHS of (18) in terms of $f$. We will do it by estimating $\langle u_+^{p\theta}\rho\rangle^{1/p\theta}$ and then applying inequality $\rho \geq c\mathbf{1}_{B_1}$ for appropriate constant $c = c_d$.

**Proposition A.6.** *There exist generic constants $C$, $k$ and $\mu_0>0$ such that for all $\mu>\mu_0$,*

$$(\mu-\mu_0)\langle u^p\rho\rangle+\langle|\nabla u^{\frac{p}{2}}|^2\rho\rangle\leq C\big\langle\big(\mathbf{1}_{|b_n|>1}+|b_n|^p\mathbf{1}_{|b_n|\leq 1}\big)|f|^p\rho\big\rangle. \tag{22}$$

*Proof.* The proof is standard, i.e. we multiply equation (11) by $u|u|^{p-2}$, integrate and then use apply to the drift term quadratic inequality and the form-boundedness condition. In the term that contain $\nabla\rho$ we apply inequality $|\nabla\rho|\leq(\frac{d}{2}+1)\sqrt{\kappa}\rho$ with $\kappa$ chosen sufficiently small; since our assumption on $\delta$ is a strict inequality $\delta<4$, this choice of $\kappa$ suffices to get rid of the terms containing $\nabla\rho$. The details can be found e.g. in [19]. □

*Proof of Theorem A.2, completed.* By Proposition A.5, for all $\mu\geq 1$,

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

where $\rho_x(y):=\rho(y-x)$, and constant $C$ is generic. Applying Proposition A.6 to the first term in the RHS (with $p\theta$ instead of $p$), we obtain for all $\mu\geq\mu_0\vee 1$

$$\sup_{y\in B_{\frac{1}{2}}(x)}|u(y)|\leq C\bigg((\mu-\mu_0)^{-\frac{1}{p\theta}}\big\langle\big(\mathbf{1}_{|b_n|>1}+|b_n|^{p\theta}\mathbf{1}_{|b_n|\leq 1}\big)|f|^{p\theta}\rho_x\big\rangle^{\frac{1}{p\theta}}$$

$$+\mu^{-\frac{\beta}{p}}\big\langle\big(\mathbf{1}_{|b_n|>1}+|b_n|^{p\theta'}\mathbf{1}_{|b_n|\leq 1}\big)|f|^{p\theta'}\rho_x\big\rangle^{\frac{1}{p\theta'}}$$

This ends the proof of Theorem A.2. □

Following the proof of Theorem A.2, we obtain

**Corollary A.1.** *In the assumptions and the notations of Theorem A.2, if $u=u_n$ solves on $\mathbb{R}^d$ $(\mu+\Lambda_n)u=f$, then, for every $x\in\mathbb{R}^d$,*

$$\sup_{B_{\frac{1}{2}}(x)}|u|\leq K\bigg((\mu-\mu_1)^{-\frac{1}{p\theta}}\langle|f|^{p\theta}\rho_x\rangle^{\frac{1}{p\theta}}+\mu^{-\frac{\beta}{p}}\langle|f|^{p\theta'}\rho_x\rangle^{\frac{1}{p\theta'}}\bigg).$$

*where $K$ does not depend on $x\in\mathbb{R}^d$ or $n$.*

## References


[1] D. Adams, Weighted nonlinear potential theory, *Trans. Amer. Math. Soc.* **297** (1986), 73-94.

[2] S. Albeverio, Yu. Kondratiev and M. Röckner, Strong Feller properties for distorted Brownian motion and applications to finite particle systems with singular interactions, "Finite and Infinite Dimensional Analysis in Honor of Leonard Gross" (H. H. Kuo et al., eds.), *Contemporary Mathematics*, **317**, Amer. Math. Soc., 2003.

[3] R. Bass and Z.-Q. Chen, Brownian motion with singular drift. *Ann. Probab.*, 31 (2003), 791-817.

[4] B. Baur and M. Grothaus, Construction and strong Feller property of distorted elliptic diffusion with reflecting boundary, *Potential Anal.* **40** (2014), 391-425.

[5] B. Baur and M. Grothaus, Skorokhod decomposition for a reflected $L^p$-strong Feller diffusion with singular drift, *Int. J. Probab. Stoch. Proc.* **90** (2018), 539-568.

[6] L. Beznea, I. Cîmpean and M. Röckner, Strong Feller semigroups and Markov processes: a counter example, arXiv:2211.14789.

[7] R.M. Blumenthal and R.K. Getoor, Markov Processes and Potential Theory. Pure and Applied Mathematics 29. Academic Press, New York.

[8] S.Y.A. Chang, J.M. Wilson and T.H. Wolff, Some weighted norm inequalities concerning the Schrödinger operator, *Comment. Math. Helvetici*, **60** (1985), 217-246.

[9] F. Chiarenza and M. Frasca, A remark on a paper by C. Fefferman, *Proc. Amer. Math. Soc.*, **108** (1990), 407-409.

[10] G. Da Prato and J. Zabczyk, Ergodicity for Infinite Dimensional Systems, London Math. Society Lecture Notes Ser. 229.

[11] C. Fefferman, The uncertainty principle, *Bull. Amer. Math. Soc.* **9** (1983), 129-206.

[12] E. Giusti, Direct Methods in the Calculus of Variations, *World Scientific*, 2003.

[13] T. Kato. Perturbation Theory for Linear Operators. *Springer-Verlag Berlin Heidelberg*, 1995.

[14] D. Kinzebulatov, Regularity theory of Kolmogorov operator revisited, *Canadian Bull. Math.* **64** (2021), 725-736 (arXiv:1807.07597).

[15] D. Kinzebulatov, A new approach to the $L^p$-theory of $-\Delta + b \cdot \nabla$, and its applications to Feller processes with general drifts, *Ann. Sc. Norm. Sup. Pisa (5)*, **17** (2017), 685-711 (arXiv:1502.07286).

[16] D. Kinzebulatov, Laplacian with singular drift in a critical borderline case, arXiv:2309.04436.

[17] D.Kinzebulatov, Form-boundedness and SDEs with singular drift, arXiv:2305.00146.

[18] D. Kinzebulatov, Parabolic equations and SDEs with time-inhomogeneous Morrey drift, arXiv:2301.13805.

[19] D. Kinzebultov, On particle systems and critical strengths of general singular interactions, arXiv:2402.17009.

[20] D.Kinzebulatov and K.R.Madou, Stochastic equations with time-dependent singular drift, *J. Differential Equations*, **337** (2022), 255-293 (arXiv:2105.07312).

[21] D. Kinzebulatov and K.R. Madou, Strong solutions of SDEs with singular (form-bounded) drift via Roeckner-Zhao approach, arXiv:2306.04825.

[22] D. Kinzebulatov and Yu.A. Semënov, Brownian motion with general drift, *Stoch. Proc. Appl.*, **130** (2020), 2737-2750 (arXiv:1710.06729).

[23] D. Kinzebulatov and Yu. A. Semënov, Heat kernel bounds for parabolic equations with singular (form-bounded) vector fields, *Math. Ann.*, **384** (2022), 1883-1929.

[24] D. Kinzebulatov and Yu. A. Semënov, Sharp solvability for singular SDEs, *Electr. J. Probab.*, **28** (2023), article no. 69, 1-15. (arXiv:2110.11232).

[25] D. Kinzebulatov and Yu. A. Semënov, On the theory of the Kolmogorov operator in the spaces $L^p$ and $C_\infty$, *Ann. Sc. Norm. Sup. Pisa (5)* **21** (2020), 1573-1647 (arXiv:1709.08598).

[26] V. F. Kovalenko and Yu. A. Semënov, $C_0$-semigroups in $L^p(\mathbb{R}^d)$ and $C_\infty(\mathbb{R}^d)$ spaces generated by differential expression $\Delta + b \cdot \nabla$. (Russian) *Teor. Veroyatnost. i Primenen.*, **35** (1990), 449-458; translation in *Theory Probab. Appl.* **35** (1990), 443-453.

[27] N.V. Krylov, On weak solutions of time inhomogeneous Itô's equations with VMO diffusion and Morrey drift, arXiv:2303.11238.

[28] G. Metafune, D. Pallara and M. Wacker, Compactness properties of Feller semigroups, *Studia Math.*, **153** (2002), 179-206.

[29] J. Pruss, A. Rhandi and R. Schnaubelt, The domain of elliptic operators on $L^p(R^d)$ with unbounded drift coefficients, *Houston J. Math.* **32** (2006), 563-576.

[30] R. Schilling, Growth and Hölder conditions for the sample paths of Feller processes, *Probab. Theory and Related Fields*, **112** (1998), 565-611.

[31] R. Schilling and J. Wang, Some theorems on Feller processes: transcience, local times and ultracontractivity, *Trans. Amer. Math. Soc.* **365** (2013), 3255-3286.

[32] Yu. A. Semënov, Regularity theorems for parabolic equations, *J. Funct. Anal.*, **231** (2006), 375-417.

Université Laval, Département de mathématiques et de statistique, Québec, QC, Canada
*Email address*: damir.kinzebulatov@mat.ulaval.ca

University of Toronto, Department of Mathematics, Toronto, ON, Canada
*Email address*: semenov.yu.a@gmail.com