FELLER GENERATORS
WITH MEASURABLE LOWER ORDER TERMS

FRANZISKA KÜHN AND MARKUS KUNZE

ABSTRACT. We study perturbations of Feller generators under 'lower order terms' with measurable coefficients. We investigate which properties of the original semigroup – such as positivity, conservativeness and the Feller property – are passed to the perturbed semigroup. We give several examples and discuss applications in the theory of martingale problems and stochastic differential equations with measurable coefficients.

1. INTRODUCTION

Given two operators A and B, it is a classical question to ask how the sum $A + B$ is defined and which properties of A are preserved under the perturbation B, see e.g. [23, 3]. In this article, we consider this problem in the context of Feller semigroups and generators. We are interested in constructing Feller generators with measurable lower order terms, and this means that the perturbation does not take values in the space of continuous functions. In consequence, classical perturbation results from semigroup theory do not apply.

Let $(T(t))_{t \geq 0}$ be a Feller semigroup, that is, a sub Markovian, strongly continuous semigroup $T(t) : \mathcal{C}_\infty(\mathbb{R}^d) \to \mathcal{C}_\infty(\mathbb{R}^d)$ on the space $\mathcal{C}_\infty(\mathbb{R}^d)$ of continuous functions vanishing at infinity. If the domain of the (infinitesimal) generator A contains the test functions $\mathcal{C}_c^\infty(\mathbb{R}^d)$, then the Courrège–van Waldenfels theorem, see e.g. [6, Thm. 2.21], shows that $\mathcal{A} := A|_{\mathcal{C}_c^\infty(\mathbb{R}^d)}$ is a *Lévy-type operator*, i.e. an integro-differential operator of the form

$$
\begin{aligned}
\mathcal{A}f(x) = {} & b(x)\nabla f(x) + \frac{1}{2}\operatorname{tr}(Q(x)\nabla^2 f(x)) \\
& + \int_{\mathbb{R}^d \setminus \{0\}} (f(x+y) - f(x) - \nabla f(x) \cdot \mathbb{1}_{(0,1)}(|y|))\,\nu(x, dy).
\end{aligned}
\tag{1.1}
$$

Here $(b(x), Q(x), \nu(x, \cdot))_{x \in \mathbb{R}^d}$ are the *(infinitesimal) characteristics* consisting of the *drift coefficients* $b = (b_j) : \mathbb{R}^d \to \mathbb{R}^d$, the *diffusion coefficients* $Q = (q_{ij}) : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ and the *jumping kernels* $\nu : \mathbb{R}^d \to \mathcal{M}^+(\mathbb{R}^d)$. We are interested in the following questions: If a Lévy-type operator B is a lower order perturbation of A, then under which conditions is (a realization of) $A + B$ the generator of a semigroup and which properties does the perturbed semigroup inherit from the original semigroup $(T(t))_{t \geq 0}$? Is the martingale problem for $A + B$ well-posed? Classical perturbations results can be used to tackle these questions if B maps $D(A)$ into $\mathcal{C}_\infty(\mathbb{R}^d)$, which in particular implies that the infinitesimal characteristics of B need to depend continuously on x. In this article, we investigate perturbations B whose characteristics depend merely measurably on x. Allowing for discontinuous characteristics of B creates a number of issues and subtleties. To give an example: If B maps $D(A)$ into $\mathcal{C}_\infty(\mathbb{R}^d)$, then the sub Markovianity of the perturbed semigroup can be verified using the positive maximum principle; this does not work any longer

if B has discontinuous coefficients, i.e. if we work on the space $\mathcal{B}_b(\mathbb{R}^d)$ of bounded measurable functions rather than $\mathcal{C}_\infty(\mathbb{R}^d)$. In fact, establishing the sub Markov property of the perturbed semigroup turns out to be rather delicate point. At the same time, being sub Markovian is crucial for the applications in the theory of stochastic processes, which we are interested in. To establish sub Markovianity, we will use an approximation argument which is of independent interest, e.g. it can be used to prove continuous dependence of solutions of certain martingale problems on the coefficients (see Theorem 7.3). Let us mention that our questions can be formulated equivalently in the context of pseudo-differential operators. Namely, the Lévy-type operator $\mathcal{A}$ can be seen alternatively as a pseudo-differential operator

$$(1.2) \qquad \mathcal{A}f(x) = -\int_{\mathbb{R}^d} q(x,\xi) e^{ix\cdot\xi}\hat{f}(\xi)\,d\xi, \qquad f \in \mathcal{C}_c^\infty(\mathbb{R}^d),\ x \in \mathbb{R}^d,$$

where $\hat{f}$ is the Fourier transform and

$$(1.3)\ \ q(x,\xi) := -ib(x)\cdot\xi + \frac{1}{2}\xi\cdot Q(x)\xi + \int_{\mathbb{R}^d\setminus\{0\}} (1 - e^{iy\cdot\xi} + iy\cdot\xi\mathbb{1}_{(0,1)}(|y|))\,\nu(x,dy)$$

is the so-called *symbol* of $\mathcal{A}$. We study under which conditions the pseudo-differential operator with symbol $q(x,\xi) + p(x,\xi)$ gives rise to a Feller semigroup if the symbol $p(x,\xi)$ of the perturbation B depends merely measurably on x.

If the symbol – or equivalently, the characteristics – of a Lévy-type operator satisfies suitable smoothness conditions, then general results from symbolic calculus show that the closure of $(\mathcal{A}, \mathcal{C}_c^\infty(\mathbb{R}^d))$ is the generator of a Feller semigroup, see e.g. [16, 20]. Under the milder assumption that the coefficients are Hölder continuous, the situation is already much more complicated and there is an immense literature on the question whether the Lévy-type operator $\mathcal{A}$ gives rise to a Feller semigroup, see e.g. [6, 16, 19, 20, 30] for a survey. None of these results applies in our framework since we are dealing with discontinuous coefficients. For the particular case that the diffusion coefficient Q is strictly elliptic, there are general results in the literature which allow discontinuous coefficients, see e.g. [19, Theorem 2.1.43] and also [54, 55] for processes on bounded domains. Moreover, the well-posedness of the $(\mathcal{A}, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem is well studied in this case, cf. [53]. No such general results are available if the operator has a vanishing diffusion component $Q \equiv 0$. There are some perturbation results for martingale problems which allow discontinuous coefficients, including bounded perturbations [13, Section 4.10] and perturbations of Lévy generators e.g. in [8, 28, 42, 43]. Note that already the existence of a solution to the martingale-problem can be highly non-trivial if the coefficients are discontinuous: While there are general existence results for Lévy-type operators with continuous coefficients, cf. [16, Theorem 3.15], there are no such results for the discontinuous framework; see [32] and the references therein. Let us point out that the existence of (unique) solutions to martingale problems associated with Lévy-type operators can be used to deduce the existence of (unique) weak solutions to Lévy-driven stochastic differential equations (SDEs)

$$dX_t = b(X_{t-})\,dt + \sigma(X_{t-})\,dL_t,$$

see Subsection 8.4 for details. In particular, drift(-type) perturbations of Lévy processes (i.e. $\sigma \equiv 1$) have been studied quite intensively, [44, 56, 8, 24, 25] to mention just a few classical and recent works.

The following general result on Feller semigroups with measurable lower order terms is obtained by combining our main results Theorem 6.4 and Theorem 7.2.

Theorem 1.1. *Let $(T(t))_{t\geq 0}$ be a Feller semigroup with generator A such that $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a core for A. Assume that there are $\rho \in (0,2)$ and $\varphi \in L^1(0,1)$ such that*

$$\|T(t)f\|_{\mathcal{C}_b^\rho} \leq \varphi(t)\|f\|_\infty, \qquad t \in (0,1),\ f \in \mathcal{B}_b(\mathbb{R}^d),$$

Let $\hat{B}$ be a Lévy-type operator

$$(1.4) \quad \hat{B}f(x) = b(x)\nabla f(x) + \int_{\mathbb{R}^d \setminus \{0\}} (f(x+y) - f(x) - \nabla f(x) \cdot y \mathbb{1}_{(0,1)}(|y|))\, \mu(x, dy),$$

where the drift b and the jumping kernels $\mu(x, dy)$ depend measurably on $x \in \mathbb{R}^d$. Assume that

$$\sup_{x \in \mathbb{R}^d} \left(|b(x)| + \int_{\mathbb{R}^d} \min\{1, |y|^\beta\} \, \mu(x, dy) \right) < \infty$$

for some $\beta \in [0, \rho)$ and that the compensated drift is zero in case that $\rho \le 1$. Then

    (a) *The bp-closure of $A + \hat{B}$ is the full generator of a Markovian $\mathcal{C}_b$-Feller semigroup S, i.e. S leaves $\mathcal{C}_b(\mathbb{R}^d)$ invariant and its restriction to that space is pointwise continuous. Moreover, S enjoys the strong Feller property;*

    (b) *The $(A + \hat{B}, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem is well-posed;*

    (c) *If additionally the tightness condition*

$$\lim_{R \to \infty} \sup_{x \in \mathbb{R}^d} \mu(x, \{y \in \mathbb{R}^d; |y| \ge R\}) = 0$$

    holds, then S is a $\mathcal{C}_\infty$-semigroup, i.e. S leaves $\mathcal{C}_\infty(\mathbb{R}^d)$ invariant and its restriction to that space is strongly continuous. Thus S is a Feller semigroup in this case.

Theorem 1.1 applies to a wide class of Feller generators A including strictly elliptic differential operators (see Section 8.1), generators of certain Lévy processes (see Section 8.2) and operators of variable order (see Section 8.3). Moreover, the well-posedness of the martingale problem in (b) yields uniqueness results for Lévy-driven SDEs, see Section 8.4.

This article is organized as follows. In Section 2, we recall some notions concerning kernel operators, in Section 3 those concerning semigroups and their generators. Here we also establish the connection between cores of of Feller generators and bp-cores of the corresponding full generators; this is important for our applications to martingale problems. In Section 4 we recall a perturbation result from [39] and prove our first main result (Theorem 4.4). Section 5 concerns convergence results for perturbed semigroups, which is crucial for establishing sub Markovianity in Theorem 1.1. Sections 2 – 5 concern abstract state spaces E rather than $\mathbb{R}^d$ and thus do not take the special structure of the operators $\mathcal{A}$ in (1.1) into account.

The remaining sections 6 – 8 apply the abstract results to Lévy-type operators as in (1.1). Section 6 concerns the actual perturbation result and Section 7 the corresponding martingale problem. The concluding Section 8 contains examples and discusses applications in the theory of stochastic differential equations.

Moreover, there are two appendices which contain results that are used in the proof of Theorem 6.4 and which we believe to be of independent interest.

2. Kernel operators

Throughout, E is a locally compact Polish space. We endow E with its Borel σ-algebra $\mathcal{B}(E)$. The spaces of bounded and measurable resp. bounded and continuous functions on E are denoted by $\mathcal{B}_b(E)$ and $\mathcal{C}_b(E)$ respectively whereas $\mathcal{C}_\infty(E)$ refers to the space of continuous functions vanishing at ∞, i.e. of those continuous functions $f : E \to \mathbb{R}$ such that for every $\varepsilon > 0$ we find a compact set $K \subset E$ with $|f(x)| \le \varepsilon$ for all $x \in E \setminus K$.

A *kernel* on E is a map $k : E \times \mathcal{B}(E) \to \mathbb{R}$ such that

    (i) the map $x \mapsto k(x, A)$ is measurable for every $A \in \mathcal{B}(E)$;

    (ii) $k(x, \cdot)$ is a (signed) measure for every $x \in E$;

    (iii) $\sup_{x \in E} |k|(x, E) < \infty$, where $|k|(x, \cdot)$ denotes the total variation of $k(x, \cdot)$.

If $k(x, \cdot)$ is a positive measure for every $x \in E$, then k is called a *positive kernel*; if every $k(x, \cdot)$ is a (sub) probability measure, k is a *(sub) Markovian kernel*.

To every kernel k on E, we can associate a bounded linear operator T on $\mathcal{B}_b(E)$ by setting

$$(2.1) \qquad (Tf)(x) := \int_E f(y) \, k(x, dy), \quad f \in \mathcal{B}_b(E), \ x \in E.$$

We call an operator T of this form a *kernel operator* on $\mathcal{B}_b(E)$. It turns out that a bounded, linear operator on $\mathcal{B}_b(E)$ is a kernel operator if, and only if, T is continuous with respect to the weak topology $\sigma := \sigma(\mathcal{B}_b(E), \mathcal{M}_b(E))$ induced by the space $\mathcal{M}_b(E)$ of bounded signed measures on E (see e.g. [38, Prop. 3.5]). For sequences, σ-convergence is nothing else than *bp-convergence* (bp is short for *bounded and pointwise*), i.e. $f_n \to f$ with respect to σ if, and only if, $\sup_{n \in \mathbb{N}} \|f_n\|_\infty < \infty$ and $f_n \to f$ pointwise. Indeed, by dominated convergence, bp-convergence implies σ-convergence; the converse follows from the uniform boundedness principle. We will write $\mathcal{L}(\mathcal{B}_b(E), \sigma)$ for the space of σ-continuous linear operators on E, i.e. kernel operators. Note that any such operator is automatically bounded. In what follows, we write $\rightharpoonup$ to indicate convergence with respect to σ while we reserve $\to$ to indicate convergence with respect to the supremum norm. If T is a bounded operator defined via (2.1) on $\mathcal{C}_b(E)$ or $\mathcal{C}_\infty(E)$, then T is also called a kernel operator (on $\mathcal{C}_b(E)$ resp. $\mathcal{C}_\infty(E)$); it can be extended in a unique way to a kernel operator on $\mathcal{B}_b(E)$. As is well known, every bounded operator on $\mathcal{C}_\infty(E)$ is a kernel operator and can thus be extended to a kernel operator on $\mathcal{B}_b(E)$, see e.g. [50, Cor. 21.12].

In applications, it is often of interest if a kernel operator on $\mathcal{B}_b(E)$ leaves one of the spaces $\mathcal{C}_b(E)$ or $\mathcal{C}_\infty(E)$ invariant. Of particular interest is the case where a kernel operator T maps $\mathcal{B}_b(E)$ into $\mathcal{C}_b(E)$. Such an operator is called *strong Feller operator*. For our perturbation results, it will be important to know under which conditions a kernel operator leaves the space

$$\mathcal{B}_\infty(E) := \big\{ f \in \mathcal{B}_b(E) : \forall \varepsilon > 0 \, \exists K \Subset E : |f(x)| \leq \varepsilon \, \forall \, x \in E \setminus K \big\}.$$

invariant. We present two results on this topic.

Lemma 2.1. *Let T be a kernel operator with associated kernel k. Then $T\mathcal{B}_\infty(E) \subset \mathcal{B}_\infty(E)$ if and only if $k(\cdot, K) \in \mathcal{B}_\infty(E)$ for every relatively compact sets K.*

Proof. As $T\mathbb{1}_K = k(\cdot, K)$ and $\mathbb{1}_K \in \mathcal{B}_\infty(E)$ for a relatively compact set K, the stated condition is certainly necessary. To prove its sufficiency, consider a function $f \in \mathcal{B}_\infty(E)$ with compact support S. We find a sequence of simple functions $\tilde{f}_n$ that converges uniformly to f. If we set $f_n := \tilde{f}_n \mathbb{1}_S$, then also the sequence f_n converges uniformly to f and, moreover, for $c \neq 0$ and every $n \in \mathbb{N}$ the set $\{f_n = c\}$ is relatively compact. Thus, $Tf_n \in \mathcal{B}_\infty(E)$. As $Tf_n \to Tf$ uniformly and since $\mathcal{B}_\infty(E)$ is closed with respect to the supremum norm, $Tf \in \mathcal{B}_\infty(E)$. The case of a general function $f \in \mathcal{B}_\infty(E)$ follows from this by approximating f uniformly with a sequence of measurable functions with compact support. $\square$

If the operator T is positive, then invariance of $\mathcal{C}_\infty(E)$ implies invariance of $\mathcal{B}_\infty(E)$.

Lemma 2.2. *Let T be a positive kernel operator with $T\mathcal{C}_\infty(E) \subset \mathcal{C}_\infty(E)$. Then we have $T\mathcal{B}_\infty(E) \subset \mathcal{B}_\infty(E)$. If, in addition, T has the strong Feller property, then $T\mathcal{B}_\infty(E) \subset \mathcal{C}_\infty(E)$.*

Proof. Let $0 \leq f \in \mathcal{B}_\infty(E)$. Given $\varepsilon > 0$, pick a compact set K such that $|f(x)| \leq \varepsilon/\|T\|$ for $x \in E \setminus K$. Then

$$0 \leq f \leq \|f\|_\infty \mathbb{1}_K + \frac{\varepsilon}{\|T\|} \mathbb{1}_{E \setminus K}$$

and, consequently,

$$0 \leq Tf \leq \|f\|_\infty k(\cdot, K) + \varepsilon.$$

Since $T\mathcal{C}_\infty(E) \subset \overline{T\mathcal{C}_\infty(E)}$, [20, Lem. 3.2.15] yields $k(\cdot, K) \in \mathcal{B}_\infty(E)$ and thus $Tf \in \mathcal{B}_\infty(E)$ follows as $\varepsilon > 0$ was arbitrary. The addendum follows from the identity $\mathcal{C}_\infty(E) = \mathcal{B}_\infty(E) \cap \mathcal{C}_b(E)$. $\qquad\square$

3. Semigroups, (pseudo)resolvents and generators

We now turn our attention to semigroups of kernel operators. A family $T = (T(t))_{t>0} \subset \mathcal{L}(\mathcal{B}_b(E), \sigma)$ is called *semigroup of kernel operators* if

(i) $T(t + s) = T(t)T(s)$ for all $s, t > 0$.
(ii) T is *exponentially bounded*, i.e. there are constants $\omega \in \mathbb{R}$, $M > 0$ such that $\|T(t)\| \leq Me^{\omega t}$. We say that T is *of type* (M, ω) to stress these constants.
(iii) The map $(t, x) \mapsto (T(t)f)(x)$ is measurable.

It is not difficult to see that if T is a semigroup of kernel operators of type (M, ω) then for every $\lambda \in \mathbb{C}$ with $\operatorname{Re} \lambda > 0$ there is an operator $R(\lambda) \in \mathcal{L}(\mathcal{B}_b(E), \sigma)$ such that

$$(3.1) \qquad \langle R(\lambda)f, \mu \rangle = \int_0^\infty e^{-\lambda t} \langle T(t)f, \mu \rangle \, dt$$

for all $f \in \mathcal{B}_b(E)$ and $\mu \in \mathcal{M}_b(E)$.

In the terms of [38, Def. 5.1], a semigroup of kernel operators is an integrable semigroup on the norming dual pair $(\mathcal{B}_b(E), \mathcal{M}_b(E))$. It turns out that the family $(R(\lambda))_{\operatorname{Re} \lambda > \omega}$ is a *pseudo-resolvent*, i.e. it satisfies the resolvent identity $R(\lambda) - R(\mu) = (\mu - \lambda)R(\lambda)R(\mu)$, cf. [38, Prop. 5.2]. For more details on pseudoresolvents we refer to [12, Sect. III.4.a].

In general, the family $(R(\lambda))_{\operatorname{Re} \lambda > \omega}$ does *not* consist of injective operators, and so it is not the resolvent of a (single-valued) operator. However, there is a *multivalued* operator $\hat{A}$ such that $R(\lambda) = (\lambda - \hat{A})^{-1}$ for $\operatorname{Re} \lambda > \omega$ (see [15, Appendix A] for more information concerning multivalued operators). We call $\hat{A}$ the *full generator* of the semigroup T. The full generator can be characterized equivalently as follows:

$$(f, g) \in \hat{A} \iff \forall t > 0 \,:\, T(t)f - f = \int_0^t T_s g \, ds,$$

cf. [38, Prop. 5.7]. In particular, our terminology is consistent with that used by Ethier–Kurtz [13, Sect. 1.5].

We now introduce additional properties that a semigroup of kernel operators might have. If T is of type $(1, 0)$, then T is called a *contraction* semigroup of kernel operators. If T is a contraction semigroup of kernel operators and every operator $T(t)$ is positive, we say that T is *sub Markovian*. If additionally $T(t)\mathbb{1} = \mathbb{1}$ for all $t > 0$, then T is called *Markovian*.

We will call T a $\mathcal{C}_\infty$-*semigroup* if $T(t)\mathcal{C}_\infty(E) \subset \mathcal{C}_\infty(E)$ and the restriction of T to $\mathcal{C}_\infty(E)$ is *strongly continuous*, i.e. for every $f \in \mathcal{C}_\infty(E)$ we have $T(t)f \to f$ with respect to $\|\cdot\|_\infty$ as $t \to 0$. A $\mathcal{C}_\infty$-semigroup that is also sub Markovian is called a *Feller semigroup*.

We will call T a $\mathcal{C}_b$-*semigroup* if $T(t)\mathcal{C}_b(E) \subset \mathcal{C}_b(E)$ and the restriction of T to $\mathcal{C}_b(E)$ is *stochastically continuous*, i.e. $T(t)f \rightharpoonup f$ as $t \to 0$ for $f \in \mathcal{C}_b(E)$. We note that, given sub Markovianity, this continuity condition can be equivalently characterized in terms of the associated kernels p_t by asking that $p_t(x, B(x, \varepsilon)) \to 1$ as $t \to 0$ for every $x \in E$ and $\varepsilon > 0$, see [20, Lem. 3.2.17]. A $\mathcal{C}_b$-semigroup that is also sub Markovian is called $\mathcal{C}_b$-*Feller semigroup*.

Let us point out that the above-defined objects are compatible with the classical notions. To wit: If T is a $\mathcal{C}_\infty$-semigroup, then the operator $R(\lambda)$ leaves the space $\mathcal{C}_\infty(E)$ invariant for $\operatorname{Re} \lambda > \omega$ and its restriction to that space is injective. Thus, the

restriction of $(R(\lambda))_{\mathrm{Re}\,\lambda>\omega}$ is the resolvent of a (single valued) operator A. Moreover, A is the *part* of $\hat{A}$ in $\mathcal{C}_\infty(E)$, i.e.

$$D(A) = \{u \in \mathcal{C}_\infty(E) : \exists f \in \mathcal{C}_\infty(E) \text{ s.t. } (u, f) \in \hat{A}\}, \qquad Au = f.$$

Identifying A with its graph, we have $A = \hat{A} \cap (\mathcal{C}_\infty(E) \times \mathcal{C}_\infty(E))$. By general semigroup theory, A coincides with the generator of the semigroup $T|_{\mathcal{C}_\infty(E)}$, defined as (norm-)derivative in 0.

Conversely, if we start with a strongly continuous semigroup on $\mathcal{C}_\infty(E)$, we may extend it to $\mathcal{B}_b(E)$ as any bounded linear operator on $\mathcal{C}_\infty(E)$ is a kernel operator. In fact, the extended semigroup is a semigroup of kernel operators in the sense defined above (this follows from [38, Lem. 6.1]). If the semigroup is sub Markovian, then the extension is also stochastically continuous, see [18, Lem. 4.8.7].

We next discuss how to recover the full generator $\hat{A}$ from the operator A. To that end, we make use of operator cores. We recall that a subset $D \subset D(A)$ is called a *core* for A if for every $u \in D(A)$ we find a sequence $(u_n) \subset D$ with $u_n \to u$ and $Au_n \to Au$. To use cores in the context of multivalued operators, we employ the notion of bp-convergence defined in Section 2. A set $M \subset \mathcal{B}_b(E)$ is called *bp-closed* if with every bp-convergent sequence, it also contains its limit. The *bp-closure* of a set M is the smallest bp-closed subset of $\mathcal{B}_b(E)$ that contains M. We use these notions, mutatis mutandis, also in $\mathcal{B}_b(E) \times \mathcal{B}_b(E)$. We call a subset $M \subset \hat{A}$ a *bp-core* if the bp-closure of M equals $\hat{A}$.

Lemma 3.1. *Let T be $\mathcal{C}_\infty$-semigroup with full generator $\hat{A}$ and let A be the part of $\hat{A}$ in $\mathcal{C}_\infty(E)$. Let D be a core for A. Then $\{(u, Au) : u \in D\}$ is a bp-core for $\hat{A}$.*

Proof. Let T be of type (M, ω) and fix $\lambda > \omega$. As D is a core for A, the set $(\lambda - A)D$ is $\|\cdot\|$-dense in $\mathcal{C}_\infty(E)$. It follows that the bp-closure of $(\lambda - A)D$ contains $\mathcal{C}_\infty(E)$ and thus equals all of $\mathcal{B}_b(E)$ (see Proposition A.1).

Put $S := \{(u, \lambda u - Au) : u \in D\}$. We note that $\hat{A}$, and hence also $\lambda - \hat{A}$, is σ-closed and thus, in particular, bp-closed. As $S \subset \lambda - \hat{A}$ it follows that the bp-closure $\bar{S}$ is a subset of $\lambda - \hat{A}$. Now define $C := \{f \in \mathcal{B}_b(E) : \exists u \text{ s.t. } (u, f) \in \bar{S}\}$. Then C is bp-closed.

Indeed, Let $(f_n) \subset C$ be bp-convergent to f. Pick u_n such that $(u_n, f_n) \in \bar{S}$. As $R(\lambda) = (\lambda - \hat{A})^{-1}$, we see that $(R(\lambda)f_n, f_n) \in \hat{B}$. It follows that $u_n - R(\lambda)f_n \in \ker(\lambda - \hat{A}) = \{0\}$, i.e. $u_n = R(\lambda)f_n$. By the σ-continuity of $R(\lambda)$, we have that $u_n \rightharpoonup u := R(\lambda)f$. This implies that $(u, f) \in \bar{S}$ whence $f \in C$, proving that C is bp-closed as claimed.

As C contains $(\lambda - A)D$, it follows that $C = \mathcal{B}_b(E)$. But this entails that $\bar{S} = \lambda - \hat{A}$. We have proved that $\lambda - \hat{A}$ is the bp-closure of S which is equivalent to the claim. $\qquad\square$

4. Perturbation of (strong) Feller semigroups

We begin by recalling a result from [39] concerning the perturbation of semigroups of kernel operators consisting of strong Feller operators. We will use the following set of assumptions.

Hypothesis 4.1. Let T be a $\mathcal{C}_b$-semigroup of type (M, ω) that consists of strong Feller operators. We denote by $\hat{A}$ the full generator of the semigroup and write $(R(\lambda))_{\mathrm{Re}\,\lambda>\omega}$ for the Laplace transform of T. Moreover, let $\hat{B} : D(\hat{B}) \to \mathcal{B}_b(E)$ be a single-valued linear operator with $D(\hat{A}) \subset D(\hat{B})$ which satisfies the following assumptions:

(i) For $t > 0$ the operator $\hat{B}T(t)$, initially defined on $D(\hat{A})$, has an extension to an operator in $\mathcal{L}(\mathcal{B}_b(E), \sigma)$ (which, by slight abuse of notation, we still denote by $\hat{B}T(t)$);

(ii) $\hat{B}R(\lambda) \in \mathcal{L}(\mathcal{B}_b(E), \sigma)$ for one/all $\lambda > \omega$;

(iii) the function $(t, x) \mapsto \hat{B}T(t)f(x)$ is measurable for $f \in \mathcal{B}_b(E)$;

(iv) there is a function φ which is integrable in a neighborhood of 0 such that $\|\hat{B}T(t)\| \leq \varphi(t)$ for $t > 0$.

Theorem 4.2. *([39, Thm. 3.3]) Assuming Hypothesis 4.1 the operator $\hat{A} + \hat{B}$ (defined on $D(\hat{A}) \subset D(\hat{B})$) is the full generator of a $\mathcal{C}_b$-semigroup S that consists of strong Feller operators. This semigroup satisfies for $t > 0$ the* Duhamel formula

$$(4.1) \qquad S(t)f = T(t)f + \int_0^t S(t - s)\hat{B}T(s)f\, ds.$$

Moreover, we can develop S in its Dyson–Phillips series*:*
(4.2)
$$S(t) = \sum_{n=0}^{\infty} S_n(t) \quad where \quad S_0(t) = T(t) \quad and \quad S_{n+1}(t)f = \int_0^t S_n(t - s)\hat{B}T(s)\, ds.$$

Here, all integrals have to be understood in the weak $\mathcal{M}_b(E)$-sense as in (3.1).

It is a rather obvious question whether the new semigroup S is a $\mathcal{C}_\infty$-semigroup if this is the case for T. Unfortunately, without further assumptions on $\hat{B}$ this is not the case.

Example 4.3. For $E = \mathbb{R}$, consider the heat semigroup T, given by

$$(T(t)f)(x) := \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} \exp\left(-\frac{(x - y)^2}{4t}\right) f(y)\, dy \quad (x \in \mathbb{R})$$

for $f \in \mathcal{B}_b(\mathbb{R})$. It is well known that T is a $\mathcal{C}_\infty$-semigroup which also consists of strong Feller operators. Now consider the operator $\hat{B} : f \mapsto f(0) \cdot \mathbb{1}_{\mathbb{R}}$. Then $\hat{B} \in \mathcal{L}(\mathcal{B}_b(\mathbb{R}), \sigma)$ (whence it satisfies Hypothesis 4.1, see [39, Ex. 3.4]). It follows that $\hat{B}T(s)f = c \cdot \mathbb{1}_{\mathbb{R}}$, for some constant $c = c(s, f)$. Consequently, $T(t - s)\hat{B}T(s)f = c$ for all $t \geq s$. If $0 < f \in \mathcal{C}_\infty(\mathbb{R})$ then $c_t(f) := \inf\{c(s, f) : 0 \leq s \leq t\} > 0$ for all $t > 0$ and we can infer from the Dyson–Phillips expansion (4.2) that the perturbed semigroup S does not leave $\mathcal{C}_\infty(\mathbb{R})$ invariant. Indeed, fixing $t > 0$, we see that $S(t)f \geq S_1(t)f \geq c_t(f)\mathbb{1}_{\mathbb{R}}$ and the latter does not vanish at infinity.

Theorem 4.4. *Assume in addition to Hypothesis 4.1 that T is a positive $\mathcal{C}_\infty$-semigroup.*

(a) *If $\hat{D}$ is a bp-core for $\hat{A}$, then $\{(u, f + \hat{B}u) : (u, f) \in \hat{D}\}$ is a bp-core for $\hat{A} + \hat{B}$.*

(b) *If $\hat{B}T(t)\mathcal{B}_\infty(E) \subset \mathcal{B}_\infty(E)$ for every $t > 0$, then S is a $\mathcal{C}_\infty$-semigroup.*

Proof. (a) Let $\hat{D}$ be a bp-core for $\hat{A}$ and denote by $\bar{S}$ the bp-closure of $\{(u, f + \hat{B}u), (u, f) \in \hat{D}\}$. As $\hat{A} + \hat{B}$ is a full generator, it is σ-closed and thus, in particular, bp-closed and it follows that $\bar{S} \subset \hat{A} + \hat{B}$. Note that this entails that for $(u, g) \in \bar{S}$ we have $u \in D(\hat{A})$. We now define $W := \{(u, g - \hat{B}u) : (u, g) \in \bar{S}\}$. Clearly, $\hat{D} \subset W \subset \hat{A}$. We prove that W is bp-closed. To that end, let $(u_n, g_n - \hat{B}u_n)$ be a sequence in W (thus $(u_n, g_n) \in \bar{S}$) such that u_n bp-converges to u and $f_n := g_n - \hat{B}u_n$ bp-converges to f. Then $(u_n, f_n) \subset \bar{A}$ bp-converges to (u, f). By the bp-closedness of $\hat{A}$, we have $(u, f) \in \hat{A}$. Picking $\lambda > \omega$, this is equivalent to $u = R(\lambda)(\lambda u - f)$ and the same equality holds with u/f replaced by u_n/f_n. Using that $\hat{B}R(\lambda)$ is σ-continuous by (ii) in Hypothesis 4.1, we infer

$$\hat{B}u_n = \hat{B}R(\lambda)(\lambda u_n - f_n) \rightharpoonup \hat{B}R(\lambda)(\lambda u - f) = \hat{B}u.$$

Setting $g := f + \hat{B}u$, the sequence $(u_n, g_n) \subset \bar{S}$ bp-converges to (u, g). As $\bar{S}$ is bp-closed, $(u, g) \in \bar{S}$ and thus $(u, f) \in W$. Since $\hat{D}$ is a bp-core for $\hat{A}$, it

follows that $W = \hat{A}$. But this is equivalent to $\bar{S} = \hat{A} + \hat{B}$ which means that $\{(u, f + \hat{B}u), (u, f) \in \hat{D}\}$ is indeed a bp-core for $\hat{A} + \hat{B}$.

(b) We first prove by induction that the operators $S_n(t)$ appearing in the Dyson–Phillips series (4.2) map $\mathcal{B}_\infty(E)$ to $\mathcal{C}_\infty(E)$. In particular, they leave the space $\mathcal{B}_\infty(E)$ invariant.

For $n = 0$ this follows from Lemma 2.2 and our assumption that T is a positive $\mathcal{C}_\infty$-semigroup that enjoys the strong Feller property.

Let us assume that, for some $n \in \mathbb{N}$, we have already proved that $S_n(t)$ maps $\mathcal{B}_\infty(E)$ to $\mathcal{C}_\infty(E)$ for every $t > 0$. Fix $f \in \mathcal{B}_\infty(E)$ and $t > 0$. By our assumption, we have $\hat{B}T(s)f \in \mathcal{B}_\infty(E)$ for every $s \in (0, t)$. By induction hypothesis, $S_n(t - s)$ maps $\mathcal{B}_\infty(E)$ to $\mathcal{C}_\infty(E)$, whence $S_n(t - s)\hat{B}T(s)f \in \mathcal{C}_\infty(E)$ for $s \in (0, t)$. Since $\mathcal{C}_\infty(E)$ is separable with dual space $\mathcal{M}_b(E)$, the Pettis measurability theorem implies that the function $s \mapsto S_n(t - s)\hat{B}T(s)f$ is Bochner integrable which, in turn, implies that $S_{n+1}(t)f = \int_0^t S_n(t - s)\hat{B}T(s)f \, ds \in \mathcal{C}_\infty(E)$.

Note that the series in (4.2) converges for small t even in the operator norm. From this it follows that for small t we have $S(t)f \in \mathcal{C}_\infty(E)$. As f was arbitrary, $S(t)\mathcal{C}_\infty(E) \subset \mathcal{C}_\infty(E)$ for small t and hence, by the semigroup law, for all $t > 0$. The strong continuity of the restriction of S to $\mathcal{C}_\infty(E)$ follows from that of T, the identity (4.1) and the observation that the integral in that formula converges to 0 in operator norm as $t \mapsto 0$. $\qquad\square$

5. Resolvent convergence of the perturbed operators

In this section, we will consider a sequence $\hat{B}_n$ of perturbations that converges, in a certain sense, to the operator $\hat{B}$. We want to know under which assumptions and in which sense the pseudoresolvents $(\lambda - (\hat{A} + \hat{B}_n))^{-1}$ converge to $(\lambda - (\hat{A} + \hat{B}))^{-1}$ as $n \to \infty$. Besides being interesting in its own right, such a convergence result will also allow us to establish dissipativity and/or resolvent positivity for large classes of operators with measurable coefficients.

The key to prove our convergence result is the following Lemma, which is taken from [39, Lem. 3.8].

Lemma 5.1. *Assuming Hypothesis 4.1, we have $\|\hat{B}R(\lambda)\| \to 0$ as $\lambda \to \infty$. Moreover, if λ is so large that $\|\hat{B}R(\lambda)\| < 1$, then*

$$(\lambda - (\hat{A} + \hat{B}))^{-1} = R(\lambda) \sum_{k=0}^{\infty} (\hat{B}R(\lambda))^k,$$

where the latter series converges in operator norm.

Proposition 5.2. *Assume Hypothesis 4.1 and let $(\hat{B}_n, D(\hat{B}_n))$ be an operator such that Hypothesis 4.1 is also fulfilled with $\hat{B}$ replaced by $\hat{B}_n$ (but the same semigroup T). Moreover assume that*

(i) $\sup_{n \in \mathbb{N}} \|\hat{B}_n R(\lambda)\| \to 0$ *as $\lambda \to \infty$,*

(ii) *for every $h \in \mathcal{B}_b(E)$ and large enough λ we have $\hat{B}_n R(\lambda)h \rightharpoonup \hat{B}R(\lambda)h$, and*

(iii) *whenever $(h_n) \subset \mathcal{B}_b(E)$ satisfies $h_n \rightharpoonup 0$, we have $\hat{B}_n R(\lambda)h_n \rightharpoonup 0$ for large enough λ.*

In this case, for large enough λ we have

(5.1) $$(\lambda - (\hat{A} + \hat{B}_n))^{-1}f \rightharpoonup (\lambda - (\hat{A} + \hat{B}))^{-1}f$$

for all $f \in \mathcal{B}_b(E)$.

Proof. By assumption (i), we can pick λ_0 so large that for $\lambda \geq \lambda_0$ we have $\|\hat{B}_n R(\lambda)\| \leq 1/2$ for every $n \in \mathbb{N}$ and $\lambda \geq \lambda_0$. By Lemma 5.1, we have

$$(5.2) \qquad (\lambda - (\hat{A} + \hat{B}_n))^{-1} = R(\lambda) \sum_{k=0}^{\infty} (\hat{B}_n R(\lambda))^k.$$

Now let $f \in \mathcal{B}_b(E)$ and $\varepsilon > 0$. We fix $N \in \mathbb{N}$ such that $\sum_{k=N+1}^{\infty} \|(\hat{B}_n R(\lambda))^k f\| \leq \varepsilon$ for all $n \in \mathbb{N}$.

We prove that, for every $k \in \mathbb{N}$, we have $(\hat{B}_n R(\lambda))^k f \rightharpoonup (\hat{B} R(\lambda))^k f$ as $n \to \infty$.

For $k = 0$, there is nothing to prove and for $k = 1$ this is exactly our assumption (ii). Assume that we know this convergence to be true for some $k \in \mathbb{N}$. We set $g_n := (\hat{B}_n R(\lambda))^k f$ and $g := (\hat{B} R(\lambda))^k f$. Then

$$(\hat{B}_n R(\lambda))^{k+1} f - (\hat{B} R(\lambda))^{k+1} f = \hat{B}_n R(\lambda)(g_n - g) + (\hat{B}_n R(\lambda)g - \hat{B} R(\lambda)g) \rightharpoonup 0.$$

Indeed, by induction hypothesis $g_n - g \rightharpoonup 0$ so that the first term converges weakly to 0 by our assumption (iii) and that the second term converges weakly to 0 follows from the case $k = 1$.

Alltogether, we find that $\sum_{k=0}^{\infty} (\hat{B}_n R(\lambda))^k f \rightharpoonup \sum_{k=0}^{\infty} (\hat{B} R(\lambda))^k f$. Given that the operator $R(\lambda)$ is σ-continuous, the representation (5.2) and Lemma 5.1 yield (5.1). $\qquad \square$

If the operator $R(\lambda)$ is positive (which is always the case in our main application) we obtain stronger convergence results.

Corollary 5.3. *In the situation of Proposition 5.2, assume additionally that $R(\lambda) \geq 0$ for large enough λ. Then the convergence in (5.1) is not only with respect to σ, but even uniform on compact subsets of E.*

Proof. Fix $f \in \mathcal{B}_b(E)$ and put $g_n := \sum_{k=0}^{\infty} (\hat{B}_n R(\lambda))^k f$ and $g := \sum_{k=0}^{\infty} (\hat{B} R(\lambda))^k f$. Fixing $m \in \mathbb{N}$, we put $h_m := \sup_{n \geq m} |g_n - g|$. We have seen in the proof of Proposition 5.2 that $g_n \rightharpoonup g$ and, consequently, $h_m \downarrow 0$ pointwise.

As $R(\lambda)$ is positive, we find for $n \geq m$

$$\left| (\lambda - (\hat{A} + \hat{B}_n))^{-1} f - (\lambda - (\hat{A} + \hat{B}))^{-1} f \right| = |R(\lambda)g_n - R(\lambda)g|$$
$$\leq R(\lambda)|g_n - g|$$
$$\leq R(\lambda)h_m.$$

The latter converges to 0 as m, hence n, tends to ∞. However, as $R(\lambda)$ is positive this is even a monotone convergence and thus the convergence is uniform on compact subsets of E by Dini's theorem. $\qquad \square$

It is natural to ask whether the perturbed semigroups also converge. In general, weak convergence of the resolvents does not imply weak convergence of the corresponding semigroups, see [11] for a concrete example. In Section 7, we will use the theory of martingale problems to establish a convergence result for semigroups associated with Lévy-type operators.

We can also prove that certain features of the pseudoresolvents are stable under the convergence described in Proposition 5.2.

Corollary 5.4. *In the situation of Proposition 5.2, the following holds true.*

(i) *Let $\lambda \in \mathbb{C}$ with $\operatorname{Re} \lambda > 0$ be in the resolvent set of $\hat{A} + \hat{B}_n$ and assume that $\|(\lambda - (\hat{A} + \hat{B}_n))^{-1}\| \leq (\operatorname{Re} \lambda)^{-1}$ for all $n \in \mathbb{N}$. Then λ is in the resolvent set of $\hat{A} + \hat{B}$ and $\|(\lambda - (\hat{A} + \hat{B}))^{-1}\| \leq (\operatorname{Re} \lambda)^{-1}$. In this case, the semigroup S, generated by $\hat{A} + \hat{B}$, is contractive.*

(ii) *Suppose that for $\lambda \geq \lambda_0$ we have $(\lambda + (\hat{A} + \hat{B}_n))^{-1} \geq 0$ for all $n \in \mathbb{N}$. Then also $(\lambda + (\hat{A} + \hat{B}))^{-1} \geq 0$ for $\lambda \geq \lambda_0$. In this case, the semigroup S generated by $\hat{A} + \hat{B}$ is positive.*

Proof. We put $F_n(\lambda) := (\lambda - (\hat{A} + \hat{B}_n))^{-1}$ and $F(\lambda) := (\lambda - (\hat{A} + \hat{B}))^{-1}$ whenever these are defined.

(i) The sequence F_n is an (operator-valued) holomorphic function that is, by assumption, locally bounded on $\{\operatorname{Re}\lambda > 0\}$. Fixing $f \in \mathcal{B}_b(E)$ and $\mu \in \mathcal{M}_b(E)$, the same is true for the scalar function $\varphi_n : \lambda \mapsto \varphi_n(\lambda) := \langle F_n(\lambda)f, \mu \rangle$. By Proposition 5.2, for real λ that are large enough, $\varphi_n(\lambda)$ converges to $\varphi(\lambda) := \langle F(\lambda)f, \mu \rangle$. It follows from Vitali's theorem (see [2, Thm. 2.1]) that φ has a holomorphic extension to the set $\{\operatorname{Re}\lambda > 0\}$ and φ_n converges locally uniformly to φ. As f and μ are arbitrary, it follows from a vector-valued analytic extension theorem [2, Thm. 3.5] that F has an analytic continuation to $\{\operatorname{Re}\lambda > 0\}$ which proves $\{\operatorname{Re}\lambda > 0\}$ is contained in the resolvent set of $\hat{A} + \hat{B}$. Moreover

$$|\langle F(\lambda)f, \mu \rangle| = \lim_{n\to\infty} |\langle F_n(\lambda)f, \mu \rangle| \le \limsup_{n\to\infty} (\operatorname{Re}\lambda)^{-1}\|f\|\|\mu\|$$

for every $\operatorname{Re}\lambda > 0$. This implies the estimate for the Laplace transform. As for the contractivity of the semigroup, we note that for $f \in \mathcal{C}_b(E)$ and $\mathcal{M}_b(E)$ the orbit $t \mapsto \langle S(t)f, \mu \rangle$ is continuous whence the (scalar) Post–Widder inversion formula (see [1, Thm. 1.7.7]) yields

$$\langle S(t)f, \mu \rangle = \lim_{n\to\infty} \left\langle \left(\frac{n}{t}\right)^n \left(\frac{n}{t} - \hat{A} - \hat{B}\right)^{-n} f, \mu \right\rangle.$$

By the resolvent estimate, the absolute value of the right-hand side is at most $\|f\|\|\mu\|$ and it follows that S is contractive on $\mathcal{C}_b(E)$. But then it is also contractive on $\mathcal{B}_b(E)$.

(ii) For $0 \le f \in \mathcal{B}_b(E)$, $0 \le \mu \in \mathcal{M}(E)$ and $\lambda \ge 0$, we have

$$\langle F(\lambda)f, \mu \rangle = \lim_{n\to\infty} \langle F_n(\lambda)f, \mu \rangle \ge 0$$

as for $\lambda \ge \lambda_0$ we have $F_n(\lambda) \ge 0$ for all $n \in N$. As f and μ were arbitrary, this proves $F(\lambda) \ge 0$ for $\lambda \ge \lambda_0$. As above, we can infer positivity of the semigroup S from this by means of the Post–Widder inversion formula. □

6. Perturbation of strong Feller semigroups by Lévy-type operators

In this section, we work in the Euclidean setting, $E = \mathbb{R}^d$, and consider perturbations of strong Feller semigroups by a class of integro-differential operators. To that end, we will make more concrete assumptions on our initial semigroup T which are tailor-made for this situation.

Hypothesis 6.1. Assume that $T = (T(t))_{t\ge 0}$ is a $\mathcal{C}_b$-semigroup of type (M, ω) with state space $E = \mathbb{R}^d$ that consists of strong Feller operators. Moreover, the following conditions hold for some $\rho > 0$:

(i) $T(t)f \in \mathcal{C}_b^\rho(\mathbb{R}^d)$ for all $t > 0$ and $f \in \mathcal{B}_b(\mathbb{R}^d)$;
(ii) There is a function $\varphi \in L^1(0, 1)$ such that

$$\|T(t)f\|_{\mathcal{C}_b^\rho} \le \varphi(t)\|f\|_\infty$$

for all $t \in (0, 1)$ and $f \in \mathcal{B}_b(\mathbb{R}^d)$;

As before, we will denote the Laplace transform of our semigroup T by $(R(\lambda))_{\operatorname{Re}\lambda > \omega}$ and the full generator of T by $\hat{A}$. For future reference, we note the following consequence of Hypothesis 6.1.

Lemma 6.2. *Assume Hypothesis 6.1. Then, for every $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$, we have $R(\lambda)f \in \mathcal{C}_b^\rho(\mathbb{R}^d)$ for all $f \in \mathcal{B}_b(\mathbb{R}^d)$ and*

$$\|R(\lambda)f\|_{\mathcal{C}_b^\rho} \le C(\lambda)\|f\|_\infty, \quad f \in \mathcal{B}_b(\mathbb{R}^d),$$

for a constant $C(\lambda)$ with $C(\lambda) \to 0$ as $\operatorname{Re}\lambda \to \infty$. In particular, $D(\hat{A}) \subset \mathcal{C}_b^\rho(\mathbb{R}^d)$.

Proof. For $r > 0$ it follows from the exponential boundedness of the semigroup and Hypothesis 6.1(ii) that

$$\|T(1+r)f\|_{\mathcal{C}_b^\rho} \leq \|T(1)\|_{\mathcal{L}(\mathcal{C}_b^\rho, \mathcal{B}_b)} \|T(r)\| \, \|f\|_\infty \leq M\varphi(1)e^{\omega r} \|f\|_\infty.$$

Setting $\psi(t) = \varphi(t)$ for $t \in (0,1)$ and $\psi(t) = M\varphi(1)e^{\omega(t-1)}$ for $t \geq 1$, we get

(6.1) $$\|T(t)f\|_{\mathcal{C}_b^\rho} \leq \psi(t)\|f\|_\infty, \quad t > 0, \ f \in \mathcal{B}_b(\mathbb{R}^d).$$

Now fix $f \in \mathcal{B}_b(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. Since $\|T(t)\| \leq Me^{\omega t}$ and $t \mapsto T(t)f(x)$ is Borel measurable, the integral $R(\lambda)f(x) = \int_{(0,\infty)} e^{-\lambda t}T(t)f(x)\, dt$ exists as a Lebesgue integral for any λ with $\operatorname{Re}\lambda > \omega$. If $\rho \in (0,1)$, then it is immediate from (6.1) and the estimate

$$|R(\lambda)f(x) - R(\lambda)f(z)| \leq \int_0^\infty e^{-\lambda t}|T(t)f(x) - T(t)f(z)|\, dt$$

that $\|R(\lambda)f\|_{\mathcal{C}_b^\rho} \leq C(\lambda)\|f\|_\infty$, where $C(\lambda) := \int_0^\infty \psi(t)e^{-\lambda t}\, dt$. If $\rho \geq 1$, then the differentiation lemma for parameter-depend integrals, see e.g. [50, Thm. 12.5] or [34, Prop. A.1], implies that $x \mapsto R(\lambda)f(x)$ is differentiable and

$$\frac{\partial}{\partial x_i}R(\lambda)f(x) = \int_0^\infty e^{-\lambda t}\frac{\partial}{\partial x_i}T(t)f(x)\, dt$$

for all $\operatorname{Re}\lambda > \omega$. Thus, by (6.1), $\|R(\lambda)f\|_{\mathcal{C}_b^\rho} \leq C(\lambda)\|f\|_\infty$ with $C(\lambda)$ as before. By dominated convergence, $C(\lambda) \to 0$ as $\operatorname{Re}\lambda \to \infty$. $\qquad\square$

We now introduce the integro-differential operator $\hat{B}$ that we will consider as a perturbation. We fix a function χ such that $\mathbb{1}_{B(0,1)} \leq \chi \leq \mathbb{1}_{B(0,2)}$ and put

(6.2) $$\hat{B}f(x) = b(x)\cdot\nabla f(x) + \int_{\mathbb{R}^d\setminus\{0\}} (f(x+y) - f(x) - y\cdot\nabla f(x)\chi(y))\, \mu(x,dy).$$

We make the following standing assumption.

Hypothesis 6.3. The function $b : \mathbb{R}^d \to \mathbb{R}^d$ and the kernel $\mu : \mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d \setminus \{0\}) \to [0,\infty]$ satisfy

    (i) b is Borel measurable and bounded;
    (ii) $x \mapsto \int_{\mathbb{R}^d\setminus\{0\}} f(y)\, \mu(x,dy)$ is Borel measurable for every $f \in \mathcal{C}_c(\mathbb{R}^d \setminus \{0\})$.
    (iii) There is a constant $\beta \in (0,2)$ such that

(6.3) $$\|\mu\|_\beta := \sup_{x\in\mathbb{R}^d} \left(\int_{\mathbb{R}^d\setminus\{0\}} \min\{|y|^\beta, 1\}\, \mu(x,dy) \right) < \infty.$$

        Moreover, β is strictly smaller than the constant ρ from Hypothesis 6.1.
    (iv) If $\rho \leq 1$, then the *compensated drift* $b(\cdot) - \int_{\mathbb{R}^d\setminus\{0\}} y\chi(y)\, \mu(\cdot,dy)$ is identically zero.

Occasionally we will additionally assume the following *tightness assumption*

(Ti) $$\sup_{x\in\mathbb{R}^d} \mu(x, \{|y| > R\}) \to 0 \quad \text{as} \quad R \to \infty.$$

Let us comment briefly on these assumptions. Assumption (iii) implies that the above operator $\hat{B}$ is well-defined on $\mathcal{C}_b^\rho(\mathbb{R}^d)$. To see this, let us first assume that $\rho > 1$. Then clearly the local part $b(\cdot)\nabla f(\cdot)$ of $\hat{B}$ is well-defined on this space. As for the integral part, the elementary estimate

$$|f(x+y) - f(x) - y\cdot\nabla f(x)\chi(|y|)| \leq 2\|f\|_{\mathcal{C}_b^\rho(\mathbb{R}^d)}\min\{1, |y|^\rho\}$$

implies that $\|\hat{B}f\|_\infty \leq 2\|\mu\|_\beta\|f\|_{\mathcal{C}_b^\rho}$. In the case $\rho \leq 1$, the additional assumption (iv) entails that the operator $\hat{B}$ simplifies to

$$\hat{B}f(x) = \int_{\mathbb{R}^d\setminus\{0\}} (f(x+y) - f(x))\, \mu(x,dy)$$

and we can argue similarly, making use of the Hölder continuity of f. The assumption $\beta < \rho$ is important to ensure compatibility (in the sense of Hypothesis 4.1) of the perturbation $\hat{B}$ with the semigroup T satisfying Hypothesis 6.1. The tightness condition (Ti) plays an important role in proving that the perturbed semigroup leaves the space $\mathcal{C}_\infty(\mathbb{R}^d)$ invariant. We note that (Ti) holds if the real part of the symbol

$$p(x,\xi) := -ib(x) \cdot \xi + \int_{\mathbb{R}^d \setminus \{0\}} (1 - e^{iy\cdot\xi} + iy \cdot \xi \chi(|y|)) \, \mu(x, dy)$$

of the operator $\hat{B}$ is equicontinuous at $\xi = 0$, i.e.

(6.4) $$\lim_{|\xi| \to 0} \sup_{x \in \mathbb{R}^d} |\operatorname{Re} p(x, \xi)| = 0,$$

see [48, (proof of) Thm. 4.4].

We can now formulate the main result of this section.

Theorem 6.4. *Assume Hypotheses 6.1 and 6.3. Then the following hold true.*
- *(a) The operator $\hat{A} + \hat{B}$ is the full generator of a $\mathcal{C}_b$-semigroup $S = (S(t))_{t\geq 0}$ that consists of strong Feller operators.*
- *(b) If T consists of sub Markovian operators then so does S.*
- *(c) Assume additionally the tightness assumption (Ti), that T is a positive $\mathcal{C}_\infty$-semigroup and that the test functions $\mathcal{C}_c^\infty(\mathbb{R}^d)$ form a core for the generator A on $\mathcal{C}_\infty(\mathbb{R}^d)$. Then S is a $\mathcal{C}_\infty$-semigroup.*

Remark 6.5. In the situation of Theorem 6.4(b), the perturbed semigroup S is Markovian if and only if the unperturbed semigroup T is Markovian. Indeed, we only need to check if the semigroup is conservative, i.e. if the semigroup leaves the function $\mathbb{1}$ invariant. By [38, Prop. 5.9] this is the case if and only if $\mathbb{1}$ belongs to the kernel of the full generator. As $\hat{B}\mathbb{1} = 0$, we see that $(\hat{A} + \hat{B})\mathbb{1} = 0$ is equivalent to $\hat{A}\mathbb{1} = 0$.

Example 6.6. Without the tightness assumption (Ti), it is not true in general that the perturbed semigroup S is a $\mathcal{C}_\infty$-semigroup even if this is the case for T. This can be seen for example by considering in dimension $d = 1$ the kernel $\mu(x, \cdot) := \delta_{-x}$, which obviously does not satisfy (Ti). Choose $b = 0$ and let T be the heat semigroup. Then $\hat{B}f = f(0)\mathbb{1}_\mathbb{R} - f$ and the perturbed semigroup S is (up to a rescaling by the factor e^{-t}) the semigroup from Example 4.3, which is not a $\mathcal{C}_\infty$-semigroup.

Corollary 6.7. *Assume Hypotheses 6.1 and 6.3 and that T is a positive $\mathcal{C}_\infty$-semigroup. If $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a core for the generator A of $T|_{\mathcal{C}_\infty(\mathbb{R}^d)}$, then $\{(f, Af + \hat{B}f) : f \in \mathcal{C}_c^\infty(\mathbb{R}^d)\}$ is a bp-core for the full generator $\hat{A} + \hat{B}$ of S.*

Proof. If $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a core for the generator on $\mathcal{C}_\infty(\mathbb{R}^d)$, then, by Lemma 3.1, $\{(f, Af); f \in \mathcal{C}_c^\infty(\mathbb{R}^d)\}$ is a bp-core for $\hat{A}$. Thus, by Theorem 4.4(a), $\{(f, Af + \hat{B}f); f \in \mathcal{C}_c^\infty(R^d)\}$ is a bp-core for the full generator $\hat{A} + \hat{B}$. $\square$

We now turn to the proof of Theorem 6.4.

Proof of parts (a) and (c) of Theorem 6.4. (a) Let us first prove prove that $\hat{B}f \in \mathcal{B}_b(\mathbb{R}^d)$ for $f \in \mathcal{C}_b^\rho(\mathbb{R}^d)$. This is obvious for the local part of $\hat{B}f$ so we focus on the integral part. From Hypothesis 6.3(ii), it follows that for any Borel subset S of $\mathbb{R}^d \setminus \{0\}$ the map $x \mapsto \int_{\mathbb{R}^d \setminus \{0\}} \mathbb{1}_S(y) \, \mu(x, dy)$ is measurable. But then so is $x \mapsto \int_{\mathbb{R}^d \setminus \{0\}} \mathbb{1}_{R \times S}(x, y) \, \mu(x, dy)$ for Borel sets $R \subset \mathbb{R}^d$ and $S \subset \mathbb{R}^d \setminus \{0\}$. An application of the monotone class theorem yields that $x \mapsto \int_{\mathbb{R}^d \setminus \{0\}} \mathbb{1}_A(x, y) \, \mu(x, dy)$ is Borel measurable for any $A \in \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ and thus, by the sombrero lemma and dominated convergence, $x \mapsto \int_{\mathbb{R}^d \setminus \{0\}} g(x, y) \, \mu(x, dy)$ is Borel measurable

for any function g which is $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^d \setminus \{0\})$-measurable and satisfies $|g(x, y)| \leq K \min\{1, |y|^\rho\}$ for some constant $K > 0$. This readily gives that $\hat{B}f$ is Borel measurable. Note that the discussion before Theorem 6.4 shows that, in fact, $\hat{B}$ is a bounded linear operator from $\mathcal{C}_b^\rho(\mathbb{R}^d)$ to $\mathcal{B}_b(\mathbb{R}^d)$.

We now verify Hypothesis 4.1, then part (a) immediately follows from Theorem 4.2. As for condition (i), the boundedness of $\hat{B}T(t)$ follows from Hypothesis 6.1(i) and the boundedness of $\hat{B}$ proved above. As for the σ-continuity, let a bounded sequence $(f_n)_{n \in \mathbb{N}} \subset \mathcal{B}_b(\mathbb{R}^d)$ be given such that $f_n \to f$ pointwise. By Hypothesis 6.1(i), $T(t)f_n$ is bounded in $\mathcal{C}_b^\rho(\mathbb{R}^d)$. By the Arzelà–Ascoli theorem, passing to a subsequence, we may and shall assume that $T(t)f_n$ converges locally uniformly to some continuous function. By the σ-continuity of $T(t)$ the sequence $T(t)f_n$ converges pointwise to $T(t)f$, whence the only possible limit is $T(t)f$ and it follows that $T(t)f_n \to T(t)f$ locally uniformly. Note that in the case where $\rho > 1$ we also obtain that $\nabla T(t)f_n \to \nabla T(t)f$ locally uniformly. From this it is immediate that the local part of $\hat{B}f_n$ converges pointwise to that of $\hat{B}f$. As for the integral part this convergence follows from dominated convergence, noting that as a consequence of the uniform boundedness in $\mathcal{C}_b^\rho(\mathbb{R}^d)$ we find an integrable majorant of the form $C \min\{1, |y|^\rho\}$.

The proof of Condition 4.1(ii) is similar, taking Lemma 6.2 into account. As for Condition 4.1(iii), we note that $(t, x) \mapsto T(t)f(x)$ is Borel measurable for every $f \in \mathcal{B}_b(\mathbb{R}^d)$ and using a reasoning similar to that in the first part of this proof, it follows that $(t, x) \mapsto \hat{B}T(t)f(x)$ is Borel measurable.

Condition 4.1(iv) is an immediate consequence of Hypothesis 6.1(ii) and the boundedness of $\hat{B}$.

(c) To prove this part, we use Theorem 4.4. We thus have to prove that $\hat{B}T(t)\mathcal{B}_\infty(\mathbb{R}^d) \subset \mathcal{B}_\infty(\mathbb{R}^d)$. We will only consider the case $\rho > 1$; for $\rho \leq 1$ the reasoning is a bit simpler because all terms involving the gradient vanish by Hypothesis 6.3(iv). Take $f \in \mathcal{C}_b^\rho(\mathbb{R}^d) \cap \mathcal{C}_c(\mathbb{R}^d)$ and choose $R > 0$ such that the support of f is contained in the ball $B(0, R)$. Taking into account that in this part we assume the tightness condition (Ti), we find that if $|x| > R + r$ for some $r > 0$, then

$$|\hat{B}f(x)| = \left| \int f(x + y)\, \mu(x, dy) \right| \leq \|f\|_\infty \int_{\{|y| \geq r\}} \mu(x, dy) \to 0$$

as $r \to \infty$. Thus, $\hat{B}f \in \mathcal{B}_\infty(\mathbb{R}^d)$.

Now let $g \in \mathcal{C}_\infty(\mathbb{R}^d)$ be given. We denote the generator of $T|_{\mathcal{C}_\infty(\mathbb{R}^d)}$ by A. As $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a core, we find a sequence $(f_n) \subset \mathcal{C}_c^\infty(\mathbb{R}^d)$ such that $g_n := \lambda f_n - A f_n \to g$, see [12, Ex. II.1.15]. It then follows that $f_n \to f = R(\lambda)g \in D(A)$. By the above, $\hat{B}R(\lambda)g_n = \hat{B}f_n \in \mathcal{B}_\infty(\mathbb{R}^d)$. As $\hat{B}R(\lambda)$ is $\|\cdot\|_\infty$-continuous, it follows that $\hat{B}f = \hat{B}R(\lambda)g = \lim_{n \to \infty} \hat{B}R(\lambda)g_n = \lim_{n \to \infty} \hat{B}f_n$ also belongs to $\mathcal{B}_\infty(\mathbb{R}^d)$. Consequently, $\hat{B}f \in \mathcal{B}_\infty(\mathbb{R}^d)$ whenever $f \in D(A)$.

As $T(t)D(A) \subset D(A)$ for every $t > 0$, it follows that $\hat{B}T(t)f \in \mathcal{B}_\infty(\mathbb{R}^d)$ for $f \in D(A)$. But as $D(A)$ is dense in $\mathcal{C}_\infty(\mathbb{R}^d)$ and $\hat{B}T(t)$ is $\|\cdot\|$-continuous, this is also true for $f \in \mathcal{C}_\infty(\mathbb{R}^d)$. □

Remark 6.8. Theorem 6.4(c) remains valid also without the assumption that $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a core for $T|_{\mathcal{C}_\infty(\mathbb{R}^d)}$, provided we assume that $b \in \mathcal{B}_\infty(\mathbb{R}^d; \mathbb{R}^d)$.

To see this, pick a sequence $\phi_n \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ with $\mathbb{1}_{B(0,n)} \leq \phi_n \leq \mathbb{1}_{B(0,2n)}$ such that $\sup_n \|\phi_n\|_{\mathcal{C}_b^2(\mathbb{R}^d)} < \infty$. Given $f \in \mathcal{C}_b^\rho(\mathbb{R}^d) \cap \mathcal{C}_\infty(\mathbb{R}^d)$, we put $f_n := f\phi_n$. Then $f_n \in \mathcal{C}_b^\rho(\mathbb{R}^d) \cap \mathcal{C}_c(\mathbb{R}^d)$ and f_n is a bounded sequence that converges to f uniformly on $\mathbb{R}^d$; moreover, ∇f_n is bounded and converges to ∇f locally uniformly.

As $b \in \mathcal{B}_\infty(\mathbb{R}^d; \mathbb{R}^d)$ and the sequence ∇f_n is uniformly bounded, it is easy to see that $b\nabla f_n \to b\nabla f$ with respect to $\|\cdot\|_\infty$. Making use of the boundedness of (f_n) in $\mathcal{C}_b^\rho(\mathbb{R}^d)$ and the tightness assumption (Ti) we can prove that also the nonlocal part converges with respect to $\|\cdot\|_\infty$. . In conclusion, $\hat{B}f_n \to \hat{B}f$ with respect to $\|\cdot\|_\infty$. As the proof of Theorem 6.4(b) shows that $\hat{B}f_n \in \mathcal{B}_\infty(\mathbb{R}^d)$, the same is true for its uniform limit $\hat{B}f$. This proves that $\hat{B}f \in \mathcal{B}_\infty(\mathbb{R}^d)$ for all $f \in \mathcal{C}_b^\rho(\mathbb{R}^d) \cap \mathcal{C}_\infty(\mathbb{R}^d)$ which is enough to ensure that $\hat{B}T(t)$ maps $\mathcal{C}_\infty(\mathbb{R}^d)$ to $\mathcal{B}_\infty(\mathbb{R}^d)$.

To prove part (b) of Theorem 6.4, we will employ the convergence result of Proposition 5.2. As this form of resolvent convergence is of independent interest, we formulate a separate lemma.

Lemma 6.9. *Assume Hypothesis 6.1 and let sequences (b_n) and (μ_n) be given that satisfy the assumptions in Hypothesis 6.3 for a common constant $\beta \in (0, \min\{2, \rho\})$. Denote the associated operators (see* (6.2)*) by $\hat{B}_n$. Moreover, we are given functions b and ν such that*

(i) *$b_n(x) \to b(x)$ and $\mu_n(x, \cdot) \to \mu(x, \cdot)$ vaguely for every $x \in \mathbb{R}^d$,*
(ii) *$\sup_{n \in \mathbb{N}} (\|b_n\|_\infty + \|\mu_n\|_\beta) < \infty$ with $\|\cdot\|_\beta$ defined in (6.3).*
(iii) *$\sup_{x \in K} \sup_{n \in \mathbb{N}} \mu_n(x, \{|y| > R\}) \to 0$ as $R \to \infty$ for $K \Subset \mathbb{R}^d$.*

Then, denoting the operator associated to b and μ by $\hat{B}$, we find

$$(\lambda - (\hat{A} + \hat{B}_n))^{-1} f \rightharpoonup (\lambda - (\hat{A} + \hat{B}))^{-1} f, \qquad f \in \mathcal{B}_b(\mathbb{R}^d),$$

for large enough λ.

Proof. We give the proof only for $\rho > 1$. If $\rho \leq 1$, then by Hypothesis 6.3(iv) all terms involving the gradient vanish in the below computations; apart from that the reasoning is analogous.

Without loss of generality, we may assume $\beta \geq 1$ (otherwise consider $\tilde{\beta} := \max\{1, \beta\}$). We first note that also the functions b and ν satisfy Hypothesis 6.3 (see Lemma A.4), whence Theorem 6.4(a) yields that $\hat{A} + \hat{B}$ and, for every $n \in \mathbb{N}$, the operator $\hat{A} + \hat{B}_n$ is the full generator of a $\mathcal{C}_b$-semigroup. The discussion following Hypothesis 6.3 shows that the operator $\hat{B}_n$ defines a bounded linear operator from $\mathcal{C}_b^\beta(\mathbb{R}^d)$ to $\mathcal{B}_b(\mathbb{R}^d)$. Using our assumption (ii) above, we actually see that there is a constant $K > 0$ such that

$$(6.5) \qquad \|\hat{B}f\|_\infty + \sup_{n \in \mathbb{N}} \|\hat{B}_n f\|_\infty \leq K \|f\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)}, \qquad f \in \mathcal{C}_b^\beta(\mathbb{R}^d).$$

We now check the assumptions of Proposition 5.2.

Assumption (i): By Lemma 6.2, we have

$$\|\hat{B}_n R(\lambda)\| \leq \|\hat{B}_n\|_{\mathcal{L}(\mathcal{C}_b^\beta, \mathcal{B}_b)} \|R(\lambda)\|_{\mathcal{L}(\mathcal{B}_b, \mathcal{C}_b^\beta)} \leq K C(\lambda) \to 0$$

as $\lambda \to \infty$.

Assumption (ii): In view of Lemma 6.2, it suffices to prove that $\hat{B}_n f \rightharpoonup \hat{B}f$ for every $f \in \mathcal{C}_b^\rho(\mathbb{R}^d)$, so fix $f \in \mathcal{C}_b^\rho(\mathbb{R}^d)$. Because of (6.5), it actually suffices to show pointwise convergence. Since b_n converges pointwise to b, it is clear that the local part of $\hat{B}_n f$ converges to the local part of $\hat{B}f$; therefore we assume in the following that the local part is zero. Write

$$\hat{B}_n f = I_n^{(r)} f + J_n^{(r,R)} f + K_n^{(R)} f,$$

where

$$I_n^{(r)} f(x) := \int_{\{|y| \leq r\}} (f(x+y) - f(x) - \nabla f(x) \cdot y)\, \mu_n(x, dy)$$

$$J_n^{(r,R)} f(x) := \int_{\{r < |y| < R\}} (f(x+y) - f(x) - \nabla f(x) \cdot y \chi(y))\, \mu_n(x, dy)$$

$$K_n^{(R)} f(x) := \int_{\{|y| \geq R\}} (f(x+y) - f(x))\, \mu_n(x, dy)$$

for $r \in (0,1)$ and $R \in (1, \infty)$. We decompose the operator $\hat{B}$ in the same way, and we study the convergence for each of the terms separately. By another application of Taylor's formula,

$$\sup_{n \in \mathbb{N}} |I_n^{(r)} f(x)| \leq \|f\|_{\mathcal{C}_b^\rho(\mathbb{R}^d)} \sup_{n \in \mathbb{N}} \int_{\{|y| \leq r\}} |y|^\rho\, \mu_n(x, dy)$$

As $\rho > \beta$, it follows from (ii) that the right-hand side is finite and converges to 0 (uniformly in x) as $r \to 0$. As $\|\mu\|_\beta < \infty$, an analogous estimate holds for $I^{(r)} f$. Using

$$|K_n^{(R)} f(x)| \leq 2\|f\|_\infty \int_{\{|y| \geq R\}} \mu_n(x, dy)$$

and the corresponding estimate for $K^{(R)} f$, assumption (iii) implies that $|K_n^{(R)} f(x)| + |K^{(R)} f(x)| \to 0$ uniformly in $n \in \mathbb{N}$ and x in compact subsets of $\mathbb{R}^d$ as $R \to \infty$. As f and ∇f are continuous, the vague convergence $\mu_n(x) \to \mu(x)$ entails that $J_n^{r,R} f(x) \to J^{r,R} f(x)$ for any $r, R > 0$ with $\mu(x, \{|y| = r\}) = 0$ and $\nu(x, \{|y| = R\}) = 0$. Since there are for each fixed x at most countably many radii with $\mu(x, \{|y| = r\}) > 0$, we can let $r \to 0$ and $R \to \infty$ along suitable sequences to deduce that $\hat{B}_n f(x) \to \hat{B} f(x)$.

Assumption (iii): Making use of Lemma 6.2 again, we see that it suffices to prove that $\hat{B}_n u_n \rightharpoonup 0$ for any sequence $(u_n)_{n \in \mathbb{N}}$ with $u_n \rightharpoonup 0$ and $\sup_{n \in \mathbb{N}} \|u_n\|_{\mathcal{C}_b^\rho} < \infty$.

Fix such a sequence. Then $\sup_{n \in \mathbb{N}} \|\hat{B}_n u_n\|_\infty < \infty$ is immediate from (6.5) and it only remains to prove pointwise convergence. Pick $\rho' \in (\beta, \rho)$. By Taylor's formula,

$$|\hat{B}_n u_n(x)| \leq \|b_n\|_\infty |\nabla u_n(x)| + \|u_n\|_{\mathcal{C}_b^{\rho'}(B[x,R])} \int_{\{|y| \leq R\}} |y|^{\rho'} \mu_n(x, dy)$$

$$+ 2\|u_n\|_\infty \int_{\{|y| > R\}} \mu_n(x, dy)$$

for any $R > 1$; here $B[x, r]$ is the closed ball around x with radius R. Using (ii), the fact that $\rho > 1$ and that $(b_n)_{n \in \mathbb{N}}$ is bounded, we find that given $\varepsilon > 0$ there is some $R \gg 1$ such that

$$|\hat{B}_n u_n(x)| \leq C \|u_n\|_{\mathcal{C}_b^{\rho'}(B[x,R])} + \varepsilon$$

for a finite constant $C = C(R)$. In order to deduce that the left-hand side converges to 0, we use the subsequence principle. As (u_n) is bounded in $\mathcal{C}_b^\rho$, by compact embedding, there is a subsequence $(u_{n''})_{n'' \in \mathbb{N}}$ which converges on the closed ball $B[x, r]$ with respect to the $C^{\rho'}$-norm. Since we know that $u_n \to 0$ pointwise, the C^ρ-limit is also zero and so

$$\limsup_{n'' \to \infty} |\hat{B}_{n''} u_{n''}(x)| \leq \varepsilon.$$

Hence, $\limsup_{n \to \infty} |\hat{B}_n u_n(x)| \leq \varepsilon$. As $\varepsilon > 0$ is arbitrary, we conclude that $\hat{B}_n u_n(x) \to 0$ for all $x \in \mathbb{R}^d$.

Now Proposition 5.2 yields the claim. $\square$

We can now finish the proof of Theorem 6.4

Proof of Theorem 6.4(b). Let us first only consider kernels μ satisfying

$$(6.6) \qquad \sup_{x \in K} \mu(x, \{|y| > R\}) \to 0 \quad \text{as} \quad R \to \infty \qquad \text{for all } K \Subset \mathbb{R}^d.$$

Given b and μ as in Hypothesis 6.3, we write $\hat{B}_{b,\mu}$ for the operator defined via (6.2) with these particular coefficients and $S_{b,\mu} = (S_{b,\mu}(t))_{t \geq 0}$ for the semigroup generated by the operator $\hat{A} + \hat{B}_{b,\mu}$, which exists by part (a) of Theorem 6.4. Denote by F the family of pairs (b, μ) for which the assertion holds, i.e.

$$F = \{(b, \mu) \text{ satisfying } (6.6) \text{ and Hypothesis } 6.3 \; : \; S_{b,\mu} \text{ is sub Markovian}\}.$$

It follows from Lemma 6.9 and Corollary 5.4 that F is bp-closed. Therefore, it suffices to show that $(b, \mu) \in F$ for any $b \in \mathcal{C}_c(\mathbb{R}^d; \mathbb{R}^d)$ and any $\mu \in \mathcal{C}_c\big(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})\big)$ with $\|\mu\|_\beta < \infty$ and (6.6). This is a consequence of the known fact that $\mathcal{C}_c(\mathbb{R}^d; \mathbb{R}^d)$ is bp-dense in $\mathcal{B}_b(\mathbb{R}^d; \mathbb{R}^d)$ and an analogous result for measures which we establish in Appendix A; see in particular Corollary A.6 (with $g(y) := \min\{|y|^\beta, 1\}$).

So fix $b \in \mathcal{C}_c(\mathbb{R}^d; \mathbb{R}^d)$ and $\mu \in \mathcal{C}_c\big(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})\big)$ with (6.6). Then, in particular, Theorem 6.4(b) and Remark 6.8 show that $S_{b,\mu}$ is a $\mathcal{C}_\infty$-semigroup; let us denote the generator of the restriction of $S_{b,\mu}$ to $\mathcal{C}_\infty(\mathbb{R}^d)$ by $L_{b,\mu}$.

It was seen in the proof of Theorem 6.4(b) that for $f \in \mathcal{C}_b^\rho(\mathbb{R}^d) \cap \mathcal{C}_\infty(\mathbb{R}^d)$ we have $\hat{B}_{b,\nu} f \in \mathcal{B}_\infty(\mathbb{R}^d)$. Moreover, $\|\mu\|_\beta < \infty$ and (6.6) give

$$(6.7) \qquad \limsup_{R \to \infty} \sup_{x \in K} \int_{\{|y| > R\}} \mu(x, dy) = 0 \qquad \lim_{r \to 0} \sup_{x \in K} \int_{\{|y| \leq r\}} |y|^2 \, \mu(x, dy) = 0,$$

for all $K \Subset \mathbb{R}^d$. This entails continuity of $\hat{B}_{b,\nu} f$, see Theorem B.1 (note that we can assume without loss of generality that χ is smooth by modifying the drift term accordingly). Hence, $\hat{B}_{b,\mu} f \in \mathcal{C}_\infty(\mathbb{R}^d)$.

Denoting by A the generator of the restriction of T to $\mathcal{C}_\infty(\mathbb{R}^d)$, we may consider $L := A + \hat{B}_{b,\mu}$, defined on $D(A)$, as an operator on the space $\mathcal{C}_\infty(\mathbb{R}^d)$. We note that $\hat{B}_{b,\mu} R(\lambda) \mathcal{C}_\infty(\mathbb{R}^d) \subset \mathcal{C}_\infty(\mathbb{R}^d)$. and from this and Lemma 5.1 it follows that, for large enough λ, the operator $\lambda - L$ is surjective.

As is well known (see e.g. [13, Thm. 4.2.2]) a strongly continuous semigroup on $\mathcal{C}_\infty(\mathbb{R}^d)$ is sub Markovian if and only if its generator satisfies the positive maximum principle. By assumption, this is certainly true for A. However, if $f \in D(A)$ satisfies $f(x_0) = \max\{f(x) : x \in \mathbb{R}^d\} \geq 0$, then $f(x_0 + y) - f(y) \geq 0$ for all $y \in \mathbb{R}^d$; if $\rho > 1$, we also see that $\nabla f(x_0) = 0$. It follows from $D(A) \subseteq \mathcal{C}_b^\rho(\mathbb{R}^d)$, cf. Lemma 6.2, and the definition of the operator $\hat{B}_{b,\mu}$ and that $\hat{B}_{b,\mu} f(x_0) \geq 0$. This shows that $L = A + \hat{B}_{b,\mu}$ satisfies the positive maximum principle.

By [13, Thm. 4.2.2], L generates a sub Markovian semigroup on $\mathcal{C}_\infty(\mathbb{R}^d)$. However, $L_{b,\nu}$ is obviously an extension of L, so we must have $L = L_{b,\nu}$ and thus the restriction of S to $\mathcal{C}_\infty(\mathbb{R}^d)$ is sub Markovian and so is then also S itself. This shows that $(b, \nu) \in F$ and finishes the proof in this case.

To remove the tightness assumption (6.6), we consider the 'small' and the 'large' jumps created by μ separately. To that end, we put

$$(6.8) \qquad \mu_s(x, dy) := \mathbb{1}_{\{|y| \leq 1\}} \mu(x, dy) \qquad \text{and} \qquad \mu_l(x, dy) := \mathbb{1}_{\{|y| > 1\}} \mu(x, dy).$$

We then split the operator $\hat{B} = \hat{B}_1 + \hat{B}_2$ where, in the notation above, $\hat{B}_1 = B_{b,\mu_s}$ and $\hat{B}_2 = \hat{B}_{0,\mu_l}$. As $\mu_s(x, \{|y| > 1\}) \equiv 0$, we can apply the above to infer that $\hat{A} + \hat{B}_1$ is the full generator of a sub Markovian semigroup. Noting that $\hat{B}_2$ is a bounded operator and taking the special structure of this operator into account, the results of [13, Sect. 4.10] (see in particular Exercise 3 on p. 261) yield that also $\hat{A} + \hat{B} = \hat{A} + \hat{B}_1 + \hat{B}_2$ is the full generator of a sub Markovian semigroup. $\qquad \square$

7. Application: Martingale problems

We apply the results from the previous section to study perturbations of martingale problems associated with Lévy-type operators. We do not strive for full generality; we rather would like to illustrate how our results can be applied in this setting. We will make the following assumptions which ensure that T and the perturbed semigroup S are Feller semigroups, provided the tightness assumption (Ti) is satisfied. Note, however, that (Ti) is not assumed in this section.

Hypothesis 7.1. Assume that $T = (T(t))_{t\geq 0}$ is a semigroup of kernel operators such that

(i) $(T(t))_{t\geq 0}$ satisfies Hypothesis 6.1 for some constant $\rho > 0$,
(ii) $(T(t))_{t\geq 0}$ is a Markovian $\mathcal{C}_\infty$-semigroup,
(iii) the test functions $\mathcal{C}_c^\infty(\mathbb{R}^d)$ are a core for the generator A of $T|_{\mathcal{C}_\infty(\mathbb{R}^d)}$.

It follows from general theory (see, e.g., [46, Sect. 3.2]) that we can associate Markov process (actually a Feller process) with càdlàg paths with the semigroup T, i.e. there is a Markov process whose transition semigroup equals T.

There is another connection between stochastic processes and (Feller) semigroups via *martingale problems*. Let us briefly recall the relevant notions. By $\mathcal{D}([0,\infty);\mathbb{R}^d)$ we denote the *Skorohod space*, i.e. the space of càdlàg functions $\omega : [0,\infty) \to \mathbb{R}^d$; see [5] for more information. Given a (possibly multi-valued) operator L with domain $D(L)$, a set $D \subset D(L)$ and a measure μ on $\mathbb{R}^d$, a measure $\mathbb{P}$ on $\mathcal{D}([0,\infty);\mathbb{R}^d)$ is called *a solution to the $(L, D; \mu)$-martingale problem* if (i) $\mathbb{P}(\omega_0 \in A) = \mu(A)$ for all Borel sets $A \subset \mathbb{R}^d$ and (ii) the process

$$M_t^{(f,g)}(\omega) := f(\omega_t) - f(\omega_0) - \int_0^t g(\omega_s)\, ds, \quad w \in \mathcal{D}[0,\infty),$$

is a martingale under $\mathbb{P}$ with respect to the canonical filtration $(\sigma(\omega_s : s \leq t))_{t\geq 0}$ for any $(f, g) \in L$ with $f \in D$. In case that L is single valued on D, we have $g = Lf$.

We say that *uniqueness in law* holds for the (L, D)-martingale problem if any two solutions $\mathbb{P}_1$, $\mathbb{P}_2$ with identical initial distribution satisfy $\mathbb{P}_1 = \mathbb{P}_2$. The (L, D)-martingale problem is *well-posed* if for every $x \in \mathbb{R}^d$ there exists a unique solution $\mathbb{P}^x$ to the (L, D, δ_x)-martingale problem.

If (X_t) is a Markov Process associated with the semigroup T, then its distribution solves the $(\hat{A}, D(\hat{A}))$-martingale problem for the full generator $\hat{A}$ of T, see [13, Prop. 4.1.7]. Conversely, if the (L, D)-martingale problem is well-posed, then under each measure $\mathbb{P}^x$, the canonical process $(\omega_t)_{t\geq 0}$ is a Markov process [13, Thm. 4.4.2]; if (L, D) is, in a sense, 'rich enough' to determine a semigroup uniquely, this semigroup is the transition semigroup of the process, see [13, Thm. 4.4.2].

Theorem 7.2. *Let $(T(t))_{t\geq 0}$ be a semigroup satisfying Hypothesis 7.1 with full generator $\hat{A}$ and denote by A the generator of $T|_{\mathcal{C}_\infty(\mathbb{R}^d)}$. If $\hat{B}$ is a Lévy-type operator*

$$\hat{B}f(x) = b(x)\nabla f(x) + \int_{\mathbb{R}^d\setminus\{0\}} (f(x+y) - f(x) - \nabla f(x) \cdot y\chi(y))\, \mu(x, dy)$$

with $b(x)$ and $\nu(x, dy)$ satisfying Hypothesis 6.3, then the $(A+\hat{B}, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem is well-posed.

Proof. By Theorem 6.4, there is a Markovian $\mathcal{C}_b$-semigroup $S = (S(t))_{t\geq 0}$ with full generator $\hat{A} + \hat{B}$. By Corollary 6.7, $\{(f, Af + \hat{B}f) : f \in \mathcal{C}_c^\infty(\mathbb{R}^d)\}$ is a bp-core for $\hat{A}+\hat{B}$. By [13, Prop. 4.3.1], a measure $\mathbb{P}$ on $\mathcal{D}([0, \infty); \mathbb{R}^d)$ solves the $(A+\hat{B}, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem if and only if it solves the $\hat{A} + \hat{B}$-martingale problem. It follows from [13, Thm. 4.4.1], applied with $L = \overline{D(\hat{A})}$, which certainly includes $\mathcal{C}_\infty(\mathbb{R}^d)$ and is thus separating, that uniqueness holds for the $\hat{A} + \hat{B}$-martingale problem (and thus for the $(A + \hat{B}, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem).

It thus remains to establish existence of solutions. To that end, we first additionally assume that the tightness assumption (Ti) is satisfied. In this case, Theorem 6.4 yields that S is actually a Feller semigroup and as is well known, see, e.g. [27, Theorem 4.10.3], this implies well-posedness for the martingale problem for the generator of $S|_{\mathcal{C}_\infty}$. In particular, there is a solution to the $(A + \hat{B}, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem. To remove the tightness condition, we proceed as in the proof of Theorem 6.4(b), i.e. we write $\hat{B} = \hat{B}_1 + \hat{B}_2$ for a bounded operator $\hat{B}_2$ and a Lévy-type operator $\hat{B}_1$, whose jumping kernels are supported in the unit ball, cf. (6.8). The existence of a solution to the $(A + \hat{B}_1, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem is clear from the first part. Since $\hat{B}_2$ is a bounded perturbation, existence of a solution to the martingale problem for $\hat{A} + \hat{B} = \hat{A} + \hat{B}_1 + \hat{B}_2$ follows from [13, Prop. 4.10.2]. □

We can now also prove continuous dependence of the solutions of our martingale problem on the 'coefficients' b and μ. To that end, we make the same assumptions as in Lemma 6.9. As a direct consequence, we obtain a convergence result for semigroups, which fits well to our earlier result on the resolvent convergence, cf. Proposition 5.2 and Corollary 5.3.

Theorem 7.3. *Let $(T(t))_{t \geq 0}$ be a semigroup satisfying Hypothesis 7.1 for some constant $\rho > 0$ and denote by $\hat{A}$ its full generator. Let $(b_n)_{n \in \mathbb{N} \cup \{\infty\}}$ and $(\mu_n)_{n \in \mathbb{N} \cup \{\infty\}}$ satisfy Hypothesis 6.3 for a common constant $\beta \in (0, \min\{2, \rho\})$. Denote the associated Lévy-type operators by $\hat{B}_n$ for $n \in \mathbb{N} \cup \{\infty\}$. Assume that*

    (i) *$b_n(x) \to b_\infty(x)$ and $\mu_n(x, \cdot) \to \mu_\infty(x, \cdot)$ vaguely for every $x \in \mathbb{R}^d$,*
    (ii) *$\sup_{n \in \mathbb{N}} (\|b_n\|_\infty + \|\mu_n\|_\beta) < \infty$ with $\| \cdot \|_\beta$ defined in (6.3).*
    (iii) *$\sup_{x \in K} \sup_{n \in \mathbb{N}} \mu_n(x, \{|y| > R\}) \to 0$ as $R \to \infty$ for $K \Subset \mathbb{R}^d$.*

Then, if $\mathbb{P}_n$ solves the $(A + \hat{B}_n, \mathcal{C}_c^\infty(\mathbb{R}^d), \delta_{x_n})$-martingale problem and $\mathbb{P}_\infty$ solves the $(A + \hat{B}_\infty, \mathcal{C}_c^\infty(\mathbb{R}^d), \delta_x)$-martingale problem for some sequence $x_n \to x$, we have weak convergence $\mathbb{P}_n \rightharpoonup \mathbb{P}_\infty$.

In particular, if $S_n(t)$ and $S_\infty(t)$ denote the perturbed semigroups with generator $\hat{A} + \hat{B}_n$ and $\hat{A} + \hat{B}_\infty$, respectively, then $S_n(t)f \to S_\infty(t)f$ pointwise for every $f \in \mathcal{C}_b(\mathbb{R}^d)$.

Proof. It was seen in the proof of Lemma 6.9 that under the assumptions above $\hat{B}_n u \rightharpoonup \hat{B}_\infty u$ for all $u \in D(\hat{A})$. Thus, if $(u, f) \in \hat{A}$, we have $(u, f + \hat{B}_n u) \in \hat{A} + \hat{B}_n$ and this bp-converges to $(u, f + \hat{B}_\infty u)$. In particular, the sequence $g_n := f + \hat{B}_n u$ is uniformly bounded. By assumption,

$$M_t^{(u,g_n)}(\omega) := u(\omega_t) - u(\omega_0) + \int_0^t g_n(\omega_s) \, ds$$

is a $\mathbb{P}_n$-martingale. Combining [13, Thm. 4.9.4] and [13, Cor. 3.9.3], it follows that the sequence $(\mathbb{P}_n)$ is tight. For an alternative argument yielding the tightness, see [32, Corollary 3.9].

We will invoke the subsequence principle, see e.g. [5, Thm. 2.6], to prove that $\mathbb{P}_n \to \mathbb{P}$ weakly, where $\mathbb{P}$ is the distribution of X.

To that end, observe that as a further consequence of Lemma 6.9, we have

$$R_n(\lambda)f := (\lambda - (\hat{A} + \hat{B}_n))^{-1}f \rightharpoonup R_\infty(\lambda)f := (\lambda - (\hat{A} + \hat{B}))^{-1}f, \quad f \in \mathcal{B}_b(\mathbb{R}^d),$$

for large enough λ. In fact, Corollary 5.3 implies that the convergence is actually uniform on compact sets.

Now, take any subsequence of $(\mathbb{P}_n)_{n \in \mathbb{N}}$, then, by tightness, there is a further subsequence converging to some measure, say, $\mathbb{Q}$. For simplicity of notation, we denote the convergent subsequence also by $\mathbb{P}_n$. For fixed $f \in \mathcal{B}_b(\mathbb{R}^d)$, $s \leq t$,

$0 \leq r_1 < \ldots < r_k \leq s$ and $h_j \in \mathcal{C}_b(\mathbb{R}^d)$, define

$$\Phi_n(\omega) := \left[R_n(\lambda)f(\omega_t) - R_n(\lambda)f(\omega_s) - \int_s^t (\lambda R_n(\lambda)f - f)(\omega_r)\,dr \right] \prod_{j=1}^k h_j(\omega_{r_j}),$$

where $n \in \mathbb{N} \cup \{\infty\}$. Since $R_n(\lambda)f \in D(\hat{A} + \hat{B}_n)$ and $\mathbb{P}_n$ solves the martingale problem for $\hat{A} + \hat{B}_n$, we find that $\mathbb{E}_{\mathbb{P}_n}(\Phi_n) := \int \Phi_n \, d\mathbb{P}_n = 0$ for all $n \in \mathbb{N}$. We claim that $\mathbb{E}_{\mathbb{Q}}(\Phi_\infty) = 0$. Because of the weak convergence, it holds that $\mathbb{E}_{\mathbb{P}_n}(\Phi_\infty) \to \mathbb{E}_{\mathbb{Q}}(\Phi_\infty)$. If we can show that

(7.1) $$\lim_{n \to \infty} |\mathbb{E}_{\mathbb{P}_n}(\Phi_n - \Phi_\infty)| = 0,$$

then it follows that $\mathbb{E}_{\mathbb{Q}}(\Phi_\infty) = \lim_n \mathbb{E}_{\mathbb{P}_n}(\Phi_n) = 0$. To prove (7.1), we note that, by the uniform boundedness of the resolvents, $M := \sup_{n \in \mathbb{N}} \|\Phi_n\|_\infty + \|\Phi_\infty\|_\infty < \infty$. By tightness, given $\varepsilon > 0$ we find $K \Subset \mathcal{D}([0, \infty); \mathbb{R}^d)$ such that $\mathbb{P}_n(K^c) \leq \varepsilon$ for all $n \in \mathbb{N}$. Note that $R := \sup_{y \in K} \sup_{s \leq t} |\omega_s| < \infty$. So, by the locally uniform convergence of the resolvents, we find $N \in \mathbb{N}$ such that

$$|R_n(\lambda)f(\omega_r) - R_\infty(\lambda)f(\omega_r)| \leq \varepsilon$$

for all $n \geq N$, $r \in [0, t]$ and $y \in K$. Thus, For some constant $C > 0$, we have

$$|\mathbb{E}_{\mathbb{P}_n}(\Phi_n - \Phi_\infty)| \leq C\varepsilon + M\varepsilon$$

for all $n \geq N$. Hence,

$$\limsup_{n \to \infty} |\mathbb{E}_{\mathbb{P}_n}(\Phi_n - \Phi_\infty)| \leq (M + C)\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this finishes the proof of (7.1) and shows that $\mathbb{E}_{\mathbb{Q}}(\Phi_\infty) = 0$. This implies that

$$M_t^{(f)}(\omega) := R_\infty(\lambda)f(\omega_t) - \int_0^t (\lambda R_\infty(\lambda)f - f)(\omega_r)\,dr, \quad \omega \in \mathcal{D}([0, \infty); \mathbb{R}^d),$$

is a martingale with respect to $\mathbb{Q}$ for any $f \in \mathcal{B}_b(\mathbb{R}^d)$. Noting that

$$\{(R_\infty(\lambda)f, \lambda R_\infty(\lambda)f - f); f \in \mathcal{B}_b(\mathbb{R}^d)\} = \hat{A} + \hat{B},$$

and $\mathbb{Q}(\omega_0 \in A) = \delta_x(A)$ this means that $\mathbb{Q}$ is a solution to the $(\hat{A} + \hat{B}, D(\hat{A}), \delta_x)$-martingale problem and thus, by well-posedness, $\mathbb{Q} = \mathbb{P}_\infty$. Hence, by the subsequence principle, $\mathbb{P}_n \to \mathbb{P}_\infty$ weakly. $\qquad \square$

8. EXAMPLES

Our perturbation results for Feller semigroups and martingale problems require two sets of assumptions: one on the original semigroup T (cf. Hypothesis 6.1 resp. 7.1) and one on the perturbation $\hat{B}$ (cf. Hypothesis 6.3). The latter ensures that $\hat{B}$ is indeed a (lower order) perturbation of the generator of $(T(t))_{t \geq 0}$ and involves the parameter $\rho \in (0, 2)$ from Hypothesis 6.1, which characterizes the regularizing properties of T. Given this parameter ρ, the conditions on $\hat{B}$ are typically easy to check, and so the main work is to verify the assumptions on T; in particular, the regularity estimate

$$\|T(t)f\|_{\mathcal{C}_b^\rho} \leq \varphi(t)\|f\|_\infty, \qquad t \in (0, 1), \ f \in \mathcal{B}_b(\mathbb{R}^d),$$

for some function $\varphi \in L^1(0, 1)$. For a brief summary of some of our main results we refer to Theorem 1.1. In this section, we present examples of semigroups satisfying our assumptions and give some applications, e.g. in the theory of stochastic differential equations.

8.1. **Differential operators.** Let $\mathcal{A}$ be a second order differential operators on $\mathbb{R}^d$ of the form

$$(8.1) \qquad \mathcal{A}f(x) := \operatorname{tr}(Q(x)\nabla^2 f(x))$$

for $Q(x) = (q_{ij}(x)) \in \mathbb{R}^{d \times d}$. Our next result gives conditions which ensure that the associated semigroup satisfies the assumptions of our main results.

Proposition 8.1. *Assume that the coefficients q_{ij} are bounded, Hölder continuous, symmetric (i.e. $q_{ij} = q_{ji}$) and strictly elliptic in the sense that there exists a constant $\eta > 0$ such that*

$$\sum_{i,j=1}^{d} q_{ij}(x)\xi_i\xi_j \geq \eta|\xi|^2, \quad x, \xi \in \mathbb{R}^d.$$

Then the semigroup T associated with (8.1) satisfies Hypothesis 6.1 and 7.1 with arbitrary $\rho \in (0, 2)$, and so Theorem 1.1 is applicable (with any $\rho \in (0, 2)$).

Proof. First of all, we note that there is a Markovian Feller semigroup T associated with (8.1); this follows from [9, Theorem 5.11], see also [13, Theorem 8.1.6]. Moreover, the second step of the proof of [9, Theorem 5.11], with (0.40) replaced by (0.41) and (0.42) from [10], actually shows that

$$(8.2) \qquad \|T(t)f\|_{\mathcal{C}_b^2(\mathbb{R}^d)} \leq ct^{-1}\|f\|_\infty, \qquad f \in \mathcal{B}_b(\mathbb{R}^d).$$

Since we also have $\|T(t)f\|_\infty \leq \|f\|_\infty$ and the Hölder space $\mathcal{C}_b^\gamma(\mathbb{R}^d)$, $\gamma \in (0, 2) \setminus \{1\}$, is a real interpolation space between $\mathcal{C}_b(\mathbb{R}^d)$ and $\mathcal{C}_b^2(\mathbb{R}^d)$, see [57, Thm. 2.7.2.1], an application of a classical interpolation theorem [57, Sec. 1.3.3] yields that $\|T(t)f\|_{\mathcal{C}_b^\gamma(\mathbb{R}^d)} \leq Mt^{-\rho/2}\|f\|_\infty$ for some constant $M = M(\rho) > 0$ and any $f \in L^\infty$, $\rho \in [0, 2] \setminus \{1\}$. Consequently, Hypothesis 6.1 is satisfied for any $\rho \in (0, 2)$. Moreover, $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a core for the generator of $T|_{\mathcal{C}_\infty(\mathbb{R}^d)}$, cf. [13, Theorem 8.1.6], and so Hypothesis 7.1 holds. $\square$

Remark 8.2. Under the milder regularity assumption that the diffusion coefficients q_{ij} are uniformly continuous, one can show with some more effort that the Hypothesis 6.1 is satisfied (for any $\rho \in (0, 2)$), e.g. using results from [41]. It is, however, not clear whether $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is still a core. In dimension $d = 1$, one can use that cores are preserved under random time changes, cf. [29, Theorem 4.1], to deduce that $\mathcal{C}_c^\infty(\mathbb{R})$ is indeed a core if the diffusion coefficient is (uniformly) continuous. Thus, Proposition 8.1 and Corollary 8.3 below, hold in dimension $d = 1$ also for uniformly continuous diffusion coefficients.

Combining Proposition 8.1 with Theorem 1.1, we get the following corollary.

Corollary 8.3. *Let $\mathcal{L}$ be a Lévy-type operator of the form*

$$(8.3) \qquad \begin{aligned} \mathcal{L}f(x) &= b(x) \cdot \nabla f(x) + \operatorname{tr}(Q(x)\nabla^2 f(x)) \\ &+ \int_{\mathbb{R}^d \setminus \{0\}} (f(x + y) - f(x) - y \cdot \nabla f(x)\mathbb{1}_{(0,1)}(|y|))\, \mu(x, dy). \end{aligned}$$

If the drift b is bounded and measurable, the diffusion coefficients are bounded, Hölder continuous, symmetric and strictly elliptic, the jumping kernels depend measurably on x and there is some $\beta \in (0, 2)$ such that

$$\sup_{x \in \mathbb{R}^d} \left(\int_{\mathbb{R}^d \setminus \{0\}} \min\{|y|^\beta, 1\}\, \mu(x, dy) \right) < \infty,$$

then there is a Markovian $\mathcal{C}_b$-Feller semigroup T whose full generator $\hat{L}$ is the bp-closure of $\{(f, \mathcal{L}f); f \in \mathcal{C}_c^\infty(\mathbb{R}^d)\}$. Moreover, the $(\mathcal{L}, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem is well-posed. If additionally the family $(\mu(x, dy))_{x \in \mathbb{R}^d}$ is tight, then T is a Feller semigroup.

As already mentioned in the introduction, the fact that second order (or even higher order) differential operators with lower order coefficients that are merely measurable generate strongly continuous semigroups on $\mathcal{C}_\infty(\mathbb{R}^d)$ is known for a long time (see [52]). Here, we also allow lower order *integral terms* that are merely measurable, providing an extension of Taira's results on domains. Also for the well-posedness of the martingale problem we require less regularity than e.g. in the classical results by Stroock [53], where continuous dependence of $\mu(x, \cdot)$ on x is assumed.

8.2. **Lévy generators.** Next we consider perturbations of operators $\mathcal{A}$ which are the generators of Lévy processes, that is, stochastic processes with càdlàg sample paths and independent and stationary increments. Recall that, by the Lévy–Khintchine formula, any Lévy process can be uniquely characterized (in distribution) by its characteristic exponent ψ or by its Lévy triplet (b, Q, ν), cf. [47] for more information. An important example is the fractional Laplacian $-(-\Delta)^{\alpha/2}$, which is the generator of the isotropic α-stable Lévy process. Its characteristic exponent is $\psi(\xi) = |\xi|^\alpha$ and the corresponding Lévy triplet is $(0, 0, \nu)$ with

$$(8.4) \qquad \nu(U) = c_{d,\alpha} \int_U \frac{1}{|y|^{d+\alpha}} \, dy$$

for a normalizing constant $c_{d,\alpha}$.

In the following, we will assume that the *Hartman–Wintner condition* is satisfied:

$$(8.5) \qquad \lim_{|\xi| \to \infty} \frac{\operatorname{Re} \psi(\xi)}{\log |\xi|} = \infty.$$

This is a mild growth condition on $\operatorname{Re} \psi$ which ensures the existence of a smooth transition density p_t for each $t > 0$ for the associated process, cf. [26] for a thorough discussion.

Proposition 8.4. *Let $(X_t)_{t \geq 0}$ be a Lévy process with transition semigroup $(T(t))_{t \geq 0}$ and characteristic exponent ψ satisfying the Hartman–Wintner condition (8.5). If there is a constant $\kappa > 0$ such that the transition density p_t of X_t satisfies*

$$(8.6) \qquad \int_0^1 \Big(\int_{\mathbb{R}^d} |\nabla p_t(x)| \, dx \Big)^\kappa dt < \infty.$$

then Hypothesis 6.1 and 7.1 hold for any $\rho \leq \kappa$, and so Theorem 1.1 is applicable (with $\rho \leq \kappa$).

Proof. It is well known that $T(t)f(x) = \int f(x+y) p_t(y) \, dy$ is a $\mathcal{C}_\infty$-semigroup and because of the existence of the density, each $T(t)$ is a strong Feller operator, see e.g. [18, pp. 438–39]. Moreover, $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a core for the generator of $T(t)|_{\mathcal{C}_\infty}$ by [6, Cor. 2.10]. Denote by $\hat{A}$ the corresponding full generator. The differentiation lemma for parameter-dependent integrals entails that $T(t)f$ is differentiable for any $t > 0$, $f \in L^\infty(\mathbb{R}^d)$, and

$$\|\nabla T(t)f\|_\infty \leq \varphi(t)\|f\|_\infty \quad \text{with} \quad \varphi(t) := \max \Big\{ 1, \int_{\mathbb{R}^d} |\nabla p_t(x)| \, dx \Big\}.$$

As p_{2t} is the convolution of p_t with itself, $\|\nabla^2 T(t)f\|_\infty \leq d^2 \varphi(t/2)^2 \|f\|_\infty$, cf. [34, Lem. 4.1]. Using that $\mathcal{C}_b^\gamma(\mathbb{R}^d)$, $\gamma \in (0, 2) \setminus \{1\}$, is a real interpolation space between $\mathcal{C}_b(\mathbb{R}^d)$ and $\mathcal{C}_b^2(\mathbb{R}^d)$, see [57, Thm. 2.7.2.1], and applying an interpolation theorem [57, Sec. 1.3.3], we find that $\|T(t)f\|_{\mathcal{C}_b^\gamma(\mathbb{R}^d)} \leq M\varphi(t/2)^\gamma \|f\|_\infty$ for some constant $M = M(\gamma) > 0$ and any $f \in L^\infty$, $\gamma \in [0, 2]$. $\qquad \square$

Gradient estimates of the form (8.6) for Lévy processes have been studied intensively in the last decade, e.g. [14, 21, 30, 36, 51] to mention but a few, and so Proposition 8.4 applies to a wide class of Lévy generators. Using the gradient

estimates from [51, Ex. 1.5], we find that the assumptions of Proposition 8.4 are satisfied for Lévy measures which are dominated from below by a stable measure.

Example 8.5. Let $(T(t))_{t\geq 0}$ be the semigroup of a pure-jump Lévy process with Lévy measure ν satisfying

$$\nu(U) \geq \int_0^{r_0} \int_{\mathbb{S}} \mathbb{1}_U(r\theta) r^{-1-\alpha}\, dr\, \varrho(d\theta), \quad U \in B(\mathbb{R}^d \setminus \{0\}),$$

for some constants $\alpha \in (0,2)$, $r_0 > 0$ and a measure ϱ on the unit sphere $\mathbb{S}$ whose support is not contained in any proper linear subspace of $\mathbb{R}^d$. Then the assumptions of Proposition 8.4 are satisfied (with $\kappa < \alpha$).

Example 8.5 covers, in particular, stable operators, i.e. operators (1.1) with $b = 0$, $Q = 0$ and Lévy measures ν of the form $\nu(U) = \int_0^\infty \int_{\mathbb{S}} \mathbb{1}_U(r\theta) r^{-1-\alpha}\, dr\, \varrho(d\theta)$. Let us mention that Komatsu [28] was one of the first to study perturbations of stable operators under Lévy-type operators; however, his result requires quite strong assumptions on the regularity the density of ν. Moreover, we would like to mention a closely related result by Peng [43] who considers perturbation of the fractional Laplacian by time-dependent kernels $\mu(t,x,dy)$. Our approach has the advantage that it is not restricted to stable operators but works in a much more general setting, as demonstrated by Proposition 8.4

8.3. **Operators of variable order.** Our next application concerns perturbations of the fractional Laplacian of variable order $-(-\Delta)^{\alpha(\cdot)/2}$, which is defined on $\mathcal{C}_c^\infty(\mathbb{R}^d)$ by

$$-(-\Delta)^{\alpha(\cdot)/2} f(x) = c_{d,\alpha(x)} \int_{\{y\neq 0\}} (f(x+y) - f(x) - y \cdot \nabla f(x) \mathbb{1}_{(0,1)}(|y|)) \frac{1}{|y|^{d+\alpha(x)}}\, dy$$

for a certain normalizing constant $c_{d,\alpha(x)}$. This operator appears as generator of so-called processes of variable order, see e.g. [4, 27, 30] for the construction of such processes and more information about the probabilistic background. Denote by $(T(t))_{t\geq 0}$ the associated $\mathcal{C}_\infty$-Feller semigroup.

Proposition 8.6. *If* $\alpha : \mathbb{R}^d \to (0,2]$ *is Hölder continuous with* $\alpha_L := \inf_x \alpha(x) > 0$, *then* $(T(t))_{t\geq 0}$ *satisfies Hypothesis 6.1 and 7.1 with* $\rho = \alpha_L$, *and so Theorem 1.1 is applicable (with* $\rho = \alpha_L$).

Proof. Let $(T(t))_{t\geq 0}$ be the Markovian Feller semigroup associated with the operator $-(-\Delta)^{\alpha(\cdot)/2}$, see e.g. [4] or [30, Thm. 5.2] for the existence, and denote by $\hat{A}$ its full generator and by A its part in $\mathcal{C}_\infty(\mathbb{R}^d)$. It is known from the construction of the semigroup that $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a core and that $T(t)f(x) = \int f(y) p_t(x,y)\, dy$ for a bounded transition function p_t, cf. [30]. In consequence, by [6, Thm. 1.9, Thm. 1.14], each $T(t)$ is a strong Feller operator. Moreover, by [33, Prop. 6.1],

$$\|T(t)f\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \leq c_\beta t^{-\beta/\alpha_L} \|f\|_\infty, \qquad f \in \mathcal{B}_b(\mathbb{R}^d),\ t \in (0,1),$$

for any $\beta < \alpha_L$. In conclusion, Hypothesis 7.1 is satisfied. $\qquad\square$

Example 8.7. Consider

$$\mathcal{L}u = -(-\Delta)^{\alpha(\cdot)/2}u - (-\Delta)^{\beta(\cdot)/2}$$

where $\alpha : \mathbb{R}^d \to (0,2]$ is Hölder continuous with $\alpha_L := \inf_x \alpha(x) > 0$ and $\beta : \mathbb{R}^d \to (0,2)$ is a measurable function with $\sup_x \beta(x) < \alpha_L$. Then, by Proposition 8.6 and Theorem 1.1, the bp-closure of $\{(f,\mathcal{L}f); f \in \mathcal{C}_c^\infty(\mathbb{R}^d)\}$ is the generator of a Markovian $\mathcal{C}_b$-Feller semigroup S and the $(\mathcal{L}, \mathcal{C}_c^\infty(\mathbb{R}^d))$-martingale problem is well-posed. If additionally $\inf_x \beta(x) > 0$, then the tightness condition (Ti) is satisfied and S is a $\mathcal{C}_\infty$-Feller semigroup.

Let us mention that Theorem 7.3 also yields continuous dependence of these solutions on the 'coefficient' β. More precisely, if β_n, $n \in \mathbb{N}$ are measurable functions

such that $\beta_n(x) \to \beta(x)$ for all $x \in \mathbb{R}^d$ and $0 < \inf_n \inf_x \beta(x) \leq \sup_n \sup_x \beta_n(x) < \alpha_L$, then $\mathbb{P}^{(\beta_n,x)} \to \mathbb{P}^{(\beta,x)}$ weakly for every $x \in \mathbb{R}^d$. In particular, the associated semigroups satisfy $S^{(\beta_n)}(t)f(x) \to S^{(\beta)}(t)f(x)$ for any $t > 0$, $x \in \mathbb{R}^d$ and $f \in \mathcal{C}_b(\mathbb{R}^d)$.

Proposition 8.6 can be extended to a wider class of non-local operators, that is, we can replace the fractional Laplacian of variable order $-(-\Delta)^{\alpha(\cdot)/2}$ by other operators of variable order, see [30, Section 5.1] for some interesting choices.

8.4. **Non-local operators and Lévy-driven SDEs.** In this subsection, we show how our perturbation results can be applied in the theory of stochastic differential equations (SDEs). The key is a result by Kurtz [40] which states that the existence of solutions to Lévy-driven SDEs can be equivalently characterized by the existence of solutions to martingale problems associated with certain Lévy-type operators (1.1). For a comprehensive study of Lévy-driven SDEs, e.g. classical existence and uniqueness results, we refer to Ikeda–Watanabe [17] and Protter [45]. Our first result concerns SDEs with additive Lévy noise.

Proposition 8.8. *Let $(L_t)_{t \geq 0}$ be a Lévy process with characteristic exponent ψ satisfying the Hartman–Wintner condition (8.5). If the transition density p_t satisfies $\int_0^1 \int_{\mathbb{R}^d} |\nabla p_t(x)| \, dx \, dt < \infty$, then the SDE*

$$(8.7) \qquad dX_t = b(X_t) \, dt + dL_t, \qquad X_0 = x,$$

has a unique weak solution. Moreover, the unique solution does not explode in finite time and gives rise to a Markovian Feller semigroup with symbol

$$q(x,\xi) = -ib(x) \cdot \xi + \psi(\xi), \qquad x, \xi \in \mathbb{R}^d.$$

Proof. This is immediate from Proposition 8.4 (with $\kappa = 1$) and the results of [40]. $\qquad\square$

Proposition 8.8 applies, in particular, to any Lévy process with characteristic exponent ψ satisfying $\operatorname{Re} \psi(\xi) \geq c|\xi|^\alpha$, $|\xi| \gg 1$, for some $\alpha > 1$; indeed, (8.5) is then trivial and the integrability condition on ∇p_t follows from [51, Thm. 3.2]. Let us mention that Tanaka et. al [56, pp. 82,83] obtained the existence of a unique solution under the milder assumption that $\frac{1}{\operatorname{Re} \psi(\xi)} = o\left(|\xi|^{-1}\right)$ as $|\xi| \to \infty$. Furthermore, there is ongoing research about the optimal assumptions on the drift b to ensure the existence of a unique weak solution to (8.7) for a given Lévy process; e.g. [8] considers isotropic α-stable drivers, $\alpha > 1$, and (generally unbounded) drifts b from some Kato class.

Proposition 8.9. *Let $(L_t)_{t \geq 0}$ be a one-dimensional isotropic α-stable Lévy process for $\alpha \in (1,2)$, i.e. $(L_t)_{t \geq 0}$ is a Lévy process with characteristic exponent $\psi(\xi) = |\xi|^\alpha$. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a Hölder continuous function with $0 < \inf_x \sigma(x) \leq \sup_x \sigma(x) < \infty$. For every $b \in \mathcal{B}_b(\mathbb{R})$, there is a unique weak solution to the SDE*

$$(8.8) \qquad dX_t = b(X_{t-}) \, dt + \sigma(X_{t-}) \, dL_t, \qquad X_0 = x.$$

The unique solution does not explode in finite time and gives rise to a Markovian $\mathcal{C}_\infty$-Feller semigroup with symbol

$$q(x,\xi) = -ib(x) \cdot \xi + |\sigma(x)|^\alpha |\xi|^\alpha.$$

Proof. Since σ is Hölder continuous, the SDE $dY_t = \sigma(Y_{t-}) \, dL_t$ has a unique weak solution which gives rise to a Feller process, see e.g. [37, Thm. 2.1] or [30, Thm. 5.23]. Denote by $(T(t))_{t \geq 0}$ the associated semigroup and by $\hat{A}$ the full generator. It is known that $\mathcal{C}_c^\infty(\mathbb{R})$ is a core for the generator of $T|_{\mathcal{C}_\infty(\mathbb{R})}$, cf. [37, 30]. The strong Feller property of $T(t)$ follows from [6, Thm. 1.14] and the fact that Y_t has a bounded transition density p_t, see e.g. [30]. Moreover, [33, (Proof of) Proposition 4.5] yields the regularity estimate

$$(8.9) \qquad \|T(t)f\|_{\mathcal{C}_b^\beta(\mathbb{R})} \leq c_\beta t^{-\beta/\alpha} \|f\|_\infty, \qquad f \in \mathcal{B}_b(\mathbb{R}), \ t > 0,$$

for $\beta < \alpha$. Choosing $\beta \in (1, \alpha)$ and setting $\hat{B}f(x) = b(x)\nabla f(x)$, we see that the assumptions of Theorem 6.4 are satisfied, and so there is a Markovian $\mathcal{C}_\infty$-semigroup with full generator $\hat{A} + \hat{B}$. Using the well-posedness of the $(A + \hat{B}, \mathcal{C}_c^\infty(\mathbb{R}))$-martingale problem, which holds by Theorem 7.2, it follows from [40] that the SDE (8.8) has a unique weak solution. $\qquad\square$

Remark 8.10. Proposition 8.9 can be extended in several ways:

(i) We can consider a wider class of driving Lévy processes $(L_t)_{t \geq 0}$. There are two key ingredients which we need. Firstly, the existence of a unique weak solution to the SDE $dY_t = \sigma(Y_{t-}) \, dL_t$; there is extensive literature on this topic, see e.g. [7, 35, 30, 37, 58] and the references therein. Secondly, a regularity estimate $\|T(t)f\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \leq \varphi(t)\|f\|_\infty$ for the corresponding semigroup, where $\varphi \in L^1(0, 1)$ and $\beta > 1$. Though there are quite some works on regularizing properties of semigroups associated with SDEs, many of them study only β-Hölder regularity for $\beta \leq 1$, and this is not enough for our purposes.

(ii) Using localization techniques for martingale problems, one can relax the growth assumptions on b and σ; see [31, Thm. 4.2] and [16, Thm. 5.3] for useful results in this direction.

Proposition 8.11. *Let $(L_t)_{t \geq 0}$ be a one-dimensional isotropic α-stable Lévy process for $\alpha \in (0, 2)$. Let $(M_t)_{t \geq 0}$ be a one-dimensional Lévy process independent of $(L_t)_{t \geq 0}$ with Lévy triplet $(b, 0, \nu)$ such that $\int_{\{|y| \leq 1\}} |y|^\beta \nu(dy) < \infty$ for some $\beta \in [0, \alpha)$. If $\alpha \leq 1$ also assume that the compensated drift $b - \int_{\{|y| < 1\}} y \nu(dy)$ is zero. Let $\sigma : \mathbb{R} \to (0, \infty)$ be a Hölder continuous mapping with $0 < \inf_x \sigma(x) \leq \sup_x \sigma(x) < \infty$. For every $\kappa \in \mathcal{B}_b(\mathbb{R})$, the SDE*

$$dX_{t-} = \kappa(X_{t-}) \, dM_t + \sigma(X_{t-}) \, dL_t, \qquad X_0 = x$$

has a unique weak solution. It does not explode in finite time and it gives rise to a Markovian Feller semigroup with symbol

$$q(x, \xi) = \psi(\kappa(x)\xi) + |\sigma(x)|^\alpha |\xi|^\alpha,$$

where ψ denotes the characteristic exponent of $(M_t)_{t \geq 0}$.

Proposition 8.11 can be extended to other driving Lévy processes $(L_t)_{t \geq 0}$, cf. Remark 8.10.

Proof of Proposition 8.11. Because of the integrability condition on the Lévy measure ν, an application of Taylor's formula shows that the operator

$$\hat{B}f(x) := b \cdot \nabla f(x) \mathbb{1}_{\{\alpha > 1\}}$$
$$+ \int_{\{y \neq 0\}} (f(x + \kappa(x)y) - f(x) - f'(x)\kappa(x)y \mathbb{1}_{(0,1)}(|y|)\mathbb{1}_{\{\alpha > 1\}}) \, \nu(dy)$$

satisfies $\|\hat{B}f\|_\infty \leq C\|f\|_{\mathcal{C}_b^\beta(\mathbb{R})}$ for some constant $C > 0$. Moreover, Hypothesis 6.3 holds. Note that the kernels $\mu(x, dy)$ associated with $\hat{B}$ satisfy the tightness condition (Ti) because

$$\mu(x, \{|y| \geq R\}) = \nu(\{y \in \mathbb{R} \setminus \{0\}; |\kappa(x)y| \geq R\})$$
$$\leq \nu(\{y \in \mathbb{R} \setminus \{0\}; |y| \geq R/\|\kappa\|_\infty\}) \xrightarrow{R \to \infty} 0.$$

Now we can proceed exactly as in the proof of Proposition 8.9; the only difference is that we consider the just-defined operator $\hat{B}$ rather than the drift operator. $\qquad\square$

Example 8.12. Let $\alpha \in (0, 2]$ and $\beta \in (0, \alpha)$. Let $(L_t)_{t \geq 0}$ be an isotropic α-stable Lévy process and $(M_t)_{t \geq 0}$ an isotropic β-stable Lévy process, which is independent of

$(L_t)_{t \geq 0}$. Let σ be a Hölder continuous function with $0 < \inf_x \sigma(x) \leq \sup_x \sigma(x) < \infty$. Then, by Proposition 8.11 and Theorem 7.3:

(i) For every $\kappa \in \mathcal{B}_b(\mathbb{R}^d)$, the SDE

$$(8.10) \qquad dX_t = \kappa(X_{t-}) \, dM_t + \sigma(X_{t-}) \, dL_t, \qquad X_0 = x,$$

has a unique weak solution. It gives rise to a Markovian Feller semigroup.

(ii) If $\kappa_n \in \mathcal{B}_b(\mathbb{R})$, $n \in \mathbb{N}$, is a sequence of measurable functions such that $\kappa_n \overset{\text{bp}}{\to} \kappa$, then the solutions $X^{(\kappa_n)}$ to (8.10) (with κ replaced by κ_n) converge weakly to $X = X^{(\kappa)}$ in the Skorohod space $\mathcal{D}[0,\infty)$.

APPENDIX A. BP-CONVERGENCE OF MEASURE-VALUED FUNCTIONS

The concept of *bp-convergence*, see Section 2, is a very useful tool in probability theory and allows for a version of Dynkin's π-λ theorem for bounded measurable functions. We recall that if E is a metric space, then a set $M \subset \mathcal{B}_b(E)$ is called *bp-closed* if whenever $(f_n) \subset M$ is a sequence with $f_n \overset{\text{bp}}{\to} f$, then also $f \in M$. The *bp-closure* of $M \subset \mathcal{B}_b(E)$ is the smallest bp-closed subset of $\mathcal{B}_b(E)$ that contains M. If the bp-closure of M is $\mathcal{B}_b(E)$ we say that M is *bp-dense* in $\mathcal{B}_b(E)$.

We recall the following general result from [13, Prop. 3.4.2].

Proposition A.1. *In a metric space E, the set $\mathcal{C}_b(E)$ is bp-dense in $\mathcal{B}_b(E)$. If E is additionally locally compact, then also $\mathcal{C}_c(E)$ is bp-dense in $\mathcal{B}_b(E)$.*

In this appendix, we aim to extend this result to functions that take values in the cone $\mathcal{M}^+(E)$ of positive, locally finite measures on E, where E is a locally compact, separable metric space. To that end, we endow $\mathcal{M}^+(E)$ with the vague topology, induced by the space $\mathcal{C}_c(E)$ of continuous functions with compact support. As is well known, see [22, Thm. A.2.3(i)], $\mathcal{M}^+(E)$ is a Polish space in the vague topology, i.e. the topology is induced by a complete, separable metric.

As a matter of fact, we will not be interested in the case of arbitrary E, but only in the case $E = \mathbb{R}^d \setminus \{0\}$. Fix a continuous function $g : \mathbb{R}^d \setminus \{0\} \to [0,\infty)$. Given a measure $\nu \in \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})$, we put

$$|\nu|_g := \int_{\mathbb{R}^d \setminus \{0\}} g(y) \, d\nu(y).$$

We denote by $K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$ the set of measurable mappings $\mu : \mathbb{R}^d \to \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})$ such that

$$\|\mu\|_g := \sup_{x \in \mathbb{R}^d} |\mu(x)|_g < \infty \quad \text{and} \quad \lim_{R \to \infty} \sup_{x \in K} \mu(x, \{|y| > R\}) = 0 \} \text{ for all } K \Subset \mathbb{R}^d.$$

We note that, as a consequence of [22, Thm. A.2.3(iv)] the map $\mu : \mathbb{R}^d \to \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})$ is measurable if and only if $x \mapsto \int_{\mathbb{R}^d} f(y)\mu(x, dy)$ is measurable for every $f \in \mathcal{C}_c(\mathbb{R}^d \setminus \{0\})$ if and only if $x \mapsto \mu(x, A)$ is measurable for every relatively compact set A.

Definition A.2. We say that a sequence $(\mu_n) \subset K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$ *converges boundedly and pointwise* to $\mu : \mathbb{R}^d \to \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})$, if

(i) $\sup_n \|\mu_n\|_g < \infty$,
(ii) $\mu_n(x) \to \nu(x)$ vaguely for every $x \in \mathbb{R}^d$ and
(iii) for every $K \Subset \mathbb{R}^d$ we have $\lim_{R \to \infty} \sup_{n \in \mathbb{N}} \sup_{x \in K} \mu_n(x, \{|y| > R\}) = 0$.

We write $\mu_n \overset{bp}{\to} \nu$ to indicate bounded and pointwise convergence.

Remark A.3. We note that if $g(x) = \varphi(|x|)$ for some function $\varphi : (0,\infty) \to (0,\infty)$ with $\varphi(t) \uparrow \infty$ as $t \to \infty$, then condition (iii) in Definition A.2 is automatically

fulfilled and follows from assumption (i). Indeed, for $R > 0$, $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$ we have

$$\mu_n(x, \{|y| > R\}) \leq \frac{1}{\varphi(R)} \int_{\{|y| > R\}} \varphi(|y|) \, \mu_n(x, dy) \leq \frac{1}{\varphi(R)} \sup_{n \in \mathbb{N}} \|\mu_n\|_g \to 0.$$

We note that we did not assume in the above definition that the limit μ belongs to $K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$. As it turns out, this holds true automatically.

Lemma A.4. *Let $(\mu_n) \subset K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$. If $\mu_n \overset{\text{bp}}{\to} \mu$, then also μ belongs to $K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$.*

Proof. We first note that μ is measurable as pointwise limit of measurable functions. We now pick $C > 0$ such that $\|\mu_n\|_g \leq C$ for all $n \in \mathbb{N}$ and fix a sequence $(f_k) \in \mathcal{C}_c(\mathbb{R}^d \setminus \{0\})$ such that $0 \leq f_k(y) \uparrow g(y)$. As $\mu_n(x) \to \mu(x)$ vaguely, we obtain for fixed $k \in \mathbb{N}$ and arbitrary $x \in \mathbb{R}^d$ that

$$\int_{\mathbb{R}^d} f_k(y) \mu(x, dy) = \lim_{n \to \infty} \int_{\mathbb{R}^d} f_k(y) \mu_n(x, dy) \leq \limsup_{n \to \infty} \int_{\mathbb{R}^d} g(y) \, \mu_n(x, dy) \leq C.$$

Taking the limit $k \to \infty$, monotone convergence implies

$$\int_{\mathbb{R}^d} g(y) \, \mu(x, dy) \leq C$$

for every $x \in \mathbb{R}^d$, proving that $\|\mu\|_g \leq C < \infty$. To prove that for a compact set K, we have $\lim_{R \to \infty} \sup_{x \in K} \mu(x, \{|y| > R\}) = 0$, we pick, given $\varepsilon > 0$, a radius $R > 0$ such that $\mu_n(x, \{|y| > R\}) \leq \varepsilon$ for all $x \in K$ and $n \in \mathbb{N}$. Approximating $\mathbb{1}_{\{|y| > R\}}$ by an increasing sequence of functions in $\mathcal{C}_c(\mathbb{R}^d)$, we can repeat the above argument to prove that $\mu(x, \{|y| > R\}) \leq \varepsilon$ for all $x \in K$. $\square$

In what follows, we use the notions of bp-closedness, bp-closure and bp-density with the obvious meaning also for subsets of $K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$. We can now prove the main result of this appendix, a version of Proposition A.1 for measure-valued functions.

Theorem A.5. *The set $\mathcal{C}_c(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})) \cap K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$ is bp-dense in $K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$.*

Proof. We write K_C^g for the set which is claimed to be bp-dense and denote by $\mathcal{F}$ its bp-closure. We proceed in several steps.

Step 1: We prove that if $\mu(x, A) = \varphi(x) \cdot \nu(A)$ for some $\varphi \in \mathcal{B}_b(\mathbb{R}^d)$ and $\nu \in \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})$ with $|\nu|_g < \infty$, then $\mu \in \mathcal{F}$.

Indeed, if $\varphi \in \mathcal{C}_c(\mathbb{R}^d)$, then a function μ of the above form belongs to K_C^g and hence to $\mathcal{F}$. However, if $\varphi_n \overset{\text{bp}}{\to} \varphi$ it is easy to see that $\varphi_n(\cdot)\nu \overset{\text{bp}}{\to} \varphi(\cdot)\nu$ and hence the latter belongs $\mathcal{F}$ if this is true for the approximating sequence. Invoking Proposition A.1, it follows that indeed for every $\varphi \in \mathcal{B}_b(\mathbb{R}^d)$ the above function μ belongs to $\mathcal{F}$.

Step 2: We prove that $\mathcal{F}$ is a cone.

To see this, fix $\mu \in \mathcal{F}$ and put

$$\mathcal{F}_\mu := \left\{ \tilde{\mu} \in \mathcal{F} : s\mu + t\tilde{\mu} \in \mathcal{F} \text{ for all } s, t \in (0, \infty) \right\}.$$

A moments thought shows that $\mathcal{F}_\mu$ is bp-closed. If $\mu \in K_C^g$, then clearly every element of K_C^g belongs to $\mathcal{F}_\mu$ as K_C^g is itself a cone. Thus $\mathcal{F} \subset \mathcal{F}_\mu$ and hence $\mathcal{F} = \mathcal{F}_\mu$. Now let $\mu \in \mathcal{F}$ be arbitrary. Then we have $K_C^g \subset \mathcal{F}_\mu$ by what was just proved. Using that K_C^g is bp-dense in $\mathcal{F}$ once again, we find $\mathcal{F}_\mu = \mathcal{F}$ for every $\mu \in \mathcal{F}$ which proves that $\mathcal{F}$ is indeed a cone.

Step 3: It follows from Step 1 and Step 2 that every function of the form

$$\mu(x, \cdot) = \sum_{k=1}^n \mathbb{1}_{A_k}(x) \nu_k(\cdot),$$

where $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R}^d)$ and $\nu_1, \ldots, \nu_n \in \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})$ with $|\nu_k|_g < \infty$ for all $k = 1, \ldots, n$, belongs to $\mathcal{F}$. Given any sequence $(A_k) \subset \mathcal{B}(\mathbb{R}^d)$ of pairwise disjoint sets and a tight sequence $\nu_k \in \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})$ with $\sup_n |\nu_n|_g < \infty$ also the function

$$\mu(x, \cdot) := \sum_{k=1}^\infty \mathbb{1}_{A_k}(x) \nu_k(\cdot)$$

belongs to $\mathcal{F}$; indeed, the partial sums are all in $\mathcal{F}$ and so their bp-limit μ belongs to $\mathcal{F}$ by the closedness of $\mathcal{F}$.

Step 4 Fix $\mu \in K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$ with $\lim_{R \to \infty} \sup_{x \in \mathbb{R}^d} \mu(x, \{|y| > R\}) = 0$. We put

$$S_\mu := \{\nu \in \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}) : |\nu|_g \leq \|\mu\|_g \text{ and}$$
$$\nu(\{|y| > R\}) \leq \sup_{x \in \mathbb{R}^d} \mu(x, \{|y| > R\}) \text{ for all } R > 0\}.$$

Arguing similar as in the proof of Lemma A.4, we see that S_μ is a closed subset of $\mathcal{M}^+(\mathbb{R}^d \setminus \{0\})$. Consequently, S_μ is separable, whence we find a dense sequence $(\nu_n) \subset S_\mu$ that is dense in S_μ.

Let us fix a metric ρ that generates the topology on S_μ. We write $B(\nu, \varepsilon)$ for the open ball in S_μ of radius ε and with center $\nu \in S_\mu$. We can now inductively define sets $B_k^{(n)}$ by $B_1^{(n)} = B(\nu_1, n^{-1})$ and

$$B_{k+1}^{(n)} := B(\nu_{k+1}, n^{-1}) \setminus (B_1^{(n)} \cup \cdots \cup B_k^{(n)}).$$

Then for every $n \in \mathbb{N}$ the sets $(B_k^{(n)})_{k \in \mathbb{N}}$ are pairwise disjoint and their union equals S_μ. Let us set

$$\mu_n(x, \cdot) := \sum_{k=1}^\infty \mathbb{1}_{\mu^{-1}(B_j^{(n)})}(x) \nu_j(\cdot).$$

By Step 3, the function μ_n belongs to $\mathcal{F}$. However, we have $\|\mu_n\|_g \leq \|\mu\|_g$ for every $n \in \mathbb{N}$ as every ν_j belongs to S_μ and for $x \in \mathbb{R}^d$ we have $\rho(\mu_n(x), \mu(x)) \leq n^{-1}$ so that $\mu_n(x) \to \mu(x)$. By the definition of the set S_μ also condition (iii) in Definition A.2 is satisfied so that $\mu_n \overset{\mathrm{bp}}{\to} \mu$. It follows that $\mu \in \mathcal{F}$.

Step 5 We finish the proof.

Given $\mu \in K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$, we set $\mu_n := \mathbb{1}_{B(0,n)}\mu$. By Step 4, we have $\mu_n \in \mathcal{F}$, as $\sup_{x \in \mathbb{R}^d} \mu_n(x, \{|y| > R\}) \to 0$ as $R \to \infty$. On the other hand, we have $\mu_n \overset{\mathrm{bp}}{\to} \mu$ and it follows that $\mu \in \mathcal{F}$. $\qquad\square$

We now combine Proposition A.1 and Theorem A.5 into a single result which will be useful in the main part of this article.

Corollary A.6. *Assume that F is a subset of $\mathcal{B}_b(\mathbb{R}^d; \mathbb{R}^d) \times K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$ such that*

(i) $\mathcal{C}_c(\mathbb{R}^d; \mathbb{R}^d) \times \left[\mathcal{C}_c\big(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})\big) \cap K^g\big(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})\big)\right] \subset F$

(ii) *whenever $((f_n, \mu_n))_{n \in \mathbb{N}} \subset F$ and $f_n \overset{\mathrm{bp}}{\to} f$ and $\mu_n \overset{\mathrm{bp}}{\to} \mu$, then also $(f, \mu) \in F$.*

Then $F = \mathcal{B}_b(\mathbb{R}^d) \times K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$.

Proof. Fixing $f \in \mathcal{C}_c(\mathbb{R}^d; \mathbb{R}^d)$ and considering sequences of the form $(f, \mu_n)_{n \in \mathbb{N}}$ bp-converging to (f, μ), Theorem A.5 shows $\mathcal{C}_c(\mathbb{R}^d) \times K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\})) \subset F$. Let us now fix $\mu \in K^g(\mathbb{R}^d, \mathcal{M}^+(\mathbb{R}^d \setminus \{0\}))$ and consider sequences (f_n, μ) bp-converging to (f, μ). Then Proposition A.1 yields the claim. $\qquad\square$

APPENDIX B. CONTINUITY OF LÉVY-TYPE OPERATORS IN TERMS OF THE
SYMBOL AND THE CHARACTERISTICS

In the proof of one of our perturbation results, Theorem 6.4, we used that the integro-differential operator

$$(B.1) \quad \hat{B}f(x) = b(x) \cdot \nabla f(x) + \int_{\mathbb{R}^d \setminus \{0\}} (f(x+y) - f(x) - y \cdot \nabla f(x)\chi(y))\, \nu(x, dy)$$

maps (sufficiently) smooth functions to continuous functions. On $\mathcal{C}_c^\infty(\mathbb{R}^d)$, the operator can be considered as a pseudo-differential operator with symbol $q(x, \xi)$, cf. (1.2), and so one can equivalently asked for conditions on the symbol ensuring that $\hat{B}f \in \mathcal{C}(\mathbb{R}^d)$ for $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. The following result answers this question; it is somewhat more refined than what we need for our proof but we believe the result to be of independent interest.

Theorem B.1. *Let $\hat{B}$ be as in* (B.1) *and denote by*

$$(B.2) \qquad q(x, \xi) = -ib(x) \cdot \xi + \int_{\mathbb{R}^d \setminus \{0\}} (1 - e^{iy \cdot \xi} + iy \cdot \xi\chi(y))\, \nu(x, dy)$$

the associated symbol; here $\chi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ is a smooth cut-off function with $\mathbb{1}_{B(0,1)} \leq \chi \leq \mathbb{1}_{B(0,2)}$. If q is locally bounded, then the following statements are equivalent.

    (i) *$\hat{B}f$ is continuous for every $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$.*
   (ii) *$x \mapsto q(x, \xi)$ is continuous for all $\xi \in \mathbb{R}^d$.*
  (iii) *Each of the following conditions is satisfied.*
      (a) *$x \mapsto b(x)$ is continuous.*
      (b) *$x \mapsto \nu(x, \cdot)$ is vaguely continuous on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$.*
      (c) *The family $(\nu(x, \cdot))_{x \in K}$, is tight for any compact set $K \Subset \mathbb{R}^d$, i.e.*
         $\lim_{R \to \infty} \sup_{x \in K} \nu(x, \{|y| \geq R\}) = 0$.
      (d) *$\lim_{r \to 0} \sup_{x \in K} \int_{\{|y| \leq r\}} |y|^2\, \nu(x, dy) = 0$ for any compact set $K \Subset \mathbb{R}^d$.*

If one (hence, all) of the conditions is satisfied and $\hat{B} : \mathcal{C}_b^\beta(\mathbb{R}^d) \to \mathcal{B}_b(\mathbb{R}^d)$ is a bounded operator for some $\beta > 0$, then $\hat{B}f$ is continuous for all $f \in \mathcal{C}_b^\beta(\mathbb{R}^d) \cap \mathcal{C}_\infty(\mathbb{R}^d)$.

Remark B.2.   (a) The implications (ii) $\iff$ (i) $\implies$ (iii) remain valid for any Borel measurable cut-off function χ such that $\mathbb{1}_{B(0,1)} \leq \chi \leq \mathbb{1}_{B(0,2)}$.
  (b) A symbol q of the form (B.2) is locally bounded if, and only if, for any compact set $K \subseteq \mathbb{R}^d$ there exists a constant $c > 0$ such that $|q(x, \xi)| \leq c(1 + |\xi|^2)$ for all $x \in K$, $\xi \in \mathbb{R}^d$. By [49, Lem. 2.1, Rem. 2.2], this is equivalent to

$$(B.3) \qquad \forall K \Subset \mathbb{R}^d \; : \; \sup_{x \in K} |b(x)| + \sup_{x \in K} \int_{\mathbb{R}^d \setminus \{0\}} \min\{|y|^2, 1\}\, \nu(x, dy) < \infty.$$

Proof of Theorem B.1. To keep notation simple, we prove the result only in dimension $d = 1$. The implication (i) $\implies$ (ii) follows from [48, Thm. 4.4]. Moreover, if $x \mapsto q(x, \xi)$ is continuous, then we find from the local boundedness of q and the dominated convergence theorem that

$$x \mapsto \hat{B}f(x) = -\int_{\mathbb{R}} q(x, \xi)\hat{f}(\xi)e^{ix\xi}\, d\xi$$

is continuous for all $f \in \mathcal{C}_c^\infty(\mathbb{R})$, and this proves (ii) $\implies$ (i). In the remainder of the proof we show that (ii) $\iff$ (iii).

(iii) $\implies$ (ii): By (iii)(a), it suffices to show that

$$x \mapsto p(x, \xi) := \int_{\mathbb{R} \setminus \{0\}} (1 - e^{iy\xi} + iy\xi\chi(y))\, \nu(x, dy)$$

is continuous for all $\xi \in \mathbb{R}$. Clearly,

$$|p(x,\xi) - p(z,\xi)| \leq I_1 + I_2 + I_3,$$

where

$$I_1 := \frac{|\xi|^2}{2} \left(\int_{\{|y| \leq r\}} |y|^2 \, \nu(x,dy) + \int_{\{|y| \leq r\}} |y|^2 \, \nu(z,dy) \right)$$

$$I_2 := \left| \int_{\{r < |y| < R\}} (1 - e^{iy\xi} + iy\xi\chi(y)) \left[\nu(x,dy) - \nu(z,dy) \right] \right|$$

$$I_3 := 2\nu(x, \{|y| \geq R\}) + 2\nu(z, \{|y| \geq R\}).$$

The vague continuity in (iii)(b) implies that $I_2 \to 0$ as $z \to x$ for fixed $r, R > 0$ with $\nu(x, \{|y| = r\}) = 0$ and $\nu(x, \{|y| = R\}) = 0$. Letting first $z \to x$ and then $r \to 0$ and $R \to \infty$, it follows from (iii)(c) and (iii)(d) that $p(\cdot, \xi)$ is continuous.

(ii) $\implies$ (iii): By [48, Thm. 4.4], (ii) implies that

$$\forall K \Subset \mathbb{R}^d \; : \; \lim_{|\xi| \to 0} \sup_{x \in K} |q(x,\xi)| = 0.$$

Moreover, exactly the same reasoning as in [48, proof of Thm. 4.4] shows that $\nu(x, \cdot)$, $x \in K$, is tight for any compact set K. For $\varphi \in \mathcal{C}_c^\infty(\mathbb{R})$ and $x \in \mathbb{R}$ define

$$S_x(\varphi) := \hat{B}(|\cdot - x|^2 \varphi(\cdot - x))(x) = \int_{\mathbb{R} \setminus \{0\}} y^2 \varphi(y) \, \nu(x,dy).$$

If we denote by $\mathcal{F}f := \hat{f}$ the Fourier transform of a function f, then

$$\mathcal{F}(|\cdot - x|^2 \varphi(\cdot - x))(\xi) = e^{-ix\xi} \mathcal{F}(|\cdot|^2 \varphi(\cdot))(\xi), \qquad \xi \in \mathbb{R}.$$

Since q is locally bounded and $x \mapsto q(x,\xi)$ is continuous for all $\xi \in \mathbb{R}$, an application of the dominated convergence theorem shows that

$$x \mapsto S_x(\varphi) = -\int_{\mathbb{R}} q(x,\xi)\mathcal{F}(|\cdot - x|^2 \varphi(\cdot - x))(\xi)e^{ix\xi} \, d\xi$$

is continuous. Choose $\varphi_k \in \mathcal{C}_c^\infty(\mathbb{R})$ such that $\mathbb{1}_{B(0,1/k)} \leq \varphi_k \leq \mathbb{1}_{B(0,2/k)}$ and $\varphi_{k+1} \leq \varphi_k$. Then $S_x(\varphi_k) \geq S_x(\varphi_{k+1})$ and

$$S_x(\varphi_k) \leq \int_{|y| \leq 2/k} |y|^2 \, \nu(x,dy) \xrightarrow{k \to \infty} 0 \quad \text{for all } x \in \mathbb{R}.$$

Applying Dini's theorem, we find that

$$\sup_{x \in K} \int_{|y| \leq 1/k} |y|^2 \, \nu(x,dy) \leq \sup_{x \in K} \int y^2 \varphi_k(y) \, \nu(x,dy) = \sup_{x \in K} |S_x(\varphi_k)| \xrightarrow{k \to \infty} 0$$

for any compact set K, and this proves (d). If we set $\mu(x,dy) := |y - x|^2 \, \nu(x, dy + x)$, then

(B.4) $\quad T_x(\varphi) := \hat{B}(|\cdot - x|^2 \varphi(\cdot))(x) = \int_{\mathbb{R} \setminus \{0\}} |y|^2 \varphi(x+y) \, \nu(x,dy) = \int \varphi(y) \, \mu(x,dy)$

for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R})$. As

$$\mathcal{F}(|\cdot - x|^2 \varphi(\cdot))(\xi)$$

$$= \frac{1}{2\pi} \left(\int_{\mathbb{R}} y^2 \varphi(y) \, e^{-iy\xi} \, dy - 2x \int_{\mathbb{R}} y\varphi(y) e^{-iy\xi} \, dy + x^2 \int_{\mathbb{R}} \varphi(y) e^{-iy\xi} \, dy \right),$$

there exists for any compact set K an integrable function g such that $\sup_{x \in K} |\mathcal{F}(|\cdot - x|^2 \varphi)(\xi)| \leq g(\xi)$ for all $\xi \in \mathbb{R}$. Since $x \mapsto q(x,\xi)$ is continuous and locally bounded, the dominated convergence theorem shows that the mapping

$$x \mapsto T_x(\varphi) = -\int_{\mathbb{R}} q(x,\xi)\mathcal{F}(|\cdot - x|^2 \varphi(\cdot))(\xi)e^{ix\xi} \, d\xi$$

is continuous for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R})$. By (B.4) this means that

$$\int_\mathbb{R} \varphi(y)\,\mu(z,dy) \xrightarrow{z\to x} \int_\mathbb{R} \varphi(y)\,\mu(x,dy) \quad \text{for all } f \in \mathcal{C}_c^\infty(\mathbb{R}), x \in \mathbb{R}.$$

Combining this with the fact that the local boundedness of q implies

$$(\text{B.5}) \qquad \sup_{z\in K} \int_{|y|\leq R} \mu(z,dy) < \infty \quad \text{for all } R > 0, K \subseteq \mathbb{R} \text{ cpt.}$$

cf. (B.3), we conclude that $\mu(\cdot,dy)$ is vaguely continuous on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Using that

$$\left| \int_\mathbb{R} \varphi(y-x)\,\mu(x,dy) - \int_\mathbb{R} \varphi(y-z)\,\mu(z,dy) \right|$$

$$\leq \int_\mathbb{R} |\varphi(y-x) - \varphi(y-z)|\,\mu(z,dy) + \left| \int_\mathbb{R} \varphi(y-x)\,[\mu(z,dy) - \mu(x,dy)] \right|$$

for all $\varphi \in \mathcal{C}_c(\mathbb{R})$, it follows from (B.5) and the vague continuity of $\mu(x,\cdot)$ that

$$\int_\mathbb{R} \varphi(y-z)\,\mu(z,dy) \xrightarrow{z\to x} \int_\mathbb{R} \varphi(y-x)\,\mu(x,dy)$$

for all $\varphi \in \mathcal{C}_c(\mathbb{R})$. Since $\nu(x,dy) = \frac{1}{|y|^2}\mu(x,dy-x)$, it is not difficult to see that this implies that $\nu(z,\cdot)$ converges vaguely on $(\mathbb{R}\setminus\{0\}, \mathcal{B}(\mathbb{R}\setminus\{0\}))$ to $\nu(x,\cdot)$ as $z \to x$. To prove continuity of the drift b, we note that

$$b(x) = \hat{B}((\cdot - x)\chi(\cdot - x))(x) = -\int_\mathbb{R} q(x,\xi)e^{ix\xi}\mathcal{F}(((\cdot - x)\chi(\cdot - x))(\xi)\,d\xi$$

$$= -\int_\mathbb{R} q(x,\xi)\mathcal{F}(\text{id}(\cdot)\chi(\cdot))(\xi)\,d\xi;$$

here $\text{id}(y) := y$. Applying the dominated convergence theorem another time, we find that $x \mapsto b(x)$ is continuous. This finishes the proof of the equivalences. Finally, if one (hence, all) conditions of the theorem are satisfied and $\|\hat{B}f\|_\infty \leq K\|f\|_{\mathcal{C}_b^\beta}$, $f \in C_c^\infty(\mathbb{R}^d)$, for some constants $K > 0$ and $\beta > 0$, then a standard approximation argument yields that $\hat{B}f$ is continuous for every $f \in \mathcal{C}_b^\beta(\mathbb{R}^d) \cap \mathcal{C}_\infty(\mathbb{R}^d)$. $\qquad\square$

Corollary B.3. *Let $\hat{B}$ be as in* (B.1) *for a smooth cut-off function χ with $\mathbb{1}_{B(0,1)} \leq \chi \leq \mathbb{1}_{B(0,2)}$, and denote by $q(x,\xi)$ the symbol of $\hat{B}$, cf.* (B.2). *If there is for every compact set $K \subseteq \mathbb{R}^d$ some constant $\alpha \in (0,2)$ such that*

$$(\text{B.6}) \qquad \sup_{x\in K} |b(x)| + \sup_{x\in K} \left(\int_{\{|y|\leq 1\}} |y|^\alpha\,\nu(x,dy) \right) < \infty,$$

and $(\nu(x,\cdot))_{x\in K}$ is tight, then the following statements are equivalent:

  (i) *$\hat{B}f$ is continuous for every $f \in C_c^\infty(\mathbb{R}^d)$,*
  (ii) *$x \mapsto q(x,\xi)$ is continuous for all $\xi \in \mathbb{R}^d$,*
  (iii) *$x \mapsto b(x)$ is continuous and $x \mapsto \nu(x,\cdot)$ is vaguely continuous.*

Proof. Because of (B.6), the symbol q is locally bounded, cf. Remark B.2, and moreover (iii)(c),(d) are clearly satisfied. Thus, the assertion is immediate from Theorem B.1. $\qquad\square$

References

[1] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace transforms and Cauchy problems*, vol. 96 of Monographs in Mathematics, Birkhäuser/Springer Basel AG, Basel, second ed., 2011.

[2] W. Arendt and N. Nikolski, *Vector-valued holomorphic functions revisited*, Mathematische Zeitschrift, 234 (2000), pp. 777–805.

[3] L. Arlotti and J. Banasiak, *Perturbations of Positive Semigroups with Applications*, Springer Monographs in Mathematics, Springer, 2006.

[4] R. F. BASS, *Uniqueness in law for pure jump Markov processes*, Probability Theory and Related Fields, 79 (1988), pp. 271–287.

[5] P. BILLINGSLEY, *Convergence of Probability Measures*, Wiley, second ed., 1999.

[6] B. BÖTTCHER, R. L. SCHILLING, AND J. WANG, *Lévy Matters III. Lévy-type processes: construction, approximation and sample path properties*, Springer, 2014.

[7] Z. CHEN, X. ZHANG, AND G. ZHAO, *Supercritical SDEs driven by multiplicative stable-like Lévy processes*, To appear: Transactions of the American Mathematical Society.

[8] Z.-Q. CHEN AND L. WANG, *Uniqueness of stable processes with drift*, Proceedings of the American Mathematical Society, 144 (2016), pp. 2661–2675.

[9] E. B. DYNKIN, *Markov Processes Vol. 1*, Springer, 1965.

[10] ———, *Markov Processes Vol. 2*, Springer, 1965.

[11] T. EISNER AND A. SERÉNY, *On the weak analogue of the Trotter-Kato theorem*, Taiwanese J. Math., 14 (2010), pp. 1411–1416.

[12] K. ENGEL AND R. NAGEL, *One-parameter semigroups for linear evolution equations*, vol. 194 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2000.

[13] S. N. ETHIER AND T. G. KURTZ, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[14] T. GRZYWNY AND K. SZCZYPKOWSKI, *Estimates of heat kernels of non-symmetric Lévy processes*, Preprint arXiv:1710.07793 [math]. To appear: Forum Mathematicum.

[15] M. HAASE, *The functional calculus for sectorial operators*, no. v. 169 in Operator theory, advances and applications, Birkhäuser Verlag, 2006.

[16] W. HOH, *Pseudo differential operators generating Markov processes.* Universität Bielefeld, 1998. Habilitationsschrift.

[17] N. IKEDA AND S. WATANABE, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, 1992.

[18] N. JACOB, *Pseudo differential operators and Markov processes. Vol. I: Fourier analysis and semigroups*, Imperial College Press, London, 2001.

[19] N. JACOB, *Pseudo Differential Operators and Markov Processes II: Generators and their potential theory*, Imperial College Press, 2002.

[20] N. JACOB, *Pseudo differential operators and Markov processes. Vol. III: Markov processes and applications*, Imperial College Press, London, 2005.

[21] K. KALETA AND P. SZTONYK, *Estimates of transition densities and their derivatives for jump Lévy processes*, Journal of Mathematical Analysis and Applications, 431 (2015), pp. 260–282.

[22] O. KALLENBERG, *Foundations of modern probability*, Probability and its Applications (New York), Springer-Verlag, New York, second ed., 2002.

[23] T. KATO, *Perturbation Theory for Linear Operators*, Springer, 1995.

[24] P. KIM AND R. SONG, *Stable process with singular drift*, Stochastic Processes and their Applications, 124 (2014), pp. 2479–2516.

[25] D. KINZEBULATOV AND K. R. MADOU, *On admissible singular drifts of symmetric α-stable process*, Preprint arXiv:2002.07001 [math].

[26] V. KNOPOVA AND R. L. SCHILLING, *A note on the existence of transition probability densities of Lévy processes*, Forum Mathematicum, 25 (2013), pp. 125–149.

[27] V. N. KOLOKOLTSOV, *Markov processes, semigroups and generators*, no. 38, De Gruyter, 2011.

[28] T. KOMATSU, *On the martingale problem for generators of stable processes with perturbations*, Osaka Journal of Mathematics, 21 (1984), pp. 113–132.

[29] F. KÜHN, *Random time changes of Feller processes*, arXiv:1705.02830 [math]. A slightly different version appeared under the title 'Perpetual integrals via random time changes' in the Bernoulli journal.

[30] ———, *Lévy Matters VI. Lévy-type processes: Moments, Construction and Heat Kernel Estimates*, Springer, 2017.

[31] ———, *On martingale problems and Feller processes*, Electronic Journal of Probability, 23 (2018).

[32] ———, *Existence of (Markovian) solutions to martingale problems associated with Lévy-type operators*, Electronic Journal of Probability, 25 (2020), pp. 1–26.

[33] ———, *Schauder estimates for Poisson equations associated with non-local Feller generators*, Journal of Theoretical Probability, 34 (2021), pp. 1506–1578.

[34] F. KÜHN AND R. L. SCHILLING, *Strong convergence of the Euler–Maruyama approximation for a class of Lévy-driven SDEs*, Stochastic Processes and their Applications, 129 (2019), pp. 2654–2680.

[35] T. KULCZYCKI, A. KULIK, AND M. RYZNAR, *On weak solution of SDE driven by inhomogeneous singular Lévy noise*, Preprint arXiv:2104.08129 [math].

[36] T. Kulczycki and M. Ryznar, *Gradient estimates of harmonic functions and transition densities for Lévy processes*, Transactions of the American Mathematical Society, 368 (2015), pp. 281–318.

[37] A. M. Kulik, *On weak uniqueness and distributional properties of a solution to an SDE with α-stable noise*, Stochastic Processes and their Applications, 129 (2019), pp. 473–506.

[38] M. Kunze, *A Pettis-type integral and applications to transition semigroups*, Czechoslovak Math. J., 61(136) (2011), pp. 437–459.

[39] ———, *Perturbation of strong Feller semigroups and well-posedness of semilinear stochastic equations on Banach spaces*, Stochastics, 85 (2013), pp. 960–986.

[40] T. G. Kurtz, *Equivalence of Stochastic Equations and Martingale Problems*, in Stochastic Analysis 2010, D. Crisan, ed., Springer, 2011, pp. 113–130.

[41] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Modern Birkhäuser Classics, Birkhäuser/Springer Basel AG, Basel, 1995. [2013 reprint of the 1995 original] [MR1329547].

[42] R. Mikulevicius and H. Pragarauskas, *On the Cauchy Problem for Integro-differential Operators in Hölder Classes and the Uniqueness of the Martingale Problem*, Potential Analysis, 40 (2014), pp. 539–563.

[43] J. Peng, *Uniqueness in law for stable-like processes of variable order*, Preprint arXiv:1802.01151 [math].

[44] N. I. Portenko, *Some perturbations of drift-type for symmetric stable processes*, Random Operators and Stochastic Equations, 2 (1994), pp. 211–224.

[45] P. E. Protter, *Stochastic Integration and Differential Equations*, vol. 21, Springer, 2005.

[46] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer, 1999.

[47] K. Sato, *Lévy processes and infinitely divisible distributions*, Cambridge University Press, Cambridge, 2013.

[48] R. L. Schilling, *Conservativeness and extensions of Feller semigroups*, Positivity, 2 (1998), pp. 239–256.

[49] ———, *Growth and Hölder conditions for the sample paths of Feller processes*, Probability Theory and Related Fields, 112 (1998), pp. 565–611.

[50] ———, *Measures, integrals and martingales*, Cambridge University Press, second ed., 2017.

[51] R. L. Schilling, P. Sztonyk, and J. Wang, *Coupling property and gradient estimates of Lévy processes via the symbol*, Bernoulli, 18 (2012), pp. 1128–1149.

[52] H. B. Stewart, *Generation of analytic semigroups by strongly elliptic operators*, Trans. Amer. Math. Soc., 199 (1974), pp. 141–162.

[53] D. Stroock, *Diffusion processes associated with Lévy-type operators*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 32 (1975), pp. 209–244.

[54] K. Taira, *On the existence of Feller semigroups with discontinuous coefficients*, Acta Math. Sin. (Engl. Ser.), 22 (2006), pp. 595–606.

[55] ———, *On the existence of Feller semigroups with discontinuous coefficients. II*, Acta Math. Sin. (Engl. Ser.), 25 (2009), pp. 715–740.

[56] H. Tanaka, M. Tsuchiya, and S. Watanabe, *Perturbation of drift-type for Lévy processes*, Kyoto Journal of Mathematics, 14 (1974), pp. 73–92.

[57] H. Triebel, *Interpolation theory, function spaces, differential operators*, North-Holland mathematical library, North-Holland Pub. Co, 1978.

[58] P. A. Zanzotto, *On stochastic differential equations driven by a Cauchy process and other stable Lévy motions*, The Annals of Probability, 30 (2002), pp. 802–825.

(F. Kühn) TU Dresden, Fakultät Mathematik, Institut für Mathematische Stochastik, 01062 Dresden, Germany. E-Mail: franziska.kuehn1@tu-dresden.de

(M. Kunze) Universität Konstanz, Fachbereich Mathematik und Statistik, Fach 193, 78357 Konstanz, Germany. E-Mail: markus.kunze@uni-konstanz.de