The Annals of Applied Probability
2018, Vol. 28, No. 5, 3028–3074
https://doi.org/10.1214/18-AAP1382

REGULARITY AND STABILITY FOR THE SEMIGROUP OF JUMP DIFFUSIONS WITH STATE-DEPENDENT INTENSITY

BY VLAD BALLY, DAN GOREAC[1] AND VICTOR RABIET

Université Paris-Est, LAMA (UMR 8050), UPEMLV, UPEC, CNRS

We consider stochastic differential systems driven by a Brownian motion and a Poisson point measure where the intensity measure of jumps depends on the solution. This behavior is natural for several physical models (such as Boltzmann equation, piecewise deterministic Markov processes, etc.). First, we give sufficient conditions guaranteeing that the semigroup associated with such an equation preserves regularity by mapping the space of k-times differentiable bounded functions into itself. Furthermore, we give an upper estimate of the operator norm. This is the key-ingredient in a quantitative Trotter–Kato-type stability result: it allows us to give an upper estimate of the distance between two semigroups associated with different sets of coefficients in terms of the difference between the corresponding infinitesimal operators. As an application, we present a method allowing to replace "small jumps" by a Brownian motion or by a drift component. The example of the 2D Boltzmann equation is also treated in all detail.

1. Introduction. We propose a quantitative analysis of the regularity of semigroups of operators associated with hybrid piecewise-diffusive systems

$$X_t = x + \sum_{l=1}^{\infty} \int_0^t \sigma_l(s, X_s)\, dB_s^l + \int_0^t b(s, X_s)\, ds + \int_{[0,t]\times E\times\mathbb{R}_+} c(s, z, X_{s-}) 1_{\{u\le\gamma(s,z,X_{s-})\}} N_\mu(ds, dz, du), \tag{1.1}$$

taking their values in some Euclidian space $\mathbb{R}^d$. Here:

- $(E, \mathcal{E})$ is a measurable space,
- $N_\mu(ds, dz, du)$ is a homogenous Poisson point measure on $E \times (0, \infty)$ with intensity measure $\mu(dz) \times 1_{(0,\infty)}(u)\, du$,
- $W_t = (W_t^l)_{l\in\mathbb{N}}$ is an infinite-dimensional Brownian motion (independent of N_μ) and
- the coefficients $\sigma_l, b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $c : \mathbb{R}_+ \times E \times \mathbb{R}^d \to \mathbb{R}^d$, $\gamma : \mathbb{R}_+ \times E \times \mathbb{R}^d \to [0, \infty)$ are assumed to be smooth enough.

Received July 2017; revised December 2017.

[1]Supported in part by he French National Research Agency Project PIECE, number ANR-12-JS01-0006.

MSC2010 subject classifications. Primary 60J75, 47D07; secondary 35Q21.

Key words and phrases. Piecewise diffusive jumps processes, trajectory-dependent jump intensity, PDMP, regularity of semigroups of operators, weak error, Boltzmann equation.

Whenever the jump intensity γ is constant, one deals with classical stochastic differential systems with jumps. Regularity of the associated flow is then immediate (see [26] or [30]). However, if γ is nonconstant, the position X_{s-} of the solution plays an important part in the intensity of jumps. This latter framework occurs in a wide variety of applications and it will receive our attention throughout the paper.

Our first result (see Theorem 15), which is the core of the paper, consists in proving that, under natural assumptions, the semigroup $\mathcal{P}_t f(x) = \mathbb{E}(f(X_t(x)))$ propagates regularity in finite time $T > 0$, that is, for some constant $Q_k(T, \mathcal{P})$ [see (5.1)],

$$\sup_{t \le T} \|\mathcal{P}_t f\|_{k,\infty} \le Q_k(T, \mathcal{P}) \|f\|_{k,\infty} \qquad \forall f \in C_b^k(R^d). \tag{1.2}$$

Here, $\|f\|_{k,\infty}$ is the infinite norm of f and its first k derivatives. In the case $k = 0$, this implies that $\mathcal{P}_t$ is a Feller semigroup.

As we have already hinted, the main difficulty to overcome is due to the presence of the jump intensity $\gamma(s, z, X_{s-})$. In classical jump equation, the indicator function $1_{\{u \le \gamma(s,z,X_{s-})\}}$ does not appear and one may construct a version of the solution such that $x \to X_t(x)$ is k-times differentiable (see [30]). Next, one proceeds with differentiating the associated semigroup and using chain rule $\partial_{x_i} \mathcal{P}_t f(x) = \sum_{j=1}^d \mathbb{E}[\partial_j f(X_t(x)) \partial_{x_i} X_t^j(x)]$ and concludes that (1.2) holds for $k = 1$ with $Q_1(T, P) = \sup_{t \le T} \sup_x \mathbb{E}[|\nabla X_t(x)|]$. Whenever the indicator function is present, the resulting stochastic differential representation of the solution in (1.1) is no longer appropriate. We will employ the alternative representation in (2.18) [known in the engineering literature as "real shock" representation, whereas (1.1) is known as the "fictive shock" representation]. The specificity of our framework is that the law of the jumps depends on the trajectory and this dependence is quantified by γ. As a consequence, the constants Q_k will depend on some quantities of type $\int_E |\partial^\alpha \ln \gamma(t, z, x)|^p \gamma(t, z, x) \mu(dz)$ (for appropriate $p \le k$ and index α; the presence of such terms is inspired by Malliavin calculus techniques).

A second result is a stability property in line with Trotter–Kato theorem (cf. [35], Theorem 4.4). We consider a sequence $(\mathcal{P}_t^n)_{n \in \mathbb{N}}$ of semigroups of operators with generators $\mathcal{L}^n$ and we assume that, for some $q \in \mathbb{N}$,

$$\|(\mathcal{L}^n - \mathcal{L}) f\|_\infty \le \varepsilon \times \|f\|_{q,\infty} \qquad \text{for all } f \in C_b^q(\mathbb{R}^d). \tag{1.3}$$

Here, $\mathcal{L}$ stands for the infinitesimal operator associated with (1.1). In Theorem 16 we prove that, under suitable hypotheses, the previous inequality yields

$$\|(\mathcal{P}_t^n - \mathcal{P}_t) f\|_\infty \le \varepsilon \times Q_q(T, \mathcal{P}) \times \|f\|_{q,\infty} \qquad \text{for all } f \in C_b^q(\mathbb{R}^d). \tag{1.4}$$

In order to understand the link between this result and the property (1.2), one writes

$$\mathcal{P}_t f(x) - \mathcal{P}_t^n f(x) = \int_0^t \partial_s \mathcal{P}_{t-s}^n \mathcal{P}_s f(x)\, ds = \int_0^t \mathcal{P}_{t-s}^n (\mathcal{L}^n - \mathcal{L}) \mathcal{P}_s f(x)\, ds$$

and notice that by (1.3) first and by (1.2) next

$$\begin{aligned}\|\mathcal{P}^n_{t-s}(\mathcal{L}^n-\mathcal{L})\mathcal{P}_s f\|_\infty &\le \|(\mathcal{L}^n-\mathcal{L})\mathcal{P}_s f\|_\infty \le \varepsilon\times\|\mathcal{P}_s f\|_{q,\infty}\\ &\le \varepsilon\times Q_q(T,\mathcal{P})\|f\|_{q,\infty}.\end{aligned}$$

We finally mention that in the paper we deal with nonhomogenous semigroups and the inequalities are written with weighted norms (for simplicity we have chosen to present the results with usual infinity norms in this Introduction).

If μ is a finite measure and σ is null, the solution of the above equation (1.1) relates to the class of Piecewise Deterministic Markov Process (abridged PDMP). These equations have been introduced in [14] and studied in detail in [15]. A wide literature is available on the subject of PDMP as they present an increasing amount of applications: on/off systems (cf. [9]), reliability (e.g., [16]), simulations and approximations of reaction networks (e.g., [2, 13, 22], with some error bounds hinted at in [28] or [21]), neuron models (e.g., [10, 11]), etc. The reader may equally take a look at the recent book [12] for an overview of some applications. In the engineering community, these equations are also known as “transport equations” (see [33] or [29]).

To the best of our knowledge, in the general case (including a diffusion component and an infinite number of jumps), under suitable assumptions, the first proof of existence and uniqueness of the solution of equation (1.1) is given in [24].

Now assume that one aims at applying some kind of numerical algorithm in order to simulate the solution of equation (1.1). Furthermore, assume for the moment, that $\sigma=b=0$ such that

$$X_t = x+\int_0^t\int_E\int_{(0,\infty)} c(s,z,X_{s-})1_{\{u\le\gamma(s,z,X_{s-})\}}N_\mu(ds,dz,du). \tag{1.5}$$

If $\mu(E)$ is finite, then one deals with a finite number of jumps in any interval of time, such that the solution X is given with respect to a compound Poisson process that can be explicitly simulated (leading, in particular in chemistry-inspired settings, to what is commonly known as Gillespie’s algorithm [22]; for other general aspects on simulation, see also [17, 32, 33]). However, even in this rather smooth case, the presence of a trajectory-triggered jump (i.e., dependence on x in the jump intensity γ) can lead, in certain regions (as γ gets large) to the accumulation of many (possibly) small jumps. In this case, the algorithm becomes very slow. One way of dealing with the problem is to replace these small jumps with an averaged motion leading (piecewise) to an ordinary differential equation (e.g., in [1]). Within the context of reaction networks, some intuitions on the partition of reactions and species to get the hybrid behavior as well as qualitative behavior (convergence to PDMP) are specified, for example, in [13]. Further heuristics can be found in [3].

In the general framework of infinite $\mu(E)$, this direct approach may fail to provide fast solutions (except particular situations, for example, in [36]). To provide an answer, the natural idea is to truncate the "small jumps" on some compatible family of sets $(E_n)_{n\in\mathbb{N}}$ and simulating X_t^n solution of

$$X_t^n = x + \int_0^t \int_{E_n^c} \int_{(0,\infty)} c_n(s, z, X_{s-}^n) 1_{\{u\le\gamma(s,z,X_{s-}^n)\}} N_\mu(ds, dz, du). \tag{1.6}$$

Here, as usual, we let $E_n^c = E \setminus E_n$, for all $n \in \mathbb{N}$. This procedure leads to a large error. To improve it, one might want to further replace the "small jumps" from E_n by a Brownian diffusion term leading to

$$\begin{aligned} X_t^n = x &+ \int_0^t \int_{E_n} \sigma_n(s, z, X_{s-}^n) W_\mu(ds, dz) + \int_0^t b_n(s, X_s^n)\, ds \\ &+ \int_0^t \int_{E_n^c} \int_{(0,\infty)} c_n(s, z, X_{s-}^n) 1_{\{u\le\gamma_n(s,z,X_{s-}^n)\}} N_\mu(ds, dz, du), \end{aligned} \tag{1.7}$$

where W_μ is a time-space Gaussian random measure [associated with $\mathbb{L}^2(\mu)$; standard procedure allows interpreting W_μ as in equation (1.1)]. The specific form of σ_n and b_n is obtained by using a second-order Taylor development in the infinitesimal operator of the initial equation.

This idea goes back to [4]. In the case of systems driven by a Lévy process (with γ fixed), [18] gives a precise estimate of the error and compares the approximation obtained by truncation as in equation (1.6) with the one obtained by adding a Gaussian noise as in equation (1.7). An enlightening discussion on the complexity of the two methods is also provided. Similar results concerning Kac's equation are obtained in [19] and for a Boltzmann-type equation in [23]. For some recent development on asymptotics of Boltzmann-type equation, we also mention [25]. Finally, it is worth mentioning that the converse approach (replacing Brownian with jump diffusions) may also be useful. The engineering literature is quite abundant in overviews of numerical methods for (continuous) diffusion processes using jump-type schemes. In this case, the stochastic integral with respect to the Brownian motion is replaced by an integral with respect to a jump process.

The aim of the present paper is to provide quantitative estimates of the weak approximation error when substituting the original system (1.1) with hybrid [piecewise diffusive Markov system (1.7)] in the general case when γ is trajectory-dependent (which constitutes the main difficulty to overcome). At intuitive level, Trotter–Kato-type results (cf. [35], Theorem 4.4) give the qualitative behavior. If $\mathcal{P}^n$ (resp., $\mathcal{L}^n$) is the semigroup (resp., infinitesimal generator) associated with (1.7) and $\mathcal{P}$ (resp., $\mathcal{L}$) is the semigroup (resp., infinitesimal generator) associated with (1.1), under Feller-type conditions, convergence of $\mathcal{L}^n$ to $\mathcal{L}$ will imply the corresponding convergence of semigroups. This type of qualitative behavior can be found, for instance, in [31] (leading to drift), [5] (leading to piecewise diffusive processes). In order to get error bounds (leading to a quantitative estimate), one employs (1.4).

A somewhat different motivation for our work comes from a method introduced in [8] (see also [7]) to study convergence to equilibrium for Markov chains. Roughly speaking, instead of looking into the long-time behavior of the Markov chain Y_n, $n \in \mathbb{N}$, one replaces this chain by a Markov process X_t sharing the same asymptotics ($t \mapsto X_t$ being an "asymptotic pseudotrajectory"). In [34], the results of our paper are used in order to extend this method (of [8]) to the piecewise deterministic Markov framework.

Finally, although many biological intuitions exist on the use of hybrid models for reaction systems (e.g., [3]), the quantitative estimates in our paper may turn out to provide a (purely mathematical) selection criterion for the components to be averaged and the contributions to be kept within the jump component. The use of diffusions punctuated by jumps (as mesoscopic approach) responds, on one hand, to the question of speeding up algorithms and, on the other, of keeping a high degree of stochasticity (needed, for example, to exhibit multistable regimes).

This paper is organized as follows. We begin with presenting the main notations used throughout the paper. We proceed, in Section 2.1 with the main elements leading to the processes involved. First, we recall some classical results on cylindrical diffusion-driven processes and the regularity of the induced flow (Section 2.1.1). Next, in Section 2.1.2, we introduce the jumping mechanism as well as the standing assumptions. We proceed with the construction of hybrid systems (piecewise diffusive with trajectory-triggered jumps) in Section 2.2. We begin with some localization estimates when the underlying measure is finite in Lemma 6. We also recall some elements on fictive and real shocks leading to some kind of Marked-point process representation of our system. These elements turn out to be of particular importance in providing the differentiability of the flow generated by our hybrid system. The norm notations and the $\mathbb{L}^p$-regularity of the solution (uniformly with respect to the initial data) are given in Section 3.

The differentiability of the associated semigroup is studied in Section 4 (with the main result being Theorem 14 whose uniform estimates extend to general underlying measures in Theorem 15).

Section 5 gives quantitative results on the distance between semigroups associated with such systems. The natural assumptions are presented in the first subsection. The main result Theorem 16 produces quantitative upper-bounds for the distance between semigroups starting from the distance between infinitesimal operators.

We present two classes of applications. In Section 6, we introduce a piecewise deterministic Markov process presenting three regimes and leading to a hybrid approximation with explicit distance on associated semigroups. First, we provide a theoretical framework describing the model, the regimes, the assumptions and the main qualitative behavior (in Theorem 19). Next, explicit measures make the object of a simple example to which our result is applied.

The second class of examples is given by a two-dimensional Boltzmann equation (following the approach in [6]) in Section 7. We begin with describing the model, its probabilistic interpretation and the (cut off) approximation given in [6] and leading to a pure-jump PDMP. In this approximated model, using our results, we replace small jumps with either a drift term (first-order approximation provided in Theorem 23) or a diffusion term (second-order approximation provided in Theorem 24).

2. Notation. Let $(E, \mathcal{E})$ be a measurable space and μ be a (fixed) nonnegative σ-finite measure on $(E, \mathcal{E})$:

- Given a standard Euclidean state-space $\mathbb{R}^m$, the spaces $\mathbb{L}^p(\mu)$ (for $1 \leq p \leq \infty$) will denote the usual space of p-power integrable, $\mathbb{R}^m$-valued functions defined on E. This space is endowed with the usual norm

$$\|\phi\|_{\mathbb{L}^p(\mu)} = \left(\int_E |\phi(z)|^p \mu(dz)\right)^{\frac{1}{p}},$$

  for all measurable function $\phi : E \to \mathbb{R}^m$. For notation purposes and by abuse of notation, the dependence on m is dropped [one should write $\mathbb{L}^2(\mu; \mathbb{R}^m)$]. The norm $|\cdot|$ denotes the classical, Euclidian norm on $\mathbb{R}^m$.
- The space $C_b^q(\mathbb{R}^m)$ is the space of real-valued bounded functions on $\mathbb{R}^m$ whose partial derivatives up to order q exist and are bounded and continuous.

Given a (fixed) probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a (fixed) time horizon $T > 0$:

- If ξ is an $\mathbb{R}^m$-valued random variable on Ω, we denote, as usual, $\|\xi\|_p = (\mathbb{E}[|\xi|^p])^{\frac{1}{p}}$.
- If Y is an adapted real-valued process and Z is an $\mathbb{L}^2(\mu)$-valued process, then we denote by

$$\|Y\|_{T,p} = \left(\mathbb{E}\Big[\sup_{t \leq T} |Y_t|^p\Big]\right)^{\frac{1}{p}} \quad \text{and} \quad \|Z\|_{T,p} = \left(\mathbb{E}\Big[\sup_{t \leq T} \|Z_t\|^p_{\mathbb{L}^2(\mu)}\Big]\right)^{\frac{1}{p}}.$$

- We use $\mathcal{M}_T$ to denote the space of the measurable functions $f : [0, T] \times E \times \mathbb{R}^d \to \mathbb{R}$ (where metric space are endowed with usual Borel fields). For $f \in \mathcal{M}_T$, we consider the norm

$$\|f\|_{(\mu,\infty)} = \sup_{t \leq T} \sup_{x \in R^d} \|f(t, \cdot, x)\|_{\mathbb{L}^2(\mu)}. \tag{2.1}$$

- Similar norm can be induced on $\mathcal{M}_T^d$ by replacing $\mathbb{L}^2(\mu; \mathbb{R})$ with $\mathbb{L}^2(\mu; \mathbb{R}^d)$ norms.
- For a multi-index $\alpha = (\alpha_1, \dots, \alpha_q) \in \{1, \dots, d\}^q$, we denote $|\alpha| = q$ the length of α and $\partial_x^\alpha = \partial_{x_{\alpha_1}} \cdots \partial_{x_{\alpha_q}}$ the corresponding derivative. To simplify notation, the variable x may be suppressed and we will use ∂^α.

- For $k \in \mathbb{N}^*$, we denote by $\mathbb{R}_{[k]}$ the family of real-valued vectors indexed by multi-indexes of at most k length, that is, $\mathbb{R}_{[k]} = \{y_{[k]} = (y_\beta)_{1\le|\beta|\le k} : y_\beta \in \mathbb{R}\}$ and, for $y_{[k]} \in \mathbb{R}_{[k]}$ we denote

$$|y_{[k]}|_{\mathbb{R}_{[k]}} = \sum_{1\le|\beta|\le k} |y_\beta|^{\frac{k+1}{|\beta|}}. \tag{2.2}$$

By convention, $|y_{[0]}|_{\mathbb{R}_{[0]}} = 0$. Similarly, $\mathbb{R}^d_{[k]}$ is defined for vectors whose components belong to $\mathbb{R}^d$ and $|y_\beta|$ is then computed with respect to the usual Euclidian norm on $\mathbb{R}^d$.

- If $x \mapsto f(t,z,x)$ is a real-valued, q times differentiable function for every $(t,z) \in [0,T] \times E$ then, for every $1 \le l \le q$ we denote

$$\begin{aligned} \|f\|_{l,q,(\mu,\infty)} &= \sum_{l\le|\alpha|\le q} \|\partial_x^\alpha f\|_{(\mu,\infty)} \quad \text{and} \\ \|f\|_{q,(\mu,\infty)} &= \|f\|_{(\mu,\infty)} + \|f\|_{1,q,(\mu,\infty)} \end{aligned} \tag{2.3}$$

REMARK 1. We emphasize that in $\|f\|_{l,q,(\mu,\infty)}$, for $l \ge 1$, only derivatives are involved ($\|f\|_{(\mu,\infty)}$ itself does not appear).

- For a measurable function $g : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ we denote by $\|g\|_\infty = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |g(t,x)|$ and, if $x \mapsto g(t,x)$ is q times differentiable for every $t \in [0,T]$, then

$$\|g\|_{l,q,\infty} = \sum_{l\le|\alpha|\le q} \|\partial_x^\alpha g\|_\infty \quad \text{and} \quad \|g\|_{q,\infty} = \|g\|_\infty + \|g\|_{1,q,\infty}.$$

2.1. *Preliminary results.*

2.1.1. *Continuous diffusion.* We assume the fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to be endowed with a Gaussian noise W_μ based on μ, as introduced by Walsh in [38]. We recall that W_μ is a family of centred Gaussian random variables $W_\mu(t,h)$ indexed by $(t,h) \in \mathbb{R}_+ \times \mathbb{L}^2(\mu)$ with covariances $\mathbb{E}[W_\mu(t,h)W_\mu(s,g)] = (t \wedge s)\langle h,g\rangle_{\mathbb{L}^2(\mu)}$. Note that whenever $(e_l)_{l\in\mathbb{N}} \in \mathbb{L}^2(\mu)$ is an orthonormal basis, the family $(W_\mu(t,e_l))_{l\in\mathbb{N}}$ is a sequence of independent standard Brownian motions.

We briefly recall the stochastic integral with respect to W_μ. One considers the natural filtration $\mathcal{F}_t^W = \sigma(W_\mu(s,h) : s \le t, h \in \mathbb{L}^2(\mu))$, for all $t \ge 0$. For a process $\phi : \mathbb{R}_+ \times \Omega \to \mathbb{L}^2(\mu)$ which is adapted [i.e., $\langle \phi_t, h\rangle_{\mathbb{L}^2(\mu)}$ is $\mathcal{F}_t^W$ measurable for every $h \in \mathbb{L}^2(\mu)$] and for which $\mathbb{E}[\int_0^T \|\phi_t\|^2_{L^2(\mu)}\,dt] < \infty$, for every $T > 0$, one defines

$$\int_0^t \int_E \phi_s(z) W_\mu(ds,dz) := \sum_{l=1}^\infty \int_0^t \langle \phi_s, e_l\rangle_{\mathbb{L}^2(\mu)} W_\mu(ds,e_l). \tag{2.4}$$

Let $0 \le s \le T$ be fixed. A nonhomogeneous continuous diffusion process $\Phi_{s,t}(x)$, $s \le t \le T$ driven by W_μ with (regular) coefficients σ and b is the solution of the stochastic equation

$$\begin{aligned}\Phi_{s,t}(x) &= x + \int_s^t \int_E \sigma\big(u, z, \Phi_{s,u}(x)\big) W_\mu(du, dz) + \int_s^t b\big(u, \Phi_{s,u}(x)\big)\, du \\ &= x + \sum_{l=1}^{\infty} \int_0^t \sigma_l\big(u, \Phi_{s,u}(x)\big)\, dB_u^l + \int_s^t b\big(u, \Phi_{s,u}(x)\big)\, du,\end{aligned} \tag{2.5}$$

with $B_s^l = W_\mu(s, e_l)$ and $\sigma_l(u, x) = \langle \sigma(u, \cdot, x), e_l \rangle_{\mathbb{L}^2(\mu)}$.

The following result is standard for finite-dimensional Brownian motions (e.g., [26, 30]) and its generalization to this setting is quite forward.

PROPOSITION 2. *Let us assume the following norm condition to hold true*:

$$\|\nabla \sigma\|_{(\mu,\infty)} + \|\nabla b\|_\infty < \infty. \tag{2.6}$$

Then, for every initial datum $x \in \mathbb{R}^d$, *equation* (2.5) *has a unique strong solution. Moreover, if* $\|\sigma\|_{1,q+1,(\mu,\infty)} + \|b\|_{1,q,\infty} < \infty$, *then there exists a version of this solution such that* $x \mapsto X_{s,t}(x)$ *is* q *times differentiable.*

REMARK 3. Let us note that the following (more popular) alternative representation for this diffusion holds: let $a^{i,j}(t, x) = \int_E \sigma^i \sigma^j(t, z, x)\mu(dz)$, $1 \le i, j \le d$ and set $\widehat{\sigma} = a^{\frac{1}{2}}$. Then the law of $\Phi_{s,t}$ coincides with the law of $\widehat{\Phi}_{s,t}$ solution of

$$\widehat{\Phi}_{s,t}(x) = x + \sum_{j=1}^{d} \int_s^t \widehat{\sigma}_j\big(u, \widehat{\Phi}_{s,u}(x)\big)\, dB_u^j + \int_s^t b\big(u, \widehat{\Phi}_{s,u}(x)\big)\, du,$$

where $B = (B^1, \dots, B^d)$ is a standard Brownian motion. We prefer working with the representation $\Phi_{s,t}$ (and not with $\widehat{\Phi}_{s,t}$) for two reasons. First, the stochastic integral with respect to $W_\mu(du, dz)$ naturally appears in our problem. Moreover, if one liked to work with $\widehat{\Phi}_{s,t}$, then one would have to compute $\widehat{\sigma} = a^{\frac{1}{2}}$ and to derive regularity properties for $\widehat{\sigma}$ from regularity properties for a, and this is more delicate (one needs some ellipticity property for a). In contrast, if one starts with equation (2.5), then the proof of the previous proposition is a straightforward extension of the classical results.

2.1.2. *Jump mechanism and further notation.* We assume the space Ω to be large enough to contain an independent Poisson point measure on $E \times \mathbb{R}_+$ denoted by N_μ and having a compensator $\widehat{N}_\mu(ds, dz, du) = ds\,\mu(dz)\, du$. (For further constructions and properties, the reader is referred to [26].) We just mention that, whenever $A_l \times I_l \in \mathcal{E} \times \mathcal{B}(\mathbb{R}_+)$, $l = 1, \dots, m$ are disjoint sets, then $t \mapsto N_\mu(t, A_l \times I_l)$ are independent Poisson processes with parameters $\mu(A_l) \times \mathrm{Leb}(I_l)$. Here, $\mathcal{B}(\mathbb{R}_+)$ stands for the family of Borel subsets of $\mathbb{R}_+$.

We consider now the coefficients $c \in \mathcal{M}_T^d$ and $\gamma \in \mathcal{M}_T$ and we assume that there exist some functions $l_c, l_\gamma : E \to \mathbb{R}_+$ such that

$$C_\mu(\gamma, c) := \sup_{t \le T} \sup_{x \in R^d} \int_E \big(l_\gamma(z)\big|c(t,z,x)\big| + l_c(z)\gamma(t,z,x)\big)\mu(dz) < \infty \tag{2.7}$$

and such that, for every $x, y \in \mathbb{R}^d$, every $t \ge 0$ and $z \in E$,

$$\begin{aligned} \big|c(t,z,x) - c(t,z,y)\big| &\le l_c(z)|x-y|, \\ \big|\gamma(t,z,x) - \gamma(t,z,y)\big| &\le l_\gamma(z)|x-y|. \end{aligned} \tag{2.8}$$

Moreover, we assume that γ takes nonnegative values and

$$\Gamma := \sup_{t \le T} \sup_{x \in \mathbb{R}^d} \sup_{z \in E} \gamma(t,z,x) < \infty. \tag{2.9}$$

We also set, for any Borel set $G \subset E$,

$$\alpha(G) := \sup_{t \le T} \sup_{x \in R^d} \int_G \big|c(t,z,x)\big|\gamma(t,z,x)\mu(dz) \tag{2.10}$$

and assume that $\alpha(E) < \infty$.

2.2. *The hybrid system.* We are interested in the (hybrid) stochastic differential equation

$$\begin{aligned} X_{s,t}(x) = x &+ \int_s^t \int_E \sigma\big(r,z,X_{s,r}(x)\big) W_\mu(dr,dz) + \int_s^t b\big(r, X_{s,r}(x)\big)\,dr \\ &+ \int_s^t \int_{E \times [0,2\Gamma]} c\big(r,z,X_{s,r-}(x)\big) 1_{\{u \le \gamma(r,z,X_{s,r-}(x))\}} N_\mu(dr,dz,du). \end{aligned} \tag{2.11}$$

REMARK 4. The stochastic components W_μ and N_μ are assumed to be associated with the same measurable space $(E, \mathcal{E}, \mu)$. This assumption is made in order to avoid heavy notation. Alternatively, one may consider W_μ on $(E, \mathcal{E}, \mu)$ and N_ν on some (independent) space $(F, \mathcal{F}, \nu)$. For most examples, the space $E = \{1, \dots, d\}$ and the uniform measure $\mu(i) = \frac{1}{d}$, for all $i \in E$ play an important role. In this setting, $W_\mu(dr,dz) = \frac{1}{d}\sum_{i=1}^d dW_r^i$, such that one comes back to a usual diffusion process driven by a finite-dimensional Brownian motion.

The following result gives the existence and uniqueness of the solution to our hybrid system in the class of càdlàg processes in $\mathbb{L}^1$.

THEOREM 5. *Suppose that* (2.7), (2.8), (2.9), (2.10) *and* (2.6) *hold. Then equation* (2.11) *has a unique* $\mathbb{L}^1$ *solution* [*i.e., a* cadlag *process* $X_{s,t}(x)$, $t \ge s$ *with* $E(|X_{s,t}(x)|) < \infty$ *which verifies* (2.11)].

The above theorem has been first proven in [24]. The main idea is that, in contrast with the standard approach to *SDEs* relying on $\mathbb{L}^2$ norms, one has to work here with $\mathbb{L}^1$ norms. This is due to the indicator function appearing in the Poisson noise. We shortly recall this argument in the following.

2.2.1. *Localization estimates for μ.* For a set $G \subset E$, we denote by $X^G_{s,t}$ the solution of the equation (2.11) in which the measure μ is restricted to G, that is, substituted by $1_G(z)\,d\mu(z)$. The first step gives the behavior of such solutions for different sets G.

LEMMA 6. *We suppose that* (2.7), (2.8), (2.9), (2.10) *and* (2.6) *hold. Let $G_1 \subset G_2 \subset E$ be two measurable sets* (*the case $G_1 = G_2 = E$ is included*) *and let $\Delta X_{s,t} = X^{G_1}_{s,t} - X^{G_2}_{s,t}$. There exists a universal constant C such that for every $T \geq 0$ one has*

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\le t\le T}|\Delta X_t|\Big] &\le \big(|\Delta X_{s,s}| + T\alpha(G_2 \setminus G_1)\big)\\ &\quad\times \exp\big(CT\big(\|\nabla\sigma\|_{(\mu,\infty)} + \|\nabla b\|_\infty + C_\mu(\gamma,c)\big)^2 + 1\big).\end{aligned} \tag{2.12}$$

The proof is quite straightforward. For our readers' sake, the elements of proof are gathered in Section 8.1.

Let us now discuss the construction of a solution of the equation (2.11) and present two alternative representations of this solution.

2.2.2. *Fictive shocks on increasing support sets.* We fix $G \subset E$ with $\mu(G) < \infty$ and we recall that γ is upper-bounded by Γ [see (2.9)]. We also fix $s \geq 0$ and we will construct $X^G_{s,t}$ solution of the equation (2.11) associated with $1_G(z)\mu(dz)$ using a compound Poisson process as follows.

One takes J_t to be a (usual) Poisson process of parameter $2\Gamma\mu(G)$ and denotes by T_k, $k \in \mathbb{N}$ the jump times of J_t. Moreover, one considers two sequences of independent random variables Z_k and U_k, $k \in \mathbb{N}$ (independent of J_t as well and supported by the set Ω assumed to be large enough). These random variables are distributed

$$\mathbb{P}(Z_k \in dz) = \frac{1}{\mu(G)} 1_G(z)\mu(dz), \qquad \mathbb{P}(U_k \in du) = \frac{1}{2\Gamma} 1_{[0,2\Gamma]}(u)\,du. \tag{2.13}$$

Finally, one defines the continuous stochastic flow $\Phi_{s,t}(x)$, $0 \leq s \leq t$ to be the solution of the *SDE* (2.5). Then the solution $X^G_{s,t}$ of the equation (2.11) associated with $1_G(z)\mu(dz)$ is constructed by setting $X^G_{s,s}(x) = x$ and

$$X^G_{s,t}(x) = \Phi_{T_k,t}\big(X^G_{s,T_k}(x)\big) \qquad \text{on } T_k \leq t < T_{k+1}$$

and

$$\begin{aligned}X^G_{s,T_{k+1}}(x) &= X^G_{s,T_{k+1}-}(x) + c\big(T_{k+1}, Z_{k+1}, X^G_{s,T_{k+1}-}(x)\big)\\ &\quad\times 1_G(Z_{k+1})1_{\{U_k \le \gamma(T_{k+1}, Z_{k+1}, X^G_{s,T_{k+1}-}(x))\}},\end{aligned}$$

where $X^G_{s,T_{k+1}-}(x) = \Phi_{T_k,T_{k+1}}(X_{s,T_k}(x))$. This gives the solution of the equation

$$
\begin{aligned}
X^G_{s,t} &= x + \int_s^t \int_E \sigma\bigl(r, z, X^G_{s,r}\bigr) W_\mu(dr, dz) + \int_s^t b\bigl(r, X^G_{s,r}\bigr)\, dr \\
&\quad + \sum_{k=J_s+1}^{J_t} c\bigl(T_k, Z_k, X^G_{s,T_k-}(x)\bigr) 1_G(Z_k) 1_{\{U_k \le \gamma(T_k, Z_k, X^G_{s,T_k-}(x))\}}
\end{aligned}
\tag{2.14}
$$

that is (2.11) associated with $1_G(z)\mu(dz)$. This is the so-called "fictive shock" representation (see [33]).

REMARK 7. To construct the global solution, one begins with considering a sequence $E_n \uparrow E$ with $\mu(E_n) < \infty$. Then one constructs $X^{E_n}_{s,t}$ as before and checks [using (2.12)] that this is a Cauchy sequence. Passing to the limit, one obtains $X_{s,t}$ solution of the general equation (2.11). Uniqueness follows directly from (2.12).

For the simplicity of the notation, in the following we will work with $s = 0$. The estimates for $s > 0$ are quite similar. As usual, we will denote $X^G_{0,t}(x)$ by $X^G_t(x)$.

2.2.3. *Real shocks*. We construct now the "real shock" representation $\overline{X}^G_t$ in the following way. We define $E_* = E \cup \{z_*\}$, where z_* is a point which does not belong to E and we extend μ to E_* by setting $\mu(z_*) = 1$. We also extend $c(t, z, x)$ to E_* by $c(t, z_*, x) = 0$, for every $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Given a sequence $(z_k)_{k\in\mathbb{N}} \subset E_*$, one denotes $z^k = (z_1, \dots, z_k)$ and constructs $x_t(x, z^{J_t})$ as follows:

$$
\begin{aligned}
x_t\bigl(x, z^k\bigr) &= \Phi_{T_k,t}\bigl(x_{T_k}\bigl(x, z^k\bigr)\bigr) \qquad \text{on } T_k \le t < T_{k+1}, \\
x_{T_{k+1}}\bigl(x, z^{k+1}\bigr) &= x_{T_{k+1}-}\bigl(x, z^k\bigr) + c\bigl(T_{k+1}, z_{k+1}, x_{T_{k+1}-}\bigl(x, z^k\bigr)\bigr) 1_G(z_{k+1}).
\end{aligned}
\tag{2.15}
$$

Next, we define, for every $(t, z, x) \in \mathbb{R}_+ \times E_* \times \mathbb{R}^d$,

$$
\begin{aligned}
q_G(t, z, x) &= \Theta_G(t, x) 1_{\{z_*\}}(z) + \frac{1}{2\Gamma\mu(G)} 1_G(z)\gamma(t, z, x) \\
&\qquad \text{where } \Theta_G(t, x) = 1 - \frac{1}{2\Gamma\mu(G)} \int_G \gamma(t, z, x)\, d\mu(z).
\end{aligned}
\tag{2.16}
$$

We consider a sequence of random variables $(\overline{Z}_k)_{k\in\mathbb{N}}$ with the laws constructed recursively by

$$
E\bigl(\overline{Z}_k \in dz \mid x_{T_k-}\bigl(x, \overline{Z}^{k-1}\bigr) = y\bigr) = q_G(T_k, z, y)\mu(dz),
\tag{2.17}
$$

where $\overline{Z}^{k-1} = (\overline{Z}_1, \dots, \overline{Z}_{k-1})$. Finally, we define $\overline{X}_t^G(x) = x_t(x, \overline{Z}^{J_t})$. This amounts to saying that

$$\begin{aligned} \overline{X}_t^G(x) = x + \int_0^t \int_E \sigma\big(s, z, \overline{X}_s^G\big)\, dW_\mu(ds, dz) \\ + \int_0^t b\big(s, \overline{X}_s^G\big)\, ds + \sum_{k=1}^{J_t} c\big(T_k, \overline{Z}_k, \overline{X}_{T_k-}^G\big) 1_G(\overline{Z}_k). \end{aligned} \tag{2.18}$$

Equation (2.18) is similar to equation (2.14) but now $1_{\{U_k \le \gamma(T_k, Z_k, X_{T_k-}(x))\}}$ no longer appears.

REMARK 8. If σ and b are smooth functions, one can choose a variant of $x \mapsto \Phi_{s,t}(x)$ that is almost surely differentiable. Moreover, if $x \mapsto c(t, z, x)$ is also smooth, then $x \mapsto x_t(x, z^{J_t})$ is smooth as well. So $x \mapsto \overline{X}_t^G(x)$ will be also differentiable. In contrast, if one represents $X_t^G(x) = y_t(x, U^{J_t}, Z^{J_t})$, then the application $x \mapsto y_t(x, z^{J_t}, u^{J_t})$ is generally no longer differentiable (which is due to the presence of the indicator function $1_{\{u_k \le \gamma(T_k, z_k, y_{T_{k-}}(x, z^k, u^k))\}}$). This explains why, in the following, we favor the real-shocks representation and we treat in a separate way the derivatives of $x \mapsto x_t(x, z^{J_t})$, respectively, the derivatives of the law of $\overline{Z}_k$ with respect to x (see the proof of Theorem 14).

Moreover, we have the following well-known identity of laws result.

LEMMA 9. *The law of* $(\overline{X}_t^G(x))_{t \ge 0}$ *coincides with the law of* $(X_t^G(x))_{t \ge 0}$ *solution to* (2.14). *In particular,* $\mathcal{P}_t^G f(x) := \mathbb{E}[f(X_t^G(x)] = \mathbb{E}[f(\overline{X}_t^G(x)]$.

3. Differentiability of the flow. We will now study the differentiability of the application $x \mapsto \overline{X}_t^G(x)$ when assuming $\mu(G) < \infty$. Let us begin with introducing some further notation. Given a regular function $g : \mathbb{R}_+ \times E \times \mathbb{R}^d \to \mathbb{R}$ that is differentiable with respect to the space variable $x \in \mathbb{R}^d$ we denote by

$$|g|_{G,p} = \sup_{0 \le t \le T} \sup_{x \in \mathbb{R}^d} \left(\int_G |g(t, z, x)|^p \gamma(t, z, x) \mu(dz) \right)^{\frac{1}{p}}, \tag{3.1}$$

$$[g]_{G,p} = \sup_{1 \le p' \le p} |g|_{G,p'}, \tag{3.2}$$

$$\theta_{q,p}(G) = 1 + \|\sigma\|_{2,q,(\mu,\infty)} + \|b\|_{2,q,\infty} + \sum_{2 \le |\alpha| \le q} \big[\partial_x^\alpha c\big]_{G,p}, \tag{3.3}$$

$$a_p(G) = \|\nabla \sigma\|_{(\mu,\infty)}^2 + \|\nabla b\|_\infty + [\nabla c]_{G,p}^p, \tag{3.4}$$

$$\alpha_{q,p}(C, G) = C \theta_{q,pq}^{\varkappa(q)}(G) \exp\big(C T \varkappa(q) a_{pq}(G)\big). \tag{3.5}$$

Here, $\varkappa(q) = q \sum_{1\le n\le q} \frac{1}{n}$ is a universal constant increasing with $q \geq 1$ and this exact form is used in proofs by recurrence. In general, one easily notes that $\varkappa(q) \leq q + \ln q$, for all $q \geq 1$. Note that if $q_1 \leq q_2$ and $p_1 \leq p_2$ then $\theta_{q_1,p_1}(G) \leq \theta_{q_2,p_2}(G)$ and $a_{p_1}(G) \leq a_{p_2}(G)$ (this is the reason of being of $\sup_{1\le p'\le p}$ in $[g]_{G,p}$).

Throughout the paper, unless mentioned otherwise, the constant C will be a generic one (assumed to be independent of both the time horizon $T \geq 0$ and the coefficients). However, the reader is invited to note that the actual upper bounds involving terms such as $\alpha_{q,p}$ actually make use of T in an exponential factor and terms such as $|g|_{G,p}$, etc. depend on the behavior of coefficients on the entire time interval $[0, T]$.

In the following, we suppose that $\mu(G) < \infty$ and $\theta_{q,p}(G) < \infty$ and consider $X_t^G(x)$ and $\overline{X}_t^G(x)$, solutions of the equations (2.14) and (2.18) constructed in the previous section. Under these hypotheses, one may choose a variant of $x \mapsto \overline{X}_t^G(x)$ which is q times differentiable. Our aim is to estimate the $\mathbb{L}^p$ norm of $\partial^\alpha \overline{X}_t^G(x)$.

LEMMA 10. *Let α be a multi-index with $|\alpha| = k$. For every $p \geq 2$, the following inequality holds true*:

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[\sup_{t\le T}\big|\partial^\alpha \overline{X}_t^G(x)\big|^p\Big]^{\frac{1}{p}} \leq \alpha_{k,p}(C, G). \tag{3.6}$$

The proof is postponed to Section 8.2.2. The main idea consists in providing estimates for the chain rule distinguishing first-order and higher-order derivatives. Subsequently, these estimates will be applied for the different components in the differential formula of $\partial^\alpha \overline{X}_t^G(x)$. Next, one provides estimates for generic equations of this type (having a linear form) and uses a recurrence argument over $|\alpha|$.

We give now some consequences of (3.6).

COROLLARY 11. **A** *Let α be a multi-index with $|\alpha| = q \geq 1$ and let $p \geq 2$ and $\eta > 0$ be given. For every $g : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ that is smooth with respect to $x \in \mathbb{R}^d$, the following inequality holds true*:

$$\left\|\sum_{k=1}^{J_t}\big|\partial^\alpha g\big(T_k, \overline{X}_{T_k-}^G(x)\big)\big|\right\|_p \leq C\|g\|_{1,q,\infty}\Gamma\mu(G)(t\vee 1)\alpha_{q,(1+\eta)pq}^q(C, G). \tag{3.7}$$

B *Let $g : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be a function that is smooth with respect to $x \in \mathbb{R}^d$. Then*

$$\begin{aligned}&\left\|\sum_{k=1}^{J_t} 1_G(\overline{Z}_k)\big|\partial^\alpha g\big(T_k, \overline{Z}_k, \overline{X}_{T_k-}^G(x)\big)\big|\right\|_p\\ &\quad\leq C\Big(\sum_{1\le|\beta|\le|\alpha|}\big[\partial^\beta g\big]_{G,(1+\eta)p}\alpha^q_{q,\frac{(1+\eta)}{\eta}pq}(C, G)\Big).\end{aligned} \tag{3.8}$$

For our readers' sake, the proof is provided in Section 8.2.3.

4. Differentiability of the semigroup. Before giving the main result on the semigroup of operators, we recall that the law of $\overline{Z}^{Jt} = (\overline{Z}_1, \dots, \overline{Z}_{J_t})$ conditional to $J_t > 0$ has, as density, $p_{J_t}(x, z^{J_t})\mu(dz_1), \dots, \mu(dz_{J_t})$ with

$$p_{J_t}(x, z^{J_t}) = \prod_{k=1}^{J_t} q_G(T_k, z_k, x_{T_k-}(x, z^{k-1})).$$

By convention, on $J_t = 0$, we set $p_{J_t} = 1$. This explicit formulation allows one to obtain a very important step in proving regularity of the semigroup.

LEMMA 12. *Let us assume that $\Gamma\mu(G) \geq 1$ and $\alpha_{q,2qp}(C, G) < \infty$ for some given $q \in \mathbb{N}$ and $p \geq 2$. Let α be a multi-index with $|\alpha| = q \geq 1$. Then there exists a universal constant C (depending on p and q but not on G) such that*

$$\begin{aligned} \|\partial^\alpha \ln p_{J_t}(x, \overline{Z}^{J_t})\|_p \leq C(t \vee 1) \times \alpha^q_{q,2pq}(C, G) \\ \times \left(\Gamma_{G,q}(\gamma) + \sum_{1 \leq |\beta| \leq q} [\partial^\beta \ln \gamma]_{G,2p}\right) \end{aligned} \tag{4.1}$$

with $[\ln \gamma]_{G,2p}$ defined in (3.2) and

$$\Gamma_{G,q}(\gamma) = \sup_{t \leq T} \sup_{x \in R^d} \sum_{h=1}^{q} \sum_{1 \leq |\rho| \leq h} \left(\int_G |\partial^\rho \ln \gamma(t, z, x)|^{\frac{h}{|\rho|}} \gamma(t, z, x)\mu(dz)\right)^{\frac{q}{h}}. \tag{4.2}$$

Before going any further, we make the following elementary remark.

REMARK 13. For every smooth function $\phi : \mathbb{R}^d \to \mathbb{R}^*_+$ and any multi-index ρ with $|\rho| = q$, one gets the existence of some function $P^\phi_\rho : \mathbb{R}^d \to \mathbb{R}^*_+$ such that

$$\partial^\rho \phi(x) = \phi(x) P^\phi_\rho(x) \quad \text{and} \quad |P^\phi_\rho(x)| \leq C \sum_{1 \leq |\beta| \leq q} |\partial^\beta \ln \phi(x)|^{\frac{q}{|\beta|}}, \tag{4.3}$$

for all $x \in \mathbb{R}^d$. In order to prove this, one first writes $\phi = \exp(\ln \phi))$ and then takes derivatives. One obtains ϕ multiplied with a polynomial applied to terms of type $\partial^\beta \ln \phi$, that is, some linear combination of products of type $\prod_{i=1}^r \partial^{\beta(i)} \ln \phi$ with $\sum_{i=1}^r |\beta(i)| = q$. Using Young's inequality with $p_i = \frac{q}{|\beta(i)|}$, we obtain

$$\left|\prod_{i=1}^r \partial^{\beta(i)} \ln \phi\right| \leq \sum_{i=1}^r \frac{|\beta(i)|}{q} |\partial^{\beta(i)} \ln \phi|^{\frac{q}{|\beta(i)|}}.$$

And this proves the upper bound for $|P^\phi_\rho(x)|$ given in (4.3).

We are now able to proceed with the proof of Lemma 12.

PROOF. We have

$$\begin{aligned}\partial^\alpha \ln p_{J_t}(x, z^{J_t}) &= \sum_{k=1}^{J_t} 1_{\{z_*\}}(z_k)\partial^\alpha\big(\ln \Theta_G\big(T_k, x_{T_k-}(x, z^{k-1})\big)\big) \\ &\quad + \sum_{k=1}^{J_t} 1_G(z_k)\partial^\alpha\big(\ln \gamma\big(T_k, z_k, x_{T_k-}(x, z^{k-1})\big)\big) \\ &=: s_1(x, z^{J_t}) + s_2(x, z^{J_t}).\end{aligned}$$

In order to estimate $s_1(x, z^{J_t})$, we will use (3.7) for $g = \ln \Theta_G$. Recalling that γ is upper bounded by Γ, we have $\Theta_G(t, x) \geq \frac{1}{2}$. Then, for every multi-index α with $|\alpha| = q$, one has

$$\begin{aligned}\big|\partial^\alpha \ln \Theta_G(t, x)\big| &\leq \sum_{r=1}^{q} \sum_{|\beta(1)|+\cdots+|\beta(r)|=q} \prod_{i=1}^{r} \big|\partial^{\beta(i)}\Theta_G(t, x)\big| \\ &\leq \sum_{r=1}^{q} \frac{C}{(\Gamma\mu(G))^r} \sum_{|\beta(1)|+\cdots+|\beta(r)|=q} \prod_{i=1}^{r} \int_G \big|\partial^{\beta(i)}\gamma(t, z, x)\big|\mu(dz).\end{aligned}$$

Using Young's inequality and (4.3) (recall that $\Gamma\mu(G) \geq 1$), one proves

$$\begin{aligned}&\prod_{i=1}^{r} \int_G \big|\partial^{\beta(i)}\gamma(t, z, x)\big|\mu(dz) \\ &\quad\leq C\sum_{i=1}^{r}\bigg(\int_G \big|\partial^{\beta(i)}\gamma(t, z, x)\big|\, d\mu(z)\bigg)^{\frac{q}{|\beta(i)|}} \\ &\quad= C\sum_{i=1}^{r}\bigg(\int_G \big|P^{\gamma}_{\beta(i)}(t, z, x)\big|\gamma(t, z, x)\mu(dz)\bigg)^{\frac{q}{|\beta(i)|}} \\ &\quad\leq C\sum_{1\leq|\rho|\leq h\leq q}\bigg(\int_G \big|\partial^{\rho}\ln\gamma(t, z, x)\big|^{\frac{h}{|\rho|}}\gamma(t, z, x)\mu(dz)\bigg)^{\frac{q}{h}} \leq C\Gamma_{G,q}(\gamma).\end{aligned}$$

We conclude that $|\partial^\alpha \ln \Theta_G(t, x)| \leq \frac{C}{\Gamma\mu(G)}\Gamma_{G,q}(\gamma)$. As a consequence of (3.7) (with $\eta = 1$), one gets

$$\big(\mathbb{E}\big[\big|s_1\big(x, \overline{Z}^{J_t}\big)\big|^p\big]\big)^{\frac{1}{p}} \leq C(t \vee 1)\Gamma_{G,q}(\gamma)\alpha^q_{q,2pq}(C, G).$$

To estimate the second term, we use (3.8) with $g(t, z, x) = \ln \gamma(t, z, x)$ and for $\eta = 1$ to get an upper bound given by

$$\begin{aligned}&\mathbb{E}\bigg[\bigg|\sum_{k=1}^{J_t} 1_G(\overline{Z}_k)\partial^\alpha \ln\gamma\big(T_k, z_k, \big(\overline{X}_{T_k-}(x)\big)\big)\bigg|^p\bigg] \\ &\qquad\leq C\alpha^q_{q,2pq}(C, G)\sum_{1\leq|\beta|\leq q}\big[\partial^\beta \ln\gamma\big]_{G,2p}.\end{aligned}$$

The proof is now complete. □

We now discuss the differentiability of the semigroup associated with our process. To this purpose, we let for f regular enough

$$\begin{aligned}\mathcal{P}_t^G f(x) &= \mathbb{E}[f(X_t^G(x))] = \mathbb{E}[f(\overline{X}_t^G(x))] \\ &= \mathbb{E}\left[\int_{E^{J_t}} f(x_t(x, z^{J_t})) p_{J_t}(x, z^{J_t}) \mu(dz_1) \cdots \mu(dz_{J_t})\right].\end{aligned}$$

THEOREM 14. *We assume* (2.7), (2.8), (2.9), (2.10) *and* (2.6) *to hold true. Then, for every $q \in \mathbb{N}$, there exists a constant $C > 0$ independent of G such that*

$$\begin{aligned}\|\mathcal{P}_t^G f\|_{q,\infty} &\le C\|f\|_{q,\infty} \times (t \vee 1)^q \times \alpha_{q,4q}^{2q}(C, G) \\ &\quad \times \left(1 + \Gamma_{G,q}(\gamma) + \sum_{1 \le |\beta| \le q} [\partial^\beta \ln \gamma]_{G,4q}\right)^q.\end{aligned} \tag{4.4}$$

PROOF. Let us begin with writing

$$\mathbb{E}[f(\overline{X}_t^G(x))] = \mathbb{E}[f(\overline{X}_t^G(x)) 1_{J_t=0}] + \mathbb{E}[f(\overline{X}_t^G(x)) 1_{J_t>0}].$$

For the first term, one notes that

$$\mathbb{E}[f(\overline{X}_t^G(x)) 1_{J_t=0}] = \mathbb{E}[\Phi_{0,t}(x) 1_{J_t=0}] = \mathbb{E}[\Phi_{0,t}(x)]\mathbb{P}[J_t = 0].$$

Estimates on $\partial^\alpha \mathbb{E}[f(\overline{X}_t^G(x)) 1_{J_t=0}]$ follow from standard arguments on diffusive (continuous) flows. To conclude, we want to estimate, for a multi-index α such that $|\alpha| \le q$, the partial derivative

$$\begin{aligned}&\partial^\alpha \mathbb{E}[f(\overline{X}_t^G(x)) 1_{J_t>0}] \\ &\quad = \sum_{(\beta,\rho)=\alpha} \mathbb{E}\left[\int_{E^{J_t}} \partial^\beta f(x_t(x, z^{J_t})) \times \partial^\rho p_{J_t}(x, z^{J_t}) \mu(dz_1) \cdots \mu(dz_{J_t})\right] \\ &\quad = \sum_{(\beta,\rho)=\alpha} \mathbb{E}\left[\int_{E^{J_t}} \partial^\beta f(x_t(x, z^{J_t})) P_\rho^{p_{J_t}(\cdot, z^{J_t})}(x, z^{J_t}) \right. \\ &\qquad \left. \times p_{J_t}(x, z^{J_t}) \mu(dz_1) \cdots \mu(dz_{J_t})\right]\end{aligned}$$

with $P_\rho^{p_{J_t}(\cdot, z^{J_t})}(x, z^{J_t})$ given by (4.3). Using the Cauchy–Schwarz inequality, we have

$$|\partial^\alpha \mathbb{E}[f(\overline{X}_t^G(x))]| \le \sum_{(\beta,\rho)=\alpha} A_\beta^{\frac{1}{2}} \times B_\rho^{\frac{1}{2}}$$

with

$$\begin{aligned}A_\beta &= \mathbb{E}[|\partial^\beta f(\overline{X}_t^G(x))|^2], \\ B_\rho &= \mathbb{E}\left[\int_{E^{J_t}} |P_\rho^{p_{J_t}(\cdot, z^{J_t})}(x, z^{J_t})|^2 \times p_{J_t}(x, z^{J_t}) \mu(dz_1) \cdots \mu(dz_{J_t})\right].\end{aligned}$$

We have (recalling that $0 \le |\beta| \le |\alpha| \le q$),

$$|A_\beta| \le C\|f\|_{q,\infty}^2 \biggl(1 + \sum_{1\le|\rho|\le q} \mathbb{E}\bigl[\bigl|\partial^\rho \overline{X}_t(x)\bigr|^{2q}\bigr]\biggr) \le C\|f\|_{q,\infty}^2 \alpha_{q,2q}^{2q}(C,G).$$

If $|\rho| = 0$, then $B_\rho = 1$. Moreover, using the estimates from (4.3) and (4.1), respectively (with $p = \frac{2|\rho|}{|\beta|} \ge 2$), one gets

$$\begin{aligned}
&\sum_{1\le|\rho|\le q} B_\rho^{\frac{1}{2}} \\
&\quad\le C \sum_{1\le|\rho|\le q} \sum_{1\le|\beta|\le|\rho|} \biggl(\mathbb{E}\biggl[\int_{E^{J_t}} \bigl|\partial^\beta \ln p_{J_t}\bigl(x, z^{J_t}\bigr)\bigr|^{\frac{2|\rho|}{|\beta|}} \\
&\qquad \times p_{J_t}\bigl(x, z^{J_t}\bigr)\, d\mu(z_1)\cdots d\mu(z_{J_t})\biggr]\biggr)^{\frac{1}{2}} \\
&\quad\le C \sum_{1\le|\beta|\le|\rho|\le q} \bigl(\mathbb{E}\bigl[\bigl|\partial^\beta \ln p_{J_t}\bigl(x, \overline{Z}^{J_t}\bigr)\bigr|^{\frac{2|\rho|}{|\beta|}}\bigr]\bigr)^{\frac{1}{2}} \\
&\quad\le C \times (t \vee 1)^q \times \alpha_{q,4q}^q(C,G)\biggl(1 + \Gamma_{G,q}(\gamma) + \sum_{1\le|\beta|\le q} \bigl[\partial^\beta \ln\gamma\bigr]_{G,4q}\biggr)^q.
\end{aligned}$$

The assertion follows from these estimates. $\square$

In the proof of the previous theorem, we need $\mu(G) < \infty$ having to argue on X_t^G and $\overline{X}_t^G$. We take an increasing sequence $E_n \uparrow E$ such that $\mu(E_n) < \infty$ and we use (4.4) and (2.12), to extend the result to (possibly) infinite total measure $\mu(E)$. For simplicity, we will write $\mathcal{P}_t$ instead of $\mathcal{P}_t^E$.

THEOREM 15. *We assume that the jump rate γ is bounded* (2.9), *the jump coefficients are Lipschitz regular* (2.8), *respectively, the diffusion coefficients are smooth* (2.6). *Moreover, we assume the integrability conditions on the jump mechanism* (2.7) *and* (2.10).

Then $\mathcal{P}_t$ maps $C_b^q(\mathbb{R}^d)$ in $C_b^q(\mathbb{R}^d)$ and there exists $C > 0$ (*independent of E*) *such that*

$$\begin{aligned}
\|\mathcal{P}_t f\|_{q,\infty} &\le C\|f\|_{q,\infty}(t\vee 1)^q \times \alpha_{q,4q}^{2q}(C,E) \\
&\qquad \times \biggl(1 + \Gamma_{E,q}(\gamma) + \sum_{1\le|\beta|\le q} \bigl[\partial^\beta \ln\gamma\bigr]_{E,4q}\biggr)^q,
\end{aligned} \tag{4.5}$$

with $\alpha_{q,p}(C,E)$, $\Gamma_{E,q}(\gamma)$ and $[\partial^\beta \ln\gamma]_{E,p}$ defined in (3.5), (4.2) *and* (3.2).

5. The distance between two semigroups. In this section, we consider two sets of coefficients σ, b, c, γ and $\widehat{\sigma}$, $\widehat{b}$, $\widehat{c}$, $\widehat{\gamma}$ on measurable space $(E, \mathcal{E}, \mu)$, respectively, $(\widehat{E}, \widehat{\mathcal{E}}, \widehat{\mu})$ and we associate the stochastic equations in (2.11). The space $(\Omega, \mathcal{F}, \mathbb{P})$ is assumed to be large enough to support the (possibly mutually independent) Poisson random measures N_μ and $N_{\widehat{\mu}}$ as well as the cylindrical Brownian processes W_μ and $W_{\widehat{\mu}}$.

We denote by $X_{t_0,t}(x)$, respectively, by $\widehat{X}_{t_0,t}(x)$ the solutions of the corresponding equations and we consider the nonhomogeneous semigroups $\mathcal{P}_{t_0,t} f(x) = \mathbb{E}[f(X_{t_0,t}(x))]$ and $\widehat{\mathcal{P}}_{t_0,t} f(x) = \mathbb{E}[f(\widehat{X}_{t_0,t}(x))]$. Our aim is to estimate the distance between these two semigroups. To begin, we give the standing assumptions.

5.1. *Standing assumptions.*

ASSUMPTION $H_1(q)$. Given the coefficients σ, b, c, γ, we denote by

$$
\begin{aligned}
Q_q(T, \mathcal{P}) &:= C(T \vee 1)^q \times \alpha_{q,4q}^{2q}(C, E) \\
&\qquad \times \Big(1 + \Gamma_{E,q}(\gamma) + \sum_{1 \le |\beta| \le q} [\partial^\beta \ln \gamma]_{E,4q}\Big)^q,
\end{aligned}
\tag{5.1}
$$

with $\alpha_{q,p}(C, E)$, $\Gamma_{E,q}(\gamma)$ and $[\partial^\beta \ln \gamma]_{E,p}$ defined in (3.5), (4.2) and (3.2) and the constant C appearing in (4.4). We assume that all these quantities are well defined and finite, so that $Q_q(T, \mathcal{P}) < \infty$.

Whenever this assumption holds true, Theorem 15 yields

$$
\sup_{t_0 \le t \le T} \|\mathcal{P}_{t_0,t} f\|_{q,\infty} \le Q_q(T, \mathcal{P}) \|f\|_{q,\infty}.
\tag{5.2}
$$

We will also need a condition on the behavior of particular (polynomial) test functions.

ASSUMPTION $H_2(k)$. For $k \in \mathbb{N}$, we denote by $\psi_k(x) = (1 + |x|^2)^{\frac{k}{2}}$ and we assume that one finds a constant $C_k(T, \mathcal{P})$ such that

$$
\sup_{t_0 \le t \le T} \left\| \frac{1}{\psi_k} \mathcal{P}_{t_0,t} \psi_k \right\|_\infty \le C_k(T, \mathcal{P}) < \infty.
\tag{5.3}
$$

Finally, we will make an assumption on the gradient of the infinitesimal operators

$$
\begin{aligned}
\mathcal{L}_t f(x) &= \frac{1}{2} Tr\big[a(t, x)\partial^2 f(x)\big] + b(t, x)\partial f(x) \\
&\qquad + \int_E \big(f\big(x + c(t, z, x)\big) - f(x)\big)\gamma(t, z, x)\mu(dz),
\end{aligned}
\tag{5.4}
$$

with

$$a^{i,j}(t,x) = \int_E \sigma^i(t,z,x)\sigma^j(t,z,x)\mu(dz),$$

for all $x \in \mathbb{R}^d$, $t \in [0,T]$ and all $1 \le i, j \le d$.

ASSUMPTION $H_3(k,q)$. We assume that there exists $C \ge 1$ such that, for all $f \in C_b^q(\mathbb{R}^d)$

$$\sup_{t \le T} \left\| \frac{1}{\psi_k} \nabla \mathcal{L}_t f \right\|_\infty \le C \|f\|_{q,\infty}. \tag{5.5}$$

5.2. *Upper bounds on the distance between semigroups*. We consider two sets of coefficients σ, b, c, γ and $\widehat{\sigma}, \widehat{b}, \widehat{c}, \widehat{\gamma}$ and the corresponding semigroups $\mathcal{P}_t$ and $\widehat{\mathcal{P}}_t$. We fix $k, q \in \mathbb{N}$.

THEOREM 16. *We assume that $\mathcal{P}_t$ satisfies $H_2(k)$ and $H_3(k,q)$ and that $\widehat{\mathcal{P}}_t$ verifies $H_1(q)$ and $H_3(k,q)$. Moreover, we assume* (2.9), (2.8), (2.7) *and* (2.6) *to hold true for $\widehat{\sigma}, \widehat{b}, \widehat{c}, \widehat{\gamma}$. Finally, we assume that there exists a function $\varepsilon(\cdot) : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ such that, for every $0 \le t \le T$,*

$$\left\| \frac{1}{\psi_k} (\mathcal{L}_t - \widehat{\mathcal{L}}_t) f \right\|_\infty \le \varepsilon(t) \|f\|_{q,\infty}. \tag{5.6}$$

Then the following inequality holds true:

$$\left\| \frac{1}{\psi_k} (\mathcal{P}_{t_0,t} - \widehat{\mathcal{P}}_{t_0,t}) f \right\|_\infty \le C_k(T,\mathcal{P}) Q_q(T,\widehat{\mathcal{P}}) \|f\|_{q,\infty} \times \int_{t_0}^t \varepsilon(s)\,ds, \tag{5.7}$$

with $C_k(\mathcal{P})$ the constant in (5.3) (*with respect to σ, b, c, γ*) *and $Q_q(T,\widehat{\mathcal{P}})$ is given in* (5.1) (*with respect to the coefficients $\widehat{\sigma}, \widehat{b}, \widehat{c}, \widehat{\gamma}$*).

PROOF. For $n \in \mathbb{N}$, we set $\delta := \frac{t-t_0}{n}$ and $t_i = t_0 + i\delta$, for all $0 \le i \le n$. With these notation,

$$\frac{1}{\psi_k} (\mathcal{P}_{t_0,t} - \widehat{\mathcal{P}}_{t_0,t}) f = \sum_{i=0}^{n-1} \frac{1}{\psi_k} \mathcal{P}_{t_{i+1},t} \psi_k \frac{1}{\psi_k} (\mathcal{P}_{t_i,t_{i+1}} - \widehat{\mathcal{P}}_{t_i,t_{i+1}}) \widehat{\mathcal{P}}_{t_0,t_i} f.$$

We make the notation $g_i = \widehat{\mathcal{P}}_{t_0,t_i} f$. Using (5.3) for $\mathcal{P}_{t_{i+1},t}$,

$$\left\| \frac{1}{\psi_k} (\mathcal{P}_{t_0,t} - \widehat{\mathcal{P}}_{t_0,t}) f \right\|_\infty \le C_k(T,\mathcal{P}) \sum_{i=0}^{n-1} \left\| \frac{1}{\psi_k} (\mathcal{P}_{t_i,t_{i+1}} - \widehat{\mathcal{P}}_{t_i,t_{i+1}}) g_i \right\|_\infty.$$

By Itô's formula,

$$\begin{aligned} &\mathcal{P}_{t_i,t_{i+1}} g_i(x) \\ &\quad = g_i(x) + \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \mathcal{L}_s g_i(X_{t_i,s}(x))\,ds \right] = g_i(x) + \int_{t_i}^{t_{i+1}} \mathcal{L}_s g_i(x)\,ds + \varepsilon_i, \end{aligned}$$

with $\varepsilon_i(x) := \mathbb{E}[\int_{t_i}^{t_{i+1}} (\mathcal{L}_s g_i(X_{t_i,s}(x)) - \mathcal{L}_s g_i(x))\,ds]$. We write the same type of formulae for $\widehat{\mathcal{P}}_{t_i,t_{i+1}} g_i$, take the difference between the two and use (5.6) in order to get

$$\begin{aligned}
&\left\| \frac{1}{\psi_k} (\mathcal{P}_{t_i,t_{i+1}} - \widehat{\mathcal{P}}_{t_i,t_{i+1}}) g_i \right\|_\infty \\
&\quad \le \int_{t_i}^{t_{i+1}} \left\| \frac{1}{\psi_k} (\mathcal{L}_s - \widehat{\mathcal{L}}_s) g_i \right\|_\infty ds + \left\| \frac{1}{\psi_k} \varepsilon_i \right\|_\infty + \left\| \frac{1}{\psi_k} \widehat{\varepsilon}_i \right\|_\infty \\
&\quad \le \|g_i\|_{q,\infty} \int_{t_i}^{t_{i+1}} \varepsilon(s)\,ds + \left\| \frac{1}{\psi_k} \varepsilon_i \right\|_\infty + \left\| \frac{1}{\psi_k} \widehat{\varepsilon}_i \right\|_\infty.
\end{aligned}$$

By (5.2), $\|g_i\|_{q,\infty} \le Q_q(T, \widehat{\mathcal{P}}) \|f\|_{q,\infty}$ so that, finally,

$$\begin{aligned}
&\left\| \frac{1}{\psi_k} (\mathcal{P}_{t_0,t} - \widehat{\mathcal{P}}_{t_0,t}) f \right\|_\infty \\
&\quad \le C_k(T, \mathcal{P}) \left[Q_q(T, \widehat{\mathcal{P}}) \|f\|_{q,\infty} \int_{t_0}^{t} \varepsilon(s)\,ds + \sum_{i=0}^{n-1} \left(\left\| \frac{1}{\psi_k} \varepsilon_i \right\|_\infty + \left\| \frac{1}{\psi_k} \widehat{\varepsilon}_i \right\|_\infty \right) \right].
\end{aligned}$$

To conclude, one still needs to estimate the terms ε_i and prove that these errors vanish as n increases. The assumption (5.5) yields

$$\begin{aligned}
&\left| \mathcal{L}_s g_i(X_{t_i,s}(x)) - \mathcal{L}_s g_i(x) \right| \\
&\quad \le \int_0^1 \left| \langle \nabla \mathcal{L}_s g_i(\lambda x + (1-\lambda) X_{t_i,s}(x)), X_{t_i,s}(x) - x \rangle \right| d\lambda \\
&\quad \le C \|g_i\|_{q,\infty} \left| X_{t_i,s}(x) - x \right| \int_0^1 \psi_k(\lambda x + (1-\lambda) X_{t_i,s}(x))\,d\lambda.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\left| \varepsilon_i(x) \right| \le C Q_q(T, \widehat{\mathcal{P}}) \|f\|_{q,\infty} \\
&\qquad \times \int_{t_i}^{t_{i+1}} \int_0^1 \mathbb{E}\left[\psi_k(\lambda x + (1-\lambda) X_{t_i,s}(x)) \left| X_{t_i,s}(x) - x \right| \right] d\lambda\,ds.
\end{aligned}$$

Using the standard trajectory estimates, $\mathbb{E}[|X_{t_i,s}(x)|^k] \le C(1+|x|^k)$. Hence,

$$\mathbb{E}\left[\psi_k^2(\lambda x + (1-\lambda) X_{t_i,s}(x)) \right] \le C \psi_k^2(x).$$

Using the Cauchy–Schwarz inequality, we get

$$\frac{1}{\psi_k(x)} \left| \varepsilon_i(x) \right| \le C Q_q(T, \widehat{\mathcal{P}}) \|f\|_{q,\infty} \int_{t_i}^{t_{i+1}} \left(\mathbb{E}\left[\left| X_{t_i,s}(x) - x \right|^2 \right] \right)^{\frac{1}{2}} ds.$$

By setting $\tau_n(s) := t_i$ for $t_i \le s < t_{i+1}$, we finally get

$$\frac{1}{\psi_k(x)} \sum_{i=1}^{n} \left| \varepsilon_i(x) \right| \le C Q_q(T, \widehat{\mathcal{P}}) \|f\|_{q,\infty} \int_0^t \left(\mathbb{E}\left[\left| X_{\tau_n(s),s}(x) - x \right|^2 \right] \right)^{\frac{1}{2}} ds$$

and the right-hand term vanishes as $n \to \infty$. Similar estimates are valid for $\widehat{\varepsilon}$ which completes our proof. □

REMARK 17. 1. This assertion is to be interpreted in connection to Trotter–Kato-type results (e.g., [35], Theorem 4.4) stating that, given $\mathcal{P}_t$ and $(\mathcal{P}_t^n)_{n\in\mathbb{N}}$ homogeneous Feller semigroups of infinitesimal operators $\mathcal{L}$, respectively, $\mathcal{L}_n$, if $\mathcal{L}_n$ converges to $\mathcal{L}$, then $\mathcal{P}_t^n \to \mathcal{P}_t$ (in an appropriate sense). The inequality (5.7) gives not only a qualitative behavior, but a quantitative one by providing estimate of the error within our framework.

2. The main difficulty and novelty in our approach is to provide (5.2). Whenever γ is constant, one deals with a usual SDE with jumps and the proof of (5.2) follows from the regularity of the flow $x \to X_t(x)$. However, since $\gamma(t, z, x)$ depends on x (which is the case for *PDMP*), the effort developed in the previous sections is necessary. Note, however, that (5.2) is needed only on one of $\mathcal{P}_t$ and $\widehat{\mathcal{P}}_t$. Hence, in a framework in which either γ or $\widehat{\gamma}$ does not depend on x, the proofs simplify considerably.

6. PDMP with three regimes. In this section, we discuss piecewise diffusive Markov processes in which three regimes are at work depending on the speed of the jumps. The intermediate regime will be purely deterministic and replaced by a drift term (corresponding to an application of the law of large numbers). The fast regime will provide a diffusive term (associated with an application of the central limit theorem). Finally, the slow regime is kept as jump-type contribution. We do not aim at treating a completely general framework but only at presenting an example in order to illustrate our approach.

6.1. *Theoretical framework.*

6.1.1. *The model.* Let us begin with fixing $\varepsilon > 0$ and a measurable space $(E, \mathcal{E}, \mu_\varepsilon)$ where μ_ε is a nonnegative finite measure. The space decomposes as follows $E = A_\varepsilon \cup B_\varepsilon \cup C_\varepsilon$, where $A_\varepsilon, B_\varepsilon, C_\varepsilon$ are mutually disjoint Borel measurable sets. Moreover, given $\Gamma_\varepsilon > 0$, to some (smooth, time-homogeneous) coefficients $c_\varepsilon, \gamma_\varepsilon : E \times \mathbb{R} \to \mathbb{R}$, $b_\varepsilon : \mathbb{R} \to \mathbb{R}$, we associate the stochastic equation

$$\begin{aligned} X_t^\varepsilon &= x + \int_0^t b_\varepsilon(X_s^\varepsilon)\,ds \\ &\quad + \int_0^t \int_{A_\varepsilon \times [0,2\Gamma_\varepsilon]} c_\varepsilon(z, X_{s-}^\varepsilon) 1_{\{u \le \gamma_\varepsilon(z, X_{s-}^\varepsilon)\}} \widetilde{N}_{\mu_\varepsilon}(ds, dz, du) \\ &\quad + \int_0^t \int_{(B_\varepsilon \cup C_\varepsilon) \times [0,2\Gamma_\varepsilon]} c_\varepsilon(z, X_{s-}^\varepsilon) 1_{\{u \le \gamma_\varepsilon(z, X_{s-}^\varepsilon)\}} N_{\mu_\varepsilon}(ds, dz, du), \end{aligned} \tag{6.1}$$

where N_{μ_ε} is a Poisson point measure on $E \times R_+$ associated with Γ_ε and μ_ε and $\widetilde{N}_{\mu_\varepsilon} = N_{\mu_\varepsilon} - \widehat{N}_{\mu_\varepsilon}$ is the associated martingale measure.

6.1.2. *The regimes*. The jumps in A_ε are assumed to occur at high frequency. They lead to a Brownian motion. The jumps in B_ε represent an intermediary regime which will be modeled by a drift term while the jumps in C_ε are rather rare and remain in the same regime. This model is expressed by the following setting. We consider a finite measure μ and the coefficients $\sigma, b : \mathbb{R} \to \mathbb{R}$ and $c, \gamma : E \times \mathbb{R} \to \mathbb{R}$. We associate the equation

$$
\begin{aligned}
X_t = x &+ \int_0^t b(X_s)\,ds + \int_0^t \sigma(X_s)\,dW_s \\
&+ \int_0^t \int_{E\times[0,2\Gamma]} c(z, X_{s-}) 1_{\{u \le \gamma(z, X_{s-})\}} N_\mu(ds, dz, du), \\
&0 \le t \le T.
\end{aligned}
\tag{6.2}
$$

Our aim is to give sufficient conditions in order to obtain the convergence of the family X^ε to X and to estimate the error.

6.1.3. *Standing* (*sufficient*) *assumptions*. Throughout the section, unless stated otherwise, we assume the following.

ASSUMPTION H_0^ε. We assume that c_ε and γ_ε satisfy integrability condition (2.7), the Lipschitz regularity assumption (2.8) and the uniform upper-bound of γ_ε assumption (2.9) (written for Γ_ε substituting Γ).

REMARK 18. Note that the constants which appear in these conditions depend on ε (so they are not uniform with respect to ε). Under these hypothesis, equation (6.1) has a unique solution (which may alternatively be constructed using a compound Poisson process).

We also need an assumption on the limit coefficients.

ASSUMPTION H_0. We assume that $\sigma, b \in C_b^3(\mathbb{R})$ and, for every $z \in E$, the functions $x \mapsto c(z,x)$ and $x \mapsto \ln \gamma(z,x)$ are three times differentiable and

$$
\begin{aligned}
\sum_{0\le|\alpha|\le 3} \sup_{x\in\mathbb{R}} \Big[\big|\partial^\alpha \sigma(x)\big| + \big|\partial^\alpha b(x)\big| + \sup_{z\in E}\big|\partial^\alpha c(z,x)\big| + \sup_{z\in E}\big|\partial^\alpha \ln\gamma(z,x)\big| \Big] \\
=: C_* < \infty.
\end{aligned}
\tag{6.3}
$$

Under this hypothesis, equation (6.2) has a unique solution (see Remark 7). Finally, we need some further assumptions in order to obtain convergence. We denote by $\nu_\varepsilon(x,dz) := \gamma_\varepsilon(z,x)\mu_\varepsilon(dz)$ and set

$$\sigma_\varepsilon(x) := \left(\int_{A_\varepsilon} c_\varepsilon^2(z,x)\nu_\varepsilon(x,dz)\right)^{\frac{1}{2}}, \qquad b^\varepsilon(x) := b_\varepsilon(x) + \int_{B_\varepsilon} c_\varepsilon(z,x)\nu_\varepsilon(x,dz),$$

$$\delta_\sigma(\varepsilon) := \|\sigma_\varepsilon^2 - \sigma^2\|_\infty, \qquad \delta_b(\varepsilon) = \|b_\varepsilon - b\|_\infty,$$

$$\delta_{c,\gamma}(\varepsilon) = \sup_{x\in\mathbb{R}} \int_{C_\varepsilon} \left|(c - c_\varepsilon)(z,x)\right|\gamma(z,x) + \left|(\gamma - \gamma_\varepsilon)(z,x)\right| d\mu(z).$$

Moreover, we denote the convenient moments by

$$\delta_A(\varepsilon) = \sup_{x\in\mathbb{R}} \int_{A_\varepsilon} \left|c_\varepsilon(z,x)\right|^3 \nu_\varepsilon(x,dz), \qquad \delta_B(\varepsilon) = \sup_{x\in\mathbb{R}} \int_{B_\varepsilon} \left|c_\varepsilon(z,x)\right|^2 \nu_\varepsilon(x,dz),$$

$$\delta_C(\varepsilon) = \sup_{x\in\mathbb{R}} \int_{E-C_\varepsilon} \left|c(z,x)\right|\gamma(z,x)\mu(dz)$$

ASSUMPTION H_1. We assume that $\delta(\varepsilon) := \delta_\sigma(\varepsilon) + \delta_b(\varepsilon) + \delta_{c,\gamma}(\varepsilon) + \delta_A(\varepsilon) + \delta_B(\varepsilon) + \delta_C(\varepsilon) \to_{\varepsilon\to 0} 0$.

ASSUMPTION H_2. Finally, we assume that the restrictions of μ_ε and μ to C_ε coincide, that is, $1_{C_\varepsilon}(z)\mu_\varepsilon(dz) = 1_{C_\varepsilon}(z)\mu(dz)$.

6.1.4. *The theoretical result.* Under these assumptions, one can state and prove the following.

THEOREM 19. *We assume that H_0^ε, H_0, H_1 and H_2 hold true. We let $\mathcal{P}_t^\varepsilon$ and $\mathcal{P}_t$ be the semigroups associated with X_t^ε, respectively, with X_t. Then there exists a universal constant C such that, for every $f \in C_b^3(R)$,*

$$\|\mathcal{P}_t^\varepsilon f - \mathcal{P}_t f\|_\infty \le (t\vee 1)^3 CC_*^{42} \exp\left((t\vee 1)CC_*^{36}\right) \times \delta(\varepsilon)\|f\|_{3,\infty}. \tag{6.4}$$

The proof follows from Theorem 16 and, for our readers' convenience a sketch is presented in Section 8.3.

REMARK 20. The notation used in the previous theorem suggests that $\mathcal{P}_t^\varepsilon$ is an approximation of $\mathcal{P}_t$. However, sometimes, the point of view is the exact opposite: the physical phenomenon is modeled by X_t^ε and X_t represents an approximation which is easier to handle. Having this in mind one may also consider the following optimization problem: given the dynamics of X_t^ε, which are the best dynamics (coefficients) of type X_t which approximates X_t^ε? In order to formulate this problem in a clean way, one has to give a criterion in order to precise the sense of "best". This would be another problem left for future work.

6.2. *A simple example.* Let us now give an explicit example.

EXAMPLE 21. To this purpose, we consider $c, \gamma \in C_b^3(\mathbb{R})$ and

$$\mu_\varepsilon(dz) = 1_{(\varepsilon,3\varepsilon]}(z)\frac{dz}{z^2} + 1_{(3\varepsilon,4\varepsilon]}(z)\frac{dz}{z^{3/2}} + 1_{(4\varepsilon,1]}(z)\frac{dz}{z},$$

$$c_\varepsilon(z,x) = c(x)\sqrt{z}\big(1_{(2\varepsilon,1]}(z) - \alpha 1_{(\varepsilon,2\varepsilon]}(z)\big) \qquad \text{with } \alpha = \frac{\sqrt{3}-\sqrt{2}}{\sqrt{6}-\sqrt{3}},$$

and we associate the equation

$$X_t^\varepsilon = x + \int_0^t\int_0^1\int_0^1 c_\varepsilon\big(z, X_{s-}^\varepsilon\big)1_{\{u\le\gamma(X_{s-}^\varepsilon)\}}N_{\mu_\varepsilon}(ds,dz,du).$$

Note that, in contrast with equation (6.1), the measure N_{μ_ε} is not compensated. But, in fact, the activity of the small jumps in $1_{(\varepsilon,2\varepsilon]}(z)$ compensate the activity of the small jumps in $1_{(2\varepsilon,3\varepsilon]}(z)$. The limit equation is

$$\begin{aligned} X_t = x &+ \int_0^t \sigma(X_s)\,dW_s + \int_0^t b(X_s)\,ds \\ &+ \int_0^t\int_0^1\int_0^1 c(X_{s-})\sqrt{z}1_{\{u\le\gamma(X_{s-}^\varepsilon)\}}N_\mu(ds,dz,du) \end{aligned}$$

with $\mu(dz) = z^{-1}\,dz$ and $\sigma(x) = \beta_1 c(x)\sqrt{\gamma}(x)$, $b(x) = \int_{B_\varepsilon} c_\varepsilon(z,x)\nu_\varepsilon(x,dz) = \beta_2 c(x)\gamma(x)$. Here, $\beta_1 = ((\alpha^2-1)\ln 2 + \ln 3)^{\frac{1}{2}}$ and $\beta_2 = \ln\frac{4}{3}$. Then, by applying Theorem 19, it follows that

$$\|\mathcal{P}_t f - \mathcal{P}_t^\varepsilon f\|_\infty \le C\sqrt{\varepsilon}\|f\|_{3,\infty}$$

with C depending on $\|c\|_{3,\infty}$ and $\|\ln\gamma\|_{3,\infty}$. It also depends on the time interval $[0,T]$. To this purpose, one only needs to check the assumption H_1 (see Section 8.3 for details on this step) and apply Theorem 19.

7. Boltzmann's equation.

7.1. *The model. Probabilistic interpretation.* In this section, we use the previous results to construct an approximation scheme for the solution of the two-dimensional Boltzmann equation taking the following form:

$$\partial_t f_t(v) = \int_{\mathbb{R}^2} dv_* \int_{-\pi/2}^{\pi/2} d\theta |v-v_*|^\kappa \theta^{-(1+\nu)}\big(f_t(v')f_t(v'_*) - f_t(v)f_t(v_*)\big). \tag{7.1}$$

Here,

- $f_t(v)$ is a nonnegative measure on $\mathbb{R}^2$ representing the density of particles with velocity v in a model for a gas in dimension two.

- R_θ is the rotation of angle θ and the new speeds after collision are $v' = \frac{v+v_*}{2} + R_\theta(\frac{v-v_*}{2})$, respectively, $v'_* = \frac{v+v_*}{2} - R_\theta(\frac{v-v_*}{2})$.
- the parameters $\nu \in (0,1)$ and $\kappa \in (0,1]$ are chosen for the cross section to model the interaction in the spirit of the assumption $\mathbf{A}$ (γ, ν) in [6].

The rigorous sense of this equation is given by integrating it against a test function [hence leading to weak solutions of (7.1)]. In [20], Corollary 2.3 and Lemma 4.1, the authors have proven that, for every $\nu \in (0,1)$ and $\kappa \in (0,1]$, the above equation admits a unique weak solution as follows. One assumes that there exists $s \in (\kappa, 2)$ such that $\int e^{|v|^s} f_0(dv) < \infty$. Then there exists a unique solution f_t of (7.1) which starts from f_0. Moreover, the solution satisfies $\sup_{t \le T} \int e^{|v|^{s'}} f_t(dv) < \infty$ for every $s' < s$.

Using the Skorohod representation theorem, we find a measurable function $v_t : [0,1] \to \mathbb{R}^2$ such that for every $\psi : \mathbb{R}^2 \to \mathbb{R}_+$:

$$\int_0^1 \psi\big(v_t(\rho)\big)\, d\rho = \int_{R^2} \psi(v) f_t(dv). \tag{7.2}$$

Throughout the section, unless stated otherwise, we fix ν, κ and $s \in (\kappa, 2)$ and the corresponding solution $f_t(v)$ [and, in particular, $v_t(\rho)$].

In [37], the author gives a probabilistic interpretation for the solutions of the classical Boltzmann equation (in dimension 3). A variant of this result in dimension two [so for the equation (7.1)], as well as an approximation result for it, is given in [6], Section 2. We briefly recall these elements.

We emphasize that throughout the section, the time horizon $T \ge 0$ is fixed and the constants C depend on the time interval.

We let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, the space $E := [-\frac{\pi}{2}, \frac{\pi}{2}] \times [0,1]$ and let $N(dt, d\theta, d\rho, du)$ be a Poisson point measure on $E \times \mathbb{R}_+$ with intensity measure $\theta^{-(1+\nu)}\, d\theta \times d\rho \times du$. We also consider the matrix

$$A(\theta) := \frac{1}{2}\begin{pmatrix} \cos\theta - 1 & -\sin\theta \\ \sin\theta & \cos\theta - 1 \end{pmatrix} = \frac{1}{2}(R_\theta - I).$$

Then we are interested in the equation

$$V_t = V_0 + \int_0^t \int_{E \times \mathbb{R}_+} A(\theta)\big(V_{s-} - v_s(\rho)\big) 1_{\{u \le |V_{s-} - v_s(\rho)|^\kappa\}} N(ds, d\theta, d\rho, du) \tag{7.3}$$

with $\mathbb{P}(V_0 \in dv) = f_0(dv)$.

In the spirit of [6], Section 2, one also constructs the following approximation. One considers a C^∞ even nonnegative function χ supported by $[-1,1]$ and such that $\int_R \chi(x)\, dx = 1$. We fix $\eta_0 \in (\frac{1}{s}, \frac{1}{\kappa \vee \nu})$. Given $\varepsilon \in (0,1]$, we denote by $\Gamma_\varepsilon = (\ln \frac{1}{\varepsilon})^{\eta_0}$ and define

$$\varphi_\varepsilon(x) = \int_R \big((y \vee 2\varepsilon) \wedge \Gamma_\varepsilon\big) \frac{\chi(\frac{x-y}{\varepsilon})}{\varepsilon}\, dy. \tag{7.4}$$

The reader is invited to note that $2\varepsilon \leq \varphi_\varepsilon(x) \leq \Gamma_\varepsilon$, for every $x \in \mathbb{R}$, $\varphi_\varepsilon(x) = x$, for $x \in (3\varepsilon, \Gamma_\varepsilon - 1)$, $\varphi_\varepsilon(x) = 2\varepsilon$ for $x \in (0, \varepsilon)$ and $\varphi_\varepsilon(x) = \Gamma_\varepsilon$ for $x \in (\Gamma_\varepsilon, \infty)$.

To the cut off function φ_ε, one associates the equation

$$V_t^\varepsilon = V_0 + \int_0^t \int_{E\times R_+} A(\theta)\big(V_{s-}^\varepsilon - v_s(\rho)\big) \times 1_{\{u \leq \varphi_\varepsilon^\kappa(|V_{s-}^\varepsilon - v_s(\rho)|)\}} N(ds, d\theta, d\rho, du). \tag{7.5}$$

Proposition 2.1 in [6] provides the following probabilistic interpretation as well as an approximation result.

PROPOSITION 22 ([6], Proposition 2.1). 1. *The equation* (7.3) *has a unique càdlàg adapted solution* $(V_t)_{t\geq 0}$ *and its law* $\mathbb{P}(V_t \in dv) = f_t(dv)$.[2]

2. *The equation* (7.5) *has a unique càdlàg solution* V^ε *and*

$$\sup_{t\leq T} \mathbb{E}\big[|V_t - V_t^\varepsilon|\big] \leq Ce^{C\Gamma_\varepsilon^\kappa} \times \varepsilon^{1+\kappa}. \tag{7.6}$$

Moreover, there exists $\varepsilon_0 > 0$ *such that, for every* $0 < s' < s$,

$$\sup_{\varepsilon\leq\varepsilon_0} \mathbb{E}\Big[\sup_{t\leq T}\big(e^{|V_t|^{s'}} + e^{|V_t^\varepsilon|^{s'}}\big)\Big] < \infty. \tag{7.7}$$

In the following, we assume that

$$\int e^{|v|^s} f_0(dv) < \infty \qquad \forall s < 2. \tag{7.8}$$

In particular, this gives the restriction $\frac{1}{2} < \eta_0 < \frac{1}{\kappa\vee\nu}$. Then, if $p \geq 1$ is such that $\kappa p < 2$, we may choose η_0 such that $\eta_0\kappa p < 1$. This guarantees that for every $a > 0$ there exists ε_a such that the quantity $a(\varepsilon) = (\ln\frac{1}{\varepsilon})^{-1+\eta_0\kappa p} \leq a$, for every $0 < \varepsilon < \varepsilon_a$. By the definition of Γ_ε and the previous inequality, it follows that

$$e^{\Gamma_\varepsilon^{\kappa p}} = \varepsilon^{-a(\varepsilon)} \leq \varepsilon^{-a} \qquad \forall 0 < \varepsilon < \varepsilon_a. \tag{7.9}$$

7.2. *First-order approximation*. The aim of this section is to construct an approximation of the solution V_t of equation (7.3) in which the small jumps, corresponding to $|\theta| \leq \delta$, are replaced by a drift term.

First, let us fix $\delta > 0$, $r = \frac{2-3\nu}{3+\kappa}(\leq \frac{1-\nu}{1-\kappa})$, set $\varepsilon = \delta^r$ and consider the solution V_t^ε of the truncated equation (7.5) associated with this ε. The inequality (7.6) provides a control of the distance between V_t and V_t^ε. Second, for this solution V_t^ε of (7.5) we apply Lemma 16 in order to replace the small jumps by a convenient drift term.

To fall in the framework given in the first part of our paper, we will denote by $V_{t_0,t}^\varepsilon(v)$ the solution of the equation (7.5) which starts from $v \in \mathbb{R}^2$ at time

[2] In this sense, V_t provides a probabilistic representation for f_t.

$t_0 \in [0, T]$ and we set $\mathcal{P}^{\varepsilon}_{t_0,t} f(v) = \mathbb{E}[f(V^{\varepsilon}_{t_0,t}(v))]$. We also denote (for $\varepsilon > 0$ fixed above),

$$\mu(d\theta, d\rho) = \theta^{-(1+\nu)}\, d\theta \times d\rho, \qquad c(t, \theta, \rho, v) = A(\theta)\bigl(v - v_t(\rho)\bigr),$$
$$\gamma_\varepsilon(t, \rho, v) = \varphi^{\kappa}_{\varepsilon}\bigl(|v - v_t(\rho)|\bigr).$$

The infinitesimal operator of $\mathcal{P}^{\varepsilon}_{t_0,t}$ is simply given by

$$\mathcal{L}^{\varepsilon}_t f(v) = \int_E \mu(d\theta, d\rho)\gamma(t, \rho, v)\bigl(f\bigl(v + c(t, \theta, \rho, v)\bigr) - f(v)\bigr).$$

We will replace the activity of small jumps (such that θ is close to 0) with a drift term. To this purpose, we denote by $E_\delta = \{(\theta, \rho) : |\theta| > \delta\}$ and we define

$$b_\delta(t, v) = \int_{\{|\theta| \le \delta\}} \gamma(t, \rho, v)c(t, \theta, \rho, v)\mu(d\theta, d\rho) \quad \text{and}$$

$$\widehat{\mathcal{L}}^{\delta}_t f(v) = b_\delta(t, v)\partial f(v) + \int_{E_\delta} \mu(d\theta, d\rho)\gamma(t, \rho, v)\bigl(f\bigl(v + c(t, \theta, \rho, v)\bigr) - f(v)\bigr). \tag{7.10}$$

The approximating equation is

$$\begin{aligned} U^{\delta}_{t_0,t}(v) = v &+ \int_{t_0}^{t} b_\delta\bigl(s, U^{\delta}_{t_0,s}(v)\bigr)\, ds \\ &+ \int_{t_0}^{t} \int_{E_\delta \times R_+} c\bigl(s, \theta, \rho, U^{\delta}_{t_0,s-}(v)\bigr) \\ &\times 1_{\{u \le \gamma(s, \rho, U^{\delta}_{t_0,s-}(v))\}} N(ds, d\theta, d\rho, du). \end{aligned}$$

We denote by $\widehat{\mathcal{P}}^{\delta}_{t_0,t}$ the semigroup associated with $\widehat{\mathcal{L}}^{\delta}_t$, that is, $\widehat{\mathcal{P}}^{\delta}_{t_0,t} f(v) := \mathbb{E}[f(U^{\delta}_{t_0,s}(v))]$.

THEOREM 23. *Suppose that $\kappa < \frac{1}{8}$ and $\nu < \frac{1}{2}$. For every $\eta < \frac{(2-3\nu)(1+\kappa)}{3+\kappa}$ there exists $C_\eta \ge 1$ and $\delta_\eta > 0$ such that for $0 < \delta \le \delta_\eta$ we have*

$$\bigl|\mathbb{E}\bigl[f(V_t)\bigr] - \mathbb{E}\bigl[f\bigl(U^{\delta}(V_0)\bigr)\bigr]\bigr| \le C_\eta \|f\|_{2,\infty} \times \delta^{\eta}. \tag{7.11}$$

The proof essentially consists in the use of Theorem 16 combined with (7.6). For our readers' sake, the complete proof is given in Section 8.4.

7.3. *Second-order approximation.* We define

$$\sigma(t, \theta, \rho, v) = c(t, \theta, \rho, v)\sqrt{\gamma_\varepsilon(t, \rho, v)},$$
$$a^{i,j}_{\delta}(t, v) = \int_{\{|\theta| \le \delta\}} \mu(d\theta, d\rho)\sigma^i\sigma^j(t, \theta, \rho, v),$$

$$\widehat{\mathcal{L}}_t^\delta f(v) = \langle b_\delta(t,v), \nabla f(v)\rangle + \frac{1}{2}\sum_{i,j=1}^d a_\delta^{i,j}(t,v)\partial_{ji}^2 f(v) + \int_{E_\delta} \mu(d\theta, d\rho)\gamma_\varepsilon(t,\rho,v)\big(f\big(v + c(t,\theta,\rho,v)\big) - f(v)\big),$$

where b_δ is given by (7.10). This is the infinitesimal operator corresponding to the semigroup $\widehat{\mathcal{P}}_{t_0,t}^\delta f(v) = \mathbb{E}[f(U_{t_0,t}^\delta(v))]$ with $U_{t_0,t}^\delta(v)$ solution to

$$\begin{aligned} U_{t_0,t}^\delta(v) = v &+ \int_{t_0}^t b_\delta\big(s, U_{t_0,s}^\delta(v)\big)\,ds \\ &+ \int_0^t \int_{E_\delta} \sigma_\delta\big(s,\theta,\rho,U_{t_0,s-}^\delta(v)\big) W_\mu(ds,d\theta,d\rho) \\ &+ \int_0^t \int_{E_\delta \times R_+} c\big(s,\theta,\rho,U_{t_0,s-}^\delta(v)\big) \\ &\times 1_{\{u \le \gamma(s,\rho,U_{t_0,s-}^\delta(v))\}} N(ds,d\theta,d\rho,du). \end{aligned}$$

The approach is quite similar to the first order. The main result is the following.

THEOREM 24. *Let us assume that* $\kappa \le \frac{1}{18}$ *and let*

$$r < \frac{1-\nu}{2-\kappa} \wedge \frac{1-\frac{\nu}{2}}{2-\frac{\kappa}{2}} \wedge \frac{3-4\nu}{4+\kappa}. \tag{7.12}$$

Then

$$\left\| \frac{1}{\psi_3}\big(\mathcal{P}_{t_0,t} f - \widehat{\mathcal{P}}_{t_0,t}^\delta\big) f \right\|_\infty \le C\delta^{r(1+\kappa)} \times \|f\|_{3,\infty}. \tag{7.13}$$

REMARK 25. It turns out that the second-order error is larger then the first-order error. This is somewhat counterintuitive. This is due to the fact that we mix two different errors: $\mathcal{P}_{t_0,t} f - \mathcal{P}_{t_0,t}^\varepsilon \sim \varepsilon^{1+\kappa}$ and $\mathcal{P}_{t_0,t}^\varepsilon f - \widehat{\mathcal{P}}_{t_0,t}^\delta \sim \delta^{3-\nu}\varepsilon^{-3}$. If ε is fixed then the second-order error is $\delta^{3-\nu}$ and the first-order error is $\delta^{2-\nu}$ and this seems coherent. But if we mix the two errors things become less obvious.

8. Proof of the results.

8.1. *Proof of the results in Section* 2.2. We begin with the estimates given in Lemma 6.

PROOF OF LEMMA 6. We follow the ideas in [24] so we just sketch the proof. Let us fix the initial time $s < T$. For every $t \in [s, T]$, one has

$$|\Delta X_{s,t}| \leq |\Delta X_{s,s}| + \left|\int_s^t \int_E h(r,z)\,dW_\mu(dr,dz)\right| + \left|\int_s^t g(r)\,dr\right| \\ + \left|\int_s^t \int_{E\times[0,2\Gamma]} H(r-,z,u)N_\mu(dr,dz,du)\right|,$$

where

$$h(r,z) = \sigma\big(r,z,X_{s,r}^{G_1}\big) - \sigma\big(r,z,X_{s,r}^{G_2}\big), \qquad g(r) = b\big(r,X_{s,r}^{G_1}\big) - b\big(r,X_{s,r}^{G_2}\big)$$

and

$$H(r,z,u) = 1_{G_1}(z)c\big(r,z,X_{s,r-}^{G_1}\big)1_{\{u\leq\gamma(r,z,X_{s,r-}^{G_1})\}} \\ - 1_{G_2}(z)c\big(r,z,X_{s,r-}^{G_2}\big)1_{\{u\leq\gamma(r,z,X_{s,r-}^{G_2})\}},$$

for all $(r,z,u) \in [s,t] \times E \times \mathbb{R}_+$. Using the inequality,

$$\big|h(r,z)\big| \leq |\Delta X_{s,r}| \times \int_0^1 \big|\nabla\sigma\big(r,z,\lambda X_{s,r-}^{G_1} + (1-\lambda)X_{s,r-}^{G_2}\big)\big|\,d\lambda,$$

we obtain

$$\left[\int_E \big|h(r,z)\big|^2 \mu(dz)\right]^{\frac{1}{2}} \leq \|\nabla\sigma\|_{(\mu,\infty)}|\Delta X_{s,r}|.$$

Burkholder's inequality yields

$$\begin{aligned} &\mathbb{E}\left[\sup_{s\leq t'\leq t}\left|\int_s^{t'}\int_E h(r,z)\,dW_\mu(dr,dz)\right|\right] \\ &\quad\leq C\mathbb{E}\left[\left(\int_s^t\int_E \big|h(r,z)\big|^2\mu(dz)\,dr\right)^{\frac{1}{2}}\right] \\ &\quad\leq C\|\nabla\sigma\|_{(\mu,\infty)}\mathbb{E}\left[\left(\int_s^t |\Delta X_{s,r}|^2\,dr\right)^{\frac{1}{2}}\right] \\ &\quad\leq C\|\nabla\sigma\|_{(\mu,\infty)}(t-s)^{\frac{1}{2}}\mathbb{E}\Big[\sup_{s\leq r\leq t}|\Delta X_{s,r}|\Big]. \end{aligned}$$

And the same inequality holds for g. Finally, since N_μ is a positive measure, one has

$$\begin{aligned} &\mathbb{E}\left[\sup_{s\leq t'\leq t}\left|\int_s^{t'}\int_{E\times[0,2\Gamma]} H(r-,z,u)N_\mu(dr\,dz\,du)\right|\right] \\ &\quad\leq \mathbb{E}\left[\int_s^T dr \int_{E\times[0,2\Gamma]}\big|H(r,z,u)\big|\mu(dz)\,du\right]. \end{aligned}$$

A careful analysis of the term $|H(r,z,u)|$ shows that the above term is upper bounded by $(t-s)\alpha(G_2\setminus G_1)+C(c,\gamma)\int_s^t \mathbb{E}[|\Delta X_{s,r}|]\,dr$. Going back to the initial inequality in our proof, one gets

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{s\le r\le t}|\Delta X_{s,r}|\Big]\\
&\qquad\le |\Delta X_{s,s}|+(t-s)\alpha(G_2\setminus G_1)\\
&\qquad\quad + C\big(\|\nabla\sigma\|_{(\mu,\infty)}+\|\nabla b\|_\infty + C(\gamma,c)\big)(t-s)^{\frac{1}{2}}\mathbb{E}\Big[\sup_{s\le r\le t}|\Delta X_{s,r}|\Big].
\end{aligned}$$

Hence, whenever $t-s\le\delta:=(2C(\|\nabla\sigma\|_{(\mu,\infty)}+\|\nabla b\|_\infty+C(\gamma,c))))^{-2}$, one gets

$$\mathbb{E}\Big[\sup_{s\le r\le t}|\Delta X_{s,r}|\Big]\le 2\big(|\Delta X_{s,s}|+(t-s)\alpha(G_2\setminus G_1)\big).$$

The argument follows by partitioning $[s,T]$ in

$$n\le 4T\big(C\big(\|\nabla\sigma\|_{(\mu,\infty)}+\|\nabla b\|_\infty+C(\gamma,c)\big)\big)^2+1$$

subintervals of length δ and iterating. □

8.2. *Proof of the results in Section* 3. The proof of Lemma 10 makes extensive use of moment estimates of some kind of linear-type stochastic system. To this purpose, we begin with briefly explaining the type of system and the estimates we have in mind.

8.2.1. *Preliminary arguments for Lemma* 10: *Moment estimates for linear SDE*. In this section, we consider the d-dimensional linear equation

$$\begin{aligned}
V_t = V_0 &+ \int_0^t\int_E \big(h(s)+\langle\nabla b(s,X_s),V_s\rangle\big)\,ds\\
&+\int_0^t\int_E\big(H(s,z)+\langle\nabla\sigma(s,z,X_s),V_s\rangle\big)W_\mu(ds,dz)\\
&+\int_0^t\int_{G\times(0,2\Gamma)}\big(Q(s-,z)+\langle\nabla_x c(s,z,X_{s-}),V_{s-}\rangle\big)\\
&\times 1_{\{u\le\gamma(s,z,X_{s-})\}}N_\mu(ds,du,dz).
\end{aligned}\tag{8.1}$$

Here, X_s is the solution of equation (1.1) and H, h and Q are predictable processes which verify

$$\mathbb{E}\Big[\int_0^T\big(\|H(s,\cdot)\|^2_{\mathbb{L}^2(\mu)}+|h(s)|\big)\,ds+\sup_{s\le T}\sup_{x\in\mathbb{R}^d}\int_G|Q(s,z)|\gamma(s,z,x)\mu(dz))\Big]<\infty.$$

This type of condition is needed in order for the corresponding stochastic (respectively Lebesgue) integrals in (8.1) to make sense.

PROPOSITION 26. *We assume that there exists some predictable process R and some measurable function $\rho : \mathbb{R}_+ \times E \times \mathbb{R}^d \to \mathbb{R}_+$ such that*

$$\big|Q(s,z)\big| \le \rho(s,z,X_s)|R_s|, \tag{8.2}$$

$\mathbb{P}$-almost everywhere on Ω, for all $(s,z) \in \mathbb{R}_+ \times E$. Then, for every $p \ge 2$ there exists a universal constant C (depending on p but not on the coefficients) such that[3]

$$\begin{aligned} \|V\|_{T,p} \le{}& C\exp\big(CT\big(\|\nabla\sigma\|^2_{(\mu,\infty)} + \|\nabla b\|_\infty + [\nabla c]^p_{G,p}\big)\big) \\ &\times \big(|V_0| + \|H\|_{T,p} + \|h\|_{T,p} + [\rho]_{G,p}\|R\|_{T,p}\big). \end{aligned} \tag{8.3}$$

PROOF. Let us begin with writing $V_t = V_0 + I_t + M_t + J_t$, where I_t designates the integral with respect to ds and so on. Using Burkholder's inequality,

$$\begin{aligned} &\mathbb{E}\Big[\sup_{t\le T}|M_t|^p\Big] \\ &\quad\le C\mathbb{E}\Big[\Big(\int_0^T\int_G\big(\big|H(s,z)\big|^2 + \big|\langle\nabla\sigma(s,z,X_s),V_s\rangle\big|^2\big)\mu(dz)\,ds\Big)^{\frac{p}{2}}\Big] \\ &\quad\le C\mathbb{E}\Big[\Big(\int_0^T\big(\|H(s,\cdot)\|^2_{\mathbb{L}^2(\mu)} + \|\nabla\sigma\|^2_{(\mu,\infty)}|V_s|^2\big)\,ds\Big)^{\frac{p}{2}}\Big]. \end{aligned}$$

Hölder's inequality then yields

$$\|M\|_{T,p} \le C\sqrt{T}\big(\|H\|_{T,p} + \|\nabla\sigma\|_{(\mu,\infty)}\|V\|_{T,p}\big).$$

A similar estimate holds true for I_t. Let us now give the estimates on the jump term J_t. To shorten notation, we write dN_μ instead of $N_\mu(ds\,du\,dz)$ and drop the dependency of the coefficients on these variables. Moreover, we consider the standard decomposition of $dN_\mu = d\widetilde{N}_\mu + d\widehat{N}_\mu$ (martingale part and compensator). Corresponding to this decomposition, we write $J_t = \widetilde{J}_t + \widehat{J}_t$. In order to estimate $\widetilde{J}_t$, we will use Burkholder's inequality for jump processes (e.g., [30], Theorem 2.11) to get

$$\begin{aligned} \mathbb{E}\Big[\sup_{t\le T}|\widetilde{J}_t|^p\Big] \le{}& C\mathbb{E}\Big[\Big(\int_0^T\int_{G\times[0,2\Gamma]}|Q + \langle\nabla c, V\rangle|^2 1_{\{u\le\gamma\}}\,d\widehat{N}_\mu\Big)^{\frac{p}{2}}\Big] \\ &+ C\mathbb{E}\Big[\int_0^T\int_{G\times[0,2\Gamma]}|Q + \langle\nabla c, V\rangle|^p 1_{\{u\le\gamma\}}\,d\widehat{N}_\mu\Big]. \end{aligned}$$

By assumption, one has $|Q + \langle\nabla c, V\rangle| \le |\rho||R| + |\nabla c||V|$. Hence (for every fixed time parameter),

$$\begin{aligned} \int_{G\times[0,2\Gamma]}|Q + \langle\nabla c, V\rangle|^2 1_{\{u\le\gamma\}}\,d\widehat{N}_\mu &\le 2\int_G|\rho|^2|R|^2\gamma\,d\mu + 2\int_G|\nabla c|^2|V|^2\gamma\,d\mu \\ &\le 2|R|^2|\rho|^2_{G,2} + 2|V|^2|\nabla c|^2_{G,2}. \end{aligned}$$

[3]We recall that $|\rho|_{G,p}$ is defined in (3.1) and $[\rho]_{G,p} = \sup_{1\le p'\le p}|\rho|_{G,p'}$.

This leads to the following inequality:

$$\mathbb{E}\Big[\Big(\int_0^T \int_{G\times[0,2\Gamma]} \big|Q + \langle \nabla c, V\rangle\big|^2 1_{\{u\le\gamma\}}\, d\widehat{N}_\mu\Big)^{\frac{p}{2}}\Big] \le CT^{\frac{p}{2}}\big(|\rho|_{G,2}^p \|R\|_{T,p}^p + |\nabla c|_{G,2}^p \|V\|_{T,p}^p\big).$$

In a similar way,

$$\mathbb{E}\Big[\int_0^T \int_{G\times[0,2\Gamma]} \big|Q + \langle \nabla c, V\rangle\big|^p 1_{\{u\le\gamma\}}\, d\widehat{N}_\mu\Big] \le CT\big(|\rho|_{G,p}^p \|R\|_{T,p}^p + |\nabla c|_{G,p}^p \|V\|_{T,p}^p\big).$$

For the term $\widehat{J}_t$, similar arguments yield

$$\mathbb{E}\Big[\sup_{t\le T} |\widehat{J}_t|^p\Big] \le \mathbb{E}\Big[\Big(\int_0^T \int_{G\times[0,2\Gamma]} \big|Q + \langle \nabla c, V\rangle\big| 1_{\{u\le\gamma\}}\, d\widehat{N}_\mu\Big)^p\Big] \le CT^p\big(|\rho|_{G,1}^p \|R\|_{T,p}^p + |\nabla c|_{G,1}^p \|V\|_{T,p}^p\big).$$

Summing up these estimates, we conclude that, if $T\le 1$, then

$$\|J\|_{T,p} \le CT^{\frac{1}{p}}\big([\rho]_{G,p}\|R\|_{T,p} + [\nabla c]_{G,p}\|V\|_{T,p}\big).$$

It follows that

$$\begin{aligned}\|V\|_{T,p} &\le \|V_0\|_p + C[\rho]_{G,p}\big(\|R\|_{T,p} + \|H\|_{T,p} + \|h\|_{T,p}\big)\\ &\quad + C\big(T^{\frac{1}{2}}\|\nabla\sigma\|_{(\mu,\infty)} + T\|\nabla b\|_\infty + T^{\frac{1}{p}}[\nabla c]_{G,p}\big)\|V\|_{T,p}.\end{aligned} \tag{8.4}$$

We will use this inequality on the successive intervals $(kT,(k+1)T)$, $k\in\mathbb{N}$ for some convenient T (see after) in order to obtain (8.3). We take

$$T = \min\Big\{\frac{1}{6C\|\nabla b\|_\infty}, \frac{1}{(6C\|\nabla\sigma\|_{(\mu,\infty)})^2}, \frac{1}{(6C[\nabla c]_{G,p})^p}, 1\Big\}$$

which implies $C(T^{\frac{1}{2}}\|\nabla\sigma\|_{(\mu,\infty)} + T\|\nabla b\|_\infty + T^{\frac{1}{p}}[\nabla c]_{G,p}) \le \frac{1}{2}$. Then the inequality (8.4) yields

$$\|V\|_{T,p} \le 2\big(\|V_0\|_p + C[\rho]_{G,p}\big(\|R\|_{T,p} + \|H\|_{T,p} + \|h\|_{T,p}\big)\big).$$

We denote $Q_k = C([\rho]_{G,p}\|R\|_{kT,p} + \|H\|_{kT,p} + \|h\|_{kT,p})$ and $v_k = \|V\|_{kT,p}$ and we obtain

$$v_{k+1} \le 2v_k + Q_k \le 2v_k + Q_n \qquad \forall k\le n$$

and as a consequence $v_n \le 2^n(|V|_0 + Q_n)$. Now, let S be fixed and let $n = [S/T]+1$. Then we get

$$\begin{aligned}\|V\|_{S,p} &\le v_n \le 2^n\big(|V|_0 + O_n\big)\\ &= e^{([S/T]+1)\ln 2}\big(|V|_0 + C\big([\rho]_{G,p}\|R\|_{S,p} + \|H\|_{S,p} + \|h\|_{S,p}\big)\big).\end{aligned}$$

We have

$$[S/T] \leq S \times \max\{6C\|\nabla b\|_\infty, (6C\|\nabla\sigma\|_{(\mu,\infty)})^2, (6C[\nabla c]_{G,p})^p\}$$

so we conclude. □

The same reasoning based on Burkholder's inequality for jump processes as in the previous proof leads to the following.

REMARK 27. For every $p \geq 2$, there exists a universal constant C (depending on p) such that for every f,

$$\begin{aligned}\Big(\mathbb{E}\Big[\Big(\int_0^t \int_{G\times(0,2\Gamma)} |f(s,z,X_{s-})| 1_{\{u\leq \gamma_\Gamma(s,z,X_{s-})\}} N_\mu(ds,du,dz)\Big)^p\Big]\Big)^{\frac{1}{p}} \\ \leq C\max\{t,1\}[f]_{G,p}.\end{aligned} \tag{8.5}$$

8.2.2. *Proof of Lemma* 10. PROOF OF LEMMA 10. We will prove, by recurrence that, for all $p \geq 2k$,

$$\begin{aligned}\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[\sup_{t\leq T} |\partial^\alpha \overline{X}_t^G(x)|^{\frac{p}{k}}\Big]^{\frac{k}{p}} &\leq \alpha_{k,p}(C,G) \\ &= C\theta_{k,p}^{k\sum_{1\leq n\leq k}\frac{1}{n}}(G)\exp\Big(CTk\Big(\sum_{1\leq n\leq k}\frac{1}{n}\Big)a_p(G)\Big).\end{aligned} \tag{8.6}$$

Step 1. (Chain Estimates) Let $f:\mathbb{R}^d\to\mathbb{R}$ and $g:\mathbb{R}^d\to\mathbb{R}^d$ be smooth functions. If $|\alpha| = k \geq 1$, then

$$\begin{aligned}|\partial^\alpha(f(g(x)))| \leq C\Big(&\sum_{i=1}^d |(\partial_i f)(g(x))| \times |\partial^\alpha g^i(x)| \\ &+ \sum_{2\leq|\alpha'|\leq k} |(\partial^{\alpha'} f)(g(x))| \times \sum_{1\leq|\beta|\leq k-1} |\partial^\beta g(x)|^{\frac{k}{|\beta|}}\Big).\end{aligned}$$

The above inequality is obtained by taking first derivatives and then by using Young's inequality in order to separate the different derivatives of g. As an immediate consequence, one gets

$$|\partial^\alpha f(g(x))| \leq C\Big(\|\nabla f\|_\infty |\partial^\alpha g(x)| + \|f\|_{2,k,\infty} \sum_{1\leq|\beta|\leq k-1} |\partial^\beta g(x)|^{\frac{k}{|\beta|}}\Big). \tag{8.7}$$

Similar reasoning for $F:\mathbb{R}_+\times E\times\mathbb{R}^d\to\mathbb{R}$ that is globally measurable and differentiable with respect to $x\in\mathbb{R}^d$ yields

$$\begin{aligned}&\left(\int_E\left|\partial_x^\alpha\big(F(t,z,g(x))\big)\right|^2\mu(dz)\right)^{\frac{1}{2}}\\ &\quad\le C\left(\|\nabla F\|_{(\mu,\infty)}\left|\partial^\alpha g(x)\right|+\|F\|_{2,k,(\mu,\infty)}\sum_{1\le|\beta|\le k-1}\left|\partial^\beta g(x)\right|^{\frac{k}{|\beta|}}\right).\end{aligned}\tag{8.8}$$

Having this inequality in mind [and the notation (2.2)], we introduce the following notation:

$$y_\alpha=\partial^\alpha g(x),\qquad y_{[k-1]}=\big(\partial^\beta g(x)\big)_{1\le|\beta|\le k-1}.\tag{8.9}$$

Using this notation, the estimate (8.7) [resp., (8.8)] reads

$$\left|\partial^\alpha f\big(g(x)\big)\right|\le C\big(\|\nabla f\|_\infty|y_\alpha|+\|f\|_{2,k,\infty}|y_{[k-1]}|_{\mathbb{R}_{[k-1]}}\big),\tag{8.10}$$

$$\begin{aligned}&\left(\int_E\left|\partial_x^\alpha\big(F(t,z,g(x))\big)\right|^2\mu(dz)\right)^{\frac{1}{2}}\\ &\quad\le C\big(\|\nabla F\|_{(\mu,\infty)}|y_\alpha|+\|F\|_{2,k,(\mu,\infty)}|y_{[k-1]}|_{\mathbb{R}_{[k-1]}}\big).\end{aligned}\tag{8.11}$$

Step 2. [Deriving the Differential Equation for $\partial^\alpha\overline{X}_t^G(x)$ and Estimates] We denote by $\overline{Y}_\alpha(t,x)=\partial^\alpha\overline{X}_t^G(x)$ and by $\overline{Y}_{[k]}(t,x)=(\overline{Y}_\alpha(t,x))_{1\le|\alpha|\le k}\in\mathbb{R}^d_{[k]}$, for all initial data x. We claim that, for every multi-index α with $|\alpha|=k\ge1$,

$$\begin{aligned}\overline{Y}_\alpha(t,x)&=\partial^\alpha\overline{X}_0^G(x)\\ &\quad+\int_0^t\big(g_\alpha\big(s,\overline{X}_s^G,\overline{Y}_{[k-1]}(s,x)\big)+\big\langle\nabla b\big(s,\overline{X}_s^G\big),\overline{Y}_\alpha(s,x)\big\rangle\big)\,ds\\ &\quad+\int_0^t\int_E\big(h_\alpha\big(s,z,\overline{X}_s^G,\overline{Y}_{[k-1]}(s,x)\big)+\big\langle\nabla\sigma_l\big(s,z,\overline{X}_s^G\big),\overline{Y}_\alpha(s,x)\big\rangle\big)\\ &\quad\times W_\mu(ds,dz)\\ &\quad+\sum_{j=1}^{J_t}Q_\alpha\big(T_j,\overline{Z}_j,\overline{X}_{T_j-}^G(x),\overline{Y}_{[k-1]}(T_j-,x)\big)1_G(\overline{Z}_j)\\ &\quad+\sum_{j=1}^{J_t}\big\langle\nabla c\big(T_j,\overline{Z}_j,\overline{X}_{T_j-}^G\big),\overline{Y}_{T_j-}^\alpha(x)\big\rangle 1_G(\overline{Z}_j),\end{aligned}\tag{8.12}$$

where g_α, h_α and Q_α satisfy

$$\left(\int_E\big(\left|h_\alpha(t,z,x,y_{[k-1]})\right|^2\mu(dz)\big)\right)^{\frac{1}{2}}\le C\|\sigma\|_{2,k,(\mu,\infty)}|y_{[k-1]}|_{\mathbb{R}^d_{[k-1]}},\tag{8.13}$$

$$\left|g_\alpha(t,x,y)\right|\le C\|b\|_{2,k,\infty}|y_{[k-1]}|_{\mathbb{R}^d_{[k-1]}}\tag{8.14}$$

and

$$|Q_\alpha(t,z,x,y)| \le C|y_{[k-1]}|_{\mathbb{R}^d_{[k-1]}} \sum_{2\le|\alpha|\le k} |\partial_x^\alpha c(t,z,x)|. \tag{8.15}$$

The equation (8.12) is obtained by taking formal derivatives in (2.18) and then, (8.13), (8.14) and (8.15) are obtained by using (8.10) and (8.11).

Step 3. Now we prove (8.6) by recurrence on k. If $k=1$, the inequality (8.6) is an immediate consequence of Proposition 26. Let us now assume (8.6) to hold true for $k-1$. In order to prove it for k, we will make use of Proposition 26. A first step is to make use the identity of laws from Lemma 9. We denote by $Y_{[k]}(t,x) = (Y_\alpha(t,x))_{1\le|\alpha|\le k} \in \mathbb{R}^d_{[k]}$ the unique solution of the system of equations

$$\begin{aligned} Y_\alpha(t,x) &= \partial^\alpha X_0^G(x) \\ &\quad + \int_0^t \big(g_\alpha\big(s, X_s^G, Y_{[k-1]}(s,x)\big) + \big\langle \nabla_x b\big(s, X_s^G\big), Y_\alpha(s,x)\big\rangle\big)\, ds \\ &\quad + \int_0^t \int_E \big(h_\alpha\big(s,z,X_s^G, Y_{[k-1]}(s,x)\big) + \big\langle \nabla_x \sigma_l\big(s,z,X_s^G\big), Y_\alpha(s,x)\big\rangle\big) \\ &\quad \times W_\mu(ds,dz) \\ &\quad + \sum_{j=1}^{J_t} \big(Q_\alpha\big(T_j, Z_j, X_{T_j-}^G(x), Y_{[k-1]}(T_j-,x)\big) \\ &\quad + \big\langle \nabla_x c\big(T_j, Z_j, X_{T_j-}^G\big), Y_\alpha(T_j-,x)\big\rangle\big) 1_{\{U_k \le \gamma(T_j,Z_j,X_{T_j-})}. \end{aligned} \tag{8.16}$$

These equations are the same as in (8.12) but we replace $\overline{X}_s^G$ by X_s^G, $\overline{Z}_j$ by Z_j and $1_G(\overline{Z}_j)$ by $1_{\{U_k\le\gamma(T_j,Z_j,X_{T_j-})\}}$. According to Lemma 9, $\overline{Y}_{[k]}(t,x)$, $t\ge 0$ has the same law as $Y_{[k]}(t,x)$, $t\ge 0$. Now we use Proposition 26 for $V_0 = \partial^\alpha X_0^G(x)$ (implying that $|V_0|\le 1$), $h(s) = g_\alpha(s, X_s^G, Y_{[k-1]}(s,x))$ and

$$H(s,z) = h_\alpha\big(s,z,X_s^G(x), Y_{[k-1]}(s,x)\big),$$
$$Q(s,z) = Q_\alpha\big(s,z,X_s^G(x), Y_{[k-1]}(s,x)\big),$$

for all $s\in[0,T]$ and all $z\in E$. In view of (8.13) and (8.14),

$$\|H(s)\|_{\mathbb{L}^2(\mu)} \le C\|\sigma\|_{2,k,(\mu,\infty)} \times \sup_{s\le t} |Y_{[k-1]}(s,x)|_{\mathbb{R}^d_{[k-1]}},$$
$$|h(s)| \le C\|b\|_{2,k,\infty} \times \sup_{s\le t} |Y_{[k-1]}(s,x)|_{\mathbb{R}^d_{[k-1]}}.$$

The estimates (8.15) give

$$|Q(s,z)| \le \rho\big(s,z,X_s^G(x)\big) R_s,$$

where

$$R_s = \left|Y_{[k-1]}(s-, x)\right|_{\mathbb{R}^d_{[k-1]}}, \qquad \rho(s, z, x) = C \sum_{2\le|\alpha'|\le k} \left|\partial_x^{\alpha'} c(s, z, x)\right|.$$

We recall that $|\alpha| = k \geq 2$ and $p \geq 2k$. Then, by Proposition 26, one has

$$\begin{aligned}
&\left\|\partial^\alpha \overline{X}^G_\cdot(x)\right\|_{T,\frac{p}{k}} \\
&\quad = \left\|\partial^\alpha X^G_\cdot(x)\right\|_{T,\frac{p}{k}} \\
&\quad \leq C \exp\bigl(CTa_{\frac{p}{k}}(G)\bigr)\theta_{k,\frac{p}{k}}(G) \sup_{x\in\mathbb{R}^d}\Bigl(\mathbb{E}\Bigl[\sup_{s\le T}\left|Y_{[k-1]}(s,x)\right|^{\frac{p}{k}}_{\mathbb{R}^d_{[k-1]}}\Bigr]\Bigr)^{\frac{k}{p}} \vee 1 \qquad (8.17)\\
&\quad = C \exp\bigl(CTa_{\frac{p}{k}}(G)\bigr)\theta_{k,\frac{p}{k}}(G) \sum_{1\le|\beta|\le k-1}\Bigl(\mathbb{E}\Bigl[\sup_{s\le T}\left|\partial^\beta X^G_s(x)\right|^{\frac{kp}{|\beta|k}}\Bigr]\Bigr)^{\frac{k}{p}}.
\end{aligned}$$

We assume that $1 \leq |\beta| = r \leq k-1$. Using the recurrence hypothesis and due to the fact that $\frac{kp}{|\beta|k} = \frac{p}{r}$, one gets

$$\begin{aligned}
\Bigl(\mathbb{E}\Bigl[\sup_{s\le T}\left|\partial^\beta X^G_s(x)\right|^{\frac{p}{|\beta|}}\Bigr]\Bigr)^{\frac{|\beta|}{p}} &= \left\|\partial^\beta \overline{X}^G_\cdot(x)\right\|_{T,\frac{p}{r}} \\
&\leq C\theta_{r,p}^{r\sum_{1\le n\le r}\frac{1}{n}}(G)\exp\biggl(CTr\biggl(\sum_{1\le n\le r}\frac{1}{n!}\biggr)a_p(G)\biggr).
\end{aligned}$$

This implies

$$\begin{aligned}
&\Bigl(\mathbb{E}\Bigl[\sup_{s\le T}\left|\partial^\beta X^G_s(x)\right|^{\frac{kp}{|\beta|k}}\Bigr]\Bigr)^{\frac{k}{p}} \\
&\quad \leq C\theta_{r,p}^{k\sum_{1\le n\le k-1}\frac{1}{n}}(G)\exp\biggl(CTk\biggl(\sum_{1\le n\le k-1}\frac{1}{n}\biggr)a_p(G)\biggr).
\end{aligned}$$

We insert this inequality in (8.17) and note that $a_{\frac{p}{k}}(G) \leq a_p(G)$ and $\theta_{k,\frac{p}{k}}(G) \leq \theta_{k,p}(G)$ to conclude

$$\begin{aligned}
\left\|\partial^\alpha \overline{X}^G_\cdot(x)\right\|_{T,\frac{p}{k}} &\leq C\theta_{k,p}^{1+k\sum_{1\le n\le k-1}\frac{1}{n}}(G)\exp\biggl(CTa_p(G)\biggl(1+k\sum_{1\le n\le k-1}\frac{1}{n}\biggr)\biggr) \\
&= C\theta_{k,p}^{k\sum_{1\le n\le k}\frac{1}{n}}(G)\exp\biggl(CTa_p(G)k\sum_{1\le n\le k}\frac{1}{n}\biggr).
\end{aligned}$$

The proof is now complete by taking pk to replace p in (8.6). $\square$

8.2.3. *Proof of Corollary 11 and Lemma 12*. We begin with the following simple remark.

REMARK 28. Whenever $n \in \mathbb{N}^*$ and $p \leq n$, one gets

$$\|J_t\|_p \leq \|J_t\|_n \quad \text{and} \quad \|J_t\|_n^n = \frac{d^n e^{2\Gamma\mu(G)t(e^s-1)}}{ds^n} /_{s=0} = P_n(\Gamma\mu(G)t),$$

an n-degree polynomial. As a consequence, for some large enough constant depending, eventually, on the upper bound n but not on Γ, $\mu(G)$ nor on t,

$$\|J_t\|_p \leq C\Gamma\mu(G)\max(t, 1). \tag{8.18}$$

We now give the proof of Corollary 11.

PROOF OF COROLLARY 11. In order to prove the first assertion, one simply writes, ($\mathbb{P}$-almost surely on $k \leq J_t$),

$$\left|\partial^\alpha\left(g\left(T_k, \overline{X}^G_{T_k-}(x)\right)\right)\right| \leq C\|g\|_{1,q,\infty}A_q \qquad \text{with } A_q = 1 \vee \sum_{1\leq|\rho|\leq q} \sup_{s\leq t}\left|\partial_x^\rho \overline{X}^G_s(x)\right|^q.$$

Using Hölder's inequality and (3.6), we upper bound the term in the left-hand side of (3.7) by

$$\begin{aligned} C\|g\|_{1,q,\infty}\mathbb{E}\left(\left[(J_t \times A_q)^p\right]\right)^{\frac{1}{p}} &\leq C\|g\|_{1,q,\infty}\|J_t\|_{\frac{(1+\eta)p}{\eta}}\left(\mathbb{E}\left[A_q^{(1+\eta)p}\right]\right)^{\frac{1}{(1+\eta)p}} \\ &\leq C\|g\|_{1,q,\infty}\Gamma\mu(G)\max(t, 1)\alpha^q_{q,(1+\eta)pq}(C, G). \end{aligned}$$

To prove the second assertion, we write

$$\sum_{k=1}^{J_t} 1_G(\overline{Z}_k)\left|\partial^\alpha g\left(T_k, \overline{Z}_k, \overline{X}^G_{T_k-}(x)\right)\right| \leq A_q \times B_q,$$

with

$$B_q = \sum_{1\leq|\beta|\leq q} B_q(\beta) \qquad \text{where } B_q(\beta) = \sum_{k=1}^{J_t} 1_G(\overline{Z}_k)\left|\left(\partial^\beta g\right)\left(T_k, \overline{Z}_k, \overline{X}^G_{T_k-}(x)\right)\right|.$$

Using Hölder's inequality and (3.6), we upper bound the term in (3.8) by

$$\|A_q\|_{\frac{(1+\eta)p}{\eta}} \times \|B_q\|_{(1+\eta)p} \leq \alpha^q_{q,\frac{(1+\eta)pq}{\eta}}(C, G) \times \|B_q\|_{(1+\eta)p}.$$

Using the identification of laws from Lemma 9 and the inequality (8.5), one has

$$\begin{aligned} &\|B_q\|_{(1+\eta)p} \\ &\quad = \left\|\int_0^t \int_{G\times[0,2T]} \left|\partial^\beta g\left(s, z, X^G_{s-}(x)\right)\right| 1_{\{u\leq\gamma(s,z,X^G_{s-}(x))\}} N_\mu(ds, dz, du)\right\|_{(1+\eta)p} \\ &\quad \leq C\left[\partial^\beta g\right]^{(1+\eta)p}_{G,(1+\eta)p}. \end{aligned}$$

The assertion follows by putting these estimates together. $\square$

8.3. *Proofs of results in Section 6*. We begin with Theorem 19. As already hinted before, the result follows from Theorem 16.

PROOF OF THEOREM 19. We use Theorem 16 with $k=0$ and $q=3$. It is easy to check that P_t^ε verifies $H_2(0)$ and $H_3(0,3)$ [note that the constant C in (5.5) depends on ε; but is not involved in the estimate (5.7)]. And $\mathcal{P}_t$ verifies $H_2(0)$ and $H_3(0,3)$ as well. Moreover, the constant $Q_3(t,P)$ defined in (5.1) verifies

$$Q_3(t,\mathcal{P}) \le C(t\vee 1)^3 C_*^{42}\exp\bigl((t\vee 1)CC_*^{36}\bigr),$$

where C_* is the constant in (6.3) and C is a universal constant. Moreover, using a Taylor expansion of order three we get

$$\bigl\|(\mathcal{L}_\varepsilon-\mathcal{L})f\bigr\|_\infty \le C\delta(\varepsilon)\|f\|_{3,\infty}$$

with C an universal constant. Then (5.7) gives (6.4). □

Next, we proceed with checking the Assumption H_1 to complete the explicit example.

PROOF OF ASSUMPTION H_1. We notice that, by the choice of α,

$$\int_\varepsilon^{3\varepsilon} c_\varepsilon(z,x)\gamma(x)\frac{dz}{z^2}=0$$

so our equation may be written as

$$\begin{aligned}X_t^\varepsilon = x &+ \int_0^t\int_\varepsilon^{3\varepsilon}\int_0^1 c_\varepsilon\bigl(z,X_{s-}^\varepsilon\bigr)1_{\{u\le\gamma(X_{s-}^\varepsilon)\}}\widetilde{N}_{\mu_\varepsilon}(ds,dz,du)\\ &+\int_0^t\int_{3\varepsilon}^1\int_0^1 c_\varepsilon\bigl(z,X_{s-}^\varepsilon\bigr)1_{\{u\le\gamma(X_{s-}^\varepsilon)\}}N_{\mu_\varepsilon}(ds,dz,du).\end{aligned}$$

This is the same as the equation (6.1). We take $E=[0,1]$, $A_\varepsilon=(0,3\varepsilon]$, $B_\varepsilon=(3\varepsilon,4\varepsilon]$ and $C_\varepsilon=(4\varepsilon,1]$ and we have

$$\int_\varepsilon^{3\varepsilon} c_\varepsilon^2(z,x)\gamma(x)\frac{dz}{z^2}=\sigma^2(x),\qquad \int_{3\varepsilon}^{4\varepsilon} c_\varepsilon(z,x)\gamma(x)\frac{dz}{z^{3/2}}=b(x),$$

so that $\delta_\sigma(\varepsilon)=\delta_b(\varepsilon)=0$. Moreover, on C_ε, $c=c_\varepsilon$ which implies $\delta_{c,\gamma}(\varepsilon)=0$. Finally, simple computations yield $\delta_A(\varepsilon)+\delta_B(\varepsilon)+\delta_C(\varepsilon)\le C\sqrt{\varepsilon}$ which leads to the desired conclusion. □

8.4. *Proofs of the results in Section 7*. Before proceeding to the proofs, we recall the following estimates for the derivatives of the above cut off function given in [6], Lemma 2.3.

LEMMA 29 ([6], Lemma 2.3). *There exists* $\varepsilon_0 > 0$ *such that, for every* $\varepsilon \in (0, \varepsilon_0)$, *every multi-index* $\alpha \in \{1,2\}^l$, $l \in \mathbb{N}^*$ *and every* $v \in \mathbb{R}^2$, *one has*

$$\bigl|\partial^\alpha \ln \varphi_\varepsilon\bigl(|v|\bigr)\bigr| \le C_l\bigl(1_{\{|v|\in(\varepsilon,\Gamma_\varepsilon-1]\}}|v|^{-l} + 1_{\{|v|\in(\Gamma_\varepsilon-1,\Gamma_\varepsilon+1)\}}\Gamma_\varepsilon^{-1}\bigr), \tag{8.19}$$

$$\bigl|\partial^\alpha \varphi_\varepsilon^\kappa\bigl(|v|\bigr)\bigr| \le C_l\bigl(1_{\{|v|\in(\varepsilon,\Gamma_\varepsilon-1]\}}|v|^{\kappa-l} + 1_{\{|v|\in(\Gamma_\varepsilon-1,\Gamma_\varepsilon+1)\}}\Gamma_\varepsilon^{\kappa-1}\bigr). \tag{8.20}$$

Moreover, for every $\beta \in (0,1]$, $\varepsilon \in (0,\varepsilon_0)$ *and* $x, y \ge 0$,

$$x^\beta\bigl|\varphi_\varepsilon^\kappa(x) - \varphi_\varepsilon^\kappa(y)\bigr| \le C_\beta \Gamma_\varepsilon^\kappa |x-y|^\beta. \tag{8.21}$$

8.4.1. *Proof of the first-order estimates in Boltzmann equation.* The proof consists in two major steps. First, we give upper-bounds for the constants in (3.3), (3.4) and (3.5). Second, we use Theorem 16 for which we check the assumptions. To conclude, we invoke (7.6) together with the estimates provided by Theorem 16.

We recall the parameters associated with $\widehat{\mathcal{P}}^\delta$ in (3.3), (3.4) and (3.5),

$$\theta_{q,p,(\delta)}(E_\delta) = 1 + \|b_\delta\|_{2,q,\infty} + \sum_{2\le|\alpha|\le q}\bigl[\partial_v^\alpha c\bigr]_{E_\delta,p} = 1 + \|b_\delta\|_{2,q,\infty},$$

$$a_{p,(\delta)}(E_\delta) = \|\nabla b_\delta\|_\infty + [\nabla c]_{G,p}^p,$$

$$\alpha_{q,p,(\delta)}(C,E_\delta) = C\theta_{q,pq,(\delta)}^{q\sum_{1\le n\le q}\frac{1}{n}}(E_\delta)\exp\biggl(CTq\sum_{1\le n\le q}\frac{1}{n}\bigl(\|\nabla b_\delta\|_\infty + [\nabla c]_{G,pq}^{pq}\bigr)\biggr)$$

the first expression following from $\partial^\alpha c = 0$ if $|\alpha| \ge 2$.

LEMMA 30. *We assume that* $\kappa < \frac{1}{8}$ *and we take* $q = 2$ (*so that* $q^2\kappa < \frac{1}{2}$). *For every* $a > 0$, *there exists* $\varepsilon_0 > 0$ *and* $C \ge 1$ *such that for every* $\varepsilon \in (0,\varepsilon_0)$ *one has*

$$\alpha_{q,4q,(\delta)}(C,E_\delta) \le C\varepsilon^{-a}. \tag{8.22}$$

Moreover [*see* (4.2) *for the notation*], *for all* $r \le \frac{1-\nu}{1-\kappa}$,

$$\Gamma_{E_\delta,q}(\gamma) + [\ln\gamma]_{E_\delta,q,4q} \le C\times\delta^{-q\nu}\times\varepsilon^{-q} = C\times\delta^{-q(\nu+r)}. \tag{8.23}$$

As a consequence, the constant in the right-hand side of (5.7) *verifies*

$$Q_q\bigl(T,\widehat{\mathcal{P}}^\delta\bigr) \le C\times\delta^{-q(\nu+r+a)}. \tag{8.24}$$

PROOF. (Throughout the proof, C will be a universal real constant and a an arbitrary small constant that may change from one line to another.)

Since $|\partial^\alpha c(t,\theta,\rho,v)| \le |\theta| \times 1_{|\alpha|=1}$ and $|\gamma| \le \Gamma_\varepsilon^\kappa$ we get, for every $p \ge 1$,

$$\int_{E_\delta}\bigl|\partial^\alpha c(t,\theta,\rho,v)\bigr|^p\gamma(t,\rho,v)\mu(d\theta,d\rho) \le C\Gamma_\varepsilon^\kappa.$$

In particular,

$$[\nabla c]_{E_\delta,4q^2}^{4q^2} = \sup_{1\le p\le 4q}|\nabla c|_{E_\delta,p}^{4q^2} \le C\Gamma_\varepsilon^{4q^2\kappa}$$

and consequently, if $4q^2\kappa < 2$ then, as a consequence of (7.9), for sufficiently small $\varepsilon > 0$,

$$\exp\big(CT[\nabla c]_{E_\delta,4q^2}^{4q^2}\big) \le \exp\big(C\Gamma_\varepsilon^{4q^2\kappa}\big) \le \varepsilon^{-a}.$$

Moreover, using (8.20) we get

$$\big|\partial^\alpha\big(c(t,\theta,\rho,v)\gamma_\varepsilon(t,\rho,v)\big)\big| \le C|\theta| \times \big(\varepsilon^{1+\kappa-|\alpha|} + \Gamma_\varepsilon^\kappa\big).$$

As a consequence,

$$\big|\partial^\alpha b_\delta(t,v)\big| \le C\delta^{1-\nu} \times \big(\varepsilon^{1+\kappa-|\alpha|} + \Gamma_\varepsilon^\kappa\big) = C\big[\delta^{1-\nu+r(1+\kappa-|\alpha|)} + \delta^{1-\nu}\Gamma_\varepsilon^\kappa\big] \le C$$

the last inequality being true for $|\alpha| = 2$, if $r \le \frac{1-\nu}{|\alpha|-1-\kappa} = \frac{1-\nu}{1-\kappa}$. We infer that $\|b_\delta\|_{1,q,\infty} \le C$ and $\theta_{q,4q^2}(E_\delta) = 1 + \|b_\delta\|_{2,q,\infty} \le C$ implying (8.22).

We now turn to the proof of the inequality (8.23). Using (8.19), we get

$$\begin{aligned}
&\gamma\big|\partial_v^\alpha \ln\gamma\big|^p \\
&\quad\le C1_{\{|v-v_t(\rho)|>\Gamma_\varepsilon-1\}} \\
&\qquad + 1_{\{|v-v_t(\rho)|\in[\varepsilon,\Gamma_\varepsilon-1]\}}\big(1+\big|v-v_t(\rho)\big|^{-|\alpha|}\big)^p\big|v-v_t(\rho)\big|^\kappa \\
&\quad\le C\big(1_{\{|v-v_t(\rho)|>\Gamma_\varepsilon-1\}} + 1_{\{|v-v_t(\rho)|\in[\varepsilon,\Gamma_\varepsilon-1]\}}\varepsilon^{-p|\alpha|+\kappa}\big).
\end{aligned}$$

For every $p \ge 1$ (and small enough δ), we infer

$$\big|\partial_v^\alpha \ln\gamma\big|_{E_\delta,p} \le C\delta^{-\nu/p}\varepsilon^{-|\alpha|} \le C\delta^{-\nu}\varepsilon^{-|\alpha|}.$$

In particular,

$$[\ln\gamma]_{E_\delta,4q} = \sum_{1\le|\alpha|\le q}\sup_{1\le p\le 4q}\big|\partial_v^\alpha\ln\gamma\big|_{E_\delta,4q} \le C\delta^{-\nu}\varepsilon^{-q} = C\delta^{-\nu-qr} \quad \text{and}$$

$$\Gamma_{G,q}(\gamma) = \sum_{h=1}^{q}\sum_{1\le|\alpha|\le h}\big|\partial_v^\alpha\ln\gamma\big|_{E_\delta,h/|\alpha|}^{q/|\alpha|} \le C\delta^{-q\nu}\varepsilon^{-q} = C\delta^{-q(\nu+r)}.$$

□

REMARK 31. The inequality (8.24) remains true (with the exact same proof) for $q = 3$ and $r \le \frac{1-\nu}{2-\kappa}$ as soon as $\kappa < \frac{1}{18}$.

PROOF. Finally, by gathering these estimates, one obtains (8.23). □

LEMMA 32. *Suppose that* (7.8) *holds and* $\kappa < \frac{1}{8}$ (*we recall that* $\varepsilon = \delta^r$ *with* $0 < r \le \frac{1-\nu}{1-\kappa}$). *For every* $a > 0$, *there exists* $\varepsilon_0 > 0$ *such that, for* $\varepsilon < \varepsilon_0$,

$$\left\|\frac{1}{\psi_2}\big(P_{t_0,t}^\varepsilon f - \widehat{P}_{t_0,t}^\delta\big)f\right\|_\infty \le C\delta^{2-3\nu-2r-3a} \times \|f\|_{2,\infty}. \tag{8.25}$$

PROOF. We will use Theorem 16 with $q = 2$, $k = 2$.

Step 1. We check that for every $p \in \mathbb{N}$ and every $a > 0$, one can find $C \geq 1$ and $\varepsilon_{p,a} > 0$ such that, for every $\varepsilon \in (0, \varepsilon_{p,a})$, the following estimate holds true:

$$\mathbb{E}\big[\big|V_t^\varepsilon(v)\big|^p\big] \leq C\psi_p(v)\varepsilon^{-a}. \tag{8.26}$$

This implies that the hypothesis $H_2(p)$ [see (5.3)] holds for the semigroup $\mathcal{P}^\varepsilon$ with $C_p(T, \mathcal{P}^\varepsilon) = \varepsilon^{-a}$.

To this purpose, we use Itô's formula for $f_p(x) = |x|^p$ to get

$$\mathbb{E}\big[\big|V_t^\varepsilon(v)\big|^p\big] = |v|^p + J_p(t),$$

where

$$\begin{aligned} J_p(t) = \mathbb{E}\bigg[\int_0^t \int_{E\times R_+} &\big[\big(\big|V_{s-}^\varepsilon + A(\theta)\big(V_{s-}^\varepsilon - v_s(\rho)\big)\big|^p - \big|V_{s-}^\varepsilon\big|^p\big) \\ &\times 1_{\{u \leq \varphi_\varepsilon^\kappa(|V_{s-}^\varepsilon - v_s(\rho)|\}}\big]\,dN\bigg]. \end{aligned}$$

Using the inequality $||a+b|^p - |a|^p| \leq C|b|(|a|^{p-1} + |b|^{p-1})$, we obtain

$$\begin{aligned} \big|J_p(t)\big| &\leq C\Gamma_\varepsilon^\kappa \mathbb{E}\bigg[\int_0^t \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^1 \big[\big|A(\theta)\big(V_s^\varepsilon - v_s(\rho)\big)\big| \\ &\qquad \times \big(\big|A(\theta)\big(V_s^\varepsilon - v_s(\rho)\big)\big|^{p-1} + \big|V_s^\varepsilon\big|^{p-1}\big)\big]\frac{d\theta}{\theta^{1+\nu}}\,d\rho\,ds\bigg] \\ &\leq C\Gamma_\varepsilon^\kappa \mathbb{E}\bigg[\int_0^t \int_0^1 \big(\big|v_s(\rho)\big| + \big|V_s^\varepsilon\big|\big)\big(\big|v_s(\rho)\big|^{p-1} + \big|V_s^\varepsilon\big|^{p-1}\big)\,d\rho\,ds\bigg] \\ &\leq C\Gamma_\varepsilon^\kappa \mathbb{E}\bigg[\int_0^t \int_0^1 \big(\big|v_s(\rho)\big|^p + \big|V_s^\varepsilon\big|^p\big)\,d\rho\,ds\bigg] \leq C\Gamma_\varepsilon^\kappa\bigg(1 + E\int_0^t \big|V_s^\varepsilon\big|^p\bigg)\,ds. \end{aligned}$$

The last inequality is a consequence of

$$\int_0^1 \big|v_s(\rho)\big|^p\,d\rho = \int_{R^2} |v|^p f_s(dv) \leq C < \infty. \tag{8.27}$$

The inequality (8.26) is then a consequence of Gronwall's lemma.

Step 2. Second, we need to estimate, for regular f, the difference between the actions of infinitesimal operators:

$$\begin{aligned} &\big(\mathcal{L}_t^\varepsilon - \widehat{\mathcal{L}}_t^\delta\big)f(v) \\ &\quad = \int_{E_\delta^c} \mu(d\theta, d\rho)\gamma_\varepsilon(t, \rho, v) \\ &\qquad \times \big(f\big(v + c(t, \theta, \rho, v)\big) - f(v) - \big\langle\nabla f(v), c(t, \theta, \rho, v)\big\rangle\big). \end{aligned}$$

One easily notes [using, again, (8.27)], that

$$\begin{aligned}\big|\big(\mathcal{L}_t^{\varepsilon}-\widehat{\mathcal{L}}_t^{\delta}\big)f(v)\big| &\leq \|f\|_{2,\infty}\Gamma_{\varepsilon}^{\kappa}\int_{E_{\delta}^{c}}\mu(d\theta,dv)\big|c(t,\theta,\rho,v)\big|^{2}\\ &\leq C\psi_2(v)\delta^{2-\nu}\|f\|_{2,\infty}\Gamma_{\varepsilon}^{\kappa}.\end{aligned}$$

We conclude that

$$\left\|\frac{1}{\psi_2}\big(\mathcal{L}_t^{\varepsilon}-\widehat{\mathcal{L}}_t^{\delta}\big)f\right\|_{\infty}\leq C\delta^{2-\nu}\|f\|_{2,\infty}\varepsilon^{-a}. \tag{8.28}$$

This proves that (5.6) holds with $k=2$ and $\varepsilon(t)=\delta^{2-\nu}\varepsilon^{-a}$.

Step 3. We check that $H_3(1,2)$ holds true for both $\widehat{\mathcal{P}}^{\delta}$ and $\mathcal{P}^{\varepsilon}$. We will only check it for the approximating semigroup $\widehat{\mathcal{P}}^{\delta}$, the remaining case being very similar. We recall that $f_t(dv)=\mathbb{P}(V_t\in dv)$ where V_t is the solution of the equation (7.3). Then, for every $x\in\mathbb{R}^2$,

$$\begin{aligned}b_{\delta}(t,x)&=\int_{E_{\delta}^{c}}\mu(d\theta,d\rho)\gamma_{\varepsilon}(t,\rho,x)c(t,\theta,\rho,x)\\ &=\int_{\{\theta\leq\delta\}}\mathbb{E}\big[A(\theta)(x-V_t)\varphi_{\varepsilon}^{\kappa}\big(|x-V_t|\big)\big]\frac{d\theta}{\theta^{1+\nu}}.\end{aligned}$$

Using (8.21) with $\beta=1$, we get

$$\big|b_{\delta}(t,x)\big|+\big|\nabla b_{\delta}(t,x)\big|\leq C\Gamma_{\varepsilon}^{\kappa}\psi_1(x)$$

such that

$$\left\|\frac{1}{\psi_1}\nabla\big(b_{\delta}(t,\circ)\nabla f\big)\right\|_{\infty}\leq C\Gamma_{\varepsilon}^{\kappa}\|f\|_{2,\infty}. \tag{8.29}$$

We write now

$$\begin{aligned}I_t(f)(x)&:=\int_{E_{\delta}}\mu(d\theta,d\rho)\gamma(t,\rho,x)\big(f\big(x+c(t,\theta,\rho,x)\big)-f(x)\big)\\ &=\int_{E_{\delta}}\mu(d\theta,d\rho)\gamma(t,\rho,x)\int_0^1 d\lambda\big\langle\nabla f\big(x+\lambda c(t,\theta,\rho,x)\big),c(t,\theta,\rho,x)\big\rangle.\end{aligned}$$

Hence, $I_t(f)(x)$ is upper bounded by

$$\mathbb{E}\bigg[\int_{\{\delta\leq|\theta|\leq\frac{\pi}{2}\}}\frac{d\theta}{\theta^{1+\nu}}\varphi_{\varepsilon}^{\kappa}\big(|x-V_t|\big)\int_0^1 d\lambda\big\langle\nabla f\big(x+\lambda A(\theta)(x-V_t)\big),A(\theta)(x-V_t)\big\rangle\bigg].$$

And, using again (8.21) with $\beta=1$, this gives

$$\left\|\frac{1}{\psi_1}I_t(f)\right\|_{\infty}+\left\|\frac{1}{\psi_1}\nabla I_t(f)\right\|_{\infty}\leq C\Gamma_{\varepsilon}^{\kappa}\|f\|_{2,\infty}. \tag{8.30}$$

Step 4. We use (5.7) with $q = 2$, $k = 2$:

$$\left\| \frac{1}{\psi_1} (\mathcal{P}^{\varepsilon}_{t_0,t} f - \widehat{\mathcal{P}}^{\delta}_{t_0,t}) f \right\|_\infty \leq C \times C_1(t, \mathcal{P}^{\varepsilon}) Q_2(t, \widehat{\mathcal{P}}^{\delta}) \int_{t_0}^t \varepsilon(s)\, ds \times \|f\|_{2,\infty}$$
$$\leq C\varepsilon^{-a} \times \delta^{-2(\nu+r+a)} \times \delta^{2-\nu} \times \|f\|_{2,\infty}$$
$$= C\delta^{2-3\nu-2r-3a} \times \|f\|_{2,\infty}.$$

Here, we have used (8.26), (8.28) and (8.24). □

We can now provide a proof for Theorem 23.

PROOF OF THEOREM 23. We recall that $r = \frac{2-3\nu}{3+\kappa}$ and $\varepsilon = \delta^r$ and we write

$$\left|\mathbb{E}[f(V_t)] - \mathbb{E}[f(U^{\delta}(V_0)]\right| \leq A + B \qquad \text{where}$$
$$A = \left|\mathbb{E}[f(V_t)] - \mathbb{E}[f(V_t^{\varepsilon}(V_0)]\right|, \qquad B = \left|\mathbb{E}[f(U_t^{\delta}(V_0))] - \mathbb{E}[f(V_t^{\varepsilon}(V_0)]\right|.$$

By (7.6),

$$A \leq \varepsilon^{-a} \times \varepsilon^{1+\kappa} \|f\|_{1,\infty} \leq \delta^{r(1+\kappa)-a} \|f\|_{1,\infty}.$$

Since $(2-3\nu)/(1-\nu) \leq 3 \leq (3+\kappa)/(1-\kappa)$, it follows that $r \leq (1-\nu)/(1-\kappa)$ and so we may use (8.25) and we obtain

$$B \leq \int_{\mathbb{R}^2} \left|\mathbb{E}[f(V_t^{\varepsilon}(v)] - \mathbb{E}[f(U^{\delta}(v)]\right| f_0(dv)$$
$$\leq C \int_{\mathbb{R}^2} (1 + |v|^2) f_0(dv) \times \delta^{2-3\nu-2r-3a} \|f\|_{2,\infty}.$$

We conclude that

$$\left|\mathbb{E}[f(V_t)] - \mathbb{E}[f(U^{\delta}(V_0)]\right| \leq C\|f\|_{2,\infty}(\delta^{r(1+\kappa)-a} + \delta^{2-3\nu-2r-3a})$$
$$\leq C\|f\|_{2,\infty}\delta^{\frac{(2-3\nu)(1+\kappa)}{3+\kappa}-3a}$$

the last inequality being a consequence of the choice of r. □

8.4.2. *Proof of the second-order estimates in the Boltzmann equation*. We begin with giving some useful estimates for the noise coefficient σ_δ.

LEMMA 33. 1. *Let $q \in \mathbb{N}^*$, $r \leq (1-\nu/2)/(q-1-\kappa/2)$ and $\varepsilon = \delta^r$. Then the following inequality holds true $\|\sigma_\delta\|_{1,q,(\mu,\infty)} \leq C$.*

2. *$\widehat{\mathcal{L}}_t^{\delta}$ verifies $H_3(2,3)$.*[4]

[4]See (5.5).

3. *Let us assume that* $\kappa \le 1/18$ *and*

$$r \le \frac{1-\nu}{2-\kappa} \wedge \frac{1-\nu/2}{2-\kappa/2}. \tag{8.31}$$

Then

$$\left\| \frac{1}{\psi_3}\big(P^\varepsilon_{t_0,t} f - \widehat{P}^\delta_{t_0,t}\big) f \right\|_\infty \le C\delta^{3-4\nu-3r-a} \times \|f\|_{3,\infty}. \tag{8.32}$$

PROOF. 1. Using (8.20) (with $\frac{\kappa}{2}$ instead of κ), we get

$$\big|\partial^\alpha \sigma(t,\theta,\rho,v)\big| = \big|\partial^\alpha\big(c(t,\theta,\rho,v)\gamma^{\frac{1}{2}}(t,\rho,v)\big)\big| \le C|\theta|\big(\varepsilon^{1+\frac{\kappa}{2}-|\alpha|} + \Gamma_\varepsilon^{\frac{\kappa}{2}}\big)$$

which gives, for $1 \le |\alpha| \le q$,

$$\int_{E_\delta^c} \big|\partial^\alpha \sigma(t,\theta,\rho,v)\big|^2 \mu(d\theta,d\rho) \le C\delta^{2-\nu}\big(\varepsilon^{2+\kappa-2|\alpha|} + \Gamma_\varepsilon^\kappa\big)$$
$$= C\delta^{2-\nu-r(2|\alpha|-2-\kappa)} \le C$$

the last inequality being true if $r \le (2-\nu)/(2q-2-\kappa)$.

2. We only check that

$$\left\| \frac{1}{\psi_2} \nabla\big(a_\delta^{i,j}(t,\circ)\partial^i\partial^j f\right\|_\infty \le C\Gamma_\varepsilon^\kappa \|f\|_{3,\infty}. \tag{8.33}$$

(The remaining estimates are similar to step 3 in the proof of Lemma 32.) One has

$$a_\delta^{i,j}(t,v) = \int_{\{|\theta|\le\delta\}} \mu(d\theta,d\rho)\big(c^i c^j\big)(t,\theta,\rho,v)\gamma(t,\rho,v)$$

so (8.33) follows from (8.21) with $\beta = 1$.

3. We will use Theorem 16 with $q = k = 3$. Using the first assertion and (8.24) (see Remark 31), we get that both $\widehat{\mathcal{P}}_t^\delta$ and $\mathcal{P}_t^\varepsilon$ verify $H_3(2,3)$ with $Q_q(t,\widehat{\mathcal{P}}^\delta) \le C \times \delta^{-q(\nu+r+a)}$. And we recall that in (8.26) we have proved that $C_3(T,\mathcal{P}^\varepsilon) = \varepsilon^{-a}$. It remains to estimate

$$\big(\mathcal{L}_t^\varepsilon - \widehat{\mathcal{L}}_t^\delta\big) f(v) = \int_{E_\delta^c} \mu(d\theta,d\rho)\gamma(t,\rho,v)\Bigg(f\big(v + c(t,\theta,\rho,v)\big) - f(v)$$
$$- \big\langle \nabla f(v), c(t,\theta,\rho,v)\big\rangle - \frac{1}{2}\sum_{i,j=1}^d c^i c^j(t,\theta,\rho,v)\partial^2_{ij} f(v)\Bigg).$$

Taylor's formula gives

$$\big|\big(\mathcal{L}_t^\varepsilon - \widehat{\mathcal{L}}_t^\delta\big) f(v)\big| \le C\int_{E_\delta^c} \mu(d\theta,dv)\big|c(t,\theta,\rho,v)\big|^3 \times \|f\|_{3,\infty}\Gamma_\varepsilon^\kappa$$
$$\le C\delta^{3-\nu}\int_0^1 \big|v - v_t(\rho)\big|^3 d\rho \times \|f\|_{3,\infty}\Gamma_\varepsilon^\kappa$$
$$\le C\psi_3(v)\delta^{3-\nu}\|f\|_{3,\infty}\Gamma_\varepsilon^\kappa$$

so that (5.6) holds with $\varepsilon(t) = \delta^{3-\nu}\varepsilon^{-a}$. We use (5.7) with $q = k = 3$ to get

$$\begin{aligned}\left\|\frac{1}{\psi_3}(P^{\varepsilon}_{t_0,t} f - \widehat{P}^{\delta}_{t_0,t})f\right\|_\infty &\le C \times C_3(P^\varepsilon)Q_3(\widehat{P}^\delta)\int_{t_0}^t \varepsilon(s)\,ds \times \|f\|_{3,\infty}\\ &\le C\varepsilon^{-a} \times \delta^{-3(\nu+r)} \times \delta^{3-\nu} \times \|f\|_{3,\infty}\\ &= C\delta^{3-4\nu-3r-a} \times \|f\|_{3,\infty}.\end{aligned}$$
□

PROOF OF THEOREM 24. We proceed as in the first-order case by combining the two errors by taking r such that $r(1+\kappa) = 3 - 4\nu - 3r$ which amounts to $r = \frac{3-4\nu}{4+\kappa}$. But we need (8.31) to hold true so we ask (7.12) to hold true. □

REFERENCES

[1] ALFONSI, A., CANCÈS, E., TURINICI, G., DI VENTURA, B. and HUISINGA, W. (2004). Exact simulation of hybrid stochastic and deterministic models for biochemical systems. Research Report No. RR-5435, INRIA. Available at https://hal.inria.fr/inria-00070572.

[2] ALFONSI, A., CANCÈS, E., TURINICI, G., DI VENTURA, B. and HUISINGA, W. (2005). Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems. In *CEMRACS* 2004—*Mathematics and Applications to Biology and Medicine*. *ESAIM Proc*. **14** 1–13. EDP Sci., Les Ulis. MR2226797

[3] ANGIUS, A., BALBO, G., BECCUTI, M., BIBBONA, E., HORVATH, A. and SIROVICH, R. (2015). Approximate analysis of biological systems by hybrid switching jump diffusion. *Theoret*. *Comput*. *Sci*. **587** 49–72. MR3354838

[4] ASMUSSEN, S. and ROSIŃSKI, J. (2001). Approximations of small jumps of Lévy processes with a view towards simulation. *J*. *Appl*. *Probab*. **38** 482–493. MR1834755

[5] BALL, K., KURTZ, T. G., POPOVIC, L. and REMPALA, G. (2006). Asymptotic analysis of multiscale approximations to reaction networks. *Ann*. *Appl*. *Probab*. **16** 1925–1961. MR2288709

[6] BALLY, V. and FOURNIER, N. (2011). Regularization properties of the 2D homogeneous Boltzmann equation without cutoff. *Probab*. *Theory Related Fields* **151** 659–704. MR2851696

[7] BENAÏM, M., BOUGUET, F. and CLOEZ, B. (2017). Ergodicity of inhomogeneous Markov chains through asymptotic pseudotrajectories. *Ann*. *Appl*. *Probab*. **27** 3004–3049. MR3719952

[8] BENAÏM, M. and HIRSCH, M. W. (1996). Asymptotic pseudotrajectories and chain recurrent flows, with applications. *J*. *Dynam*. *Differential Equations* **8** 141–176. MR1388167

[9] BOXMA, O., KASPI, H., KELLA, O. and PERRY, D. (2005). On/off storage systems with state-dependent input, output, and switching rates. *Probab*. *Engrg*. *Inform*. *Sci*. **19** 1–14. MR2104547

[10] BRESSLOFF, P. C. and NEWBY, J. M. (2013). Metastability in a stochastic neural network modeled as a velocity jump Markov process. *SIAM J*. *Appl*. *Dyn*. *Syst*. **12** 1394–1435. MR3085109

[11] BUCKWAR, E. and RIEDLER, M. G. (2011). An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *J*. *Math*. *Biol*. **63** 1051–1093. MR2855804

[12] COCOZZA-THIEVENT, C. (2017). Renouvellement Markovien et PDMP. Available at http://perso-math.univ-mlv.fr/users/cocozza.christiane/recherche-page-perso/PresentationRMetPDMP.html.

[13] CRUDU, A., DEBUSSCHE, A., MULLER, A. and RADULESCU, O. (2012). Convergence of stochastic gene networks to hybrid piecewise deterministic processes. *Ann. Appl. Probab.* **22** 1822–1859. MR3025682

[14] DAVIS, M. H. A. (1984). Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *J. Roy. Statist. Soc. Ser. B* **46** 353–388. MR0790622

[15] DAVIS, M. H. A. (1993). *Markov Models and Optimization. Monographs on Statistics and Applied Probability* **49**. Chapman & Hall, London. MR1283589

[16] DUFOUR, F., DUTUIT, Y., GONZALEZ, K. and ZHANG, H. (2008). Piecewise deterministic Markov processes and dynamic reliability. *J. Risk Reliab.* **222** 545–551.

[17] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York. MR0838085

[18] FOURNIER, N. (2011). Simulation and approximation of Lévy-driven stochastic differential equations. *ESAIM Probab. Stat.* **15** 233–248. MR2870514

[19] FOURNIER, N. and GODINHO, D. (2012). Asymptotic of grazing collisions and particle approximation for the Kac equation without cutoff. *Comm. Math. Phys.* **316** 307–344. MR2993918

[20] FOURNIER, N. and MOUHOT, C. (2009). On the well-posedness of the spatially homogeneous Boltzmann equation with a moderate angular singularity. *Comm. Math. Phys.* **289** 803–824. MR2511651

[21] GANGULY, A., ALTINTAN, D. and KOEPPL, H. (2015). Jump-diffusion approximation of stochastic reaction dynamics: Error bounds and algorithms. *Multiscale Model. Simul.* **13** 1390–1419. MR3431122

[22] GILLESPIE, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22** 403–434. MR0503370

[23] GODINHO, D. (2013). Asymptotic of grazing collisions for the spatially homogeneous Boltzmann equation for soft and Coulomb potentials. *Stochastic Process. Appl.* **123** 3987–4039. MR3091097

[24] GRAHAM, C. (1992). McKean–Vlasov Itô–Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Process. Appl.* **40** 69–82. MR1145460

[25] HE, L. (2014). Asymptotic analysis of the spatially homogeneous Boltzmann equation: Grazing collisions limit. *J. Stat. Phys.* **155** 151–210. MR3180973

[26] IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. North-Holland, Amsterdam. MR1011252

[27] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877

[28] JAHNKE, T. and KREIM, M. (2012). Error bound for piecewise deterministic processes modeling stochastic reaction systems. *Multiscale Model. Simul.* **10** 1119–1147. MR3022034

[29] KOLOKOLTSOV, V. N. (2011). *Markov Processes, Semigroups and Generators. De Gruyter Studies in Mathematics* **38**. de Gruyter, Berlin. MR2780345

[30] KUNITA, H. (2004). Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In *Real and Stochastic Analysis*. 305–373. Birkhäuser, Boston, MA. MR2090755

[31] KURTZ, T. G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* **8** 344–356. MR0287609

[32] KURTZ, T. G. (1977/78). Strong approximation theorems for density dependent Markov chains. *Stochastic Process. Appl.* **6** 223–240. MR0464414

[33] LAPEYRE, B., PARDOUX, É. and SENTIS, R. (1998). *Méthodes de Monte-Carlo Pour les équations de Transport et de Diffusion. Mathématiques & Applications* (*Berlin*) [*Mathematics & Applications*] **29**. Springer, Berlin. MR1621249

[34] LÖCHERBACH, E. (2017). Convergence to equilibrium for time inhomogeneous jump diffusions with state dependent jump intensity. Preprint. Available at arXiv:1712.03507.

[35] PAZY, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. *Applied Mathematical Sciences* **44**. Springer, New York. MR0710486

[36] PROTTER, P. and TALAY, D. (1997). The Euler scheme for Lévy driven stochastic differential equations. *Ann*. *Probab.* **25** 393–423. MR1428514

[37] TANAKA, H. (1978/79). Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch*. *Verw. Gebiete* **46** 67–105. MR0512334

[38] WALSH, J. B. (1986). An introduction to stochastic partial differential equations. In *École D'été de Probabilités de Saint-Flour, XIV*—1984. *Lecture Notes in Math*. **1180** 265–439. Springer, Berlin. MR0876085

V. BALLY
D. GOREAC
V. RABIET
UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE
5 BOULEVARD DESCARTES, CITÉ DESCARTES
77420, CHAMPS-SUR-MARNE
FRANCE
E-MAIL: vlad.bally@u-pem.fr
dan.goreac@u-pem.fr
victor_rabiet@yahoo.fr