Decay rates of convergence for Fokker-Planck equations with confining drift

Alessio Porretta∗

November 1, 2023

Abstract

We consider Fokker-Planck equations in the whole Euclidean space, driven by Levy processes, under the action of confining drifts, as in the classical Ornstein-Uhlenbeck model. We introduce a new PDE method to get exponential or sub-exponential decay rates, as time goes to infinity, of zero average solutions, under some diffusivity condition on the Levy process, which includes the fractional Laplace operator as a model example. Our approach relies on the long time oscillation estimates of the adjoint problem and applies to (the possible superposition of) both local and nonlocal diffusions, as well as to strongly or weakly confining drifts. Our results extend, with a unifying perspective, many previous works based on different analytic or probabilistic methods, with several interesting connections. On one hand, we make a link between the (nonlinear) PDE methods used for the long time behavior of Hamilton-Jacobi equations and the decay estimates of Fokker-Planck equations; on another hand, we give a purely analytical approach towards some oscillation decay estimates which were obtained so far only with probabilistic coupling methods.

Keywords: Fokker-Planck equations, long time decay, rate of convergence, fractional Laplacian, drift-diffusion equations, coupling methods.

MSC: 35K15, 47G20

1 Introduction

This article is concerned with Fokker-Planck equations in $\mathbb{R}^d$ and with the issue of time-decay rates of zero average solutions, which results as effect of the presence of a confining drift. To fix the ideas, the model problem that we will discuss is the following

$$\begin{cases}
m_t + (-\Delta)^{\frac{\sigma}{2}} m - \text{div}(b(t, x)m) = 0 & \text{in } (0, T) \times \mathbb{R}^d, \\
m(0, x) = m_0(x), & \text{in } \mathbb{R}^d
\end{cases}$$

(1.1)

where $(-\Delta)^{\frac{\sigma}{2}} m$ is the usual Laplace operator for $\sigma = 2$ or the fractional Laplacian for $\sigma \in (0, 2)$.

If $b(t, x) = b(x)$, the decay of zero average solutions is equivalent, in particular, to the problem of long time convergence of unit mass solutions towards stationary invariant measures. This is a fundamental question in the study of Fokker-Planck equations, which attracted a huge interest all over the years. As is well-known, the confining property of the drift plays a crucial role in this convergence, being related to the ergodicity property of the underlying stochastic process. The typical result which is proved in this context, for $m_0$ with zero average, is that

$$\|m(t)\|_X \leq \varpi(t)\|m_0\|_X, \quad \forall t > 0, \quad \text{with } \varpi(t) \downarrow 0, \quad t \to \infty, \quad (1.2)$$

∗Department of Mathematics, University of Rome Tor Vergata. Via della Ricerca Scientifica 1, 00133 Roma, Italy. Email: porretta@mat.uniroma2.it. The author is member of GNAMPA research group of Indam. Partially supported by the MIUR Excellence Department Project Math@TOV awarded to the Department of Mathematics, University of Rome Tor Vergata.
where \( \varpi(t) \) decays exponentially if the confinement force is sufficiently strong \((b \cdot x \gtrsim |x|^2 \text{ at infinity})\), or it possibly decays sub-exponentially if \( b \cdot x \gtrsim |x|^{\gamma} \) with \( \gamma \in (0, 2) \). The space \( X \) is typically a \( L^p \) weighted space, with \( p = 1 \) being the most natural choice, in a way that (1.2) implies decay of the total mass.

It is impossible here to mention all relevant contributions in the literature, where an estimate such as (1.2) has been proved. Let us just point out two main approaches that were developed so far for this kind of analysis.

Starting from the case of Laplace operator and drift \( b = \nabla V(x) \), a first approach relies on functional analytic methods, exploiting spectral gap results in \( L^2 \) (weighted) spaces, logarithmic Sobolev (and Poincaré) inequalities and, more generally, entropy-energy methods. We refer the reader to many influential papers such as [2], [24], [41], [42]. Further refinements of spectral gap methods, relying on ideas in semigroup theory, were developed more recently, see e.g. [23], [31]. Even if the above results are mostly related to local operators, methods based on entropy decay or spectral gaps were also used for nonlocal diffusions, see [11], [24], [43].

A second approach, which is different in spirit, stands on ideas and probabilistic methods developed in the study of recurrence and decay properties of Markov processes. This is often referred to as the Meyn & Tweedie approach (see e.g. [21], [38]) and this is intrinsically related to Harris’ ergodic theorem on Markov chains, as presented in the enlightening version of Hairer and Mattingly [27]. The main tools in this approach are the existence of a Lyapunov function at infinity and the (local) strict positivity of the operator, see e.g. [18], [12]. Extensions to nonlocal diffusions were recently developed in [32]. We also address the reader to the influential paper [3] for a comparison between the different approaches mentioned above. Finally, it is to mention that similar estimates can be also deduced from bilateral bounds on the fundamental solution (see e.g. [36], [39], [46] for divergence free drift terms) or, with a probabilistic approach, from the study of decay rates of the transition probabilities associated to the process, see e.g. [19], [20], [37]. We will comment more about that, in what follows.

The purpose of this article is to introduce a new method for proving decay estimates such as (1.2), for solutions of (1.1). This is an entirely PDE method, relatively simple, which applies to a variety of different regimes, ranging from local to nonlocal diffusions and including both geometrical and subgeometrical rates of convergence.

The motivation of our work comes from the study of the long time behavior of mean-field game systems [22]; those are systems of PDEs coupling Fokker-Planck and Hamilton-Jacobi equations, which describe Nash equilibria in large populations of interacting agents (see [33], [28], [13]). In that context, the Fokker-Planck equation is driven by the optimal strategies chosen by the agents, which results in the fact that the drift term in (1.1) depends on the solution of the coupled Hamilton-Jacobi equation. With this in mind, the goal is to obtain decay estimates for Fokker-Planck equations under hopefully cheap conditions on the drift term \( b(t, x) \); in particular, our results apply to time-dependent fields \( b \) and avoid any direct assumption on \( \text{div}(b) \).

Before explaining the novelty of our approach, let us give a sample of the kind of results that we prove, when specialized to the model problem (1.1). We start with the case of purely local diffusion, where the results are slightly more general.

**Theorem 1.1** Let us set \( \langle x \rangle := \sqrt{1 + |x|^2} \). For \( k > 0 \), assume that \( m_0 \in L^1(\mathbb{R}^d, \langle x \rangle^k \, dx) \) be such that \( \int_{\mathbb{R}^d} m_0 \, dx = 0 \). Let \( m \) be a weak solution to the Cauchy problem

\[
\begin{aligned}
  m_t - \Delta m - \text{div}(b(t, x)m) &= 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \\
  m(0, x) &= m_0(x), \quad \text{in } \mathbb{R}^d
\end{aligned}
\]

where we assume that \( b(t, x) \) is a locally bounded function satisfying

\[
\exists \alpha, R > 0 : \quad b(t, x) \cdot x \geq \alpha |x|^{\gamma} \quad \forall x \in \mathbb{R}^d : |x| \geq R, \forall t > 0, \tag{1.4}
\]

and

\[
\exists c_0 > 0 : \quad (b(t, x) \cdot x - b(t, y)) \cdot (x - y) \geq -c_0|x - y| \quad \forall x, y \in \mathbb{R}^d, \forall t > 0. \tag{1.5}
\]

Then we have:
(a) If $\gamma \geq 2$ in (1.3), then
\[ \|m(t)\|_{L^1(\mathbb{R}^d)} \leq K e^{-\omega t} \|m_0\|_{L^1(\mathbb{R}^d)} \]
for some constant $K, \omega$ only depending on $\alpha, c_0, k, d$.

(b) If $\gamma \in (0, 2)$ in (1.4), then for any $\bar{k} > k$ such that $m_0 \in L^1(\mathbb{R}^d, \langle x \rangle^{\bar{k}} dx)$ we have
\[ \|m(t)\|_{L^1(\mathbb{R}^d)} \leq K (1 + t)^{-q} \|m_0\|_{L^1(\mathbb{R}^d)} \]
where $q = \frac{k-k}{2-\gamma}$ for some $K$ depending on $\alpha, \gamma, c_0, k, \bar{k}, d$.

Theorem 1.1 recovers, with a unifying approach, the results which are known in the literature, including the case of slow, or degenerate, confining drift, which was studied previously both with the entropy method by Toscani and Villani [42] and with more sophisticated tools of semigroup theory by Kavian. We stress however that we avoid any method by Toscani and Villani [42] and with more sophisticated tools of semigroup theory by Kavian, Mischler and Dao [31]. We stress however that we avoid any $C^1$ regularity of the drift $\ell$ and any requirement on $\text{div}(b)$ or $Db$: the one-sided condition in (1.4) plays somehow a similar role, although it appears to be a milder requirement which, for instance, allows for any (space-time) bounded perturbation of dissipative vector fields. We also mention that, at the expense of a mild restriction on the parameter $k$ in the weight function, the previous result holds even replacing (1.5) with more general conditions, e.g. assuming $(b(t, x) - b(t, y)) \cdot (x-y) \geq -c_0 |x-y| \vee |x-y|^\gamma$, see Remark 3.7.

Whilst Theorem 1.1 only gives a sample of the results that we prove, the main outcome of our approach appears in the extension of this kind of estimates to general Levy processes, including possibly nonlocal diffusions. This is the model result that we prove for the specific example of the fractional Laplacian, to be compared with [32, Thm 1.4].

**Theorem 1.2** Let $\sigma \in (0, 2)$ and $k \in (0, \sigma)$. Assume that $m_0 \in L^1(\mathbb{R}^d, \langle x \rangle^k dx)$ be such that $\int_{\mathbb{R}^d} m_0 dx = 0$. Let $m$ be a weak solution to the Cauchy problem (1.7), where we assume that $b$ is a locally Lipschitz vector field satisfying (1.4) and either of the following conditions:

(i) $\sigma \in (1, 2)$ and (1.5) holds true.

(ii) $\sigma \in (0, 1]$ and $(b(t, x) - b(t, y)) \cdot (x-y) \geq -c_0 |x-y| \vee |x-y|^\gamma$, for some $\delta > 0, c_0 > 0$. Then we have:

(a) If $\gamma \geq 2$ in (1.4), then
\[ \|m(t)\|_{L^1(\mathbb{R}^d)} \leq K e^{-\omega t} \|m_0\|_{L^1(\mathbb{R}^d)} \]
for some constant $K, \omega$ only depending on $\alpha, \sigma, \delta, c_0, k, d$.

(b) If $\gamma \in ((2-\sigma) \vee 1, 2) \in (1.4)$, then for any $k < \bar{k}$ such that $m_0 \in L^1(\mathbb{R}^d, \langle x \rangle^{\bar{k}} dx)$ we have
\[ \|m(t)\|_{L^1(\mathbb{R}^d)} \leq K (1 + t)^{-q} \|m_0\|_{L^1(\mathbb{R}^d)} \]
where $q = \frac{k-k}{2-\gamma}$ for some $K$ depending on $\alpha, \gamma, \sigma, \delta, c_0, k, \bar{k}, d$.

Let us stress that the local Lipschitz character of the drift, which is assumed in the above statement, was only used to make sure that $m(t) \in L^1$. More general statements, which extend Theorem 1.2 will be given in Section 4 for Levy operators: we refer the reader to Theorem 4.2 in Section 4. In that case, the decay is formulated in the total variation norm for measure-valued solutions. Correspondingly, we will adopt a suitable formulation of solution which can be applied to a far more general setting; as an example, one can consider the case of fractional Laplacian with additional second order degenerate diffusions.

Compared with the existing literature, our strategy is certainly close to (and even inspired from) the probabilistic approach mentioned above, although we develop a purely PDE method which has an interest in its own. There are two main cornerstones in the approach that we suggest:
(i) the decay estimates for the Fokker-Planck equation are entirely deduced by duality from oscillation decay estimates obtained on the solutions of the adjoint problem

\[
\begin{align*}
&\left\{ u_t + (-\Delta)^\sigma u + b(t, x) \cdot Du = 0 \right. \\
&\left. u(0, x) = u_0(x), \right. \\
&\text{in } (0, T) \times \mathbb{R}^d, \\
&\text{in } \mathbb{R}^d.
\end{align*}
\]  

(1.6)

(ii) for solutions of (1.6), we prove decay estimates which are dual to estimates of the type (1.2). Namely, we show that

\[
[u(t)]_{(x)^k} \leq \varpi(t) [u_0]_{(x)^k} \quad \text{with } \varpi(t) \uparrow \infty,
\]

where

\[
[u]_{(x)^k} = \sup_{x, y \in \mathbb{R}^d} \frac{|u(x) - u(y)|}{(x)^k + (y)^k}.
\]  

(1.7)

Estimates such as (1.7) are evidence of the ergodicity of the underlying process. In the PDE perspective, global (in time) bounds for the oscillation of solutions are usually exploited in the study of long time behavior of Hamilton-Jacobi equations and the convergence to the ergodic problem (see e.g. [10, 14]). Those results typically stand on either classical gradient estimates (through the celebrated Bernstein method) or doubling variables techniques developed in the theory of viscosity solutions, specifically by Ishii and Lions [29]. This kind of methods was later extended to nonlocal operators in [5, 30]. However, quantitative estimates like (1.7) are new in this framework, compared to those earlier works. Otherwise, estimates similar to (1.7) can be found in the probabilistic literature in the form of contraction estimates for transition probabilities in Wasserstein’s metrics, see e.g. [19], [20], [37], [44]. In that context, the proofs rely on probabilistic tools based on coupling methods ([15], [35]).

In our approach, the key idea is to obtain (1.7) by proving that

\[
u(t, x) - u(t, y) \leq \varpi(t) [u_0]_{(x)^k} [(x)^k + (y)^k + \psi(|x - y|)]
\]

(1.8)

for some bounded function \(\psi(r)\) which is locally Hölder near \(r = 0\). Notice that this estimate combines the effects of the Lyapunov function at infinity (used to control long-range interactions) with the local ellipticity used to control short-range interactions, when \(|x - y|\) is small. This is the role of the function \(\psi\), namely to handle the small oscillations; at the same time, since \(\psi\) is bounded, (1.8) readily implies (1.7) up to multiplying the decay rate \(\varpi(t)\) by a bounded factor. Finally, by a simple duality argument, an estimate like (1.7) yields the decay \(\|m(t)\|_{L^1((x)^k)} \leq \varpi(t) \|m_0\|_{L^1((x)^k)}\) for the solution of (1.4).

It seems remarkable that the decay of Fokker-Planck equations will be deduced from the single estimate (1.7) proved on the dual problem (1.6). This is the main difference of our method compared to the Meyn-Tweed, or Harris-type approach, where the local ellipticity is exploited directly in Harnack-type properties of the Fokker-Planck equation (or strong positivity, say irreducibility conditions, on the associated semigroup). By contrast, (1.8) is a pointwise estimate that we derive from maximum principle, to be applied to a suitable operator in the product space for the variables \((x, y)\). This kind of application of the maximum principle contains ideas that connect the doubling variables methods of viscosity solutions to probabilistic coupling methods. We will explain more about that in Section 3.1 where the key steps are given in Lemmas 3.3 and 3.4.

Let us mention that the idea of an analytic version of the coupling method was already suggested in [10] for local diffusions (some related ideas appeared later in [23], in a rather informal way). On that occasion, we explained that the doubling variables approach developed in the theory of viscosity solutions was parallel, and conceptually equivalent, to the coupling method used in probability (see [10], Appendix). In the present article, we come back on this viewpoint and introduce similar ideas to handle nonlocal diffusions as well, getting a crucial improvement in the doubling variable method which is used for Levy operators. This allows us to give entirely PDE proofs of some contraction estimates that were only proved with probabilistic methods so far, with the advantage of relaxing the assumptions and making them ready to several possible extensions.

In the present paper, in order not to increase the level of technicalities which could be induced by generalizations, we confine ourselves to the study of Levy processes, which contain both local diffusion and jump processes, giving evidence of the generality of the method.
We also show how the suggested strategy can be adapted to different Lyapunov functions in order to obtain exponential or sub-exponential rates of convergence. We refer the reader to Theorem 5.5 for a general version of our results, which is not limited to power-type Lyapunov functions. In particular, we show how different decay rates are proved for the weighted norm \( \| m(t) \|_{\mathcal{M}_\varphi} \) (the total variation of \( m(t) \varphi \)) according to whether the Lyapunov function \( \varphi \) satisfies stronger, or weaker, conditions of super-solution. This way, Theorem 5.5 will provide with a general recipe which may please the readers inclined towards statements in a more abstract form. As an example of application of similar generalizations, we will obtain exponential convergence of solutions to (1.3) even for slowly confining drifts, up to using stronger Lyapunov weights rather than powers: we refer to Corollary 5.6 which extends recent results obtained in [31].

To conclude, let us briefly mention the organization of the paper. In Section 2, we introduce the notation and the assumptions that will be used, as well as the notions of solutions for the evolution problems to be considered. Section 3 is devoted to the oscillation decay estimates for problems like (1.6), with power-type weight functions. In Section 4, we deduce by duality the decay of Fokker-Planck equations, which generalize Theorem 1.1 and 1.2. Section 5 is devoted to a more general version of the problems to be considered. Section 3 is devoted to the oscillation decay estimates for problems like (1.6); we refer to Corollary 5.6, which extends recent results obtained in [31].

### 2 Notation, assumptions and preliminary tools

We denote by \( \mathbb{R}^d \), \( d \geq 1 \), the \( d \)-dimensional Euclidean space, with Lebesgue measure \( dx \). \( I_d \) denotes the identity matrix in \( \mathbb{R}^d \). Given two real numbers \( a, b \), we use the standard notation \( a \vee b = \max(a, b) \) and \( a \wedge b = \min(a, b) \). For a positive continuous function \( \varphi(x) \), we denote by \( L^1(\varphi(x)) \) the usual Lebesgue space \( L^1(\mathbb{R}^d, \mu) \) defined in terms of the measure \( \mu = \varphi(x)dx \), and by \( L^\infty(\varphi(x)) \) the space of measurable functions \( u : \mathbb{R}^d \rightarrow \mathbb{R} \) such that \( u \varphi \) is essentially bounded (i.e. \( u \varphi \in L^\infty(\mathbb{R}^d) \)).

We will use systematically the weight function

\[
\langle x \rangle := \sqrt{1 + |x|^2}
\]

replacing the role of \(|x|\). Notice that \( D\langle x \rangle = \frac{x}{|x|} \).

The space \( \mathcal{M}(\mathbb{R}^d) \) denotes the space of finite (signed) Borel measures \( m \) endowed with the total variation norm \( \| m \|_{TV} = |m|(\mathbb{R}^d) \); here \( |m| = m^+ + m^- \), where \( m^\pm \) are positive measures giving the Kantorovich-Rubinstein distance

\[
d_1(m, \tilde{m}) := \sup\{ \int_{\mathbb{R}^d} \varphi d(m - \tilde{m}), \ \varphi \text{ Lipschitz}, \ |\varphi|_\infty, \| D\varphi \|_\infty \leq 1 \}.
\]

Finally, we denote by \( \mathcal{M}_k(\mathbb{R}^d) \) the subset of measures \( m \) with finite \( k \) moments, i.e. \( \langle x \rangle^k \in L^1(\mathbb{R}^d, d|m|) \). For \( m \in \mathcal{M}_k \), we use the notation

\[
\| m \|_{\mathcal{M}_k} := \int_{\mathbb{R}^d} \langle x \rangle^k d|m|
\]

to denote the total variation of \( \langle x \rangle^k m \). We use a similar notation with possibly different weights, setting \( \mathcal{M}_\varphi(\mathbb{R}^d) \) the subset of measures such that \( \varphi \in L^1(\mathbb{R}^d, d|m|) \) for some positive function \( \varphi \), with \( \| m \|_{\mathcal{M}_\varphi} = \int_{\mathbb{R}^d} \varphi d|m| \).

For given \( T > 0 \), we set \( Q_T := (0, T) \times \mathbb{R}^d \). We denote by \( C^{1,2}(Q_T) \) the functions defined on \( Q_T \) which are \( C^1 \) in \( t \) and \( C^2 \) in \( x \). The notation \( C_b(U) \) refers to continuous bounded functions on a set \( U \), and \( C_c(U) \) to continuous compactly supported functions. We denote by \( C([0, T]; \mathcal{M}(\mathbb{R}^d)^*) \) the functions
m from \([0, T]\) into \(\mathcal{M}(\mathbb{R}^d)\) which are continuous with respect to the weak* topology of measures. When working with probability measures, we denote \(C([0, T]; \mathcal{P}(\mathbb{R}^d))\) the functions \(m\) which are continuous from \([0, T]\) into \(\mathcal{P}(\mathbb{R}^d)\), endowed with the \(d_1\) distance.

Associated to a continuous function \(\varphi\) such that \(\varphi(x) \geq 1\) for all \(x \in \mathbb{R}^d\), we introduce suitable weighted seminorms, defined as

\[
[u]_\varphi = \sup_{x, y \in \mathbb{R}^d} \frac{|u(x) - u(y)|}{\varphi(x) + \varphi(y)}.
\]

In particular, we will set

\[
[u]_{(x)}^k := \sup_{x, y \in \mathbb{R}^d} \frac{|u(x) - u(y)|}{(x)_k^k + (y)_k^k}.
\]

With a slight abuse of notation, when considering the case \(k = 0\), we denote

\[
[u]_0 := \sup_{x, y \in \mathbb{R}^d} (|u(x) - u(y)|)
\]

the seminorm measuring the oscillation of a bounded function \(u\).

Given a continuous matrix \(\Sigma : \mathbb{R}^d \to \mathbb{R}^{d \times \ell}\), \(\Sigma^*\) denotes its transposed matrix; then we define the second order operator

\[
\mathcal{L}_0(u) := -\lambda_0 \Delta u - \text{tr}(\Sigma(x)\Sigma^*(x)D^2u)
\]

\[
= -\lambda_0 \sum_{i=1}^d u_{x_i,x_i} - \sum_{i,j=1}^d \sum_{k=1}^\ell \Sigma_{ik}(x)\Sigma_{jk}(x)u_{x_i,x_j}
\]

where \(\lambda_0 \geq 0\). Without going in search of the most generality, we will assume \(\Sigma\) to be bounded and Lipschitz continuous:

\[
\exists \sigma_0, \sigma_1 > 0 : \|\Sigma(x)\| \leq \sigma_0, \quad \|\Sigma(x) - \Sigma(y)\| \leq \sigma_1|x - y| \quad \forall x, y \in \mathbb{R}^d.
\]

Now we define a pure jump Levy operator as

\[
\mathcal{I}(x, [u]) := \int_{\mathbb{R}^d} \{u(x + z) - u(x) - (Du(x) \cdot z)\mathbb{1}_{|z| \leq 1}\} \nu(dz)
\]

in which the measure \(\nu\) is absolutely continuous and satisfies

\[
\exists \lambda, \Lambda \geq 0, \sigma \in (0, 2) : \frac{\lambda}{|z|^{d+\sigma}} \leq \frac{d\nu}{dz} \leq \frac{\Lambda}{|z|^{d+\sigma}} \quad \forall z \in \mathbb{R}^d.
\]

Typical examples of operators which satisfy (2.3)-(2.4) are given by the fractional Laplacian \((-\Delta)^{\frac{\sigma}{2}}\) \((\sigma \in (0, 2))\) or more generally by the generators of \(\alpha\)-stable jump processes. We define the operator \(\mathcal{L}\) as

\[
\mathcal{L}[u] := \mathcal{L}_0(u) - \mathcal{I}(x, [u]).
\]

It is well known (see e.g. [1]) that \(\mathcal{L}[u] + b \cdot Du\) is the infinitesimal generator of a general Levy process where the drift-diffusion part is generated by \(\mathcal{L}_0 + b \cdot D\) and the jump part is represented by \(\mathcal{I}\). This is associated with the stochastic SDE

\[
\begin{align*}
\frac{dX_t}{dt} &= -b(X_t)dt + \sqrt{2\lambda_0}dB_t + \sqrt{2\Sigma(X_t)}dB_t + dP_t \\
X_0 &= x \in \mathbb{R}^d
\end{align*}
\]

where \(B_t\) is a standard Brownian motion and \(P_t\) is a compound Poisson process with measure \(\nu\).

We point out that more general conditions upon the diffusion matrix \(\Sigma(x)\) and the Levy measure \(\nu(z)\) would be possible, but we keep things simpler here for the reader’s convenience.
Under the assumptions (2.2), (2.3), and for a continuous vector field \( b(t, x) \), we consider the evolution problem

\[
\begin{align*}
\partial_t u + \mathcal{L}[u] + b(t, x) \cdot Du &= 0 & \text{in } (0, T) \times \mathbb{R}^d, \\
\nu(0, x) &= u_0(x), & \text{in } \mathbb{R}^d.
\end{align*}
\]

(2.7)

It is well known that, if \( b \) is sufficiently regular, then the unique solution of (2.7) is characterized as \( u(t, x) = \mathbb{E}(X_t) \), where \( X_t \) is the solution of (2.6).

The key-assumption, throughout the whole paper, is the following confining condition on the drift:

\[
\text{for some } c \geq 0, \exists \gamma > 1. \text{ Namely, we suppose that }
\]

\[

\nu(x) > \gamma
\]

where we assume that \( \nu \) where we define

\[

\nu^\ast(B) := \nu(-B), \text{ for every Borel set } B \subseteq \mathbb{R}^d. \text{ Then the operator } \mathcal{L}^\ast \text{ is defined as}
\]

\[
\mathcal{L}^\ast[m] := \mathcal{L}_0^\ast(m) - \mathcal{I}^\ast(x, [m])
\]

\[
=-\lambda_0 m - \sum_{i,j=1}^d \sum_{k,l=1} \Sigma_{ik}\Sigma_{jk} m \partial_{ij} - \int_{\mathbb{R}^d} \{m(x + z) - m(x) - (Dm(x) \cdot z) 1_{|z| \leq 1}\} \nu^\ast(dz).
\]

(2.8)

This allows us to introduce the Fokker-Planck equation

\[
\begin{align*}
\partial_t m + \mathcal{L}^\ast[m] - \text{div}(b(t, x) m) &= 0 & \text{in } (0, T) \times \mathbb{R}^d, \\
m(0) &= m_0, & \text{in } \mathbb{R}^d.
\end{align*}
\]

(2.9)

where \( m_0 \in \mathcal{M}(\mathbb{R}^d) \).

Suitable formulations of problems (2.7) and (2.10) are given below. In order to ease notations, in the following we set, for \( \varphi \in C^2(\mathbb{R}^d) \),

\[
\mathcal{L}^0[\varphi] := \mathcal{L}[\varphi] + b(t, x) \cdot D\varphi \quad ; \quad \mathcal{L}_0^0[\varphi] := \mathcal{L}_0[\varphi] + b(t, x) \cdot D\varphi.
\]

(2.11)

The key-assumption, throughout the whole paper, is the following confining condition on the drift:

\[
\exists \alpha, R > 0: \quad b(t, x) \cdot x \geq \alpha |x|^{\gamma} \quad \forall x \in \mathbb{R}^d: |x| \geq R, \forall t > 0,
\]

(2.12)

where we assume that \( \gamma > 0 \). Three different regimes will be considered: \( \gamma \geq 2 \), which is the coercive case, \( \gamma \in [1, 2) \), which is often referred to in the literature as a slowly confining case, and \( \gamma \in (0, 1) \), which may be called a degenerate confining case. Our goal is to give a unifying approach which provides with optimal decay estimates in all such ranges of \( \gamma \).

We also assume a one-sided control for the accretivity of \( b \). Two slightly different conditions will be assumed depending on the ellipticity of the nonlocal operator, say whether \( \sigma \) in (2.3) is bigger or smaller than 1. Namely, we suppose that \( b \) satisfies either

\[
\exists c_0 > 0: \quad (b(t, x) - b(t, y)) \cdot (x - y) \geq -c_0 |x - y| \quad \forall x, y \in \mathbb{R}^d, \forall t > 0,
\]

(2.13)

or the strongest condition (when \( \sigma \leq 1 \)):

\[
\sigma \in (0, 1] \text{ and } \exists \delta \in (0, 1): \quad (b(t, x) - b(t, y)) \cdot (x - y) \geq -c_0 |x - y| (|x - y| \wedge 1)^{1-\sigma + \delta},
\]

(2.14)

for some \( c_0 > 0 \). Let us mention that further generalizations of (2.13), (2.14) are also possible; we refer to Remark 3.7.
2.1 Viscosity solutions

It is convenient to use the framework of viscosity solutions for problem (2.7). In our setting, where the operator is defined through (2.3), we could use slightly different formulations of the notion of viscosity solution, which all turn out to be equivalent; we refer the reader to [8] for a discussion of this issue, and we pick just one among those possible versions. Moreover, since our focus is on regularity estimates, we will simplify some requirement by only considering continuous solutions. For the interested reader, upper semi-continuous (USC) subsolutions, and lower semi-continuous (LSC) supersolutions, are defined accordingly in [8].

**Definition 2.1** A continuous function \(u \in C^0([0,T] \times \mathbb{R}^d)\) is a viscosity solution of (2.7) if \(u \in L^\infty((0,T);L^\infty(x)^{m})\) for some \(m \in (0,\sigma)\), if \(u(0,x) = u_0(x)\) in \(\mathbb{R}^d\) and if \(u\) satisfies the following two requirements:

(i) for every \(\varphi \in C^{1,2}(Q_T)\) such that \(\varphi \in L^\infty((0,T);L^\infty(x)^{k})\) for some \(k \in (0,\sigma)\), if \((t_0,x_0) \in Q_T\) is a global maximum point of \(u - \varphi\), then

\[
\partial_t \varphi(t_0,x_0) + \mathcal{L}[\varphi](t_0,x_0) + b(t_0,x_0) \cdot D\varphi(t_0,x_0) \leq 0.
\]

(ii) for every \(\varphi \in C^{1,2}(Q_T)\) such that \(\varphi \in L^\infty((0,T);L^\infty(x)^{k})\) for some \(k \in (0,\sigma)\), if \((t_0,x_0) \in Q_T\) is a global minimum point of \(u - \varphi\), then

\[
\partial_t \varphi(t_0,x_0) + \mathcal{L}[\varphi](t_0,x_0) + b(t_0,x_0) \cdot D\varphi(t_0,x_0) \geq 0.
\]

**Remark 2.2** We stress that the requirement made in Definition 2.1 concerning the growth at infinity, for both the solution and the test function, is only due to the nonlocal term \(\mathcal{I}\). Due to assumption (2.4), some limitation in the growth at infinity is required in order that \(\mathcal{I}(x, u)\), \(\mathcal{I}(x, \varphi)\) be well-defined. Alternative definitions, formulated with a more local character, are suggested in [8], which turn out to be equivalent.

Another alternative definition of solution can be given using sub-jets and super-jets of the function \(u\), as introduced in [8], denoted by \(J_{Q_T}^{2,2-} u\), \(J_{Q_T}^{2,2+} u\) respectively. The notation \(J_{Q_T}^{2,2} u\) denotes the closure of elements in the sub/super-jets.

The following lemma is the extension to nonlocal operators of the classical Ishii-Jensen’s lemma in viscosity solutions’ theory [13] Thm 8.3. The following nonlocal version follows from [5, Corollary 2]; we use here a slightly more readable version, which is obtained asymptotically, see e.g. [6, Proposition 2].

**Theorem 2.3** Let \(u\) be an USC sub-solution of (2.7), let \(v\) be a LSC super-solution of (2.7), and let \((\tilde{t}, \tilde{x}, \tilde{y}) \in (0,T) \times (\mathbb{R}^d)^2\) be a local maximum point of the map \((t,x,y) \mapsto [u(t,x) - v(t,y) - \Phi(t,x,y)]\), for some function \(\Phi \in C^{1,2}((0,T) \times (\mathbb{R}^d)^2)\).

Then, for every \(n\) sufficiently large, there exist real numbers \(a, b \in \mathbb{R}\) and matrices \(X, Y \in S_d\) such that

\[
a - b = \partial_t \Phi(\tilde{t}, \tilde{x}, \tilde{y}), \quad (a, D_x \Phi(\tilde{t}, \tilde{x}, \tilde{y}), X) \in J_{Q_{\tilde{t}}}^{2,2+} u(\tilde{t}, \tilde{x}), \quad (b, -D_y \Phi(\tilde{t}, \tilde{x}, \tilde{y}), Y) \in J_{Q_{\tilde{t}}}^{2,2-} v(\tilde{t}, \tilde{y}),
\]

\[
-(n + c_d\|A\|)I \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq A + \frac{1}{n} A^2
\]

where \(A = D_{x,y}^2 \Phi(\tilde{t}, \tilde{x}, \tilde{y})\). Moreover, we have

\[
a - \text{tr}(Q(\tilde{x})X) + b(\tilde{t}, \tilde{x}) \cdot D_x \Phi(\tilde{t}, \tilde{x}, \tilde{y}) - \mathcal{I}(\tilde{x}, u(\tilde{t}), D_x \Phi(\tilde{t}, \tilde{x}, \tilde{y})) \leq 0
\]

\[
b - \text{tr}(Q(\tilde{y})Y) - b(\tilde{t}, \tilde{y}) \cdot D_y \Phi(\tilde{t}, \tilde{x}, \tilde{y}) - \mathcal{I}(\tilde{y}, v(\tilde{t}), -D_y \Phi(\tilde{t}, \tilde{x}, \tilde{y})) \geq 0
\]

where we have denoted, for \(x, p \in \mathbb{R}^d\):

\[
Q(x) = \lambda_0 I_d + \Sigma \Sigma^*(x)
\]

\[
\mathcal{I}(x, u, p) = \int_{\mathbb{R}^d} [u(x + z) - u(x) - (p \cdot z)1_{|z| \leq 1}] d\nu(z).
\]
Let us stress that the previous theorem also implies the consistency of the viscosity formulation; if $u \in C^{1,2}(Q_T)$ is a viscosity solution, then it is also a classical solution. Otherwise, if $u_0$ is a continuous function with polynomial growth, problem (2.7) admits a unique viscosity solution; we refer to Proposition 6.1 for a statement of this kind.

**Remark 2.4** The reader which is not familiar with viscosity solutions should not be puzzled by the statement of Theorem 2.6. In fact, if $u, v$ are smooth, the above statement would be straightforward: if $u(t, x) - v(t, y) = \Phi(t, x, y)$ has a local maximum at $(\hat{t}, \hat{x}, \hat{y})$, then by elementary calculus one has

$$\partial_t u - \partial_t v = \partial_t \Phi,$$

where everything is computed on $(\hat{t}, \hat{x}, \hat{y})$. In this case the statement obviously holds (and can be read for $n \to \infty$) with $a = \partial_t u, b = \partial_t v, X = D^2 u$ and $Y = D^2 v$, where the two inequalities (2.13) are just the conditions $[\partial_t u + \mathcal{L}[u] + b \cdot Du]((\hat{t}, \hat{x}) \leq 0$ (respectively, $[\partial_t v + \mathcal{L}[v] + b \cdot Dv]((\hat{t}, \hat{y}) \geq 0$) having replaced $D_x u(\hat{t}, \hat{x}), D_y v(\hat{t}, \hat{y})$ by $D_x \Phi, -D_y \Phi$ respectively.

The relevance of Theorem 2.3 is to guarantee that a similar conclusion holds for functions $u, v$ that are not differentiable (even possibly just semi-continuous); in that case, $X, Y$ should be considered as approximations (as $n \to \infty$) of the second derivatives $D^2_x u, D^2_y v$, and similarly $a, b$ as replacements of $\partial_t u, \partial_t v$.

Let us mention here that $X, Y$ in principle depend on $n$, but a standard compactness argument (see e.g. [41, Rmk 2.3]) allows eventually to let $n \to \infty$ and to obtain new matrices $X, Y$ that satisfy the upper bound $\left(\begin{array}{cc} X & 0 \\ 0 & -Y \end{array}\right) \leq A$ in the limit. Thus, there is no loss in comprehension for a reader that would consider $n \to \infty$ in a rough reading of the above statement.

### 2.2 Duality solutions of Fokker-Planck equations

Solutions of the Fokker-Planck equation (2.10) will be defined here in duality with the notion of viscosity solution for $\mathcal{L}^b$.

**Definition 2.5** Let $m_0 \in \mathcal{M}(\mathbb{R}^d)$. A function $m \in C^0([0, T); \mathcal{M}(\mathbb{R}^d)^\ast)$ is a solution to (2.10) if

$$\int_{\mathbb{R}^d} \xi \, dm(t) + \int_0^t \int_{\mathbb{R}^d} \xi \, dm(\tau) \, d\tau = \int_{\mathbb{R}^d} \varphi(0, x) \, dm_0$$

$$\forall t \in (0, T), \xi \in C_b(\mathbb{R}^d), f \in C_b(Q_T), \varphi \in C_b([0, t] \times \mathbb{R}^d) \text{ such that}$$

$$\varphi \text{ is a viscosity solution of } \left\{ \begin{array}{ll} -\partial_t \varphi + \mathcal{L}^b[\varphi] = f & \text{in } (0, t) \times \mathbb{R}^d, \\ \varphi(t, x) = \xi(x) & \text{in } \mathbb{R}^d \end{array} \right.$$

Let us point out that, since classical solutions are viscosity solutions, then a solution $m$ according to Definition 2.5 is also a weak solution in the usual distributional sense, which means

$$\int_0^T \int_{\mathbb{R}^d} -\partial_t \varphi + \mathcal{L}^b[\varphi] \, dm(t) \, dt = \int_{\mathbb{R}^d} \varphi(0, x) \, dm_0$$

$$\forall \varphi \in C^1_c([0, T) \times \mathbb{R}^d).$$

The converse statement, namely that weak solutions enjoy the duality property of Definition 2.5, is more delicate under the general setting in which we consider the drift term $b$ and the nonlocal diffusion $\mathcal{I}$. However, this is certainly true for second order operators (i.e. when $\lambda_0 > 0$ in (2.4)) or for locally Lipschitz vector fields $b$. In fact, weak solutions satisfy Definition 2.5 whenever they are unique. This is a consequence of the following existence and uniqueness result, which fully justifies the previous notion of solution in the present setting. We recall that $\mathcal{M}_k(\mathbb{R}^d)$ denotes the subset of measures in $\mathbb{R}^d$ having finite $k$-th moment.

**Theorem 2.6** Let $m_0 \in \mathcal{M}_k(\mathbb{R}^d)$, for some $0 < k < \sigma$. Let $\mathcal{L}$ be defined by (2.7) where $\Sigma$ satisfies (2.2), the Levy measure in $\mathcal{L}$ satisfies (2.4), and $\lambda_0 + \lambda > 0$. Assume that $b \in C^0(Q_T)$ is such that $b(t, x) \cdot x$ is bounded below and either of the following conditions hold:
(i) \( \lambda_0 > 0 \) or \( \sigma \in (1, 2) \), and (2.13) holds true whenever \( |x - y| \leq 1 \).

(ii) \( \lambda_0 = 0 \), \( \sigma \in (0, 1] \) and (2.14) holds true whenever \( |x - y| \leq 1 \).

Then there exists a unique solution \( m \) of (2.10) in the sense of Definition 2.5 and \( m(t) \in M_k \) for all \( t > 0 \). Moreover, \( m \) is obtained as the limit of solutions \( m_\varepsilon \) corresponding to the operator \((\mathcal{L}^b)^* - \varepsilon \Delta\).

Finally, the application \( m_0 \mapsto m(t) \) preserves positivity and mass.

The proof of Theorem 2.6 will be given in the Appendix, for the sake of completeness. It stems from the simple idea that, in duality by what happens for problem (B), the notion of solution given in Definition 2.5 sounds a bit unusual, the reader can see that this is actually a quite natural setting in which problem (2.5) is stable by vanishing viscosity limits. Thus, even if at first sight the formulation of Definition 2.5 or classical solutions. In particular, one can readily deduce that

(2.10) can be formulated (and proved to be well-posed) outside a standard set of assumptions. In fact, the uniqueness result, and the characterization of the solution as the limit of Fokker-Planck equations with vanishing viscosity, show the consistency of this formulation with all known results in terms of weak or classical solutions. In particular, one can readily deduce that \( m \in C([0, T]; L^1(\mathbb{R}^d)) \) when \( \lambda_0 > 0 \) (second order problems), or for the case of fractional Laplacian (or more generally, symmetric kernels) under some extra condition on \( \text{div}(b) \) (see e.g. [32, Thm 1.2]).

Finally, we point out that the Definition 2.5 has been given for signed measures, because we will deal with zero average measures. Due to the Hahn decomposition, and the linearity of the formulation, every solution \( m \) can be split as \( m = m_1 - m_2 \), where \( m_1, m_2 \) are the solutions corresponding to \( m_0^+ \), \( m_0^- \) respectively.

3 Oscillation decay estimates

In this Section we give the oscillation estimates for problem (2.7), where we assume that \( u_0 \) is a continuous function in \( \mathbb{R}^d \), and \( \mathcal{L}, b \) satisfy the assumptions (2.2), (2.4), (2.12) and (2.13) (or (2.14)).

A crucial role here is played by the explicit Lyapunov functions \( \langle x \rangle^\beta := (\sqrt{1 + |x|^2})^\beta \). The properties of those super-solutions are listed below.

Lemma 3.1 Assume (2.12). Let us denote \( \varphi(x) = \langle x \rangle^\beta \), with \( \beta > 0 \). Then we have:

(i) for every \( \varepsilon > 0 \) there exists \( K_\varepsilon > 0 \) such that

\[
\mathcal{L}_0[\varphi] + b(t, x) \cdot D\varphi \geq (\alpha - \varepsilon)\beta \frac{\psi(x)}{\langle x \rangle^{2-\gamma}} - K_\varepsilon \quad \forall x \in \mathbb{R}^d, t > 0.
\]

(ii) Let the Levy measure \( \nu \) satisfy (2.4). If \( 0 < \beta < \sigma \), then for every \( \varepsilon > 0 \) there exists \( K_\varepsilon > 0 \) such that

\[
\mathcal{L}[\varphi] + b(t, x) \cdot D\varphi \geq (\alpha - \varepsilon)\beta \frac{\psi(x)}{\langle x \rangle^{2-\gamma}} - K_\varepsilon \quad \forall x \in \mathbb{R}^d, t > 0,
\]

provided \( \beta \leq 1 \) or \( \gamma > 1 \) in (2.12).

Proof. We compute \( D\varphi = \beta \langle x \rangle^{\beta - 1} \frac{x}{\langle x \rangle}, D^2\varphi = \beta \langle x \rangle^{\beta - 2} \left( I_d + (\beta - 1) \frac{x}{\langle x \rangle} \otimes \frac{x}{\langle x \rangle} \right) \) and so, if \( Q(x) = \lambda_0 I_d + \Sigma \Sigma^*(x) \), we have from (2.12)

\[
|\text{tr}(Q(x)D^2\varphi)| \leq \beta c_\beta (\sigma_0^2 + \lambda_0) \langle x \rangle^{\beta - 2}
\]

for some \( c_\beta \) only depending on \( \beta \) and \( d \). Hence

\[
\mathcal{L}_0[\varphi] \geq \beta \langle x \rangle^{\beta - 2} (b(t, x) \cdot x - c_\beta (\sigma_0^2 + \lambda_0))
\]

Using (2.12), we get (3.1).
As for the Levy operator $\mathcal{I}$, the difference is that we need to handle the nonlocal term for large values of $|z|$. Indeed, we have

$$
\mathcal{I}(x, [\varphi]) = \int_{|z|>1} \{\varphi(x + z) - \varphi(x)\} \nu(dz) + \int_{|z|\leq 1} \{\varphi(x + z) - \varphi(x) - (D\varphi(x) \cdot z)\} \nu(dz)
$$

$$
\leq c_\beta \int_{|z|>1} \max((x)^{\beta-1} |z|, |z|^\beta) \nu(dz) + c_\beta(x)^{\beta-2} \int_{|z|\leq 1} |z|^2 \nu(dz)
$$

and, due to (2.3), the first term is integrable provided $\beta < \sigma$. If $\sigma \leq 1$, this also implies that $\beta \leq 1$, so we deduce that $\mathcal{I}(x, [\varphi])$ is bounded above. Hence we get

$$
\mathcal{L}^b[\varphi] \geq \beta(x)^{\beta-2} b(t, x) \cdot x - C_\beta
$$

for some different constant $C_\beta$, possibly depending on $\lambda_0$ and $\sigma_0$ as well. Using (2.12), we conclude with (3.2). The same remains true if $\sigma > 1$ and $\beta \leq 1$.

If $\sigma > 1$ and $\beta > 1$, we estimate, for some $c_\beta > 0$,

$$
\mathcal{I}(x, [\varphi]) \leq c_\beta((x)^{\beta-1} + 1)
$$

and we get

$$
\mathcal{L}^b[\varphi] \geq \beta(x)^{\beta-2} b(t, x) \cdot x - C_\beta(1 + (x)^{\beta-1}).
$$

Using (2.12), the conclusion follows provided $\gamma > 1$.

\textbf{Remark 3.2} Let us observe, from the previous Lemma, the coercive case given by $\gamma \geq 2$. In this range, we have that $\lim \inf_{|x| \to \infty} \frac{\mathcal{L}^b[\varphi]}{\nu} > 0$. This holds for every $\beta > 0$ in the local case $\mathcal{L} = \mathcal{L}_0$, otherwise it holds for every $\beta < \sigma$ if $\mathcal{I}$ satisfies (2.4).

By contrast, if $\gamma < 2$, then $\mathcal{L}^b[\varphi] \to +\infty$ as $|x| \to \infty$ only for $\beta > 2 - \gamma$. This introduces a natural threshold for nonlocal operators satisfying (2.4): since $\beta < \sigma$ is required for the behavior of $\nu$ at infinity, this will lead to require $\gamma > 2 - \sigma$ for the confinement force to play a significant role.

Notice that $\gamma + \beta - 2 > 0$ already implies $\gamma > 1$ whenever $\beta \leq 1$. This always occurs if $\sigma \leq 1$, since $\beta < \sigma$. Otherwise, when $\sigma > 1$, it is possible to consider $\beta > 1$; nevertheless the restriction $\gamma > 1$ is still needed to obtain (3.2), because one needs the drift term to dominate the diffusion coming from the jumps involving large values of $x$ (see (3.3)). An interpretation is that the jump part contains terms which are comparable to drift components; so they would be controlled by a confining force only if this blows-up at infinity (say, $\gamma > 1$).

To sum up this remark, we observe that requiring $(x)^{\beta}$ to be a Lyapunov function (i.e. $\mathcal{L}^b[(x)^{\beta}] \to +\infty$ as $|x| \to \infty$) leads to the restriction that $\gamma > 1$ and $\gamma > 2 - \sigma$.

However, we stress that the restriction $\beta < \sigma$ in the nonlocal case only comes from the behavior at infinity of the Levy measure $\nu(z)$, and the reader may easily think at generalizing this result to different settings for $\nu$.

\subsection{Doubling variables methods}

As we mentioned in the Introduction, two ingredients yield the oscillation estimates of this Section. While Lemma 3.1 exploits the existence of Lyapunov functions using the confinement property of the drift, the second ingredient comes out by handling the local oscillation of solutions through the maximum principle. Here we exploit in full strength the Ishii-Lions doubling variables method, revisited through the ideas of [40].

Since this is the main ingredient in our approach, let us give first a rough idea of how it works, letting aside any technicality. As we explained in the Introduction, our purpose is to show an oscillation estimate of the form (1.8), for some quantitative rate $\omega(t)$. Rephrasing (1.8), this amounts to show that

$$
\mathcal{U}(t, x, y) \leq \zeta(t, x, y)
$$

(3.4)
for any matrix in the product space for \((x, y)\) where choice \((3.6)\) becomes and a clever choice of the matrix suitable choice of \(C\) will give the estimate \((3.9)\) of Lemma 3.3 below, where \(\nu = \nu(A, x, y)\) is any orthogonal matrix (which will play the role of the reflection coupling in the probabilistic \(U \leq C\) from \([40]\), where the reader may also find (in the Appendix) an extended discussion of the analytic and probabilistic coupling approach. From \([40]\), where the reader may also find (in the Appendix) an extended discussion of the analytic and probabilistic coupling approach.

Somehow we will use a similar idea for nonlocal operators, that we roughly explain here in a model case. Suppose that the Levy measure \(\nu\) is symmetric and is normalized with unit mass; then for any coupling \(\pi\) of the measures \(\nu(dz), \nu(dz')\); on a maximum point of \(U - \zeta\) this yields

\[
\mathcal{I}(x, [u(x)]) - \mathcal{I}(y, [u(y)]) \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [\mathcal{U}(t, x + z, y + z') - \mathcal{U}(t, x, y)]d\pi(z, z')
\]

for any coupling \(\pi\) of the measures \(\nu(dz), \nu(dz')\). The choice of \(\pi\) can now provide with suitable estimates. For instance, if \(\nu\) is the fractional Laplacian, a simple choice would be \(\pi = (I, A)\) for any orthogonal matrix \(A\). This will yield the simplified estimate

\[
\mathcal{I}(x, [u(x)]) - \mathcal{I}(y, [u(y)]) \leq \int_{\mathbb{R}^d} [\mathcal{I}(t, x + z, y + Az) - \mathcal{I}(t, x, y)]\nu(dz)
\]

where \(A\) is any orthogonal matrix (which will play the role of the reflection coupling in the probabilistic approach).

Finally, using \((3.7)\) and \((3.8)\) in \((3.9)\), where we can replace \(\partial_t\mathcal{U}, D_{(x, y)}\mathcal{U}\) with \(\partial_t\zeta, D_{(x, y)}\zeta\), we will obtain estimates on the maximum point of \(U - \zeta\) and, eventually, conclude that \(U \leq \zeta\).

To sum up, the strategy towards the oscillation estimate \((1.8)\) entirely relies on the maximum principle for an (augmented) auxiliary problem in the product space. The justification of this comparison argument finds a natural place in the framework of viscosity solutions, which provides with a rigorous proof in a context of possibly non smooth solutions. It is worth pointing out that the contributions of local and nonlocal diffusions (namely, \((3.7)\) and \((3.8)\)) is independently dealt with; this is why we give two corresponding separate statements, the first one being already developed in our previous work \([40]\).
Lemma 3.3 Let \( u \) be a viscosity solution of (2.7) and suppose that \( (x, y) \) is a local maximum point of the function
\[
    u(t, x) - u(t, y) - \zeta(t, x, y), \quad \zeta(t, \cdot, \cdot) = [\phi(x) + \phi(y)] + \psi(|x - y|)
\]
where \( \phi \) is a \( C^2 \) function and \( \psi : [0, \infty) \to \mathbb{R}_+ \) is a positive, increasing function such that \( \psi(|\xi|) \) is \( C^2 \) in a neighborhood of \( \xi = x - y \).

For \( n > 0 \), let \( X_n, Y_n \in S^N \) be the matrices given by Theorem 2.3 where \( A = D^2_{(x,y)} \zeta \). Then, setting \( Q(x) := \lambda_0 I_d + \Sigma(x) \Sigma^*(x) \), we have
\[
    \limsup_{n \to \infty} \text{tr}(Q(x)X_n - Q(y)Y_n) \leq 4\lambda_0 \psi''(|x - y|) + \psi'(|x - y|) \|\Sigma(x) - \Sigma(y)\|^2 \tag{3.9}
\]
\[
    + \text{tr}(Q(x)D^2\phi(x)) + \text{tr}(Q(y)D^2\phi(y)).
\]

Proof. This is a consequence of the method introduced in [40]. In the above precise form, inequality (3.9) stems from [41] Prop. 2.4] up to a specific computation involving the very definition of \( D^2 \zeta \) (see [40, formula (3.20)]).

The next tool is an extension of the previous lemma to nonlocal operators, as defined in (2.8). Generalizations of the Ishii-Lions method to nonlocal operators have been developed in [3, 7], but we need here an improvement which is crucial for the purpose of our results. This improvement relies on the idea of coupling (by reflection) for jump operators, as we roughly explained above. However, compared to the probabilistic approach, here the coupling will appear as naturally inserted in the analytical doubling variables argument.

In the following, we omit the time dependence of the function \( u \), since it is irrelevant here. Moreover, we assume that all functions involved are integrable at infinity with respect to the Levy measure \( \nu \) (in our context, this is satisfied by all functions which have power growth of order less than \( \sigma \)).

Lemma 3.4 Let the operator \( \mathcal{I} \) be defined by (2.8), where \( \nu \) satisfies (2.4). Suppose that \( (x, y) \) is a maximum point of the function
\[
    u(x) - u(y) - \zeta(x, y), \quad \zeta := [\phi(x) + \phi(y)] + \psi(|x - y|)
\]
where \( \phi \) is a \( C^2 \) function and \( \psi : [0, \infty) \to \mathbb{R}_+ \) is a positive, increasing function such that \( \psi(|\xi|) \) is \( C^2 \) in a neighborhood of \( \xi = x - y \).

Then, for every \( \delta \leq \left( \frac{|x - y|}{2} \right) \), we have (recall the notations in (2.16)):
\[
    \mathcal{I}(x, u(x), D_x \zeta(x, y)) - \mathcal{I}(y, u(y), -D_y \zeta(x, y)) \leq \mathcal{I}(x, \phi(x), D\phi(x)) + \mathcal{I}(y, \phi(y), D\phi(y))
\]
\[
    + 4\lambda \int_0^1 (1 - s) \int_{B_{\delta}} \psi''(|x - y| + 2s(x - y \cdot z)) |x - y \cdot z|^2 dz ds \tag{3.10}
\]
where \( B_{\delta} := \{ z \in \mathbb{R}^d : |z| < \delta \} \) and \( x - y = \frac{x - y}{|x - y|} \).

Proof. Let us set
\[
    H(x, y) := u(x) - u(y) - \zeta(x, y) = u(x) - u(y) - [\phi(x) + \phi(y)] - \psi(|x - y|).
\]
Since \( H(x, y) \geq H(x + z, y + z) \) for every \( z \), we have
\[
    [u(x + z) - u(x) - [u(y + z) - u(y)] \\
    \leq \{ \phi(x + z) - \phi(x) + \phi(y + z) - \phi(y) \} \quad \forall z \in \mathbb{R}^d. \tag{3.11}
\]

In addition, we have
\[
    D_x \zeta + D_y \zeta = D\phi(x) + D\phi(y).
\]
Therefore, for any $A \subseteq \mathbb{R}^d$ and any Levy measure $\nu$, we have
\[
\int_A \{u(x+z) - u(x) - \langle D_x \zeta, z \rangle \mathbb{1}_{|z| \leq 1}\} \nu(dz) - \int_A \{u(y+z) - u(y) + \langle D_y \zeta, z \rangle \mathbb{1}_{|z| \leq 1}\} \nu(dz)
\leq \int_A \{\varphi(x+z) - \varphi(x) - \langle D \varphi(x), z \rangle \mathbb{1}_{|z| \leq 1}\} \nu(dz) + \int_A \{\varphi(y+z) - \varphi(y) - \langle D \varphi(y), z \rangle \mathbb{1}_{|z| \leq 1}\} \nu(dz).
\] (3.12)

Notice in particular that (3.12) with $A = \mathbb{R}^d$ implies (3.10) when $\lambda = 0$. But now we wish to exploit the ellipticity, and to this purpose we split the nonlocal terms for $|z| \geq r/2$ and $|z| < r/2$, where $r = |x - y|$. Moreover, in order to handle the small jumps we decompose the measure $\nu$ as
\[\nu = \lambda \frac{dz}{|z|^{d+\sigma}} + \mu.\]

By assumption (2.4), $\mu$ is nonnegative, and is itself a Levy measure. We define the ball
\[B := \{z \in \mathbb{R}^d : |z| < \left(1 + \frac{r}{2}\right)\}\]
and we decompose the nonlocal integrals in three parts, using that
\[\nu = \nu_B + \mu_B + \lambda \frac{dz}{|z|^{d+\sigma}} \mathbb{1}_B.\]

For the first two measures, we only use (3.12), applied for both $\tilde{\nu} = \nu_B$ and $\tilde{\nu} = \mu_B$. Hence
\[
\int_{\mathbb{R}^d} \{u(x+z) - u(x) - \langle D_x \zeta, z \rangle \mathbb{1}_{|z| \leq 1}\} \nu(dz) - \int_{\mathbb{R}^d} \{u(y+z) - u(y) + \langle D_y \zeta, z \rangle \mathbb{1}_{|z| \leq 1}\} \nu(dz)
\leq \int_{\mathbb{R}^d} \{\varphi(x+z) - \varphi(x) - \langle D \varphi(x), z \rangle \mathbb{1}_{|z| \leq 1}\} \nu(dz) + \int_{\mathbb{R}^d} \{\varphi(y+z) - \varphi(y) - \langle D \varphi(y), z \rangle \mathbb{1}_{|z| \leq 1}\} \nu(dz)
+ \lambda \int_B \{u(x+z) - u(x) - \langle D_x \zeta, z \rangle\} \frac{dz}{|z|^{d+\sigma}} - \lambda \int_B \{u(y+z) - u(y) + \langle D_y \zeta, z \rangle\} \frac{dz}{|z|^{d+\sigma}}
- \lambda \int_B \{\varphi(x+z) - \varphi(x) - \langle D \varphi(x), z \rangle\} \frac{dz}{|z|^{d+\sigma}} - \lambda \int_B \{\varphi(y+z) - \varphi(y) - \langle D \varphi(y), z \rangle\} \frac{dz}{|z|^{d+\sigma}}.\] (3.13)

We will use a different strategy for the measure $\lambda \frac{dz}{|z|^{d+\sigma}} \mathbb{1}_B$, which provides with the desired ellipticity. Since $H$ achieves its maximum at $(x, y)$, then $H(x, y) \geq H(x + z, y + Az)$ where $A$ is the $d \times d$ matrix defined as
\[A := A_d - 2(\overline{x - y} \otimes \overline{x - y})\].

We notice that, for a given $z \in \mathbb{R}^d$ (corresponding to a jump in the $x$-process), $Az$ yields the reflection of $z$ through the hyperplane $\{z \cdot \overline{x - y} = 0\}$. This encodes the idea to use a reflected process in the $y$-variable to optimize the estimate, as in the probabilistic coupling methods.

Thus we have, by the maximality condition,
\[
[u(x+z) - u(x)] - [u(y+Az) - u(y)]
\leq \langle \psi(|x-y+z + Az|) - \psi(|x-y|)|
+ \{\varphi(x+z) - \varphi(x) + \varphi(y+Az) - \varphi(y)\}\forall z \in \mathbb{R}^d.\] (3.14)

Integrating (3.14) on $B$ against the measure $\frac{dz}{|z|^{d+\sigma}}$ we get
\[
\int_B \{u(x+z) - u(x) - D_x \zeta \cdot z\} \frac{dz}{|z|^{d+\sigma}} - \int_B \{u(y+Az) - u(y) + D_y \zeta \cdot Az\} \frac{dz}{|z|^{d+\sigma}}
\leq \int_B \{\varphi(x+z) - \varphi(x) - D \varphi(x) \cdot z\} \frac{dz}{|z|^{d+\sigma}} + \int_B \{\varphi(y+Az) - \varphi(y) - D \varphi(y) \cdot Az\} \frac{dz}{|z|^{d+\sigma}}
+ \int_B \{\psi(|x-y+z + Az|) - \psi(|x-y|)\} \frac{dz}{|z|^{d+\sigma}}\frac{dz}{|z|^{d+\sigma}}.
\]
Now we use that $A$ is an orthogonal matrix and the measure $\frac{dz}{|z|^{d+\sigma}}$ is invariant by rotation. Therefore, the previous inequality reads as follows:

$$\int_B \{u(x+z) - u(x) - D_x \zeta \cdot z\} \frac{dz}{|z|^{d+\sigma}} - \int_B \{u(y+z) - u(y) + D_y \zeta \cdot z\} \frac{dz}{|z|^{d+\sigma}}$$

$$\leq \int_B \{\varphi(x+z) - \varphi(x) - D_x \varphi(x) \cdot z\} \frac{dz}{|z|^{d+\sigma}} + \int_B \{\varphi(y+z) - \varphi(y) - D_y \varphi(y) \cdot z\} \frac{dz}{|z|^{d+\sigma}}$$

$$+ \int_B \left\{\psi(|x - y + z - Az|) - \psi(|x - y|) - \psi'(|x - y|) \overline{x-y} \cdot (z - Az)\right\} \frac{dz}{|z|^{d+\sigma}}.\quad (3.15)$$

Putting together (3.14) and (3.15), we deduce the following inequality

$$\mathcal{I}(x, u(x), D_x \zeta(x, y)) - \mathcal{I}(y, u(y), -D_y \zeta(x, y)) \leq \mathcal{I}(x, \varphi(x), D\varphi(x)) + \mathcal{I}(y, \varphi(y), D\varphi(y))$$

$$+ \lambda \int_B \left\{\psi(|x - y + z - Az|) - \psi(|x - y|) - \psi'(|x - y|) \overline{x-y} \cdot (z - Az)\right\} \frac{dz}{|z|^{d+\sigma}}. \quad (3.16)$$

Recalling the definition of $A$, we compute

$$\int_B \left\{\psi(|x - y + z - Az|) - \psi(|x - y|) - \psi'(|x - y|) \overline{x-y} \cdot (z - Az)\right\} \frac{dz}{|z|^{d+\sigma}}$$

$$= \int_B \left\{\psi(|r + 2(\overline{x-y} \cdot z)|) - \psi(r) - 2 \psi'(r) \overline{x-y} \cdot z\right\} \frac{dz}{|z|^{d+\sigma}}$$

where, we recall, $r = |x - y|$. Now we estimate the right-hand side. To this purpose, notice that for $|z| < \frac{r}{2}$ we have $r + 2 \overline{x-y} \cdot z \geq 0$. Hence, Taylor’s expansion implies

$$\int_B \left\{\psi(|r + 2(\overline{x-y} \cdot z)|) - \psi(r) - 2 \psi'(r) \overline{x-y} \cdot z\right\} \frac{dz}{|z|^{d+\sigma}}$$

$$\leq 4 \int_0^1 (1 - s) \int_B \psi''(r + 2s(\overline{x-y} \cdot z)) |\overline{x-y} \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} ds.$$

Finally, coming back to (3.14), we deduce (3.10).

### 3.2 Exponential rate

We consider here the case that the drift is strongly confining, in the sense that $\gamma \geq 2$ in (2.12). This provides with an exponential decay estimate for a weighted oscillation of the solution of (2.7). For the sake of clarity, we distinguish two different statements for the case of local operators and the case of general Levy processes with jumps.

**Theorem 3.5** Let $\mathcal{L}_0$ be defined by (2.1), where $\lambda_0 > 0$ and $\Sigma$ satisfies (2.2). Assume that $b$ satisfies (2.13) with $\gamma \geq 2$, and (2.14). Let $u$ be a (viscosity) solution of

$$\begin{align*}
\partial_t u + \mathcal{L}_0[u] + b(t, x) \cdot Du &= 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \\
u(0, x) &= u_0(x),
\end{align*}$$

such that $|u(t, x)| \leq C(1 + |x|)^m$ for some $m > 0$ and $C > 0$ (uniformly in time).

For any $k > 0$ such that $u_0 \in L^\infty((x)^{-k})$, there exist positive constants $K, \omega > 0$ (depending on $\alpha, \gamma, \lambda_0, \sigma_0, \sigma_1, c_0, d, k$) such that $u$ satisfies

$$[u(t)](x) \leq K e^{-\omega t} [u_0](x).$$

In case of nonlocal operators with $\sigma$ fractional growth, the statement is similar except for a restriction in the weight which is used (and the growth of solutions).
2.3. In particular, we deduce the existence of matrices the positive maximum cannot either be attained at \( t, x, y \)

\[ u(T) \text{ has a positive maximum in } [0, 1], \]

for every \( \varepsilon > 0 \) sufficiently small.

We claim that, for suitably chosen large constants \( K, C_1, C_2 \), and for a convenient choice of \( \omega \), we have

\[ u(t, x) - u(t, y) \leq e^{-\omega t} \{ K \langle \varphi(x) + \varphi(y) \rangle + \psi(|x - y|) + \varepsilon [\Phi(x) + \Phi(y)] \} + \frac{\varepsilon}{T - t} \]

for every \( \varepsilon > 0 \) sufficiently small.

As usual, we observe that if (3.21) were not true, then the function

\[ u(t, x) - u(t, y) - e^{-\omega t} \{ K [\varphi(x) + \varphi(y)] + \psi(|x - y|) + \varepsilon [\Phi(x) + \Phi(y)] \} \]

has a positive maximum in \([0, T] \times \mathbb{R}^d\) (because \( u(t, x) = o(\Phi(x)) \) as \( |x| \to \infty \)), and this positive maximum is attained at some point \((t, x, y)\) such that, obviously, \( t < T \) and \( x \neq y \). In addition, if \( K \geq |u_0| \langle x \rangle \), then the positive maximum cannot be attained at \( t = 0 \). As a consequence, the point \((t, x, y)\) is a local maximum and we can use the viscosity inequalities satisfied by \( u(t, x) \) and \( u(t, y) \), thanks to Theorem 2.2. In particular, we deduce the existence of matrices \( X_n, Y_n \in \mathcal{S}^N \) such that

\[ \begin{pmatrix} X_n & 0 \\ 0 & -Y_n \end{pmatrix} \leq e^{-\omega t} D^2_{(x,y)} \zeta + \frac{1}{n} \left(e^{-\omega t} D^2_{(x,y)} \zeta\right)^2 \]

and we have

\[ \frac{\varepsilon}{(T - t)^2} - \omega e^{-\omega t} (K \varphi(x) + \varphi(y)) + \psi(|x - y|) + \varepsilon [\Phi(x) + \Phi(y)] \]

\[ + e^{-\omega t} \left(b(t, x) - b(t, y)\right) \cdot \dot{x} - \dot{y} \]

\[ + e^{-\omega t} \left\{ b(t, x) \cdot (K D\varphi(x) + \varepsilon D\Phi(x)) + b(t, y) \cdot (K D\varphi(y) + \varepsilon D\Phi(y)) \right\} \]

\[ \leq \text{tr}(Q(x) X_n - Q(y) Y_n) + I(x, u(t), e^{-\omega t} D_{\zeta} \zeta) - I(y, u(t), e^{-\omega t} D_{\zeta} \zeta) \]

where \( \dot{x} - \dot{y} = \frac{x - y}{|x - y|} \) and \( Q(x) = \lambda_0 I_d + \Sigma \Sigma^*(x) \).

The right-hand side is estimated through Lemma 3.3 for the local diffusion and Lemma 3.4 for the
nonlocal part. Therefore, letting $n \to \infty$, we obtain
\[
\frac{\epsilon}{(I - t)^2} - \omega e^{-\omega t} (K[\varphi(x) + \varphi(y)] + \psi(|x - y|)) + e[L^b + \Phi(y)]
+ e^{-\omega t} \left\{ \sqrt{(b(t, x) - b(t, y)) \cdot \frac{x - y}{x - y}} \right\} \psi'(|x - y|)
+ e^{-\omega t} \left\{ \left\{ K \left[ \sqrt{d} (\varphi(x) + \varphi(y)) + b(t, y) \cdot (K D \varphi(y) + \epsilon D \Phi(y)) \right] \right\}
\leq e^{-\omega t} \left\{ 4 \lambda_0 \psi''(|x - y|) \right\} + \frac{\psi'(|x - y|)}{2} \| \Sigma(x) - \Omega(y) \| ^2 \}
+ e^{-\omega t} \left\{ \text{tr}(Q(x) (K D^2 \varphi + \epsilon D^2 \Phi(y)) + \text{tr}(Q(y) (K D^2 \varphi + \epsilon D^2 \Phi(y))) \right\}
+ e^{-\omega t} K \left\{ I(x, \varphi(x), D \varphi(x)) + I(y, \varphi(y), D \varphi(y)) \right\} + e^{-\omega t} e \left\{ I(x, \Phi(x), D \Phi(x)) + I(y, \Phi(y), D \Phi(y)) \right\}
+ e^{-\omega t} 4 \lambda \int_0^1 (1 - s) \int_B \psi''(|x - y| + 2s \frac{x - y}{z}) |x - y| \cdot z |z| z \cdot d \xi \int_0^s \right\}
\]
where
\[
B := \left\{ z \in \mathbb{R}^d : |z| < \left( 1 + \frac{|x - y|}{4} \right) \right\}
\]
Henceforth we set $r = |x - y|$, we use (2.2) and we rearrange terms, reconstructing the operator $L^b$ on the auxiliary functions $\varphi, \Phi$. Then we get
\[
K \left\{ L^b[\varphi(x) - \varphi(y) + L^b[\varphi(x) - \varphi(y)] \right\} + \epsilon (L^b[\varphi(x) - \varphi(y) + L^b[\varphi(x) - \varphi(y)]]
+ \left\{ \left\{ (b(t, x) - b(t, y)) \cdot \frac{x - y}{x - y} \right\} \right\} \psi'(|x - y|) < 4 \lambda_0 \psi''(|x - y|) + (2 \sigma_0 \land \sigma_1) \psi'(|x - y|) + \omega \psi(|x - y|)
+ 4 \lambda \int_0^1 (1 - s) \int_B \psi''(|x - y| + 2s \frac{x - y}{z}) |x - y| \cdot z |z| z \cdot d \xi \int_0^s \right\}
\]
Now we show that the previous inequality leads to a contradiction for suitable choices of $K, C_1, C_2, \omega$.
We first analyze (3.22) for the case that $r = |x - y| \geq 1$. Using (2.13) or (2.14) (they reduce to the same for $|x - y| \geq 1$), and dropping the terms with $\psi''$ (which is negative), we obtain
\[
K \left\{ L^b[\varphi(x) - \varphi(y) + L^b[\varphi(x) - \varphi(y)] \right\} + \epsilon (L^b[\varphi(x) - \varphi(y) + L^b[\varphi(x) - \varphi(y)]]
< (c_0 + 2 \sigma_0 \sigma_1) \psi'(|x - y|) + \omega \psi(|x - y|)
\]
Let us suppose that $r = |x - y| \geq r_1$, for a sufficiently large $r_1 > 1$. This implies that $\max(|x|, |y|) \geq \frac{1}{2}$. Due to Lemma 3.1 choosing $\omega$ sufficiently small (depending only on $\alpha, k$) we have
\[
L^b[\varphi(x) - \varphi(x)] \geq \frac{\alpha k}{2} \varphi(x) \quad \text{for large } |x|\]
\[
L^b[\varphi(x) - \varphi(x)] \geq -k_0 \quad \text{for every } x,
\]
for some $k_0 > 0$, and so is for the function $\Phi$. Notice that $k_0$ also depends on $\alpha, k$ (and possibly on $\sigma, d, \sigma_0$, but this will not be recalled). Therefore, we can choose $r_1$ sufficiently large (so far, only depending on $\alpha, k$) such that
\[
L^b[\varphi(x) - \varphi(y) + L^b[\varphi(x) - \varphi(y)] \geq 1 \quad \forall x, y : |x - y| \geq r_1.
\]
If we also estimate $\psi'(|x - y|) \leq \psi'(r_1)$ for $r \geq r_1$ (by concavity) and $\psi(|x - y|) \leq C_1$, we deduce from $\psi'(|x - y|) \leq \psi'(r_1)$
\[
K = 2 \epsilon_0 - (c_0 + 2 \sigma_0 \sigma_1) \psi'(|x - y|) - \omega C_1 < 0 \quad \forall r \geq r_1.
\]
We now fix
\[
K := \omega C_1 + (c_0 + 2 \sigma_0 \sigma_1) \psi'(|x - y|) + 2 \epsilon k_0 + [u_0]_{\omega}^0
\]
and we see that inequality (3.22) cannot hold. Therefore, we are left with the possibility that $r < r_1$. In this case we go back to (3.22) and we need to distinguish two cases according to the diffusivity of the nonlocal term $I$. 17
(A) Case $\lambda_0 > 0$ or (2.4) with $\sigma \in (1,2)$. Here we use condition (2.4), and we deduce from (3.22) (using also $L^b[\varphi] - \omega \varphi \geq -k_0$, $L^b[\Phi] - \omega \Phi \geq -k_0$)

$$
-(2Kk_0 + 2\varepsilon k_0) < 4\lambda_0 \psi''(r) + (c_0 + 2\sigma_0 \sigma_1)\psi'(r) + \omega \psi(r)
+ 4\lambda \int_0^1 (1 - s) \int_B \psi''(r + 2s(\widehat{x - y \cdot z})) |\widehat{x - y \cdot z}|^2 \frac{dz}{|z|^{d+\sigma}} ds.
$$

(3.26)

We observe that, using once more the concavity of $\psi$, we have $\psi'(r_1) \leq \psi'(r)$ for $r < r_1$, so the choice of $K$ in (3.25) implies

$$
K \leq \omega C_1 + 2\varepsilon k_0 + [u_0(x)]^s + (c_0 + 2\sigma_0 \sigma_1)\psi'(r) \quad \forall r < r_1.
$$

Therefore (3.26) yields

$$
-(2k_0(\omega C_1 + [u_0(x)]^s) + 2\varepsilon k_0(1 + 2k_0)) < 4\lambda_0 \psi''(r) + (1 + 2k_0)(c_0 + 2\sigma_0 \sigma_1)\psi'(r) + \omega \psi(r)
+ 4\lambda \int_0^1 (1 - s) \int_B \psi''(r + 2s(\widehat{x - y \cdot z})) |\widehat{x - y \cdot z}|^2 \frac{dz}{|z|^{d+\sigma}} ds.
$$

(3.27)

We point out that, if $\lambda_0 > 0$, we could conclude from here using only the local term $\psi''(r)$ (this would be similar as in (40)). So here the main novelty is in the use of nonlocal diffusion, neglecting the term with $\lambda_0$. Hence, using also $\psi \leq C_1$, we reduce (3.27) to the following:

$$
-(1 + 2k_0)\omega C_1 + 2\varepsilon k_0[1 + 2k_0]) < (1 + 2k_0)(c_0 + 2\sigma_0 \sigma_1)\psi'(r)
+ 4\lambda \int_0^1 (1 - s) \int_B \psi''(r + 2s(\widehat{x - y \cdot z})) |\widehat{x - y \cdot z}|^2 \frac{dz}{|z|^{d+\sigma}} ds.
$$

(3.28)

Now we estimate the nonlocal term as follows:

$$
\int_0^1 (1 - s) \int_B \psi''(r + 2s(\widehat{x - y \cdot z})) |\widehat{x - y \cdot z}|^2 \frac{dz}{|z|^{d+\sigma}} ds \leq \int_0^1 (1 - s) \int_B \psi''(r + 2s(\widehat{x - y \cdot z})) |\widehat{x - y \cdot z}|^2 \frac{dz}{|z|^{d+\sigma}} ds
$$

where

$$
B^- := \{ z \in B : (\widehat{x - y \cdot z}) \leq 0 \}.
$$

(3.29)

Setting $e := \widehat{x - y}$, and using polar coordinates, we observe that

$$
\int_{B^-} |e \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} \geq C_{d,\sigma} (r \wedge 1)^{2-\sigma} > 0.
$$

Hence we also estimate, on account of the concavity of $\psi$,

$$
\psi'(r) \leq \frac{2}{C_{d,\sigma}} \int_0^1 (1 - s) \int_B \frac{\psi'(r)}{(r \wedge 1)^{2-\sigma}} |e \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} ds
$$

(3.30)

$$
\leq \frac{2}{C_{d,\sigma}} \int_0^1 (1 - s) \int_B \frac{\psi'(r + 2s(e \cdot z))}{(r + 2s(e \cdot z))^2} |e \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} ds
$$

where we also used, in the last step, that for $z \in B^-$ we have $0 < r + 2s(e \cdot z) \leq r$.

Using (3.30) in (3.28), we obtain

$$
-(1 + 2k_0)\omega C_1 + 2k_0[1 + 2k_0]) \leq \int_0^1 (1 - s) \int_B \frac{4\lambda \psi''(\xi_{s,z}) + C}{(\xi_{s,z} \wedge 1)^{2-\sigma}} |e \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} ds
$$

(3.31)

where $\xi_{s,z} = r + 2s(e \cdot z)$, and $C = \frac{2}{C_{d,\sigma}}(1 + 2k_0)(c_0 + 2\sigma_0 \sigma_1)$. 

18
Recall that for $z \in B^-$ we have $r/2 \leq \xi_{s,z} \leq r$. In particular, we have $\xi_{s,z} \leq r_1$ whenever $r < r_1$. Using the definition of $\psi$ we compute
\[
4\lambda \psi' (\xi) + C \frac{\psi'(\xi)}{(\xi \wedge 1)^{2-\sigma}} = -C_1 \theta C_2 e^{-C_2 \xi^\theta} \xi^{\theta-2} \left\{ 4\lambda(1-\theta) + 4\lambda \theta C_2 \xi^\theta - C(\xi^{\sigma-1} \vee \xi) \right\}.
\]
Now we choose $\theta < \sigma - 1$, and we fix $C_2$ sufficiently large so that
\[
4\lambda \theta C_2 \xi^\theta - C(\xi^{\sigma-1} \vee \xi) \geq 0 \quad \forall \xi \in (0,r_1).
\]
With this choice of $C_2$ (which only depends on $\lambda, \theta, \sigma, C, r_1$), we finally estimate
\[
\forall \xi \leq r \leq r_1, \quad 4\lambda \psi''(\xi) + C \frac{\psi'(\xi)}{(\xi \wedge 1)^{2-\sigma}} \leq -C_1 \theta C_2 e^{-C_2 \xi^\theta} \xi^{\theta-2} 4\lambda(1-\theta) \leq -C_1 L r^{\theta-2} \quad \forall r < r_1
\]
for some constant $L > 0$ independent of $C_1$. Then from (3.31) (and using again (3.29)) we deduce
\[
C_1 (c_{d,\sigma} L r_1^{\theta-\sigma} - \omega(1+2k_0)) - 2k_0[u_0]_{(x)^k} - 2\varepsilon k_0(1+2k_0) \leq 0.
\]
Choosing $\omega$ sufficiently small and $C_3$ suitably large ($C_3 \gtrsim n([u_0]_{(x)^k} + \varepsilon k_0)$ for some large $n$) we get a contradiction. This shows that (3.21) holds true; then letting $\varepsilon \to 0$ and recalling our choices of the parameters $K, C_1, C_2$, we proved that
\[
u(t, x) - u(t, y) \leq e^{-\omega t} \left\{ K (\varphi(x) + \varphi(y)) + \psi(|x-y|) \right\} \leq e^{-\omega t} (K + C_1) (\varphi(x) + \varphi(y)) \leq e^{-\omega t} K [u_0]_{(x)^k} (\varphi(x) + \varphi(y))
\]
for some constant $\bar{K}$, which yields (3.19).

(B) Case $\lambda = 0$ and (2.4) with $\sigma \in (0,1]$. In this case we use (2.14) in (3.22) and, rather than (3.21), we deduce:
\[
- (2Kk_0 + 2\varepsilon k_0) \leq (c_0(r \wedge 1)^{1-\sigma+\delta} + \sigma_1(\sigma_1 r \wedge 2\sigma_0)) \psi'(r) + \omega \psi(r)
+ 4\lambda \int_0^1 (1-s) \int_B \psi''(r + 2s(x-y \cdot z)) |x-y \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} ds.
\]
(3.32)

Similarly as in (3.30), setting $e := x-y$ we estimate
\[
(c_0(r \wedge 1)^{1-\sigma+\delta} + \sigma_1(\sigma_1 r \wedge 2\sigma_0)) \psi'(r) \leq C \int_0^1 (1-s) \int_{B_r} \frac{\psi'(r + 2s(e \cdot z))}{|r + 2s(e \cdot z)|^{1-\delta}} |e \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} ds.
\]
for some constant $C$ only depending on $d, \sigma, \sigma_1, c_0$. From (3.32) we get, using also $\psi \leq C_1$,
\[
- (2Kk_0 + 2\varepsilon k_0 + \omega C_1) \leq \int_0^1 (1-s) \int_{B_r} \left\{ 4\lambda \psi''(\xi_{s,z}) + C \frac{\psi'(\xi_{s,z})}{(\xi_{s,z} \wedge 1)^{1-\delta}} \right\} |e \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} ds
\]
(3.33)

where $\xi_{s,z} = r + 2s(x-y \cdot z)$. Now we compute
\[
4\lambda \psi''(\xi) + C \frac{\psi'(\xi)}{(\xi \wedge 1)^{2-\sigma}} = -C_1 \theta C_2 e^{-C_2 \xi^\theta} \xi^{\theta-2} \left\{ 4\lambda(1-\theta) + 4\lambda \theta C_2 \xi^\theta - C(\xi^{\sigma-1} \vee \xi) \right\}.
\]
and we fix $C_2$ sufficiently large so that $4\lambda \theta C_2 \xi^\theta - C(\xi^{\sigma-1} \vee \xi) \geq 0$ for every $\xi \in (0,r_1)$. To this purpose it is enough to choose $\theta < \delta$ and $C_2$ sufficiently large depending on $r_1, C, \lambda, \theta$.

With this choice of $C_2$, we finally estimate
\[
4\lambda \psi''(\xi_{s,z}) + C \frac{\psi'(\xi_{s,z})}{(\xi_{s,z} \wedge 1)^{1-\delta}} \leq -4\lambda(1-\theta)C_1 \theta C_2 e^{-C_2 \xi^\theta} \xi^{\theta-2}
\]
\[
\leq -4\lambda(1-\theta)C_1 \theta C_2 e^{-C_2 \xi^\theta} \xi^{\theta-2} \quad \forall r < r_1
\]
since we have \( r/2 \leq \xi, z \leq r < r_1 \). Then from (3.33) we deduce

\[-(2Kk_0 + 2ck_0 + \omega C_1) \leq -c_{d, \sigma} 4\lambda (1 - \theta) C_1 \theta C_2 e^{-\epsilon r_1^{1-\sigma}}.\]

Recall from (3.33), and the definition of \( \psi \), that we have

\[K = \omega C_1 + (c_0 + 2\sigma_0\sigma_1) C_1 C_2 e^{-\epsilon r_1^{1-\sigma}} + 2ck_0 + [u_0]_{(x)^k}\]

so that we can rephrase what we obtained as

\[C_1 \left( 4\lambda (1 - \theta) c_{d, \sigma} r_1^{1-\sigma} - 2k_0 (c_0 + 2\sigma_0\sigma_1) C_2 e^{-\epsilon r_1^{1-\sigma}} - \omega (1 + 2k_0) \right) \leq 2k_0 [u_0]_{(x)^k} + 2ck_0 (1 + 2k_0).\]

Suppose now that \( \sigma < 1 \); then we first choose \( r_1 \) sufficiently large so that \( 4\lambda (1 - \theta) c_{d, \sigma} r_1^{1-\sigma} - 2k_0 (c_0 + 2\sigma_0\sigma_1) > 0 \). Next we choose \( C_2 \) as we said, depending on \( r_1 \), and finally, choosing \( \omega \) sufficiently small and \( C_1 \) suitably large \( (C_1 \geq n (1 + 2k_0) ([u_0]_{(x)^k} + 2ck_0)) \) we get a contradiction. This shows that (3.24) holds true; then letting \( \epsilon \to 0 \) and recalling our choices of the parameters \( K, C_1 \), we proved (3.19).

Finally, we are left with the case \( \sigma = 1 \). To handle this case we observe that, with the strategy used above, we have \( 4\lambda \psi''(\xi) + C \frac{\psi'(\xi)}{\xi^{1+\sigma}} \leq 0 \), hence

\[\int_0^1 (1 - s) \int_{B_r} \left\{ 4\lambda \psi''(\xi, z) + C \frac{\psi'(\xi, z)}{\xi, z \wedge 1} \right\} |e \cdot z|^2 \frac{dz}{|z|^{d+\sigma}} ds \leq \int_0^1 (1 - s) \int_{B_r} \left\{ 4\lambda \psi''(\xi, z) + C \frac{\psi'(\xi, z)}{\xi, z \wedge 1} \right\} |e \cdot z|^2 \frac{dz}{|z|^{d+\sigma-\eta}} ds \]

for any \( \eta > 0 \). This allows us to follow the same steps as before, replacing \( r_1^{1-\sigma} \) with \( r_1^q \) in (3.34).

**Remark 3.7** We point out that the conditions assumed on the drift \( b(t, x) \) can be slightly relaxed, at the expense of a mild restriction on the power \( (x)^k \) which is used in the weighted seminorm. To be precise, we can replace assumptions (2.13) with

\[\exists c_0 > 0 : \quad (b(t, x) - b(t, y)) \cdot (x - y) \geq -c_0 |x - y|(1 \vee |x - y|)^q \quad \forall x, y \in \mathbb{R}^d, \forall t > 0, \quad (3.35)\]

for some \( c_0 > 0, q > 0 \), and respectively (2.12) with

\[\sigma \in (0, 1) \text{ and } \exists \delta \in (0, 1) : \quad (b(t, x) - b(t, y)) \cdot (x - y) \geq -c_0 |x - y|(1 \wedge |x - y|)^{1-\sigma+\delta} (1 \vee |x - y|)^q. \quad (3.36)\]

Then the conclusion of Theorems 3.5 and Theorem 3.6 remain valid for any \( k \) such that \( \gamma + k - 2 \geq q \) (notice that the above Theorems correspond to \( q = 0 \), where any \( k > 0, \gamma \geq 2 \) satisfy this condition).

To provide with this generalization, the only difference in the proof occurs for large values of \( |x - y| \). This is a straightforward variation, since using

\[L[\varphi](x) + b(t, x) \cdot D\varphi(x) - \omega \varphi(x) \geq \frac{\alpha k}{2} |x|^{\gamma - 2} \varphi(x) \quad \text{for large } |x|\]

one can change (3.24) into the following estimate

\[K \frac{\alpha k}{2} \left( \frac{r}{2} \right)^{\gamma - 2 + k} - 2ck_0 - (c_0 r^q + 2\sigma_0\sigma_1) \psi'(r_1) - \omega C_1 < 0 \quad \forall r \geq r_1.\]

which, whenever \( \gamma + k - 2 \geq q \), still implies a choice of \( K \) of the same order as (3.25).

This remark shows that the classical condition which is mostly used in the probabilistic approach, say

\[(b(t, x) - b(t, y)) \cdot (x - y) \geq -c_0 |x - y|^2 \quad \forall x, y \in \mathbb{R}^d, \forall t > 0\]

is not only allowed in our setting, but can be improved e.g. into

\[(b(t, x) - b(t, y)) \cdot (x - y) \geq -c_0 (|x - y| \vee |x - y|^2) \quad \forall x, y \in \mathbb{R}^d, \forall t > 0.\]
Remark 3.8 An inspection of the proof of Theorem 3.8 and 3.9 may give some information on how $K, \omega$ in \eqref{e:K1} depend on the constants appearing in assumptions \eqref{e:2.2}, \eqref{e:2.4}, \eqref{e:2.12}-\eqref{e:2.14}, as well as on other relevant parameters such as the dimension $d$. In fact, an explicit choice of $K$ is provided by \eqref{e:2.20} and a quantitative estimate of $K$, as well as of $\omega$, will mostly depend on the parameter $r_1$ and on the choice of the function $\psi(\cdot)$.

Notice that the parameter $r_1$ only relies on the Lyapunov function $\varphi$, and possibly depends on the dimension $d$ (in the model case with Ornstein-Uhlenbeck drift $b(x) = x$, we have $r_1 = O(\sqrt{d})$). However, the explicit choice of $\psi$ made in \eqref{e:2.20} may not lead to optimal estimates of $\omega$ (neither of $K$, in turn). For example, in case that $L$ is the Laplace operator and $b = x$ is the Ornstein-Uhlenbeck drift, a better choice for $\psi$ would be to take the solution of the ODE $4\psi''(r) + \omega \psi(r) + rL = 0$ for $r \in (0, r_1)$, where $L$ is some suitable constant and $\psi(0) = 0, \psi'(r_1) \approx 0$. This would lead to $\omega = O(r_1^{-2}) = O(d^{-1})$, as it is expected for this model case, but the same estimate would not be achieved if $\psi$ is given by \eqref{e:2.20}.

This is to mention that, while we opted for the explicit choice \eqref{e:2.20} in order to give a unifying proof, sufficiently simple, which works in all different settings (local and nonlocal operators, diffusive or sub-diffusive, slowly and strongly confining drifts, etc...), the same approach could yield better bounds if the auxiliary function $\psi$ is built through more sophisticated constructions (which involve possibly auxiliary eigenvalue problems). A similar discussion is done in \cite{20}, where quantitative explicit bounds are obtained using the probabilistic coupling method. This could also be possible here, but at the expense of either restricting to more specific examples or adding an extra source of technical details.

3.3 Sub-exponential rate

We now come at the slowly confining case, where, for Lyapunov functions $\varphi$ of power type, we have $\mathcal{L}^0(\varphi) \to \infty$ but $\frac{\mathcal{L}^0(\varphi)}{\varphi} \to 0$. This occurs when $\gamma < 2$ in \eqref{e:2.12}. In this case one can prove a sub-exponential decay rate (so-called sub-geometrical decay, from discrete models). We start with a general statement of polynomial decay of the weighted oscillation. A similar strategy as in Theorem 3.9 will work, but provided the weight is sufficiently strong. As in Section 3.2 we give first a separate statement for the purely local case: indeed, if the diffusion is local the full range of $\gamma \in (0, 2)$ can be exploited and we have estimates even in the regime of degenerate confining, say when $\gamma \in (0, 1)$ in \eqref{e:2.12}.

Theorem 3.9 Let $\mathcal{L}_0$ be defined by \eqref{e:2.1} where $\Sigma$ satisfies \eqref{e:2.2} and $\lambda_0 > 0$. Assume that $b$ satisfies \eqref{e:2.12}, \eqref{e:2.13} with $\gamma \in (0, 2)$. Let $u$ be a (viscosity) solution of

$$
\begin{align*}
\begin{cases}
\partial_t u + \mathcal{L}_0[u] + b(t,x) \cdot Du = 0 & \text{in } (0,T) \times \mathbb{R}^d, \\
u(0,x) = u_0(x), & \text{in } \mathbb{R}^d
\end{cases}
\end{align*}
$$

\number{3.37}

such that $|u(t,x)| \leq C(1 + |x|)^m$ for some $m > 0$ and $C > 0$ (uniformly in time).

Then, for any $k > 2 - \gamma$ and $u_0 \in L^\infty((x)^{-k})$, and for any $k > k_0$, there exists a positive constant $M > 0$ (depending on $\alpha, \gamma, c_0, d, \lambda_0, \sigma_0, \sigma_1, k, k_0$) such that $u$ satisfies

$$
|u(t)|_{(\varphi)^k} \leq M (1 + t)^{-\gamma} |u_0|_{(\varphi)^k} \text{ where } q = \frac{k + k_0}{2 - \gamma}.
$$

\number{3.38}

For nonlocal diffusions we have a similar result, although the presence of the jump part (in particular, the behavior of the kernel at infinity) restricts the result to the case that $1 < \gamma < 2$, as observed in Remark 3.2.

Theorem 3.10 Let $\mathcal{L}$ be defined by \eqref{e:2.1} where $\Sigma$ satisfies \eqref{e:2.2}, the Levy measure in $\mathcal{I}$ satisfies \eqref{e:2.4}, and $\lambda_0 + \lambda > 0$. Assume that $b$ satisfies \eqref{e:2.13} with $\gamma \in (2 - \sigma, 2)$ such that $\gamma > 1$, and that either of the following conditions hold:

(i) $\lambda_0 > 0$ or $\sigma \in (1, 2)$, and \eqref{e:2.13} holds true.

(ii) $\lambda_0 = 0$ and $\sigma \in (0, 1]$, and \eqref{e:2.13} holds true.

Let $u$ be a (viscosity) solution of \eqref{e:2.1} such that $|u(t,x)| \leq C(1 + |x|)^m$ for some $m \in (0, \sigma)$ and $C > 0$ (uniformly in time).
Then, for any $k \in (0, \sigma)$ with $k > 2 - \gamma$ and $u_0 \in L^\infty((x)^{-k})$, and for any $\bar{k} > k$ with $k < \bar{k} < \sigma$, there exists a positive constant $M > 0$ (depending on $\alpha, \gamma, \sigma, (\Lambda + \lambda_0)^{-1}, \sigma_0, \sigma_1, \Lambda, c_0, d, k, \bar{k}$) such that $u$ satisfies

$$[u(t)]_{(x)^k} \leq M (1 + t)^{-q} [u_0]_{(x)^k} \quad \text{where } q = \frac{k - \gamma}{\sigma}.$$

(3.39)

**Proof of Theorem 3.9 and Theorem 3.10.** We give the proof for the nonlocal case, the other being similar, except for the range of $\gamma$ which is allowed in the Lyapunov function (from (3.11)).

Let us set $\varphi(x) = (x)^k (k < \sigma)$, $\Phi(x) = (x)^p$ $(m < p < \sigma)$. With the same notations of Theorem 3.6 and $\psi$ given by (3.20), we aim at proving that

$$u(t, x) - u(t, y) \leq (1 + t)^{-q} (k (x) + \varphi(y)) + \psi(|x - y|) + \varepsilon [\Phi(t, x) + \Phi(t, y)] + \frac{\varepsilon}{T - t}$$

for every $x, y \in \mathbb{R}^d$, $t > 0$, and for every $\varepsilon > 0$ sufficiently small.

As usual, we observe that if (3.40) were not true, then the function

$$u(t, x) = u(t, y) = \frac{\varepsilon}{T - t} + (1 + t)^{-q} [K \varphi(x) + \varphi(y)] + \psi(|x - y|) + \varepsilon [\Phi(t, x) + \Phi(t, y)]$$

admits a point $(t, x, y)$ of global maximum with $x \neq y$, $t < T$ and $t > 0$, provided $K \geq [u_0]_{(x)^k}$.

Then we use Theorem 2.3 writing the viscosity inequalities for $u$, and we proceed as in Theorem 3.6 using Lemma 3.3 and Lemma 3.4. We end up with a similar kind of inequality as (3.22).

$$\frac{\varepsilon}{(T - t)^2} + (1 + t)^{-q} K \left\{L^b[\varphi] (x) - \frac{q}{(1 + t)} \varphi(x) + L^b [\varphi] (y) - \frac{q}{(1 + t)} \varphi(y)\right\}$$

$$+ (1 + t)^{-q} \varphi \left(\frac{b(t, x) - b(t, y)}{1 - s} \cdot x - y\right) \psi(r) - q (1 + t)^{-q} \psi(r)$$

$$\leq (1 + t)^{-q} \left\{4 \lambda_0 \psi''(r) + (2\sigma_0 + \sigma_1 r) \psi'(r)\right\}$$

$$+ (1 + t)^{-q} 4 \lambda \int_0^1 (1 - s) \int_B \psi''(r + 2s(x - y \cdot z)) |x - y \cdot z|^2 dz |z|^{d + \sigma} ds.$$

where $r = |x - y|$. We start by considering the case $q = 0$. As we did in Theorem 3.6, we first suppose that $r = |x - y| \geq r_1$ for a sufficiently large $r_1 > 1$; in this range assumptions (2.13) and (2.14) reduce to the same. Moreover, since one between $|x|, |y|$ is larger than $r_1/2$, and since $k > 2 - \gamma$ implies $L^b[\varphi] \to +\infty$ as $|x| \to \infty$ (from Lemma 3.1), we can choose $r_1 > 1$ such that

$$L^b[\varphi] (x) + L^b[\varphi] (y) \geq \frac{\alpha k}{2} (|x| + |y|)^{k + \gamma - 2} \geq 1.$$

Moreover, there exists $k_0$ such that $L^b[\varphi] \geq -k_0$, and so is for $\Phi$. Therefore, from (3.11) written with $q = 0$ we deduce (using (2.13) or (2.14) and dropping the terms with $\psi''$)

$$K \geq 2k_0 \leq |c_0 + 2\sigma_1 \sigma_1 \psi'(r) \quad \text{if } r \geq r_1 > 1.$$

Since, by concavity, $\psi'(r) \leq \psi'(r_1)$ for $r \geq r_1$, the above inequality cannot hold if we choose

$$K = 2k_0 + (c_0 + 2\sigma_0 \sigma_1) \psi'(r_1) + [u_0]_{(x)^k}.$$

On the other hand, for $r < r_1$, just using $L^b[\varphi] \geq -k_0$, $L^b[\varphi] \geq -k_0$, we get

$$\left(\frac{b(t, x) - b(t, y)}{1 - s} \cdot x - y\right) \psi'(r) < 4 \lambda_0 \psi''(r) + (2\sigma_0 + \sigma_1 r) \sigma_1 \psi'(r)$$

$$+ 4 \lambda \int_0^1 (1 - s) \int_B \psi''(r + 2s(x - y \cdot z)|x - y \cdot z|^{2} dz |z|^{d + \sigma} ds + 2K k_0 + 2\varepsilon k_0.$$
Henceforth, using \( \lambda_0 + \lambda > 0 \), we argue as in Theorem 5.6 in order to get at a contradiction with suitable choices of \( C_2 \) (depending on \( r_1 \)) and \( C_1 \geq [u_0]_{(x)k}^\gamma \). This proves (3.30) for the case \( q = 0 \). After letting \( \varepsilon \to 0 \), we deduce that, for any \( k > 2 - \gamma \), we have

\[
u(t,x) - \nu(t,y) \leq K ([x]^k + \langle y \rangle^k) + \psi(|x-y|) \leq K ([x]^k + \langle y \rangle^k) + C_1
\]

which implies (after the choices of \( K, C_1 \))

\[
[u(t)]_{(x)k} \leq C [u_0]_{(x)k} \quad \forall t > 0 ,
\]

(3.42)

where \( C \) is a constant only depending on \( \alpha, \gamma, (\lambda + \lambda_0)^{-1}, c_0, d, \sigma_0, \sigma_1, \sigma, \delta \) and \( k \). In particular, we have proved that, for all weights \( [x]^k \) with \( 2 - \gamma < k < \sigma \), the weighted oscillation of \( u \) at time \( t \) is controlled by the initial one.

Now we improve the estimate, by proving (3.44) with \( \frac{q}{\varepsilon} \). To this purpose, we observe that, on the maximum point \( (t,x,y) \), we have

\[
(1 + t)^{-q}[\varphi(x) + \varphi(y)] \leq u(t,x) - u(t,y) \leq ([x]^k + \langle y \rangle^k)[u(t)]_{(x)k} .
\]

Choosing \( \varphi(x) = [x]^k \), with \( k < \bar{k} < \sigma \), due to (3.42) we estimate

\[
(1 + t)^{-q}K \leq C \frac{[x]^k + \langle y \rangle^k}{[x]^k + \langle y \rangle^k} [u_0]_{(x)k}
\]

and choosing \( q = \frac{\bar{k} - k}{2 - \gamma} \) we get

\[
\frac{1}{1 + t} \leq C \left( \frac{[u_0]_{(x)k}}{K} \right) \frac{1}{[x]^k + \langle y \rangle^k}
\]

(3.43)

for a possibly different \( C \) depending on the same constants as before, as well as on \( q \). Thanks to (3.43), there exists a sufficiently large constant \( L \) such that if we choose \( K > L [u_0]_{(x)k} \) then we have

\[
\mathcal{L}^b[\varphi](x) - \frac{q}{1 + t} \varphi(x) \to +\infty \quad \text{as } |x| \to \infty ,
\]

where we used Lemma 4.1 (and the fact that \( \bar{k} > 2 - \gamma \)). In particular we have \( \mathcal{L}^b[\varphi](x) - \frac{q}{1 + t} \varphi(x) \geq -k_0 \) for some \( k_0 > 0 \). Similarly we may reason for \( \Phi(x) = [x]^p \), where we can assume, without loss of generality, that \( p > 2 - \gamma \). Therefore, we can choose \( r_1 \) sufficiently large so that, whenever \( r = |x - y| \geq r_1 \) we have

\[
\mathcal{L}^b[\varphi](x) - \frac{q}{1 + t} \varphi(x) + \mathcal{L}^b[\varphi](y) - \frac{q}{1 + t} \varphi(y) \geq 1 \quad \text{and} \quad \mathcal{L}^b[\Phi](x) - \frac{q}{1 + t} \Phi(x) \geq -k_0 .
\]

Hence we deduce from (3.31)

\[
K < 2 \varepsilon k_0 + [c_0 + 2 \sigma_0 \sigma_1] \psi'(r) + \psi(r) \frac{q}{1 + t}
\]

\[
< 2 \varepsilon k_0 + [c_0 + 2 \sigma_0 \sigma_1] \psi'(r_1) + C_1 \left( \frac{[u_0]_{(x)k}}{K} \right)^\frac{1}{\gamma} \quad \forall r \geq r_1 ,
\]

(3.44)

where we used estimate (3.30), and that \( \psi(r) \leq C_1 , \psi'(r) \leq \psi'(r_1) \) for \( r > r_1 \). Here \( C \) is a generic constant independent of \( r_1 , K , C_1 , C_2 , u_0 \). Henceforth, we fix

\[
C_1 := n [u_0]_{(x)k} + n^{\frac{1}{\gamma}} \varepsilon k_0 (1 + 4 k_0)
\]

(3.45)

where \( n \) is a sufficiently large number to be chosen later. Moreover, we also fix

\[
K := L (1 + n^{\frac{1}{\gamma}}) [u_0]_{(x)k} + 4 \varepsilon k_0 + [c_0 + 2 \sigma_0 \sigma_1] \psi'(r_1)
\]

(3.46)
where $L$ is a sufficiently large constant. With this choice of $C_1, K$, we have

$$C_1 C \left( \frac{\|u_0(x)\|_K}{K} \right)^\frac{1}{\lambda} \leq C_1 C \left( \frac{1}{L (1 + n^{\frac{\lambda}{n+\lambda}})} \right)^\frac{1}{\lambda} \leq C \left( \frac{n^{\frac{\lambda}{n+\lambda}}[u_0(x) + \varepsilon k_0(1 + 4k_0)]}{L} \right) \right.$$ (3.47)

and (3.34) implies

\[ K \leq 2\varepsilon k_0 + \left[ c_0 + 2\sigma_0\sigma_1 \right] \psi'(r_1) + \frac{C}{L} \left( \frac{n^{\frac{\lambda}{n+\lambda}}[u_0(x) + \varepsilon k_0(1 + 4k_0)]}{L} \right) \text{.} \]

On account of (3.46), this inequality cannot hold if $L$ is chosen suitably large (only depending on the constant $C$ and $k_0$).

We are left to analyse the case that $r < r_1$, which we split in the two ranges of $\sigma$.

**(A) Case** $\lambda_0 > 0$ or (2.4) with $\sigma \in (1, 2)$.

In this case we use (2.13) and we deduce from (3.41)

\[-(2Kk_0 + 2\varepsilon k_0) \leq 4\lambda_0 \psi''(r) + (c_0 + (2\sigma_0 \wedge \sigma_1 r)\sigma_1) \psi'(r) + \frac{q}{(1 + t)} \psi(r) \]

and (3.48)

\[-(2k_0(1 + 4k_0) + 2k_0 L (1 + n^{\frac{\lambda}{n+\lambda}}))[u_0(x) + \varepsilon k_0(1 + 4k_0)] \leq (1 + 2k_0)[c_0 + 2\sigma_0\sigma_1 \psi'(r) + \frac{q}{(1 + t)} \psi(r) \]

+ $4\lambda \int_0^1 (1 - s) \int_B \psi''(r + 2s(x - y \cdot z))(x - y \cdot z)^2 \frac{dz}{|z|^{\lambda+\sigma}} ds \text{, } \forall r < r_1.$ (3.49)

Now we observe that (3.48) and (3.47), together with $\psi(r) \leq C_1$, imply

\[ \frac{q}{1 + t} \psi \leq q C_1 \left( \frac{\|u_0(x)\|_K}{K} \right)^\frac{1}{\lambda} \leq q C \left( \frac{n^{\frac{\lambda}{n+\lambda}}[u_0(x) + \varepsilon k_0(1 + 4k_0)]}{L} \right) \leq 2k_0 L (1 + n^{\frac{\lambda}{n+\lambda}})[u_0(x) + 2\varepsilon k_0(1 + 4k_0)] \]

where last inequality follows by choosing $L$ large enough. Using this estimate, we deduce from (3.49)

\[-4\varepsilon k_0(1 + 4k_0) - 4k_0 L (1 + n^{\frac{\lambda}{n+\lambda}})[u_0(x) + \varepsilon k_0(1 + 4k_0)] \leq (1 + 2k_0)[c_0 + 2\sigma_0\sigma_1 \psi'(r) \]

+ $4\lambda \int_0^1 (1 - s) \int_B \psi''(r + 2s(x - y \cdot z))(x - y \cdot z)^2 \frac{dz}{|z|^{\lambda+\sigma}} ds \text{, } \forall r < r_1.$ (3.50)

We handle the right-hand side exactly as in Theorem 4.6 In particular, with a convenient choice of $C_2$ we estimate

\[ (1 + 2k_0)[c_0 + 2\sigma_0\sigma_1 \psi'(r) + 4\lambda \int_0^1 (1 - s) \int_B \psi''(r + 2s(x - y \cdot z))(x - y \cdot z)^2 \frac{dz}{|z|^{\lambda+\sigma}} ds \leq -\tilde{C} C_1 r_1^{\sigma-\sigma} \]

\[ \forall r < r_1 \]

for some constant $\tilde{C}$ independent of $C_1, K$. Inserting this inequality in (3.50), we obtain

\[-4\varepsilon k_0(1 + 4k_0) - 4k_0 L (1 + n^{\frac{\lambda}{n+\lambda}})[u_0(x) + \varepsilon k_0(1 + 4k_0)] \leq -\tilde{C} C_1 r_1^{\sigma-\sigma} \text{, } \forall r < r_1.$ (3.51)
Recalling the value of $C_1$ from (3.43), we see that the above inequality cannot hold, if we finally choose $n$ sufficiently large (depending on all previous parameters).

Finally, we proved that (5.30) holds true, where letting $ε → 0$ and recalling (5.35) – (5.40), we deduce that

$$u(t, x) - u(t, y) ≤ (1 + t)^{−q} \{ K (φ(x) + φ(y)) + C_1 \} ≤ (1 + t)^{−q} M \{ ω_0(x, y) + |x|^k + |y|^k \} \quad ∀ x, y ∈ \mathbb{R}^d, \; t > 0.$$  \hspace{2em} (3.52)

Hence (3.50) is proved.

**B. Case** $λ_0 = 0$ and (2.4) with $σ \in (0, 1)$. If $σ \in (0, 1)$, we use (2.14) and we get from (3.41)

$$- (2K k_0 + 2ε k_0) ≤ (c_0 (r + 1)^{1−σ+δ} + (2σ_0 \land σ_1 r) σ_1)ψ'(r) + \frac{q}{1 + t} \psi(r) + 4λ \int_0^1 (1 - s) \int_B ψ''(r + 2s(x - y \cdot z)) |x - y \cdot z|^2 \frac{dz}{|z|^{d+σ}} ds.$$  \hspace{2em} (3.53)

Here we recall the explicit choice of $K$ (3.16), which (using the explicit value of $ψ'$) reads as

$$K = L (1 + n \sqrt{r}) |ω_0(x)|^{k} + 4ε k_0 + [c_0 + 2σ_0 σ_1] C_2 e^{-C_2 r_1^θ - 1}.$$  \hspace{2em} (3.54)

Moreover, we estimate as before $\frac{q}{1 + t} \psi$ using (3.43) and (3.47), and we obtain from (3.55)

$$\theta C_2 e^{-C_2 r_1^θ - 1} \left[ 4λ(1 - \theta) c_{d,σ} r_1^{1−σ} - 2k_0 (c_0 + 2σ_0 σ_1) \right] ≤ \frac{C}{L^+(1 + n \sqrt{r})} \quad (3.56)$$

for some $\tilde{C}$ independent of $C_1, C_2, r_1$. Now we observe that $r_1$ can be chosen sufficiently large so that

$$4λ(1 - \theta) c_{d,σ} r_1^{1−σ} - 2k_0 (c_0 + 2σ_0 σ_1) > 0.$$  \hspace{2em} (3.57)

Finally, having fixed $r_1$, and then $C_2$ accordingly, and recalling the choice of $C_1$ from (3.43), we see that (3.40) cannot hold for $n$ sufficiently large. This concludes the argument and proves that (3.40) holds true. From that, letting $ε → 0$, we deduce (3.52), and then (3.59).

Finally, we only mention that the case $σ = 1$ is dealt with as in Theorem 3.6 obtaining a slight variation of (3.54) where $1 - σ$ is replaced by $η$ for some small $η > 0$. This is enough to conclude exactly as before.

**Remark 3.11** We point out that the same generalization as mentioned in Remark 3.7 on the growth condition of the drift also applies for the content of Theorem 3.6.

### 4 Decay of Fokker-Planck equations

We deduce now the decay estimates for the general Fokker-Planck equation (2.10). Let us recall that, in the setting of our assumptions, problem (2.10) admits a unique solution, according to Theorem 2.6.

As a first step, we recall the following characterization of weighted seminorms. The following lemma is taken from [27], we give a slightly different statement for our purposes (and a proof for the reader’s convenience).
Lemma 4.1 Let $\phi$ be a continuous function such that $\phi(x) \geq 1$ for all $x \in \mathbb{R}^d$. For any function $u$ such that $u \in L^\infty(\phi^{-1})$, we have
\[
\sup_{x,y \in \mathbb{R}^d} \frac{|u(x) - u(y)|}{\phi(x) + \phi(y)} = \inf_{c \in \mathbb{R}} \|u + c\|_{L^\infty(\phi^{-1})}
\]

Proof. Let us call $M := \sup_{x,y \in \mathbb{R}^d} \frac{|u(x) - u(y)|}{\phi(x) + \phi(y)}$ and $m := \inf_{c \in \mathbb{R}} \|u + c\|_{L^\infty(\phi^{-1})}$. Since
\[
|u(x) - u(y)| \leq |u(x) + c| + |u(y) + c| \\
\leq \phi(x)\|u + c\|_{L^\infty(\phi^{-1})} + \phi(y)\|u + c\|_{L^\infty(\phi^{-1})}
\]
we immediately deduce that $M \leq m$.

For the reverse inequality, define $c := \inf_{x,y \in \mathbb{R}^d} [M\phi - u]$. Notice that $c$ is finite because $M\phi(x) - u(x) \geq -M\phi(y) - u(y)$ for any $y \in \mathbb{R}^d$. By definition, we have
\[
\frac{u(x) + c}{\phi(x)} \leq M.
\]
We also estimate, using the definition of $M$ and $c$,
\[
\frac{u(x) + c}{\phi(x)} = \inf_{y \in \mathbb{R}^d} \left[ \frac{u(x) + M\phi(y) - u(y)}{\phi(x)} \right] \\
\geq \inf_{y \in \mathbb{R}^d} \left[ M\phi(y) - M\left( \frac{\phi(x) + \phi(y)}{\phi(x)} \right) \right] = -M.
\]
Hence we deduce that $\|u + c\|_{L^\infty(\phi^{-1})} \leq M$, which yields $m \leq M$. \qed

Thanks to Lemma 4.1, the decay estimates follow by duality using the results of the previous Section.

Theorem 4.2 Let $\mathcal{L}$ be defined by (2.7) where $\Sigma$ satisfies (2.2), the Levy measure in $\mathcal{I}$ satisfies (2.4), and $\lambda_0 + \lambda > 0$. Assume that $b \in C^0(Q_\mathcal{T})$ satisfies (2.12) and either of the following conditions:

(i) $\lambda_0 > 0$ or $\sigma \in (1, 2)$, and $\Sigma$ holds true.

(ii) $\lambda_0 = 0$, $\sigma \in (0, 1]$ and $\Sigma$ holds true.

Assume that $m_0 \in \mathcal{M}_k(\mathbb{R}^d)$, for some $0 < k < \sigma$, and that $\int_{\mathbb{R}^d} dm_0 = 0$, and let $m$ be the unique solution of (2.17) (in the sense of Definition 2.7). Then we have:

(a) If $\gamma \geq 2$ in (2.17), then
\[
\|m(t)\|_{\mathcal{M}_k} \leq K e^{-\omega t} \|m_0\|_{\mathcal{M}_k}
\]
for some constant $K, \omega$ only depending on $\alpha, \sigma, (\lambda_0 + \lambda)^{-1}, \sigma_0, \sigma_1, \Lambda, \delta, d, c_0, k$.

(b) If $\gamma \in \left( \max(2(2 - \sigma), 1), 2 \right)$ in (2.17), then for any $k < \tilde{k} < \sigma$ such that $m_0 \in \mathcal{M}_{\tilde{k}}$ we have
\[
\|m(t)\|_{\mathcal{M}_{\tilde{k}}} \leq K (1 + t)^{-q} \|m_0\|_{\mathcal{M}_{\tilde{k}}} \quad \text{where } q = \frac{k + \tilde{k}}{2 - \gamma}
\]
for some $K$ depending on $\alpha, \gamma, \sigma, (\lambda_0 + \lambda)^{-1}, \sigma_0, \sigma_1, \Lambda, \delta, d, c_0, k, \tilde{k}$.

Proof. For $t > 0$ we consider the unique viscosity solution of the adjoint equation
\[
\begin{align*}
-\partial_t u + \mathcal{L}[u] + b(t, x) \cdot Du &= 0 \quad \text{in } (0, t) \times \mathbb{R}^d, \\
u(t, x) &= \xi, \quad \text{in } \mathbb{R}^d
\end{align*}
\]
where $\xi \in C(\mathbb{R}^d) \cap L^\infty((x)^{-k})$. By Definition 2.5 $m$ satisfies
\[
\int_{\mathbb{R}^d} \xi dm(t) = \int_{\mathbb{R}^d} u(0) dm_0 = \int_{\mathbb{R}^d} (u(0) + c) dm_0 \\
\leq \|m_0\|_{\mathcal{M}_k} \|u(0) + c\|_{L^\infty((x)^{-k})}
\]

26
Theorem 4.4

If \( \gamma \geq 2 \), we estimate \([u(0)]_{(x)^k}\) from Theorem 3.6 and we obtain
\[
\int_{\mathbb{R}^d} \xi \, dm(t) \leq K e^{-\omega t} [m_0]_{(x)^k} [\xi]_{(x)^k}^k \\
\leq K e^{-\omega t} [m_0]_{(x)^k} \| \xi \|_{L^\infty((x)^{-k})}.
\]

Since \( \xi \) is arbitrary, we deduce (4.1).

If \( \gamma \in (1, 2) \) with \( \gamma > 2 - \sigma \), we argue similarly but rather than (4.1) we estimate as follows:
\[
\int_{\mathbb{R}^d} \xi \, dm(t) \leq [m_0]_{(x)^k} \| u(0) \|_{(x)^k}
\]

for \( k > k > 2 - \gamma \) (notice that \( u \in L^\infty((x)^{-k}) \) implies \( u \in L^\infty((x)^{-k}) \)). Applying now Theorem 3.10 we obtain
\[
\int_{\mathbb{R}^d} \xi \, dm(t) \leq [m_0]_{(x)^k} K (1 + t)^{-q} \| \xi \|_{L^\infty((x)^{-k})} \quad \text{where } q = \frac{k-k}{2-\gamma}
\]

which yields (4.2).

Remark 4.3

The previous result for Levy operators includes the particular case stated for the fractional Laplacian in the Introduction (Theorem 1.3). We only need to observe that, if the drift \( b \) is locally Lipschitz in \( x \) (or alternatively, under milder conditions involving only \( \text{div}(b) \)), the weak solution is unique (so it coincides with the unique solution given by Theorem 2.7) and belongs to \( L^1 \) (see e.g. [10], [33], [45]).

In the same spirit of the above remark, the solutions are \( L^1 \) whenever the operator is second order. In addition, if the operator is local, one can consider the full range of \( \gamma \), including \( \gamma \in (0, 1) \). The next result extends Theorem 1.1 which was stated in the Introduction for the case of Laplace operator. The proof is the same as above, using duality and Theorem 3.5 or Theorem 3.9. This extends results as in [12] Thm 3], or [31] Thm 1.2.

Theorem 4.4

Let \( m_0 \in L^1((x)^k) \) be such that \( \int_{\mathbb{R}^d} m_0 \, dx = 0 \). Let \( m \) be a weak solution to the Cauchy problem
\[
\begin{align*}
\partial_t m + \mathcal{L}_0^t[m] - \text{div}(b(t,x)m) &= 0 \quad \text{in } (0,T) \times \mathbb{R}^d, \\
m(0,x) &= m_0(x), \quad \text{in } \mathbb{R}^d
\end{align*}
\]

where we assume that \( \lambda_0 > 0 \) and (2.2) holds true, and \( b(t,x) \) is a continuous vector field satisfying (2.13) and (2.14).

Then we have:

(a) If \( \gamma \geq 2 \) in (2.12), then
\[
\| m(t) \|_{L^1((x)^k)} \leq K e^{-\omega t} \| m_0 \|_{L^1((x)^k)}
\]

for some constant \( K, \omega \) only depending on \( \alpha, \lambda_0^{-1}, \sigma_0, \sigma_1, d, c_0, k \).

(b) If \( \gamma \in (0, 2) \) in (2.12), then for any \( k < \bar{k} < \sigma \) such that \( m_0 \in L^1((x)^{\bar{k}}) \) we have
\[
\| m(t) \|_{L^1((x)^k)} \leq K (1 + t)^{-q} \| m_0 \|_{L^1((x)^\bar{k})} \quad \text{where } q = \frac{\bar{k}-k}{2-\gamma}
\]

for some \( K \) depending on \( \alpha, \gamma, \lambda_0^{-1}, \sigma_0, \sigma_1, d, c_0, k, \bar{k} \).
5 Generalizations and further results

We presented so far the results that can be obtained using Lyapunov functions of power type, resulting in moments decay of measure-valued solutions. Here we review our strategy in a more abstract form, showing that it can be applied to different choices of Lyapunov functions. The Lyapunov function will determine the weighted space for the oscillation estimate and, what is more important, the rate of convergence is consequential to the weighted norm which is used.

There are two main settings that we consider, corresponding to exponential or sub-exponential decay. Here and below, \( \varphi \) denotes a generic positive smooth function such that \( \varphi(x) \to \infty \) as \( |x| \to \infty \). There is no loss of generality in assuming that \( \varphi(x) \geq 1 \) in \( \mathbb{R}^d \). We first consider the simplest case of exponential decay in some weighted norm, namely when some \( \varphi \) exists satisfying

\[
\exists \omega_0 > 0 : \liminf_{|x| \to \infty} \frac{\mathcal{L}^b[\varphi]}{\varphi} \geq \omega_0. \tag{H1}
\]

We denote

\[
\Lambda_1 := \{ \varphi \in C^2(\mathbb{R}^d) : \varphi(x)^{|x| \to \infty} \to \infty, \mathcal{I}(x,[\varphi]) \text{ is well defined and } \varphi \text{ satisfies (H1)} \}. \]

Clearly, for every \( \varphi \in \Lambda_1 \), one can choose \( \omega \) sufficiently small so that

\[
\mathcal{L}^b[\varphi] - \omega \varphi \geq \frac{\omega_0}{2} \varphi - k_0 \quad \forall x \in \mathbb{R}^d
\]

for some \( k_0 > 0 \). Using the above property, the proof of Theorem 3.6 follows replacing \( (x)^k \) with \( \varphi(x) \), so that we can generalize the exponential decay estimate for the oscillation, as follows.

**Theorem 5.1** Let \( \mathcal{L} \) be defined by (2.7) where \( \Sigma \) satisfies (2.2), the Levy measure in \( \mathcal{I} \) satisfies (2.4), and \( \lambda_0 + \lambda > 0 \). Assume that \( b \) satisfies (2.12) and either of the following conditions:

(i) \( \lambda_0 > 0 \) or \( \sigma \in (1, 2) \), and (2.13) holds true.

(ii) \( \lambda_0 = 0 \) and \( \sigma \in (0, 1] \), and (2.14) holds true.

Assume that there exists \( \varphi \in \Lambda_1 \). Let \( u \) be a (viscosity) solution of (2.7) such that, for some \( \Phi \in \Lambda_1 \), one has \( |u(t,x)| = o(\Phi) \) as \( |x| \to \infty \), uniformly for \( t \in (0,T) \). Then there exist positive constants \( K, \omega > 0 \) (depending on \( \alpha, \gamma, \sigma, (\lambda + \lambda_0)^{-1}, \sigma_0, \sigma_1, c_0, d, \varphi \)) such that \( u \) satisfies

\[
[u(t)]_\varphi \leq K e^{-\omega t} [u_0]_\varphi. \tag{5.1}
\]

The case of sub-exponential growth is more subtle. This corresponds to the case that the Lyapunov function \( \varphi \) satisfies a condition as follows:

\[
\mathcal{L}^b[\varphi] \underset{|x| \to \infty}{\to} \infty, \quad \text{and } \exists \text{ a decreasing function } h : \mathbb{R}^+ \to \mathbb{R}^+, \text{ with } h(r)^{r \to \infty} 0 : \mathcal{L}^b[\varphi] \geq h(\varphi) \varphi \quad \text{as } |x| \to \infty. \tag{H2}
\]

Correspondingly, we set

\[
\Lambda_2 := \{ \varphi \in C^2(\mathbb{R}^d) : \varphi(x)^{|x| \to \infty} \to \infty, \mathcal{I}(x,[\varphi]) \text{ is well defined and } \varphi \text{ satisfies (H2)} \}. \]

As in the estimate (3.39), the sub-exponential decay is observed only when using different weights in the seminorms of \( u(t) \) and \( u(0) \). One possible choice, among others, is to use different powers of \( \varphi \), as we did in Theorem 5.1.

**Theorem 5.2** Let \( \mathcal{L} \) be defined by (2.7) where \( \Sigma \) satisfies (2.2), the Levy measure in \( \mathcal{I} \) satisfies (2.4), and \( \lambda_0 + \lambda > 0 \). Assume that \( b \) satisfies (2.12) with \( \gamma \in (0, 2) \), and either of the following conditions:

(i) \( \lambda_0 > 0 \) or \( \sigma \in (1, 2) \), and (2.13) holds true.
(ii) \( \lambda_0 = 0 \) and \( \sigma \in (0,1] \), and \( \mathcal{H} \) holds true.

Assume that there exists \( \varphi \) satisfying \( \mathcal{H} \). Let \( u \) be a (viscosity) solution of (2.7) such that \( |u(t,x)| = o(\Phi) \) as \( |x| \to \infty \) (uniformly for \( t \in (0,T) \)) for some \( \Phi \in C^2(\mathbb{R}^d) \) such that \( \Phi(x) \to \infty \), \( L_b[\Phi] \to \infty \) and \( \liminf_{|x| \to \infty} L_b[\Phi] - \kappa h(\varphi) \Phi > -\infty \) for some \( \kappa > 0 \).

Then, for any \( \theta < 1 \) such that \( h(\varphi) \varphi^\theta \to \infty \) as \( |x| \to \infty \), there exist positive constants \( L,M > 0 \) (depending on \( \alpha, \gamma, \sigma, (\lambda_0 + \lambda)^{-1}, \sigma_0, \sigma_1, c_0, d, \varphi, \theta \)) such that \( u \) satisfies

\[
|u(t)| \leq M \varpi(t) |u_0| \psi \quad \text{where } \varpi(t) \text{ is implicitly defined as } \int_{|z| > \varpi(t)^{-1} \varpi} \frac{dz}{\varpi(L \varpi^{-1} \varpi)} = \frac{t}{2}, \tag{5.2}
\]

with the function \( h \) given by \( \mathcal{H} \).

**Remark 5.3** The interesting part of (5.2) is that the decay rate \( \varpi(t) \) is given in terms of the condition \( \mathcal{H} \) involving the Lyapunov function \( \varphi \). Notice that \( \varpi \) is sub-exponential precisely because \( h(r) \to 0 \) as \( r \to \infty \) (\( h \) gives a quantification of the fact that \( L_b[\varphi] \to \infty \)). The implicit condition for \( \varpi(t) \) comes from the requirement that \( \varpi \) satisfies the ODE

\[
\begin{cases}
\varpi' = -\frac{1}{\varpi} h(L \varpi^{-1} \varpi), & t > 0 \\
\varpi(0) = 1
\end{cases} \tag{5.3}
\]

Here, we did not look for sharp conditions in terms of \( \theta \), and the choice of the constant \( L \) is not meant to be optimal, either. Thus, there could be room for sharper conditions. However, the rate function which stems from (5.3) looks consistent with all previous results; in particular, in the range of power functions \( \varphi = (x)^k \), the above statement reduces to Theorem 3.10.

**Remark 5.4** Despite the technical assumption appearing in the statement, the growth of solutions \( u \) plays a minor role in the above result. One could readily use the statement with \( \Phi = \varphi \), and apply the estimate to any bounded solution \( u \). This is how we are going to use this result later.

**Proof.** The proof is a (slightly more abstract) version of Theorem 3.10.

As a first step, we observe that, for all functions \( \varphi \in C^2(\mathbb{R}^d) \) such that \( L_b[\varphi] \to \infty \) as \( |x| \to \infty \), it holds

\[
|u(t)| \leq C |u_0| \varphi
\]

for a suitable constant \( C > 0 \). This is obtained exactly as (3.42) in Theorem 3.10. Now, if \( \theta < 1 \), we have

\[
D^2 \varphi^\theta \geq \theta \varphi^\theta - 1 \left( D^2 \varphi + (\theta - 1) \frac{D\varphi \otimes D\varphi}{\varphi} \right) \leq \theta \varphi^\theta - 1 D^2 \varphi
\]

and, by concavity,

\[
I(x, [\varphi^\theta]) = \int_{|z| > 1} \{ \varphi^\theta(x + z) - \varphi^\theta(x) \} \nu(dz) + \int_{|z| \leq 1} \{ \varphi^\theta(x + z) - \varphi^\theta(x) - \theta \varphi^\theta - 1 (D\varphi(x) \cdot z) \} \nu(dz)
\]

\[
\leq \theta \varphi(x)^\theta - 1 I(x, [\varphi]).
\]

Hence we have

\[
L_b[\varphi^\theta] \geq \theta \varphi^\theta - 1 L_b[\varphi] \geq \theta \varphi^\theta h(\varphi) \to \infty \quad \text{as } |x| \to \infty
\]

by assumption. This implies that \( \varphi^\theta \) is itself a Lyapunov function and we have, by what proved before, \( [u(t)]_{\varphi^\theta} \leq C |u_0|_{\varphi^\theta} \). Now we go back to the maximum point of

\[
u(t,x) - u(t,y) - \varpi(t) \left\{ K[\varphi(x) + \varphi(y)] + \psi(|x - y|) + \varepsilon[\Phi(x) + \Phi(y)] \right\} = \frac{\varepsilon}{\zeta(x,y)}
\]

29
where $K \geq |u_0|_{\varphi^s} \geq |u_0|_{\varphi}$ and $\varpi$ satisfies \((5.3)\). On the maximum point we have, similarly as in \((3.11)\):

$$
\frac{\varepsilon}{\varepsilon (T-t)} + \varpi(t) K \left\{ L^b[\varphi](x) + \frac{\varpi'}{x} \varphi(x) + L^b[\varphi](y) + \frac{\varpi'}{x} \varphi(y) \right\} 
+ \varpi(t) \varepsilon \left( L^b[\Phi](x) + \frac{\varpi'}{x} \Phi(x) + L^b[\Phi](y) + \frac{\varpi'}{x} \Phi(y) \right) 
+ \varpi(t) \left( (b(t,x) - b(t,y)) \cdot x - y \right) \psi'(r) + \varpi(t) \psi(r)
\leq \varpi(t) \{ 4\lambda_0 \psi''(r) + (2\sigma_0 \wedge \sigma_1 r) \sigma_1 \psi'(r) \}
+ \varpi(t) 4\lambda \int_0^1 0 (1-s) \int_B \psi''(r + 2(s(x-y \cdot z))) |x-y \cdot z|^2 \frac{dz}{|z|} ds
$$

where $r = |x-y|$. We consider the case that $r = |x-y| > r_1$, for a sufficiently large $r_1$. In particular, we have $|x| \wedge |y| > \frac{r_1}{2}$. We drop the terms with $\psi''$, we use \((2.13)\) (or \((2.14)\)) for the drift term, and we obtain

$$
K \left\{ L^b[\varphi](x) + \frac{\varpi'}{x} \varphi(x) + L^b[\varphi](y) + \frac{\varpi'}{x} \varphi(y) \right\} + \varepsilon \left( L^b[\Phi](x) + \frac{\varpi'}{x} \Phi(x) + L^b[\Phi](y) + \frac{\varpi'}{x} \Phi(y) \right)
+ \varpi'(r) < (c_0 + 2\sigma_0 \sigma_1) \psi'(r).
$$

Using $\psi \leq C_1$ and $\psi'(r) \leq \psi'(r_1)$ we get

$$
K \left\{ L^b[\varphi](x) + \frac{\varpi'}{x} \varphi(x) + L^b[\varphi](y) + \frac{\varpi'}{x} \varphi(y) \right\} + \varepsilon \left( L^b[\Phi](x) + \frac{\varpi'}{x} \Phi(x) + L^b[\Phi](y) + \frac{\varpi'}{x} \Phi(y) \right)
< (c_0 + 2\sigma_0 \sigma_1) \psi'(r_1) - \frac{\varpi'}{x} C_1.
$$

We observe that, since the maximum is positive, we have

$$
K \varpi(t)[\varphi(x) + \varphi(y)] \leq u(t,x) - u(t,y) \leq |\varphi^s(x) + \varphi^s(y)| |u(t)||_{\varphi^s} \leq C |\varphi^s(x) + \varphi^s(y)| |u_0|_{\varphi^s}
$$

which yields

$$
\varpi(t) \leq C \frac{|u_0|_{\varphi^s}}{K} \varphi(x) + \varphi(y) \leq C \frac{|u_0|_{\varphi^s}}{K} \frac{1}{(\varphi(x) + \varphi(y))^{1-\gamma}}.
$$

In particular, since $h$ is decreasing, we have

$$
h(\varphi) \geq h \left( \frac{C |u_0|_{\varphi^s}}{K} \right)^{1-\gamma} \varphi^{1-\gamma}.
$$

For a constant $L$ to be chosen below, such that $L \geq \left( \frac{C |u_0|_{\varphi^s}}{K} \right)^{1-\gamma}$, we fix $\varpi$ as the solution of \((5.3)\). Then

$$
\frac{\varpi'}{x} = -\frac{1}{2} h \left( L \varphi^{1-\gamma} \right) \geq -\frac{1}{2} h(\varphi)
$$

which implies

$$
L^b[\varphi] + \frac{\varpi'}{x} \varphi \geq \frac{1}{2} h(\varphi) \varphi \to \infty \quad \text{as} \quad |x| \to \infty.
$$

Since $|x| \wedge |y| > \frac{r_1}{2}$, choosing $r_1$ large enough we have that

$$
L^b[\varphi](x) + \frac{\varpi'}{x} \varphi(x) + L^b[\varphi](y) + \frac{\varpi'}{x} \varphi(y) \geq 1
$$

for every $x, y : |x - y| > r_1$. Thus we deduce that

$$
K + \varepsilon \left( L^b[\Phi](x) - \frac{1}{2} h(\varphi(x)) \Phi(x) + L^b[\Phi](y) - \frac{1}{2} h(\varphi(y)) \Phi(y) \right) < (c_0 + 2\sigma_0 \sigma_1) \psi'(r_1) - \frac{\varpi'}{x} C_1. \quad (5.4)
$$
Notice that \( \varpi \leq 1 \) implies \( h \left( t^{1/\varpi} \right) \leq h(L) \), hence \( -\frac{h'}{h} C_1 \leq \frac{1}{2} h(L) C_1 \). As for the function \( \Phi \), we know by assumption that \( \liminf_{|x| \to \infty} L^b[\Phi] - k h(\varphi) \Phi > -\infty \), for some \( \kappa > 0 \) which we can assume, without loss of generality, such that \( \kappa \geq \frac{1}{2} \). Hence \( L^b[\Phi](x) - \frac{1}{2} h(\varphi(x)) \Phi(x) \geq -k_0 \) for some constant \( k_0 > 0 \). Then, inequality (5.4) implies

\[
K < 2k_0 \varepsilon + (c_0 + 2\sigma_0 \sigma_1) \psi'(r_1) + \frac{1}{2} h(L) C_1.
\]

We conclude that the maximum point cannot be reached with \( r > r_1 \) provided

\[
K \geq 2k_0 \varepsilon + (c_0 + 2\sigma_0 \sigma_1) \psi'(r_1) + \frac{1}{2} h(L) C_1.
\]

This replaces (5.48) in Theorem 3.10. Once \( K \) is chosen, the rest of the proof is the same, by simply using that \( L^b[\varphi], L^b[\Phi] \geq -k_0 \) for some \( k_0 > 0 \), and the conclusion is obtained by choosing \( L \) and \( C_1 \) sufficiently large.

As in Section 4, we deduce by duality the decay estimate for Fokker-Planck equations.

**Theorem 5.5** Let \( L \) be defined by (2.2) where \( \Sigma \) satisfies (2.4), the Levy measure in \( \tilde{I} \) satisfies (2.4), and \( \lambda_0 + \lambda > 0 \). Assume that \( b \) satisfies (2.12) and either of the following conditions:

(i) \( \lambda_1 > 0 \) or \( \sigma \in (1,2) \), and (2.13) holds true.
(ii) \( \lambda_0 = 0 \) and \( \sigma \in (0,1) \), and (2.14) holds true.

For \( m_0 \in M(\mathbb{R}^d) \) with \( \int_{\mathbb{R}^d} dm_0 = 0 \), let \( m \) be the solution of (2.10). Then we have:

(a) If \( \varphi \in \Lambda_1 \) and \( m_0 \in M_\varphi \), then

\[
\|m(t)\|_{M_\varphi} \leq K e^{-\omega t} \|m_0\|_{M_\varphi}
\]

for some \( K, \omega \) depending on \( \alpha, \gamma, (\lambda_1 + \lambda)^{-1}, \sigma, \sigma_0, \sigma_1, c_0, \varphi, d \).

(b) If \( \varphi \in \Lambda_2 \) and \( m_0 \in M_\varphi \), then for all \( \theta \in (0,1) \) such that \( h(\varphi) \varphi^\theta \to \infty \) as \( |x| \to \infty \), we have

\[
\|m(t)\|_{M_\varphi} \leq K \varpi(t) \|m_0\|_{M_\varphi}
\]

where \( \varpi \) is defined as in (5.3), for some \( K, L \) depending on \( \theta, \alpha, \gamma, (\lambda_1 + \lambda)^{-1}, \sigma, \sigma_0, \sigma_1, c_0, \varphi, d \).

As an example of application of Theorem 5.5, when we consider the Lyapunov function

\[
\varphi = e^{\mu(x)^k}
\]

(5.5)

with parameters \( \mu, k > 0 \), we extend to the local operator \( L_0 \) similar results obtained in the recent paper [31] for the Laplace operator.

**Corollary 5.6** Let \( L_0 \) be defined by (2.1), where we assume \( \lambda_0 > 0 \) and (2.2). Assume that \( h(t,x) \) is continuous and satisfies (2.12) and (2.13).

Let \( \varphi \) be defined by (2.3), where we assume that either \( k < \gamma \) (and \( \mu > 0 \) is arbitrary) or \( k = \gamma \) and \( \mu < \frac{a}{\gamma} \) (\( \alpha \) appears in (2.12)). Let \( m_0 \in L^1(\varphi) \) be such that \( \int_{\mathbb{R}^d} dm_0 dx = 0 \), and let \( m \) be the solution to the Cauchy problem (4.7). Then we have:

(a) If \( \gamma \geq 2 \), then

\[
\|m(t)\|_{L^1(\varphi)} \leq K e^{-\omega t} \|m_0\|_{L^1(\varphi)}
\]

for some constant \( K, \omega \) only depending on \( \alpha, \lambda_0, \sigma_0, \sigma_1, d, c_0, k, \mu \).

(b) If \( \gamma \in [1,2) \), then the exponential decay (5.6) holds for \( k \geq 2 - \gamma \).
(c) If $\gamma \in (0, 1)$, for any $\theta \in (0, 1)$ we have

$$\|m(t)\|_{L^1(\varphi)} \leq K e^{-C t}\|m_0\|_{L^1(\varphi)}$$

(5.7)

for some $K, C$ depending on $\theta, \alpha, \gamma, \lambda_0, \sigma_0, \sigma_1, d, c_0, k, \mu$.

**Proof.** We observe the following, through standard computations:

(i) \( \lim_{|x| \to \infty} \mathcal{L}_0[\varphi] = \infty \) provided $k < \gamma$ or $k = \gamma$ and $\mu < \alpha / \gamma$ (appear in 2.12)

(ii) if $\gamma \geq 2$, we have \( \liminf_{|x| \to \infty} \frac{\mathcal{L}_0[\varphi]}{\varphi} > 0 \). Hence \( \|m(t)\|_{L^1(\varphi)} \) decays exponentially, according to Theorem 5.5.

(iii) if $\gamma \in [1, 2)$, \( \liminf_{|x| \to \infty} \frac{\mathcal{L}_0[\varphi]}{\varphi} > 0 \) provided $k \in [2 - \gamma, \gamma)$ (or even $k = \gamma$ if $\mu < \frac{\alpha}{\gamma}$)

(iv) If $\gamma \in (0, 1)$, then we have $\mathcal{L}_0[\varphi] \to \infty$ but $\mathcal{L}_0[\varphi] \simeq h(\varphi)\varphi$ for $h = O\left(\frac{1}{(\log \varphi)^{k-1}}\right)$. Following the recipe of Theorem 5.5, we deduce that, for any $\theta \in (0, 1)$, there exists some $C, K \ (\text{depending on } \theta)$ such that (5.7) holds true.

In particular, the above result shows that, in the local case, there is exponential decay of solutions, even for $\gamma \in [1, 2)$, provided one uses a suitable weighted norm. For a milder confining property of the drift, say $\gamma \in (0, 1)$, compared to Theorem 4.4, we have proved that the solution decays much faster than polynomially when considered in the above weighted norm.

Let us observe that, in view of assumption (2.4), the Lyapunov function (5.5) grows too fast at infinity for the nonlocal part, this is why the above result was restricted to the case of local operator. However, by simply modifying the condition at infinity on the Levy measure $\nu$, the reader can readily generalize the above statement to other classes of Levy operators, with fast decay of the kernel at infinity.

### 5.1 Convergence to stationary solutions

We stress that the above results can also be used to provide existence of stationary solutions, namely invariant measures of Fokker-Planck equations. We give here a sample of this application.

**Theorem 5.7** Under the assumptions of Theorem 5.5, assume that either $\varphi \in \Lambda_1$, or $\varphi \in \Lambda_2$ and $\varphi$ defined in (5.2) satisfies $\varphi \in L^1(0, \infty)$. Assume that $b = b(x)$ is a continuous, time-independent vector field. Then there exists a unique stationary measure $\bar{m} \in P(\mathbb{R}^d)$, which is a solution of

$$\mathcal{L}^*[\bar{m}] - \text{div}(b(x)\bar{m}) = 0 \quad \text{in } \mathbb{R}^d,$$

(5.8)

and such that, if $m_0 \in P(\mathbb{R}^d) \cap M_\varphi$ and $m$ is the solution of (2.10), we have, for some $K > 0$,

- if $\varphi \in \Lambda_1$, then \( \|m(t) - \bar{m}\|_{M_\varphi} \leq K e^{-\omega t} \)
- if $\varphi \in \Lambda_2$, then \( \|m(t) - \bar{m}\|_{M_\varphi} \leq K \varphi(t) \)

(5.9) (5.10)

for any $\theta < 1$ such that $h(\varphi)\varphi^\theta \to \infty$ as $|x| \to \infty$.

**Proof.** First of all we observe that, if $m$ is the solution of (2.10) (in the sense of Definition 2.5), then $m_0 \in M_\varphi$ implies $m(t) \in M_\varphi$ (reasoning exactly as in (6.3) for the function $(x)^k$).

Suppose now that $\varphi \in \Lambda_1$, or $\varphi \in \Lambda_2$ is such that $\varphi$ defined in (5.2) satisfies $\varphi \in L^1(0, \infty)$. For any $s > 0$, we can apply Theorem 5.5 to $m(t + s) - m(t)$, which yields

- if $\varphi \in \Lambda_1$, then \( \|m(t + s) - m(t)\|_{M_\varphi} \leq K e^{-\omega t} \|m(s) - m_0\|_{M_\varphi} \)
- if $\varphi \in \Lambda_2$, then \( \|m(t + s) - m(t)\|_{M_\varphi} \leq K \varphi(t) \|m(s) - m_0\|_{M_\varphi} \).

(5.11)

In particular, setting $m_n(x) := m(n, x)$, we have

$$m_n = m_0 + \sum_{k=0}^{n-1} [m_{k+1} - m_k]$$

(5.12)
where last series is strongly convergent in \( \mathcal{M}_\varphi \) (or in \( \mathcal{M}_{\varphi^\sigma} \) if \( \varphi \in \Lambda_2 \)), due to \((5.11)\). Hence there exists a measure \( \tilde{m} \in \mathcal{P}(\mathbb{R}^d) \) such that \( m_n \to \tilde{m} \) in \( \mathcal{M}_\varphi \) (or in \( \mathcal{M}_{\varphi^\sigma} \) if \( \varphi \in \Lambda_2 \)). In fact, \( \tilde{m} \) is itself a probability measure. Now we observe that, for any stationary solution \( \bar{\varphi} \), we have

\[
\int_{\mathbb{R}^d} \xi \, dm_n - m_{n-1} + \int_{n-1}^n \int_{\mathbb{R}^d} \mathcal{L}^b[\xi] \, dm(t) \, dt = 0
\]

which implies

\[
\int_{\mathbb{R}^d} \mathcal{L}^b[\xi] \, dm_n = \int_{n-1}^n \int_{\mathbb{R}^d} \mathcal{L}^b[\xi] \, dm(t) \, dt - \int_{\mathbb{R}^d} \xi \, dm - m_{n-1} \cdot \tilde{m}.
\]

If \( \varphi \in \Lambda_1 \), by \((5.11)\) we have \( \|m_n - m_{n-1}\|_{\mathcal{M}_\varphi} \to 0 \). The first integral in the right-hand side also goes to zero; indeed, again by \((5.11)\) we have

\[
\left| \int_{n-1}^n \int_{\mathbb{R}^d} \mathcal{L}^b[\xi] \, dm(t) \, dt \right| \leq \|\mathcal{L}^b[\xi]\|_{L^\infty(\varnothing^\sigma)} \int_{n-1}^n \|m(t) - m_n\|_{\mathcal{M}_\varphi} \, dt \leq C \int_{n-1}^n e^{-\omega t} \, dt \to 0.
\]

If \( \varphi \in \Lambda_2 \), the same argument applies in \( \mathcal{M}_{\varphi^\sigma} \) and replacing \( e^{-\omega t} \) with \( \varpi(t) \), according to \((5.11)\). Therefore, passing to the limit in \((5.13)\), recalling the convergence of \( m_n \), we conclude that \( \tilde{m} \) is a stationary solution. Finally, we can now apply Theorem \(5.5\) to \( m - \tilde{m} \), and we deduce that the whole sequence \( m(t) \) converges to \( \tilde{m} \). Let us point out that, in case \( \varphi \in \Lambda_2 \), one should apply Theorem \(5.5\) between the spaces \( \mathcal{M}_{\varphi^\sigma} \) and \( \mathcal{M}_{\varphi^\sigma + \varepsilon} \), since a priori one only knows that \( \tilde{m} \in \mathcal{M}_{\varphi^\sigma} \) for any \( \theta < 1 \).

---

6 Appendix: existence and uniqueness of solutions

In this section we give a proof of the existence and uniqueness of solutions of the Fokker-Planck equation \((2.10)\). Once more, this can be essentially deduced by duality from the well-posedness of viscosity solutions of the adjoint problem. This is why we start by giving a proof of the comparison principle, and uniqueness, of viscosity solutions.

**Proposition 6.1** Let \( u_0 \in C(\mathbb{R}^d) \cap L^\infty((x)^{-k}) \), \( f \in C(Q_T) \cap L^\infty(0,T;L^\infty((x)^{-k})) \), for some \( k < \sigma \). Let \( \mathcal{L} \) be defined by \((2.9)\) where \( \Sigma \) satisfies \((2.2)\), the Levy measure in \( \mathcal{I} \) satisfies \((2.7)\), and \( \lambda_0 + \lambda > 0 \). Assume that \( b \in C(Q_T) \) is such that \( b(t,x) \cdot x \) is bounded below and either of the following conditions hold:

1. \( \lambda_0 > 0 \) or \( \sigma \in (1,2) \), and \((2.13)\) holds true whenever \( |x-y| \leq 1 \).
2. \( \lambda_0 = 0 \), \( \sigma \in (0,1] \) and \((2.14)\) holds true whenever \( |x-y| \leq 1 \).

Then there exists a unique viscosity solution of

\[
\begin{cases}
\partial_t u + \mathcal{L}[u] + b(t,x) \cdot Du = f & \text{in } (0,T) \times \mathbb{R}^d, \\
u(0,x) = u_0(x), & \text{in } \mathbb{R}^d
\end{cases}
\]

such that \( |u(t,x)| \leq C(1 + |x|)^m \) for some \( m < \sigma \).
Proof. We prove first the uniqueness of solutions. Let us take some $\beta < \sigma$ such that $\beta \geq k \lor m$. We first remark that (see the computations in Lemma 3.1) $\mathcal{L}^b([x]^{\beta}) \geq \beta ([x]^{\beta} - (b \cdot x) - c_\beta (1 + [x]^{\beta-1})$; hence there exists $K_0 > 0$ such that

$$\langle x \rangle^\beta + \mathcal{L}^b([x]^{\beta}) \geq \frac{1}{2} \langle x \rangle^\beta - K_0$$

where $K_0$ only depends on $\sigma, d, ||\Sigma(x)||, \Lambda$ and the lower bound of $b \cdot x$. We deduce that, choosing $K$ large enough, the function $W := e^t [x]^{\beta} + Kt + K_1 + \frac{1}{t^\epsilon}$ is a supersolution, provided $\langle x \rangle^\beta + K_1 \geq |u_0|$. Let now $u, v$ be, respectively, a sub and super-solution of (6.1). By linearity, we have that $u_\epsilon := u - \epsilon W$ and $v_\epsilon := v + \epsilon W$ are still sub/super-solutions, and, if $\beta > m$, there exists a compact set $K_\epsilon \subset \mathbb{R}^d$ such that $u_\epsilon \leq v_\epsilon$ for $t \in (0, T), x \in \mathbb{R}^d \setminus K_\epsilon$. Of course, if we prove the comparison for $u_\epsilon, v_\epsilon$, then we get the comparison for $u, v$ by letting $\epsilon \rightarrow 0$. Therefore, up to replacing $u, v$ with $u_\epsilon, v_\epsilon$, we can assume that $u \leq 0 \leq v$ outside a compact set $K$. Now, suppose that $M := \text{sup}_{Q_t}(u - v) > 0$, which means that $M$ is attained at some point in $(0, T) \times K$, with $t > 0$. Doubling variables, we consider

$$M_\epsilon := \text{sup}_{(x,y)} \{ u(t, x) - v(t, y) - \frac{|x - y|^2}{2\epsilon} - A_\epsilon v^\sigma \}$$

for some $\eta, A_\epsilon > 0$ to be chosen, with $A_\epsilon \rightarrow 0$. Since $M_\epsilon \geq M - A_\epsilon T^\sigma$, then $M_\epsilon$ is positive for $\epsilon$ small. Recalling that $u \leq 0 \leq v$ outside $K$, this implies that $M_\epsilon$ is assumed at points $(x_\epsilon, y_\epsilon)$ which lie in a compact set, which is actually an $\epsilon$-neighborhood of $K$, since $\frac{|x_\epsilon - y_\epsilon|^2}{2\epsilon} \leq M = \text{sup} u - \text{inf} v < \infty$. In particular we have

$$|x_\epsilon - y_\epsilon| \leq \sqrt{2\epsilon M} \quad (6.2)$$

and by usual arguments, since $M_\epsilon \rightarrow M$ as $\epsilon \rightarrow 0$, we also get $\frac{|x_\epsilon - y_\epsilon|^2}{\epsilon} \rightarrow 0$. Using Theorem 2.3, the Lipschitz character of $\Sigma(x)$ and standard arguments in the doubling variable method, we get

$$\eta A_\epsilon t^{\sigma - 1} + (b(t, x_\epsilon) - b(t, y_\epsilon)) \cdot \frac{x_\epsilon - y_\epsilon}{\epsilon} \leq f(t, x_\epsilon) - f(t, y_\epsilon) + C \sigma_1 \frac{|x_\epsilon - y_\epsilon|^2}{\epsilon}. \quad (6.3)$$

At this stage we use the following regularity result, which can be obtained with similar arguments as used in Section 3 (see the Appendix in [22]): using conditions (2.13), or (2.14), for $|x - y| \leq 1$, every viscosity solution of (6.1) satisfies

$$|u(t, x) - u(t, y)| \leq C_T \frac{[x]^{\beta} + [y]^{\beta}}{t^\beta} |x - y|, \quad \forall x, y, \quad \text{if } \sigma \in (1, 2), \quad (6.4)$$

$$|u(t, x) - u(t, y)| \leq C_T \frac{[x]^{\beta} + [y]^{\beta}}{t^\beta} |x - y|^{\theta}, \quad \forall x, y, \quad \text{if } \sigma \in (0, 1], \theta < \sigma.$$

Therefore, assuming that at least one between $u, v$ is a solution, estimate $6.2$ can be improved accordingly. Recalling that the points $(x_\epsilon, y_\epsilon)$ lie in a compact set, we deduce that they satisfy, if $\sigma > 1$,

$$\frac{|x - y|^2}{2\epsilon} \leq C_{T,K} |x - y| \leq C \frac{\sqrt{\epsilon}}{t^\beta}.$$ 

Notice that this constant $C_{T,K}$ only depends on the local bounds of $u, v$, and in particular is determined independently of $A_\epsilon$. Hence, (6.3) implies

$$\eta A_\epsilon t^{\sigma - 1} \leq \frac{2C_{T,K}}{t^\beta} |b(t, x_\epsilon) - b(t, y_\epsilon)| + f(t, x_\epsilon) - f(t, y_\epsilon) + C \sigma_1 \frac{|x_\epsilon - y_\epsilon|^2}{\epsilon} \leq C \omega(\epsilon) \left( \frac{1}{t^\beta} + 1 \right) + C \frac{\sqrt{\epsilon}}{t^\beta}$$

for some $C > 0$ and some modulus of continuity $\omega(\cdot)$ depending on $b, f$. Choosing $\eta = 1 - \frac{1}{\sqrt{\epsilon}}$, if $A_\epsilon = L (\omega(\epsilon) + \sqrt{\epsilon})$ we obtain a contradiction for $L$ sufficiently large.

If $\sigma \in (0, 1]$, we estimate

$$\frac{|x - y|^2}{2\epsilon} \leq C \frac{|x - y|^\theta \epsilon}{t^\beta}$$

34
then we use (2.14) and we get

\[\eta A_{\varepsilon} t^{\alpha - 1} \leq c_0 |x_\varepsilon - y_\varepsilon|^{1 - \sigma + \delta} \frac{|x_\varepsilon - y_\varepsilon|}{\varepsilon} + f(t, x_\varepsilon) - f(t, y_\varepsilon) + C \sigma \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon} \leq \frac{C}{t^\theta} |x_\varepsilon - y_\varepsilon|^{\theta - \sigma + \delta} + f(t, x_\varepsilon) - f(t, y_\varepsilon) + \sigma t \frac{C}{t^\theta} |x_\varepsilon - y_\varepsilon|^\theta.\]

Choosing \(\theta = \sigma - \frac{\alpha}{2}\) and \(\eta = 1 - \frac{\alpha}{2}\), we conclude as before with a suitable choice of \(A_{\varepsilon}\). Having obtained a contradiction, we infer that \(sup_{Q_{\varepsilon}}(u - v) \leq 0\), i.e. \(u \leq v\). Reversing the role, we get the desired uniqueness. The existence of solutions can be proved through the vanishing viscosity method (using (6.4) to infer local compactness, see also the proof of Theorem 2.4 below).

The following lemma is a classical statement in viscosity solutions theory to ensure equi-continuity in time from continuity in space (see [4, Lemma 9.1]). We give a proof for the reader’s convenience, since the setting of our assumptions does not fall into previous versions.

**Lemma 6.2** Let \(K \subset \mathbb{R}^d\) be a compact set, and let \(u\) be a viscosity solution of (6.7), where \(b, f, u_0\) are continuous. For any \(\varepsilon > 0\), there exists \(\delta > 0\) such that

\[\|u(t) - u_0\|_{L^\infty(K)} \leq \varepsilon \quad \forall t < \delta,\]

where \(\delta\) depends only on the bounds of \(b, f, u\) and on the modulus of continuity of \(u_0\), for \((t, x)\) in a neighborhood of \(\{0\} \times K\) (and of course on \(\sigma, d, ||\Sigma(x)||_{\infty}, \Lambda\) as well).

**Proof.** For \(\eta > 0\) small enough, \(y \in B_\eta, x \in B_{2\eta}\), let us consider the function

\[\Phi(x) := u_0(y) + C \zeta(|x - y|) + kt + \eta\]

where \(\zeta(s) = \tau_R(s^2)\); here \(\tau_R(\cdot)\) denotes a \(C^2\), bounded and concave function such that \(\tau_R(s) = s\) for \(|s| < (3R)^2\). By concavity of \(\tau_R\), we have

\[I(x, [\zeta(|x - y|)]) \leq \int_{|z| > 1} \{\zeta(|x + z - y|) - \zeta(|x - y|)\} \nu(dz) + \tau_R^\theta(|x - y|^2) \int_{|z| \leq 1} \{|x + z - y|^2 - |x - y|^2| - 2((x - y) \cdot z)I_{|z| \leq 1}\} \nu(dz) \leq M_R.\]

In the local terms, we just observe that \(\zeta(|x - y|) = |x - y|^2\); so, using the local bound of \(b\), we have

\[L^0(\zeta) \geq -M_R\]

for a possibly different constant, only depending on \(R\), and local bounds of \(Q(x), b\). Hence we have

\[\Phi_t + L^0(\Phi) - f(x) \geq k - C M_R - \|f\|_{L^\infty(B_{2\eta})}\]

which implies that \(\Phi\) is a supersolution for \(x \in B_{2\eta}\), up to choosing \(k \geq C M_R + \|f\|_{L^\infty(B_{2\eta})}\). If \(x \in \partial B_{2\eta}\), we have \(|x - y| > R\), and \(\Phi \geq u\) provided \(C R^2 + u_0(y) \geq \|u(t)\|_{L^\infty(B_{2\eta})}\); in particular, this holds if

\[C R^2 \geq 2\|u\|_{L^\infty((0, T) \times B_{2\eta})} \]

At \(t = 0\), we have \(u_0 \leq \Phi\) if \(u_0(x_\varepsilon) - u_0(y) \leq C|x - y|^2 + \eta\). This holds if \(\omega(|x - y|) - \eta \leq C|x - y|^2\), where \(\omega(\cdot)\) is a modulus of continuity of \(u_0\) in \(B_{2\eta}\). Without loss of generality, we can assume that \(\omega(r) \leq L r\) near the origin, so there exists \(\eta_C := sup_{r \in (0, 3R)} \omega(r) - C r^2 > 0\) and \(\eta_C \to 0\) as \(C \to \infty\). Choosing \(\eta = \eta_C\), we have the comparison between \(u\) and \(\Phi\) and we conclude that \(u(t, x) \leq \Phi(t, x)\) for \(x \in B_{2\eta}\). Choosing \(x = y\) gives

\[u(t, y) \leq u_0(y) + C M_R t + \|f\|_{L^\infty(B_{2\eta})} t + \eta_C.\]
Using $\eta C \to 0$ as $C \to \infty$, we deduce that
\[ \|u(t) - u_0\|_{L^\infty(B_R)} \xrightarrow{t \to 0} 0, \]
and this limit is uniform for all solutions $u$ which are uniformly bounded in $B_{2R}$.

As a Corollary of the previous lemma, we infer a compactness principle for solutions of the Fokker-Planck equation.

**Lemma 6.3** Let $m \in C^0([0, T]; P(\mathbb{R}^d))$ be a solution of (2.10) in the sense of Definition 2.5 where $m_0 \in P(\mathbb{R}^d) \cap M_k(\mathbb{R}^d)$, and $L, b$ satisfy the assumptions of Theorem 2.6. Then, for every $\varepsilon > 0$ there exists $\delta > 0$ such that
\[ d_1(m(t), m(s)) \leq \varepsilon \quad \forall t, s, |t - s| \leq \delta \quad (6.5) \]
where $\delta$ only depends on $\|m_0\|_{M_k}$, on local uniform bounds of $b$, on lower bound of $b(t, x) \cdot x$, as well as on the upper bounds $\|\Sigma(x)\|_{L^\infty}$.

In particular, any sequence $\{m_n\} \subset C^0([0, T]; P(\mathbb{R}^d))$ of solutions of (2.10), corresponding to $b_n, m_{0n}$, is relatively compact in $C^0([0, T]; P(\mathbb{R}^d))$ as soon as $b_n$ is locally uniformly bounded in $Q_T$, $b_n \cdot x$ is uniformly bounded below, and $m_{0n}$ is bounded in $M_k(\mathbb{R}^d)$.

**Proof.** As in Proposition 6.1, we take the function $W \simeq e^{(t-s)\langle x \rangle^k} + K(t-s)$ so that $-W_s + \mathcal{L}^b[W] \geq 0$ in $(0, t)$. Any (smooth) truncation of this function remains a supersolution and can be used as test function in Definition 2.5. We deduce
\[ \int_{\mathbb{R}^d} \langle x \rangle^k dm(t) \leq C_t \int_{\mathbb{R}^d} \langle x \rangle^k dm_0, \quad (6.6) \]
so we have estimates of $m(t)$ in $M_k(\mathbb{R}^d)$ whenever $m_0 \in M_k(\mathbb{R}^d)$.

Now, let us estimate $d_1(m(t), m(s))$. To this purpose, take $\xi : \|\xi\|_{W^{1, \infty}(\mathbb{R}^d)} \leq 1$ and assume $t > s$. In Definition 2.5 we consider the solution $\varphi$ of (B) with $f = 0$. Then we have
\[ \int_{\mathbb{R}^d} \xi d(m(t) - m(s)) = \int_{\mathbb{R}^d} (\varphi(s) - \xi) dm(s) \leq 2 \int_{B_R} dm(s) + \int_{B_R} (\varphi(s) - \xi) dm(s) \leq 2 \frac{C_t}{R^k} \|m_0\|_{M_k} + \int_{B_R} (\varphi(s) - \xi) dm(s) \]
where we used (6.6). By Lemma 6.2 for fixed $R$ we have
\[ \int_{B_R} (\varphi(s) - \xi) dm(s) \leq \omega_R(t-s) \]
for some quantity $\omega_R(r)$ converging to zero as $r \to 0$. Moreover $\omega_R(\cdot)$ only depends on $R$ and $\|b\|_{L^\infty((0, T) \times B_{2R})}$, and in particular it is uniform for all $\xi$ with $\|\xi\|_{W^{1, \infty}(\mathbb{R}^d)} \leq 1$. We conclude that
\[ d_1(m(t), m(s)) \leq \frac{C}{R^k} \|m_0\|_{M_k} + \omega_R(t-s). \]
Hence (6.5) is proved. By Ascoli-Arzelà theorem, this means that $m$ lies in a relatively compact set of $C^0([0, T]; P(\mathbb{R}^d))$ as $m_0$ is in a bounded set of $M_k$, and $b$ varies in a set of functions which are locally uniformly bounded with $b \cdot x$ uniformly bounded below.

Now we are ready to give the proof of Theorem 2.6.

**Proof of Theorem 2.6.** Since any $m_0$ can be decomposed as $m_0^+ - m_0^-$, using the Hahn decomposition, and since the problem is linear, it is enough to consider the case that $m_0^+$ is a positive measure. Without loss of generality, we assume $m_0 \in P(\mathbb{R}^d)$. Indeed, it is immediate to check from Definition 2.5 that $\int_{\mathbb{R}^d} dm(t) = \int_{\mathbb{R}^d} dm_0$, and $m_0 \geq 0$ implies $m(t) \geq 0$; hence $m(t) \in P(\mathbb{R}^d)$.

We construct a solution through vanishing viscosity approximation. So we consider the problem
\[
\begin{cases}
\partial_t m - \varepsilon \Delta m + \mathcal{L}^*[m] - \text{div}(b(t, x)m) = 0 & \text{in } (0, T) \times \mathbb{R}^d, \\
m(0, x) = m_0(x), & \text{in } \mathbb{R}^d
\end{cases}
\]
\[(FP)_{\varepsilon}^b\]
where \( \varepsilon > 0 \). If \( b \) is a smooth bounded vector field (e.g. \( b \in C^2(\mathbb{R}^d) \) with \( b, Db \) bounded), then for any smooth \( m_0 \in \mathcal{P}(\mathbb{R}^d) \) there exists a smooth solution \( m_\varepsilon \in \mathcal{P}(\mathbb{R}^d) \). Of course \( m_\varepsilon \) is a weak as well as a duality solution, and satisfies

\[
\int_{\mathbb{R}^d} m_\varepsilon(t) \xi + \int_0^t \int_{\mathbb{R}^d} f m_\varepsilon \, dx \, dt = \int_{\mathbb{R}^d} \varphi(0,x) \, dm_0
\]

for every \( t \in (0,T) \), \( \xi \in C_0(\mathbb{R}^d) \), \( f \in C_0(\Omega_t) \), where \( \varphi \) is the (unique) viscosity solution of

\[
\begin{cases}
-\partial_t \varphi - \varepsilon \Delta \varphi + L[\varphi] + b \cdot D \varphi = f & \text{in } (0,t) \times \mathbb{R}^d, \\
\varphi(t,x) = \xi(x), & \text{in } \mathbb{R}^d.
\end{cases}
\]

Now, we consider \( b \in C^0_\varepsilon \), and a sequence \( \{b_n\} \) of smooth functions, uniformly bounded, which converges to \( b \) locally uniformly. By standard results on the stability of viscosity solutions, the solutions \( \varphi_n^\varepsilon \) of \((H)_{b_n}^\varepsilon \) are relatively compact and converge locally uniformly in \([0,t] \times \mathbb{R}^d\) towards a function \( \varphi \) which is a viscosity solution of \((H)^\varepsilon\). This implies \( \int_{\mathbb{R}^d} \varphi_n^\varepsilon(0,x) \, dm_0 \to \int_{\mathbb{R}^d} \varphi(0,x) \, dm_0 \). As for the solutions \( m_n^\varepsilon \), Lemma 6.3 implies that the sequence \( m_n^\varepsilon \) is relatively compact in \( C^0([0,T];\mathcal{P}(\mathbb{R}^d)) \). Hence, one can pass to the limit in the equality (6.7), so we build a solution \( m \) for \( b \in C^0_\varepsilon \). In a similar way, we can get rid of the global boundedness of \( b \); it is enough to take a sequence \( \zeta_R \) of cut-off functions converging to \( 1 \) and to consider \( b_R := b \zeta_R \). This is a sequence of bounded functions converging to \( b \) locally uniformly, and such that \( b_R \cdot x \) is uniformly bounded. By Lemma 6.3, we have again that \( m_R \), the corresponding solutions, are relatively compact in \( C^0([0,T];\mathcal{P}(\mathbb{R}^d)) \). In parallel, the viscosity solutions \( \varphi_R \) of \((H)^\varepsilon_{b_R} \), are uniformly bounded and relatively compact in \( C^0([0,t] \times K) \) for all compact subsets \( K \); hence, up to subsequences, \( \varphi_R \) converges to a viscosity solution of \((H)^\varepsilon \).

Now we can let \( \varepsilon \to 0 \). The viscosity solutions of \((H)^\varepsilon \) are still uniformly bounded, and they are also relatively compact in \( C^0([0,t] \times K) \) for all compact subsets \( K \), by using estimates (6.4) and Lemma 6.2 for the time equi-continuity. By Lemma 6.3, the sequence \( m_\varepsilon \) is also relatively compact in \( C^0([0,T];\mathcal{P}(\mathbb{R}^d)) \) and converges to some \( m \in C^0([0,T];\mathcal{P}(\mathbb{R}^d)) \). We deduce that \( m \) satisfies Definition 2.3 where \( \varphi \) is the viscosity solution of \((B) \) which has been obtained in the vanishing viscosity limit. But this is the unique viscosity solution of problem \((B) \), from Proposition 6.1. Thus, the existence of \( m \), satisfying Definition 2.3, is proved. The uniqueness is straightforward, because problem \((B) \) is solvable for all \( \xi \in C^0_\varepsilon \) (we just proved it, through vanishing viscosity limit).

Finally, since \( m_0 \in \mathcal{M}_K \), we have \( m(t) \in \mathcal{M}_K(\mathbb{R}^d) \) from 6.5. Then the formulation of \( m \) extends to test functions \( f \in L^\infty((0,T);\mathcal{L}^\infty((x)^-k)) \), \( \xi \in L^\infty((x)^-k) \) with an easy truncation argument, using dominated convergence theorem; it is enough, once more, to observe that the viscosity solutions of \((B) \) are stable as bounded sequences \( f_n, \xi_n \) approximate \( f \in L^\infty((0,T);\mathcal{L}^\infty((x)^-k)) \), \( \xi \in L^\infty((x)^-k) \), and in addition the limit solution is unique.

\section*{References}

[1] D. Applebaum: \textit{Lévy Processes and Stochastic Calculus}. Cambridge Univ. Press. (2004).

[2] A. Arnold, P. Markowich, G. Toscani, A. Unterreiter, \textit{On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations}, Comm. Partial Differential Equations 26 (2001) 43-100.

[3] D. Bakry, P. Cattiaux, A. Guillin, \textit{Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré}, J. Funct. Anal. 254 (2008) 727-759.

[4] G. Barles, S. Biton, E. Ley, \textit{A geometrical approach to the study of unbounded solutions of quasilinear parabolic equations}, Arch. Ration. Mech. Anal. 162 (2002), 287-325.

[5] G. Barles, E. Chasseigne, A. Ciomaga, C. Imbert, \textit{Lipschitz regularity of solutions for mixed integro-differential equations}, J. Differential Equations 252 (2012), 6012-6060.
[6] G. Barles, E. Chasseigne, A. Ciomaga, C. Imbert, \textit{Large time behavior of periodic viscosity solutions for uniformly parabolic integro-differential equations}, Calc. Var. PDE 50 (2014), 283-304.

[7] G. Barles, E. Chasseigne, C. Imbert, \textit{Hölder continuity of solutions of second-order non-linear elliptic integro-differential equations}, J. Eur. Math. Soc. 13 (2011), 1-26.

[8] G. Barles, C. Imbert, \textit{Second-order elliptic integro-differential equations: viscosity solutions’ theory revisited}, Ann. I. H. Poincaré - AN 25 (2008), 567-585.

[9] G. Barles, O. Ley, E. Topp, \textit{Lipschitz regularity for integro-differential equations with coercive Hamiltonians with application to large time behavior}, Nonlinearity 30 (2017) 703-734.

[10] G. Barles, P. Souganidis, \textit{Space-time periodic solutions and long-time behavior of solutions of quasilinear parabolic equations}, SIAM J. Math. Anal. 32 (2001), 1311-1323.

[11] P. Biler, G. Karch, \textit{Generalized Fokker-Planck equations and convergence to their equilibria}, Banach Center Publ., 60 (2003), 307-318.

[12] J. A. Canizo, S. Mischler, \textit{Harris-type results on geometric and subgeometric convergence to equilibrium for stochastic semigroups}, preprint arXiv:2110.09650 (2021).

[13] P. Cardaliaguet and A. Porretta, \textit{An introduction to Mean Field Game theory}, in: Mean field games, Lectures Notes in Mathematics (CIME - series), Springer, (2021).

[14] E. Chasseigne, O. Ley, T.-T. Nguyen, \textit{A priori Lipschitz estimates for solutions of local and nonlocal Hamilton-Jacobi equations with Ornstein-Uhlenbeck operator}, Revista Matemática Iberoamericana 35 (2019), 1415-1449.

[15] M.F. Chen, S.F. Li, \textit{Coupling methods for multidimensional diffusion processes}, Ann. Probab. 17 (1989) 151-177.

[16] M. Cirant, A. Goïfi, \textit{On the existence and uniqueness of solutions to time dependent fractional MFG}, SIAM J. Math. Anal., 51 (2019), 913-954.

[17] M. G. Crandall, H. Ishii, P. L. Lions, \textit{User’s guide to viscosity solutions of second order partial differential equations.} Bull. Amer. Math. Soc. (N.S.) 27 (1992), 1-67.

[18] R. Douc, G. Fort, A. Guillin, \textit{Subgeometric rates of convergence of f-ergodic strong Markov processes}, Stochastic Process. Appl., 119 (2009), 897-923.

[19] A. Eberle, \textit{Reflection couplings and contraction rates for diffusions}, Probab. Theory Related Fields 166 (2016) 851-886.

[20] A. Eberle, A. Guillin, R. Zimmze, \textit{Quantitative Harris type theorems for diffusions and McKean-Vlasov processes}. Trans. American Math. Soc. 371 (2019), 7135-7173.

[21] D. Down, S.P. Meyn, R.L. Tweedie, \textit{Exponential and uniform ergodicity of Markov processes}, Ann. Probab. 23 (1996) 1671-1691.

[22] O. Ersland, E.R. Jakobsen, A. Porretta, \textit{Long time behaviour of fractional Mean Field Games}, in preparation.

[23] N. Fournier, B. Perthame, \textit{Transport distances for PDEs: the coupling method}, EMS Surv. Math. Sci. 7 (2020), 1-31.

[24] I. Gentil and C. Imbert, \textit{The Lévy-Fokker-Planck equation: Phi-entropies and convergence to equilibrium}, Asymptot. Anal. 59 (2008), 125-138.

[25] M.P. Guadagni, S. Mischler, C. Mouhot, \textit{Factorization of non-symmetric operators and exponential H-Theorem}. Mém. Soc. Math. Fr. 153, (2017), 1-137.
[26] L. Gross, *Logarithmic Sobolev inequalities*, American Journal of Math. 97 (1975), 1061-1083.

[27] M. Hairer, J.C. Mattingly, *Yet another look at Harris’ ergodic theorem for Markov chains*, in: Seminar on Stochastic Analysis, Random Fields and Applications VI, in: Progress in Probab., vol. 63, 2011, pp. 109-117.

[28] M. Huang, R. P. Malhamé, P. E. Caines, *Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Comm. in Information & Systems, 6 (2006), 221-252.

[29] H. Ishii, P. L. Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations*, J. Differential Equations 83 (1990), 26-78.

[30] E.R. Jakobsen, K.H. Karlsen, *A “maximum principle for semicontinuous functions” applicable to integro-partial differential equations*, NoDEA Nonlinear Differential Equations Appl. 13 (2006), 137-165.

[31] O. Kavian, S. Mischler, M. Ndao, *The Fokker-Planck equation with subcritical confinement force*, J. Math. Pures Appl. 151 (2021), 171-211.

[32] L. Lafleche, *Fractional Fokker-Planck equations with general confinement force*, Siam J. Math. Anal. 52 (2020), 164-196.

[33] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. II-Horizon fini et contrôle optimal*, Comptes Rendus Mathématique, 343 (2006), 679-684.

[34] T. M. Liggett, *$L^2$ rates of convergence for attractive reversible nearest particle systems*, Ann. Probab. 14 (1991), 935-959.

[35] T. Lindvall, C. Rogers, *Coupling of multidimensional diffusions by reflection*, Ann. Probab. 14 (1986), 860-872.

[36] Y. Maekawa, H. Miura, *On fundamental solutions for non-local parabolic equations with divergence free drift*, Adv. Math. 247 (2013), 123-191.

[37] M.B. Majka, *Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes*, Stochastic Processes and their Applications 127 (2017) 4083-4125.

[38] S. P. Meyn and R. L. Tweedie, *Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes*, Adv. Appl. Prob., 25 (1993), 518-548.

[39] H. Osada, *Diffusion processes with generators of generalized divergence form*, J. Math. Kyoto Univ 27(4) (1987), 579–619.

[40] A. Porretta, E. Priola, *Global Lipschitz regularizing effects for linear and nonlinear parabolic equations*, J. Math. Pures Appl. (9) 100 (2013), 633-686.

[41] M. Röckner, F.-Y. Wang, *Weak Poincaré inequalities and $L^2$-convergence rates of Markov semigroups*. J. Funct. Anal. 185 (2001), 564-603.

[42] G. Toscani, C. Villani, *On the trend to equilibrium for some dissipative systems with slowly increasing a priori bounds*. J. Statist. Phys. 98 (2000), 1279-1309.

[43] I. Tristani, *Fractional Fokker-Planck equation*, Comm. Math. Sci. 13 (2015), 1243-1260.

[44] F.Y. Wang, *Coupling for Ornstein-Uhlenbeck processes with jumps*, Bernoulli 17 (2011), 1136-1158.

[45] J. Wei, R. Tian, *Well-posedness for the fractional Fokker-Planck equations*, J. Math. Phys. 56 (2015).

[46] Q. Zhang, *A strong regularity result for parabolic equations*, Comm. Math. Phys. 244 (2004), 245–260.