

# On the equivalence between some jumping SDEs with rough coefficients and some non-local PDEs

Nicolas Fournier and Liping Xu

Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, UPMC, Case 188, 4 pl. Jussieu, F-75252 Paris Cedex 5, France.

E-mail: [nicolas.fournier@upmc.fr](mailto:nicolas.fournier@upmc.fr); [liping.xu@upmc.fr](mailto:liping.xu@upmc.fr)

Received 26 December 2016; revised 1 May 2018; accepted 15 May 2018

**Abstract.** We study some jumping SDE and the corresponding Fokker–Planck (or Kolmogorov forward) equation, which is a non-local PDE. We assume only some measurability and growth conditions on the coefficients. We prove that for any weak solution  $(f_t)_{t \in [0, T]}$  of the PDE, there exists a weak solution to the SDE of which the time marginals are given by  $(f_t)_{t \in [0, T]}$ . As a corollary, we deduce that for any given initial condition, existence for the PDE is equivalent to weak existence for the SDE and uniqueness in law for the SDE implies uniqueness for the PDE. This extends some ideas of Figalli (*J. Funct. Anal.* **254** (2008) 109–153) concerning continuous SDEs and local PDEs.

**Résumé.** On étudie certaines EDS à sauts et les équations de Fokker–Planck (ou Kolmogorov progressives) correspondantes, qui sont des EDP non-locales. On suppose seulement que les coefficients sont mesurables et à croissance au plus linéaire. On montre que pour toute solution faible  $(f_t)_{t \in [0, T]}$  de l'EDP, il existe une solution faible à l'EDS, dont les lois marginales sont données par  $(f_t)_{t \in [0, T]}$ . On en déduit que pour toute donnée initiale, l'existence pour l'EDP est équivalente à l'existence faible pour l'EDS, et que l'unicité en loi pour l'EDS implique l'unicité pour l'EDP. Nous étendons ainsi des idées de Figalli (*J. Funct. Anal.* **254** (2008) 109–153) concernant des EDS continues et des EDP locales.

MSC: 60H10; 60J75; 40K05

Keywords: Existence and uniqueness; Weak solution; Jumping SDEs; Non-local PDEs

## 1. Introduction

We consider the  $d$ -dimensional stochastic differential equation posed on some time interval  $[0, T]$

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s + \int_0^t \int_E h(s, z, X_{s-}) N(ds, dz), \quad (1)$$

where  $(B_t)_{t \in [0, T]}$  is a  $d$ -dimensional Brownian motion and  $N(ds, dz)$  is a Poisson measure on  $[0, T] \times E$  with intensity measure  $ds \mu(dz)$ . The coefficients  $b : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ ,  $\sigma : [0, T] \times \mathbb{R}^d \mapsto \mathcal{S}_d^+$  and  $h : [0, T] \times E \times \mathbb{R}^d \mapsto \mathbb{R}^d$  are supposed to be at least measurable. The space  $E$  is endowed with a  $\sigma$ -field  $\mathcal{E}$  and with a  $\sigma$ -finite measure  $\mu$  and  $\mathcal{S}_d^+$  is the set of nonnegative symmetric  $d \times d$  real matrices. The Fokker–Planck (or Kolmogorov forward) equation associated to (1) is

$$\partial_t f_t + \operatorname{div}(b(t, \cdot) f_t) = \frac{1}{2} \sum_{i, j=1}^d \partial_{ij}([\sigma(t, \cdot) \sigma^*(t, \cdot)]_{i, j} f_t) + \mathcal{L}_t f_t, \quad (2)$$

where  $\mathcal{L}_t f_t : \mathbb{R}^d \mapsto \mathbb{R}$  is defined by

$$\int_{\mathbb{R}^d} (\mathcal{L}_t f_t)(x) \varphi(x) dx = \int_{\mathbb{R}^d} \int_E [\varphi(x + h(t, z, x)) - \varphi(x)] f_t(x) \mu(dz) dx$$

for any reasonable  $\varphi : \mathbb{R}^d \mapsto \mathbb{R}$ . We use the notation  $\nabla = \nabla_x$ ,  $\text{div} = \text{div}_x$  and  $\partial_{ij} = \partial_{x_i x_j}^2$ .

Let  $\mathcal{P}(\mathbb{R}^d)$  be the set of probability measures on  $\mathbb{R}^d$  and

$$\mathcal{P}_1(\mathbb{R}^d) = \{f \in \mathcal{P}(\mathbb{R}^d) : m_1(f) < \infty\} \quad \text{with } m_1(f) := \int_{\mathbb{R}^d} |x| f(dx).$$

We define  $L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$  as the set of all measurable families  $(f_t)_{t \in [0, T]}$  of probability measures on  $\mathbb{R}^d$  such that  $\sup_{[0, T]} m_1(f_t) < \infty$ .

1.1. *Main result*

We will suppose the following conditions.

**Assumption 1.1.** The functions  $\sigma : [0, T] \times \mathbb{R}^d \mapsto \mathcal{S}_d^+$ ,  $b : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$  and  $h : [0, T] \times E \times \mathbb{R}^d \mapsto \mathbb{R}^d$  are measurable and there is a constant  $C$  such that for all  $(t, x) \in [0, T] \times \mathbb{R}^d$ ,

$$|\sigma(t, x)| + |b(t, x)| + \int_E |h(t, z, x)| \mu(dz) \leq C(1 + |x|).$$

We set  $a(t, x) = \sigma(t, x) \sigma^*(t, x)$ , which satisfies  $|a(t, x)| \leq C(1 + |x|^2)$ .

**Definition 1.2.** Suppose Assumption 1.1. A measurable family  $(f_t)_{t \in [0, T]}$  of probability measures on  $\mathbb{R}^d$  is called a weak solution to (2) if for all  $\varphi \in C_c^2(\mathbb{R}^d)$ , all  $t \in [0, T]$ ,

$$\int_{\mathbb{R}^d} \varphi(x) f_t(dx) = \int_{\mathbb{R}^d} \varphi(x) f_0(dx) + \int_0^t \int_{\mathbb{R}^d} [\mathcal{A}_s \varphi(x) + \mathcal{B}_s \varphi(x)] f_s(dx) ds, \tag{3}$$

with the diffusion operator  $\mathcal{A}_s \varphi(x) := b(s, x) \cdot \nabla \varphi(x) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}(s, x) \partial_{ij} \varphi(x)$  and the jump operator  $\mathcal{B}_s \varphi(x) := \int_E [\varphi(x + h(s, z, x)) - \varphi(x)] \mu(dz)$ .

We will check the following facts in the [Appendix](#), implying in particular that (3) makes sense.

**Remark 1.3.** Suppose Assumption 1.1.

- (i) For  $\varphi \in C_c^2(\mathbb{R}^d)$ ,  $\sup_{[0, T] \times \mathbb{R}^d} (|\mathcal{A}_s \varphi(x)| + |\mathcal{B}_s \varphi(x)|) < \infty$ .
- (ii) Any weak solution  $(f_t)_{t \in [0, T]}$  to (2) starting from  $f_0 \in \mathcal{P}_1(\mathbb{R}^d)$  belongs to  $L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$ .
- (iii) If  $f_0 \in \mathcal{P}_1(\mathbb{R}^d)$ , the weak formulation (3) automatically extends to all functions  $\varphi \in C^2(\mathbb{R}^d)$  such that  $(1 + |x|)[|\varphi(x)| + |\nabla \varphi(x)| + |D^2 \varphi(x)|]$  is bounded.

Point (iii) is far from optimal, but sufficient for our purpose. Our main result reads as follows.

**Theorem 1.4.** *Suppose Assumption 1.1 and consider any weak solution  $(f_t)_{t \in [0, T]}$  to (2) such that  $f_0 \in \mathcal{P}_1(\mathbb{R}^d)$ . There exist, on some probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ , a  $d$ -dimensional  $(\mathcal{F}_t)_{t \in [0, T]}$ -Brownian motion  $(B_t)_{t \in [0, T]}$ , a  $(\mathcal{F}_t)_{t \in [0, T]}$ -Poisson measure  $N(dt, dz)$  on  $[0, T] \times E$  with intensity measure  $dt \mu(dz)$ , these two objects being independent, as well as a càdlàg  $(\mathcal{F}_t)_{t \in [0, T]}$ -adapted process  $(X_t)_{t \in [0, T]}$  solving (1) and such that  $\mathcal{L}(X_t) = f_t$  for all  $t \in [0, T]$ .*

For  $(X_t)_{t \in [0, T]}$  a solution to (1) and for  $f_t = \mathcal{L}(X_t)$ , a simple application of the Itô formula (to compute  $\int_{\mathbb{R}^d} \varphi(x) f_t(dx) = \mathbb{E}[\varphi(X_t)]$  with  $\varphi \in C_c^2(\mathbb{R}^d)$ ) shows that the family  $(f_t)_{t \in [0, T]}$  is a weak solution to (2). The following corollary is thus immediately deduced from Theorem 1.4.

**Corollary 1.5.** *Suppose Assumption 1.1 and fix  $f_0 \in \mathcal{P}_1(\mathbb{R}^d)$ .*

- (i) *The existence of a (weak) solution  $(X_t)_{t \in [0, T]}$  to (1) such that  $\mathcal{L}(X_0) = f_0$  is equivalent to the existence of a weak solution  $(f_t)_{t \in [0, T]}$  to (2) starting from  $f_0$ .*
- (ii) *The uniqueness (in law) of the solution  $(X_t)_{t \in [0, T]}$  to (1) with  $\mathcal{L}(X_0) = f_0$  implies the uniqueness of the weak solution  $(f_t)_{t \in [0, T]}$  to (2) starting from  $f_0$ .*

In almost all models arising from applied sciences, the jump operator is given under the form  $\mathcal{B}_s \varphi(x) = \int_F [\varphi(x + g(s, y, x)) - \varphi(x)] \kappa(s, y, x) \nu(dy)$ , meaning that when in the position  $x$  at time  $s$ , the process jumps to  $x + g(s, y, x)$  at rate  $\kappa(s, y, x) \nu(dy)$ . Here  $F$  is a measurable space endowed with a  $\sigma$ -finite measure  $\nu$  and we have two measurable functions  $g : [0, T] \times F \times \mathbb{R}^d \mapsto \mathbb{R}^d$  and  $\kappa : [0, T] \times F \times \mathbb{R}^d \mapsto \mathbb{R}_+$ . Introducing  $E = F \times \mathbb{R}_+$ ,  $\mu(dy, du) = \nu(dy) du$  and  $h(s, (y, u), x) = g(s, y, x) \mathbf{1}_{\{u \leq \kappa(s, y, x)\}}$ , one easily verifies that  $\mathcal{B}_s \varphi(x) = \int_E [\varphi(x + h(s, (y, u), x)) - \varphi(x)] \mu(dy, du)$ . Our results thus apply if  $\int_F |g(s, y, x)| \kappa(s, y, x) \nu(dy) \leq C(1 + |x|)$ .

### 1.2. Motivation

Stochastic differential equations with jumps are now playing an important role in modeling and applied sciences. We refer to the book of Situ [11] for all basic results and a lot of possible applications. The book of Jacod [8] contains many important results about weak and strong existence and uniqueness, relations between SDEs and martingale problems, etc. See also the survey paper of Bass [2].

Existence for PDEs is often more developed than for SDEs, so Theorem 1.4 might be useful to derive some new weak existence results for the SDE (1).

Our main motivation is the uniqueness for some nonlinear PDEs, for which the use of nonlinear (in the sense of McKean) SDEs has proved to be a powerful tool. For example, the first (partial) uniqueness result concerning the homogeneous Boltzmann for long range interactions was derived by Tanaka [13]. He was studying the simplest case of Maxwell molecules. Unfortunately, he was only able to prove the uniqueness in law of the nonlinear SDE associated to the Boltzmann equation. Horowitz and Karandikar [7] were able to deduce the uniqueness for the (same) Boltzmann equation proceeding as follows. Let us recall that the original equation writes  $\partial_t f_t = Q(f_t, f_t)$ , for some quadratic nonlocal operator  $Q$ . For  $f$  a solution, they consider the linear PDE  $\partial_t g_t = Q(g_t, f_t)$ , with unknown  $g$  satisfying  $g_0 = f_0$ . They prove uniqueness in law for the (linear) SDE associated to this PDE (for any initial condition). They deduce, extending some results of Ethier and Kurtz [4, Chapter 4, Propositions 9.18 and 9.19], the uniqueness for the linear PDE (for any initial condition). So the unique solution (with  $g_0 = f_0$ ) to  $\partial_t g_t = Q(g_t, f_t)$  is  $f$  itself. Consequently, the time marginals of the solution  $X$  to the linear SDE (when  $X_0 \sim f_0$ ), which solve  $\partial_t g_t = Q(g_t, f_t)$  are necessarily  $(f_t)_{t \in [0, T]}$ . Thus  $X$  actually solves the nonlinear SDE. Since uniqueness in law holds for the nonlinear SDE by Tanaka [13], they deduce that there is at most one solution to the Boltzmann equation  $\partial_t f_t = Q(f_t, f_t)$ , for some given reasonable initial condition  $f_0$ .

Let us recall that the above mentioned results of Ethier and Kurtz (extended by Horowitz and Karandikar [7, Theorem B1] and by Bhatt and Karandikar [3, e.g. Theorems 4.1 and 5.2]) state in spirit that if some SDE has a unique solution (in law) for any deterministic initial condition, then the corresponding PDE has a unique weak solution for any reasonable initial condition.

Our result is much stronger, since it does not require at all uniqueness for (1). If, for example, studying the Boltzmann equation, it directly implies that, to any solution  $f$  to the nonlinear equation (seen here as a solution to the linear equation  $\partial_t g_t = Q(g_t, f_t)$ ), we can associate a solution  $X$  to the corresponding linear SDE with additionally  $X_t \sim f_t$  for all  $t$ . In other words,  $X$  solves the nonlinear SDE. This might look anodyne, but this was crucial when studying more singular nonlinear equations, such as the Landau or Boltzmann equations for moderately soft potentials, see [6] and [14]. Indeed, in such cases, we really need to use some physical symmetries to prove uniqueness: it is absolutely not clear that uniqueness holds for the linear PDE  $\partial_t g_t = Q(g_t, f_t)$ , since one really uses that the two arguments of  $Q$  are the same.

We hope the above discussion shows that Theorem 1.4 is an interesting variation of the mentioned results of Ethier and Kurtz [4]. As already said, the method we use was initiated by Figalli [5] for continuous SDEs ( $h = 0$ ) with bounded coefficients. The boundedness assumption was relaxed in [6, Appendix B]. A special jumping SDE (with  $a = b = 0$  and a special jump operator) was considered in [14] to study a singular homogeneous Boltzmann equation.

We decided to write down the general case in the present paper. We did not want to assume some boundedness of the coefficients, although it complicates the proofs without introducing new deep ideas, because it is very useful for practical purposes.

Finally, as explained in the next subsection, we are not able to prove a general result when the jump part of the SDE has infinite variations, and this is a rather important limitation.

### 1.3. Strategy of the proof and plan of the paper

At many places, the situation is technically more involved, but the global strategy is exactly the same as that introduced by Figalli [5, Theorem 2.6]. Let  $(f_t)_{t \in [0, T]}$  be a given weak solution to (2).

- I. In Section 2, we introduce  $f_t^\varepsilon = f_t \star \phi_\varepsilon$ , where  $\phi_\varepsilon$  is the centered Gaussian density with covariance matrix  $\varepsilon I_d$ . We compute the PDE satisfied by  $f_t^\varepsilon$ : we find that  $\partial_t f_t^\varepsilon + \operatorname{div}(b^\varepsilon(t, \cdot) f_t^\varepsilon) = \frac{1}{2} \sum_{i,j} \partial_{i,j} (a_{i,j}^\varepsilon(t, \cdot) f_t^\varepsilon) + \mathcal{L}_t^\varepsilon f_t^\varepsilon$ , for some coefficients  $a^\varepsilon, b^\varepsilon$  and some jump operator  $\mathcal{L}_t^\varepsilon$ . Let us mention that  $a^\varepsilon(t, \cdot), b^\varepsilon(t, \cdot)$  and  $\mathcal{L}_t^\varepsilon$  of course depend on  $f_t$ .
- II. Still in Section 2, we prove that  $a^\varepsilon, b^\varepsilon$  and the coefficient of the jump operator  $\mathcal{L}^\varepsilon$  satisfy
  - (i) the same linear growth conditions as  $a, b, \mathcal{L}$ , uniformly in  $\varepsilon \in (0, 1)$ ,
  - (ii) some (non-uniform) local Lipschitz conditions.
- III. In Section 3, we use II to build, for each  $\varepsilon \in (0, 1)$ , a solution  $(X_t^\varepsilon)_{t \in [0, T]}$  to some SDE of which the Fokker–Planck equation is the PDE satisfied by  $(f_t^\varepsilon)_{t \in [0, T]}$ . Since both the SDE and the PDE (with  $\varepsilon \in (0, 1)$  fixed) are well-posed (because the coefficients are regular enough), we conclude that  $\mathcal{L}(X_t^\varepsilon) = f_t^\varepsilon$ . Indeed, the time marginals of  $(X_t^\varepsilon)_{t \in [0, T]}$  satisfy the same PDE as  $(f_t^\varepsilon)_{t \in [0, T]}$ .
- IV. Still in Section 3, we prove that the family  $\{(X_t^\varepsilon)_{t \in [0, T]}, \varepsilon \in (0, 1)\}$  is tight. This is rather easy from the Aldous criterion [1], using only II-(ii).
- V. In Section 4, we finally consider a limit point  $(X_t)_{t \in [0, T]}$ , as  $\varepsilon \rightarrow 0$ , of  $\{(X_t^\varepsilon)_{t \in [0, T]}, \varepsilon \in (0, 1)\}$ . Since  $\mathcal{L}(X_t^\varepsilon) = f_t^\varepsilon$  by III, we deduce that  $\mathcal{L}(X_t) = f_t$  for each  $t \in [0, T]$ . It then remains to show that  $(X_t)_{t \in [0, T]}$  is a weak solution to (1) and we classically make use of martingale problems. Since the coefficients  $a, b, h$  are possibly rough, we have to approximate them by some continuous (in  $x$ ) coefficients  $\tilde{a}, \tilde{b}, \tilde{h}$ . We use that we already know the time marginals of  $(X_t)_{t \in [0, T]}$ : we can take  $\tilde{a}(t, \cdot), \tilde{b}(t, \cdot)$  and  $\tilde{h}(t, \cdot, z)$  close to  $a(t, \cdot), b(t, \cdot)$  and  $h(t, \cdot, z)$  in  $L^1(f_t)$ .

The proof of Remark 1.3 is written in the [Appendix](#).

To conclude this paragraph, let us mention a few difficulties. The regularized jump operator, in its weak form writes  $\int_{\mathbb{R}^d} \mathcal{L}_t^\varepsilon f_t^\varepsilon(y) \varphi(y) dy = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_E [\varphi(y + h(t, z, x)) - \varphi(y)] \phi_\varepsilon(x - y) f_t^\varepsilon(dx) dy$ . We found no *regular* Poisson representation of the associated SDE. We use an indicator function, see (4). This is why we are not able to treat the case of an infinite variation jump term: we do not know how to prove that a SDE like (4), with a compensated Poisson measure and some weaker condition on  $h$  (something like  $\int_E |h(s, z, x)|^2 \mu(dz) \leq C(1 + |x|^2)$ ), is well-posed.

Although this should be classical since the coefficients are rather regular for  $\varepsilon \in (0, 1)$  fixed, we found no reference about the uniqueness for the PDE satisfied by  $(f_t^\varepsilon)_{t \in [0, T]}$  (see Lemma 2.1). We have not been able to write down a deterministic proof. We thus use that the corresponding SDE is well-posed (for any deterministic initial condition) and we apply a result of Horowitz and Karandikar [7].

### 1.4. Convention

During the whole paper, we always suppose Assumption 1.1 and that  $f_0 \in \mathcal{P}_1(\mathbb{R}^d)$ . We use the generic notation  $C$  for a positive finite constant, of which the value may change from line to line. It is allowed to depend only on the dimension  $d$ , on the parameters  $a, b, h, E, \mu, T$  of our equations, and on the weak solution  $(f_t)_{t \in [0, T]}$  to (2) under study. When a constant depends on another parameter, we indicate it in subscript. For example,  $C_\varepsilon$  is a constant allowed to depend only on  $a, b, h, E, \mu, T, (f_t)_{t \in [0, T]}$  and on  $\varepsilon$ .

## 2. Regularization

We introduce the following regularization procedure, as Figalli in [5], see also [14].

**Lemma 2.1.** For  $(f_t)_{t \in [0, T]} \in L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$  a weak solution to (2) and  $\varepsilon \in (0, 1)$ , we set

$$f_t^\varepsilon(y) := \int_{\mathbb{R}^d} \phi_\varepsilon(x - y) f_t(dx) = (f_t \star \phi_\varepsilon)(y) \quad \text{with } \phi_\varepsilon(x) = (2\pi\varepsilon)^{-d/2} e^{-|x|^2/(2\varepsilon)}.$$

Then for any test function  $\psi \in C_c^2(\mathbb{R}^d)$ , any  $t \in [0, T]$ ,

$$\int_{\mathbb{R}^d} \psi(y) f_t^\varepsilon(y) dy = \int_{\mathbb{R}^d} \psi(y) f_0^\varepsilon(y) dy + \int_0^t \int_{\mathbb{R}^d} [\mathcal{A}_{s,\varepsilon} \psi(y) + \mathcal{B}_{s,\varepsilon} \psi(y)] f_s^\varepsilon(y) dy ds,$$

with

$$\mathcal{A}_{t,\varepsilon} \psi(y) = b^\varepsilon(t, y) \cdot \nabla \psi(y) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}^\varepsilon(t, y) \partial_{ij} \psi(y),$$

$$\mathcal{B}_{t,\varepsilon} \psi(y) = \int_E \int_{\mathbb{R}^d} [\psi(y + h(t, z, x)) - \psi(y)] F_t^\varepsilon(x, y) f_t(dx) \mu(dz),$$

where

$$a^\varepsilon(t, y) = \frac{\int_{\mathbb{R}^d} \phi_\varepsilon(x - y) a(t, x) f_t(dx)}{f_t^\varepsilon(y)}, \quad b^\varepsilon(t, y) = \frac{\int_{\mathbb{R}^d} \phi_\varepsilon(x - y) b(t, x) f_t(dx)}{f_t^\varepsilon(y)},$$

$$F_t^\varepsilon(x, y) := \frac{\phi_\varepsilon(x - y)}{f_t^\varepsilon(y)}.$$

**Proof.** It is obvious that  $f_t^\varepsilon(y) > 0$  for each  $(t, y) \in [0, T] \times \mathbb{R}^d$ . We first apply (3) with the choice  $\varphi(x) = \phi_\varepsilon(x - y)$  (with some fixed  $y \in \mathbb{R}^d$ ), which is licit by Remark 1.3-(iii). We then integrate the obtained equality against  $\psi \in C_c^2(\mathbb{R}^d)$ . This gives

$$\int_{\mathbb{R}^d} \psi(y) f_t^\varepsilon(y) dy = \int_{\mathbb{R}^d} \psi(y) f_0^\varepsilon(y) dy + \int_0^t (I_s + J_s) ds,$$

where

$$I_t := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \psi(y) \mathcal{A}_t \phi_\varepsilon(x - y) f_t(dx) dy \quad \text{and} \quad J_t := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \psi(y) \mathcal{B}_t \phi_\varepsilon(x - y) f_t(dx) dy.$$

First,

$$I_t = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \psi(y) b(t, x) \cdot \nabla \phi_\varepsilon(x - y) f_t(dx) dy + \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \sum_{i,j=1}^d \psi(y) a_{ij}(t, x) \partial_{ij} \phi_\varepsilon(x - y) f_t(dx) dy.$$

But we have  $\int_{\mathbb{R}^d} \psi(y) \nabla \phi_\varepsilon(x - y) dy = \int_{\mathbb{R}^d} \phi_\varepsilon(x - y) \nabla \psi(y) dy$  as well as  $\int_{\mathbb{R}^d} \psi(y) \partial_{ij} \phi_\varepsilon(x - y) dy = \int_{\mathbb{R}^d} \phi_\varepsilon(x - y) \partial_{ij} \psi(y) dy$ , so that

$$\begin{aligned} I_t &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi_\varepsilon(x - y) b(t, x) \cdot \nabla \psi(y) f_t(dx) dy + \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \sum_{i,j=1}^d a_{ij}(t, x) \phi_\varepsilon(x - y) \partial_{ij} \psi(y) f_t(dx) dy \\ &= \int_{\mathbb{R}^d} b^\varepsilon(t, y) \cdot \nabla \psi(y) f_t^\varepsilon(y) dy + \frac{1}{2} \int_{\mathbb{R}^d} \sum_{i,j=1}^d a_{ij}^\varepsilon(t, y) \partial_{ij} \psi(y) f_t^\varepsilon(y) dy \\ &= \int_{\mathbb{R}^d} \mathcal{A}_{t,\varepsilon} \psi(y) f_t^\varepsilon(y) dy \end{aligned}$$

as desired. For the jump term, we use a similar computation as in [14, Proposition 3.1]. Since  $\mu$  is  $\sigma$ -finite, there exists a non-decreasing sequence  $(E_n)_{n \geq 1} \subset E$  such that  $\bigcup_{n=1}^\infty E_n = E$  and  $\mu(E_n) < \infty$  for each  $n \geq 1$ . We fix  $n$  and write

$$J_t = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{E_n} \psi(y) \phi_\varepsilon(x - y + h(t, z, x)) \mu(dz) f_t(dx) dy - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{E_n} \psi(y) \phi_\varepsilon(x - y) \mu(dz) f_t(dx) dy + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{E \setminus E_n} \psi(y) [\phi_\varepsilon(x - y + h(t, z, x)) - \phi_\varepsilon(x - y)] \mu(dz) f_t(dx) dy.$$

Using the change of variables  $y - h(t, z, x) \mapsto y$ , we see that

$$\int_{\mathbb{R}^d} \psi(y) \phi_\varepsilon(x - y + h(t, z, x)) dy = \int_{\mathbb{R}^d} \psi(y + h(t, z, x)) \phi_\varepsilon(x - y) dy,$$

and consequently,

$$J_t = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{E_n} [\psi(y + h(t, z, x)) - \psi(y)] \phi_\varepsilon(x - y) \mu(dz) f_t(dx) dy + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{E \setminus E_n} \psi(y) [\phi_\varepsilon(x - y + h(t, z, x)) - \phi_\varepsilon(x - y)] \mu(dz) f_t(dx) dy.$$

Observe now that  $|\psi(y + h(t, z, x)) - \psi(y)| \phi_\varepsilon(x - y) \leq C|h(t, z, x)| \phi_\varepsilon(x - y) \in L^1(\mu(dz) f_t(dx) dy)$  and  $|\psi(y) [\phi_\varepsilon(x - y + h(t, z, x)) - \phi_\varepsilon(x - y)]| \leq C_\varepsilon |\psi(y)| |h(t, z, x)| \in L^1(\mu(dz) f_t(dx) dy)$ : this uses that  $\psi \in C_c^2(\mathbb{R}^d)$ , Assumption 1.1 and that  $f_t \in \mathcal{P}_1(\mathbb{R}^d)$ . We thus can let  $n \rightarrow \infty$ :

$$J_t = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_E [\psi(y + h(t, z, x)) - \psi(y)] \phi_\varepsilon(x - y) \mu(dz) f_t(dx) dy = \int_{\mathbb{R}^d} \mathcal{B}_{t,\varepsilon} \psi(y) f_t^\varepsilon(y) dy,$$

which completes the proof. □

Let us now give some growth and regularity estimates on the regularized coefficients.

**Lemma 2.2.** *Let  $(f_t)_{t \in [0, T]} \in L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$  be a weak solution to (2) and recall that  $a^\varepsilon, b^\varepsilon, F^\varepsilon$  were introduced in Lemma 2.1.*

(i) *There exists a constant  $C > 0$  such that for all  $\varepsilon \in (0, 1)$ , all  $y \in \mathbb{R}^d$ , all  $t \in [0, T]$ ,*

$$|b^\varepsilon(t, y)| + |a^\varepsilon(t, y)|^{1/2} + \int_{\mathbb{R}^d} \int_E |h(t, z, x)| F_t^\varepsilon(x, y) \mu(dz) f_t(dx) \leq C(1 + |y|).$$

(ii) *For all  $\varepsilon \in (0, 1)$  and  $R > 0$ , there is  $C_{R,\varepsilon} > 0$  such that for all  $y_1, y_2 \in B(0, R)$ , all  $t \in [0, T]$ ,*

$$|b^\varepsilon(t, y_1) - b^\varepsilon(t, y_2)| + |a^\varepsilon(t, y_1) - a^\varepsilon(t, y_2)| + |[a^\varepsilon(t, y_1)]^{1/2} - [a^\varepsilon(t, y_2)]^{1/2}| + \int_{\mathbb{R}^d} \int_E |h(t, z, x)| |F_t^\varepsilon(x, y_1) - F_t^\varepsilon(x, y_2)| \mu(dz) f_t(dx) \leq C_{R,\varepsilon} |y_1 - y_2|.$$

**Proof.** We start with (i). By Assumption 1.1,

$$\begin{aligned} & |b^\varepsilon(t, y)| + |a^\varepsilon(t, y)|^{1/2} + \int_{\mathbb{R}^d} \int_E |h(t, z, x)| F_t^\varepsilon(x, y) \mu(dz) f_t(dx) \\ & \leq C \frac{\int_{\mathbb{R}^d} \phi_\varepsilon(x - y) (1 + |x|) f_t(dx)}{f_t^\varepsilon(y)} + C \left[ \frac{\int_{\mathbb{R}^d} \phi_\varepsilon(x - y) (1 + |x|)^2 f_t(dx)}{f_t^\varepsilon(y)} \right]^{1/2} \\ & =: CI_\varepsilon(t, y) + CJ_\varepsilon(t, y). \end{aligned}$$

Since for  $y$  fixed,  $[f_t^\varepsilon(y)]^{-1}\phi_\varepsilon(x-y)f_t(dx)$  is a probability measure, we infer from Cauchy–Schwarz that  $I_\varepsilon(t, y) \leq J_\varepsilon(t, y)$ . We thus only have to prove that  $[J_\varepsilon(t, y)]^2 \leq C(1 + |y|^2)$ . Let  $L := 2 \sup_{[0, T]} m_1(f_t) + 2$ . We use that

$$1 + |x| \leq 1 + |y| + |x - y| \leq 1 + 2|y| + L + |x - y|\mathbf{1}_{\{|x-y|>|y|+L\}}$$

to write

$$\begin{aligned} [J_\varepsilon(t, y)]^2 &\leq 2 \frac{\int_{\mathbb{R}^d} (1 + 2|y| + L)^2 \phi_\varepsilon(x - y) f_t(dx)}{f_t^\varepsilon(y)} + 2 \frac{\int_{|x-y|\geq|y|+L} |x - y|^2 \phi_\varepsilon(x - y) f_t(dx)}{f_t^\varepsilon(y)} \\ &\leq 2(1 + 2|y| + L)^2 + 2 \frac{(|y| + L)^2 \phi_\varepsilon(|y| + L)}{f_t^\varepsilon(y)}. \end{aligned}$$

For the second term, we used that  $|y| + L \geq 2 \geq \sqrt{2\varepsilon}$  and that  $z \mapsto |z|^2 \phi_\varepsilon(z)$  is radially symmetric and decreasing on  $\{|z| \geq \sqrt{2\varepsilon}\}$ . To conclude the proof of (i), it suffices to note that

$$f_t^\varepsilon(y) \geq \int_{|x-y|\leq|y|+L} \phi_\varepsilon(x - y) f_t(dx) \geq \phi_\varepsilon(|y| + L) f_t(B(y, |y| + L)) \geq \phi_\varepsilon(|y| + L)/2$$

because  $z \mapsto \phi_\varepsilon(z)$  is radially symmetric decreasing and because  $f_t(B(y, |y| + L)) \geq f_t(B(0, L)) \geq 1/2$ , since  $f_t(B(0, L)^c) \leq m_1(f_t)/L \leq 1/2$ .

For point (ii), it suffices to prove that  $\nabla_y b^\varepsilon(t, y)$ ,  $\nabla_y a^\varepsilon(t, y)$ ,  $D_y^2 a^\varepsilon(t, y)$  are locally bounded on  $[0, T] \times \mathbb{R}^d$ , as well as  $G^\varepsilon(t, y) := \int_{\mathbb{R}^d} \int_E |h(t, z, x)| |\nabla_y F_t^\varepsilon(x, y)| \mu(dz) f_t(dx)$ . No uniformity in  $\varepsilon$  is required here. By Stroock and Varadhan [12, Theorem 5.2.3], the local boundedness of  $D_y^2 a^\varepsilon(t, y)$  implies that of  $\nabla_y ([a^\varepsilon(t, y)]^{1/2})$ .

First, one easily checks that  $y \mapsto (f_t^\varepsilon(y))^{-1}$  is of class  $C^\infty$  for each  $t \in [0, T]$  and that it is locally bounded, as well as its derivatives of order 1 and 2, on  $[0, T] \times \mathbb{R}^d$ . This uses in particular the lower bound  $f_t^\varepsilon(y) \geq \phi_\varepsilon(|y| + L)/2$  proved a few lines above.

Recall that by definition, we have  $a^\varepsilon(t, y) = (f_t^\varepsilon(y))^{-1} \int_{\mathbb{R}^d} \phi_\varepsilon(x - y) a(t, x) f_t(dx)$  and  $b^\varepsilon(t, y) = (f_t^\varepsilon(y))^{-1} \times \int_{\mathbb{R}^d} \phi_\varepsilon(x - y) b(t, x) f_t(dx)$ . Recall finally that  $|a(t, x)| + |b(t, x)| \leq C(1 + |x|^2)$ . So concerning  $a^\varepsilon$  and  $b^\varepsilon$ , our goal is only to check that

$$K_\varepsilon(t, y) := \int_{\mathbb{R}^d} [|\nabla_y \phi_\varepsilon(x - y)| + |D_y^2 \phi_\varepsilon(x - y)|] (1 + |x|^2) f_t(dx)$$

is locally bounded on  $[0, T] \times \mathbb{R}^d$ . But using that  $(1 + |z|^2)[|\nabla \phi_\varepsilon(z)| + |D^2 \phi_\varepsilon(z)|]$  is bounded on  $\mathbb{R}^d$ , we deduce that  $[|\nabla_y \phi_\varepsilon(x - y)| + |D_y^2 \phi_\varepsilon(x - y)|] (1 + |x|^2) \leq C_\varepsilon(1 + |y|^2)$ , whence  $K_\varepsilon(t, y) \leq C_\varepsilon(1 + |y|^2)$ .

Next, one has  $|\nabla_y F_t^\varepsilon(x, y)| \leq C_\varepsilon (f_t^\varepsilon(y))^{-2} [\phi_\varepsilon(x - y) |\nabla f_t^\varepsilon(y)| + f_t^\varepsilon(y) |\nabla \phi_\varepsilon(x - y)|]$ . Using again that  $f_t^\varepsilon$  is smooth and positive, the goal concerning  $G^\varepsilon$  is to verify that

$$L_\varepsilon(t, y) := \int_{\mathbb{R}^d} \int_E |h(t, z, x)| [\phi_\varepsilon(x - y) + |\nabla \phi_\varepsilon(x - y)|] \mu(dz) f_t(dx)$$

is locally bounded. By Assumption 1.1,

$$L_\varepsilon(t, y) \leq \int_{\mathbb{R}^d} [\phi_\varepsilon(x - y) + |\nabla \phi_\varepsilon(x - y)|] (1 + |x|) f_t(dx) \leq C_\varepsilon(1 + |y|)$$

as previously, because  $(1 + |z|)[\phi_\varepsilon(z) + |\nabla \phi_\varepsilon(z)|]$  is bounded. □

### 3. Study of the regularized equations

In this section, we build a realization of the regularized weak solution  $(f_t^\varepsilon)_{t \in [0, T]}$ .

**Proposition 3.1.** Let  $(f_t)_{t \in [0, T]} \in L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$  be a weak solution to (2) and fix  $\varepsilon \in (0, 1)$ . Consider  $(f_t^\varepsilon)_{t \in [0, T]}$  and  $a^\varepsilon, b^\varepsilon, F^\varepsilon$  defined in Lemma 2.1 and put  $\sigma^\varepsilon(t, y) := (a^\varepsilon(t, y))^{1/2}$ . Consider a random variable  $X_0^\varepsilon$ , a  $d$ -dimensional Brownian motion  $(B_s)_{s \in [0, T]}$  and a Poisson measure  $N(ds, dz, dx, du)$  on  $[0, T] \times E \times \mathbb{R}^d \times [0, \infty)$  with intensity measure  $ds\mu(dz)f_s(dx)du$ , these three objects being independent. We work with the filtration generated by  $X_0^\varepsilon, B, N$ .

(i) There is a pathwise unique càdlàg adapted solution  $(X_t^\varepsilon)_{t \in [0, T]}$  to

$$X_t^\varepsilon = X_0^\varepsilon + \int_0^t b^\varepsilon(s, X_s^\varepsilon) ds + \int_0^t \sigma^\varepsilon(s, X_s^\varepsilon) dB_s + \int_0^t \int_E \int_{\mathbb{R}^d} \int_0^\infty h(s, z, x) \mathbf{1}_{\{u \leq F_s^\varepsilon(x, X_{s-}^\varepsilon)\}} N(ds, dz, dx, du). \tag{4}$$

(ii) There is a constant  $C$  (not depending on  $\varepsilon$ ) such that  $\mathbb{E}[\sup_{[0, T]} |X_t^\varepsilon|] \leq C(1 + \mathbb{E}[|X_0^\varepsilon|])$ .

(iii) If  $\mathcal{L}(X_0^\varepsilon) = f_0^\varepsilon$ , then  $\mathcal{L}(X_t^\varepsilon) = f_t^\varepsilon$  for all  $t \in [0, T]$ .

**Proof.** (i) The existence of a pathwise unique solution to (4) is more or less standard, because of the linear growth and local Lipschitz properties of the coefficients proved in Lemma 2.2. We only prove pathwise uniqueness, the existence being shown similarly, using a localization procedure (to make the coefficients globally Lipschitz continuous) and a Picard iteration. Consider two solutions  $(X_t^\varepsilon)_{t \in [0, T]}$  and  $(\tilde{X}_t^\varepsilon)_{t \in [0, T]}$  to (4) with  $X_0^\varepsilon = \tilde{X}_0^\varepsilon$  and introduce the stopping time  $\tau_R := \inf\{t \in [0, T] : |X_t^\varepsilon| \vee |\tilde{X}_t^\varepsilon| \geq R\}$ , for  $R > 0$ , with the convention that  $\inf \emptyset = T$ . Using the Burkholder–Davis–Gundy inequality for the Brownian part, we find

$$\mathbb{E} \left[ \sup_{[0, t \wedge \tau_R]} |X_s^\varepsilon - \tilde{X}_s^\varepsilon| \right] \leq \mathbb{E} \left[ \int_0^{t \wedge \tau_R} |b^\varepsilon(s, X_s^\varepsilon) - b^\varepsilon(s, \tilde{X}_s^\varepsilon)| ds + C \left( \int_0^{t \wedge \tau_R} |\sigma^\varepsilon(s, X_s^\varepsilon) - \sigma^\varepsilon(s, \tilde{X}_s^\varepsilon)|^2 ds \right)^{1/2} + \int_0^{t \wedge \tau_R} \int_E \int_{\mathbb{R}^d} |h(s, z, x)| |F_s^\varepsilon(x, X_s^\varepsilon) - F_s^\varepsilon(x, \tilde{X}_s^\varepsilon)| f_s(dx) \mu(dz) ds \right].$$

By Lemma 2.2-(ii), we deduce that

$$\mathbb{E} \left[ \sup_{[0, t \wedge \tau_R]} |X_s^\varepsilon - \tilde{X}_s^\varepsilon| \right] \leq C_{R, \varepsilon} \mathbb{E} \left[ \int_0^{t \wedge \tau_R} |X_s^\varepsilon - \tilde{X}_s^\varepsilon| ds + \left( \int_0^{t \wedge \tau_R} |X_s^\varepsilon - \tilde{X}_s^\varepsilon|^2 ds \right)^{1/2} \right] \leq C_{R, \varepsilon} (t + \sqrt{t}) \mathbb{E} \left[ \sup_{[0, t \wedge \tau_R]} |X_s^\varepsilon - \tilde{X}_s^\varepsilon| \right].$$

We deduce that  $\mathbb{E}[\sup_{[0, t_R \wedge \tau_R]} |X_s^\varepsilon - \tilde{X}_s^\varepsilon|] = 0$ , where  $t_R > 0$  is such that  $C_{R, \varepsilon}(t_R + \sqrt{t_R}) = 1/2$ . But then, the same computation allows us to prove that  $\mathbb{E}[\sup_{[t_R \wedge \tau_R, (2t_R) \wedge \tau_R]} |X_s^\varepsilon - \tilde{X}_s^\varepsilon|] = 0$ , etc, so that we end with  $\mathbb{E}[\sup_{[0, T \wedge \tau_R]} |X_s^\varepsilon - \tilde{X}_s^\varepsilon|] = 0$  for each  $R > 0$ . Since  $\lim_{R \rightarrow \infty} \tau_R = T$  a.s. (because  $(X_t^\varepsilon)_{t \in [0, T]}$  and  $(\tilde{X}_t^\varepsilon)_{t \in [0, T]}$  are assumed to be a.s. càdlàg and thus locally bounded on  $[0, T]$ ), we conclude that  $\mathbb{E}[\sup_{[0, T]} |X_s^\varepsilon - \tilde{X}_s^\varepsilon|] = 0$ , which was our goal.

(ii) Using the Burkholder–Davis–Gundy inequality for the Brownian part, we find, for  $t \in [0, T]$ ,

$$u_t^\varepsilon := \mathbb{E} \left[ \sup_{[0, t]} |X_s^\varepsilon| \right] \leq \mathbb{E}[|X_0^\varepsilon|] + \mathbb{E} \left[ \int_0^t |b^\varepsilon(s, X_s^\varepsilon)| ds \right] + C \mathbb{E} \left[ \left( \int_0^t |\sigma^\varepsilon(s, X_s^\varepsilon)|^2 ds \right)^{1/2} \right] + \mathbb{E} \left[ \int_0^t \int_E \int_{\mathbb{R}^d} |h(s, z, x)| F_s^\varepsilon(x, X_{s-}^\varepsilon) f_s(dx) \mu(dz) ds \right].$$

Inserting the estimates proved in Lemma 2.2-(i), we find, for some constant  $C$  not depending on  $\varepsilon \in (0, 1)$  nor on  $\mathbb{E}[|X_0^\varepsilon|]$ ,

$$u_t^\varepsilon \leq \mathbb{E}[|X_0^\varepsilon|] + C \mathbb{E} \left[ \int_0^t (1 + |X_s^\varepsilon|) ds + \left( \int_0^t (1 + |X_s^\varepsilon|^2) ds \right)^{1/2} \right] \leq u_0^\varepsilon + C(t + \sqrt{t})(1 + u_t^\varepsilon).$$

With  $t_0 > 0$  such that  $C(t_0 + \sqrt{t_0}) = 1/2$ , we conclude that  $u_{t_0}^\varepsilon \leq 2u_0^\varepsilon + 1$ . One checks similarly that  $u_{2t_0}^\varepsilon \leq 2u_{t_0}^\varepsilon + 1 \leq 4u_0^\varepsilon + 3$ . Repeating the argument, we end with  $u_T^\varepsilon \leq 2^{\lfloor T/t_0 \rfloor + 1} u_0^\varepsilon + 2^{\lfloor T/t_0 \rfloor + 1} - 1$ .

(iii) We now assume that  $\mathcal{L}(X_0^\varepsilon) = f_0^\varepsilon$  and we set  $g_t^\varepsilon := \mathcal{L}(X_t^\varepsilon)$ . A direct application of the Itô formula shows that for all  $t \in [0, T]$ , recalling the notation of Lemma 2.1,

$$\int_{\mathbb{R}^d} \psi(y) g_t^\varepsilon(dy) = \int_{\mathbb{R}^d} \psi(y) f_0^\varepsilon(dy) + \int_0^t \int_{\mathbb{R}^d} [\mathcal{A}_{s,\varepsilon} \psi(y) + \mathcal{B}_{s,\varepsilon} \psi(y)] g_s^\varepsilon(dy) ds.$$

Recalling Lemma 2.1 again,  $(f_t^\varepsilon)_{t \in [0, T]}$  solves the same equation. The following uniqueness result will thus complete the proof of (iii): for any  $\nu_0 \in \mathcal{P}(\mathbb{R}^d)$ , there exists at most one measurable family  $(\nu_t)_{t \in [0, T]}$  of probability measures such that for all  $\psi \in C_c^2(\mathbb{R}^d)$  and all  $t \in [0, T]$ ,

$$\int_{\mathbb{R}^d} \psi(y) \nu_t(dy) = \int_{\mathbb{R}^d} \psi(y) \nu_0(dy) + \int_0^t \int_{\mathbb{R}^d} \nu_s(dy) [\mathcal{A}_{s,\varepsilon} \psi(y) + \mathcal{B}_{s,\varepsilon} \psi(y)]. \quad (5)$$

This must be classical (because the coefficients are rather regular), but we found no reference and thus make use of martingale problems. A càdlàg adapted  $\mathbb{R}^d$ -valued process  $(Y_t)_{t \in [0, T]}$  on some filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$  is said to solve  $MP_\varepsilon(\nu_0)$  if  $\mathcal{L}(Y_0) = \nu_0$  and if

$$\psi(Y_t) - \int_0^t [\mathcal{A}_{s,\varepsilon} \psi(Y_s) + \mathcal{B}_{s,\varepsilon} \psi(Y_s)] ds$$

is a martingale for all  $\psi \in C_c^2(\mathbb{R}^d)$ . Due to Horowitz and Karandikar [7, Theorem B1], the following points imply uniqueness for (5). Here  $C_0(\mathbb{R}^d)$  is the set of continuous functions from  $\mathbb{R}^d$  to  $\mathbb{R}$  vanishing at infinity.

- (a)  $C_c^2(\mathbb{R}^d)$  is dense in  $C_0(\mathbb{R}^d)$  for the uniform convergence topology,
- (b)  $(t, y) \mapsto \mathcal{A}_{t,\varepsilon} \psi(y) + \mathcal{B}_{t,\varepsilon} \psi(y)$  is measurable for all  $\psi \in C_c^2(\mathbb{R}^d)$ ,
- (c) for each  $t \in [0, T]$ ,  $\mathcal{A}_{t,\varepsilon} + \mathcal{B}_{t,\varepsilon}$  satisfies the maximum principle,
- (d) there exists a countable family  $(\psi_k)_{k \geq 1} \subset C_c^2(\mathbb{R}^d)$  such that for all  $t \in [0, T]$ ,

$$\overline{\{(\psi_k, \mathcal{A}_{t,\varepsilon} \psi_k + \mathcal{B}_{t,\varepsilon} \psi_k), k \geq 1\}} \supset \{(\psi, \mathcal{A}_{t,\varepsilon} \psi + \mathcal{B}_{t,\varepsilon} \psi), \psi \in C_c^2(\mathbb{R}^d)\},$$

where the closure in the left-hand side is under the bounded pointwise convergence,

- (e) for each  $y_0 \in \mathbb{R}^d$ , there exists a unique (in law) solution to  $MP_\varepsilon(\delta_{y_0})$ .

Points (a) and (b) are obvious. The SDE associated to  $MP_\varepsilon$  is precisely (4):  $(Y_t)_{t \in [0, T]}$  solves  $MP_\varepsilon(\nu_0)$  if and only if it is a weak solution to (4) and  $\mathcal{L}(Y_0) = \nu_0$ , see Jacod [8, Theorem 13.55], see also [7, Theorem A1]. Thus (e) follows from (i). For (c), assume that  $\psi \in C_c^2(\mathbb{R}^d)$  attains its maximum at  $y_0$ . Then  $\mathcal{B}_{t,\varepsilon} \psi(y_0) \leq 0$  (this is immediate) and  $\mathcal{A}_{t,\varepsilon} \psi(y_0) \leq 0$  (because  $\nabla \psi(y_0) = 0$  and, since  $a(t, y_0)$  is symmetry and nonnegative,  $\sum_{i,j} a_{ij}(t, y_0) \partial_{ij} \psi(y_0) \leq 0$ ). It only remains to prove (d). Consider any countable subset  $(\psi_k)_{k \geq 1} \subset C_c^2(\mathbb{R}^d)$  dense in  $C_c^2(\mathbb{R}^d)$ : for  $\psi \in C_c^2(\mathbb{R}^d)$  with  $\text{Supp } \psi \subset B(0, M)$ , there exists  $(\psi_{k_n})_{n \geq 1}$  with  $\text{Supp } \psi_{k_n} \subset B(0, 2M)$  such that

$$\lim_{n \rightarrow \infty} (\|\psi - \psi_{k_n}\|_\infty + \|\nabla(\psi - \psi_{k_n})\|_\infty + \|D^2(\psi - \psi_{k_n})\|_\infty) = 0.$$

We will prove that (i)  $\lim_{n \rightarrow \infty} \sup_{[0, T]} \|\mathcal{A}_{t,\varepsilon} \psi_{k_n} - \mathcal{A}_{t,\varepsilon} \psi\|_\infty = 0$ , and (ii)  $\lim_{n \rightarrow \infty} \sup_{[0, T]} \|\mathcal{B}_{t,\varepsilon} \psi_{k_n} - \mathcal{B}_{t,\varepsilon} \psi\|_\infty = 0$ , which are more than needed.

By Lemma 2.2,

$$\begin{aligned} |\mathcal{A}_{t,\varepsilon}(\psi_{k_n} - \psi)(y)| &\leq \|\nabla(\psi_{k_n} - \psi)\|_\infty |b^\varepsilon(t, y)| \mathbf{1}_{\{|y| \leq 2M\}} + \frac{1}{2} \|D^2(\psi_{k_n} - \psi)\|_\infty \|a^\varepsilon(t, y)\| \mathbf{1}_{\{|y| \leq 2M\}} \\ &\leq C \|\nabla(\psi_{k_n} - \psi)\|_\infty + C \|D^2(\psi_{k_n} - \psi)\|_\infty, \end{aligned}$$

which tends to 0, implying (i). We next write, using that  $\text{Supp}(\psi_{k_n} - \psi) \subset B(0, 2M)$ ,

$$\begin{aligned} |(\psi_{k_n} - \psi)(y + h(t, z, x)) - (\psi_{k_n} - \psi)(y)| &\leq \mathbf{1}_{\{|y| \leq 4M\}} \|\nabla(\psi_{k_n} - \psi)\|_\infty |h(t, z, x)| \\ &\quad + \mathbf{1}_{\{|y| \geq 4M\}} \|\psi_{k_n} - \psi\|_\infty \mathbf{1}_{\{|y+h(t,z,x,z)| \leq 2M\}}. \end{aligned}$$

Observing that

$$\mathbf{1}_{\{|y| \geq 4M, |y+h(t,z,x,z)| \leq 2M\}} \leq \mathbf{1}_{\{|y| \geq 4M, |h(t,z,x,z)| \geq |y|/2\}} \leq \mathbf{1}_{\{|y| \geq 4M\}} \frac{2|h(t, z, x)|}{|y|},$$

we deduce that

$$\begin{aligned} |\mathcal{B}_{t,\varepsilon}(\psi_{k_n} - \psi)(y)| &\leq \mathbf{1}_{\{|y| \leq 4M\}} \|\nabla(\psi_{k_n} - \psi)\|_\infty \int_E \int_{\mathbb{R}^d} |h(t, z, x)| F_t^\varepsilon(x, y) f_t(dx) \mu(dz) \\ &\quad + \mathbf{1}_{\{|y| \geq 4M\}} \|\psi_{k_n} - \psi\|_\infty \int_E \int_{\mathbb{R}^d} \frac{2|h(t, z, x)|}{|y|} F_t^\varepsilon(x, y) f_t(dx) \mu(dz). \end{aligned}$$

Recalling that  $\int_E \int_{\mathbb{R}^d} |h(t, z, x)| F_t^\varepsilon(x, y) f_t(dx) \mu(dz) \leq C(1 + |y|)$  by Lemma 2.2, we find

$$\begin{aligned} |\mathcal{B}_{t,\varepsilon}(\psi_{k_n} - \psi)(y)| &\leq \mathbf{1}_{\{|y| \leq 4M\}} C(1 + |y|) \|\nabla(\psi_{k_n} - \psi)\|_\infty + \mathbf{1}_{\{|y| \geq 4M\}} C \frac{1 + |y|}{|y|} \|\psi_{k_n} - \psi\|_\infty \\ &\leq C \|\nabla(\psi_{k_n} - \psi)\|_\infty + C \|\psi_{k_n} - \psi\|_\infty \end{aligned}$$

and the conclusion follows. □

**Lemma 3.2.** For  $(f_t)_{t \in [0, T]} \in L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$  a weak solution to (2) and  $\varepsilon \in (0, 1)$ , consider the process  $(X_t^\varepsilon)_{t \in [0, T]}$ , with  $X_0^\varepsilon \sim f_0^\varepsilon$ , introduced in Lemma 3.1. The family  $\{(X_t^\varepsilon)_{t \in [0, T]}, \varepsilon > 0\}$  is tight in  $\mathbb{D}([0, T], \mathbb{R}^d)$  and any limit point  $(X_t)_{t \in [0, T]}$  satisfies  $\mathbb{P}(\Delta X_t \neq 0) = 0$  for all  $t \in [0, T]$ .

**Proof.** We use the Aldous criterion [1], see also Jacod and Shiryaev [9, p. 356], which implies tightness and that any limit point  $(X_t)_{t \in [0, T]}$  is quasi-left-continuous and thus has no deterministic jump time. It suffices to check that

- (i)  $\sup_{\varepsilon \in (0, 1)} \mathbb{E}[\sup_{[0, T]} |X_t^\varepsilon|] < \infty$ ,
- (ii)  $\lim_{\beta \rightarrow 0} \sup_{\varepsilon \in (0, 1)} \sup_{(S, S') \in \mathcal{S}_T(\beta)} \mathbb{E}[|X_{S'}^\varepsilon - X_S^\varepsilon|] = 0$ , where  $\mathcal{S}_T(\beta)$  is the set of all pairs of stopping times  $(S, S')$  satisfying  $0 \leq S \leq S' \leq S + \beta \leq T$  a.s.

Point (i) has already been checked in Lemma 3.1-(ii), since  $\mathbb{E}[|X_0^\varepsilon|] = m_1(f_0^\varepsilon) \leq m_1(f_0) + \sqrt{d\varepsilon}$ . Next, for  $S, S' \in \mathcal{S}_T(\beta)$  and  $\varepsilon \in (0, 1)$ , we have

$$\begin{aligned} \mathbb{E}[|X_{S'}^\varepsilon - X_S^\varepsilon|] &\leq \mathbb{E}\left[\int_S^{S+\beta} |b^\varepsilon(s, X_s^\varepsilon)| ds\right] + \mathbb{E}\left[\left|\int_S^{S'} \sigma^\varepsilon(s, X_s^\varepsilon) dB_s\right|\right] \\ &\quad + \mathbb{E}\left[\int_S^{S+\beta} \int_E \int_{\mathbb{R}^d} |h(s, z, x)| F_s^\varepsilon(x, X_s^\varepsilon) f_s(dx) \mu(dz) ds\right] \\ &\leq C \mathbb{E}\left[\int_S^{S+\beta} (1 + |X_s^\varepsilon|) ds\right] + C \mathbb{E}\left[\left(\int_S^{S'} |\sigma^\varepsilon(s, X_s^\varepsilon)|^2 ds\right)^{1/2}\right], \end{aligned}$$

where the last inequality follows from Lemma 2.2-(i) and the Burkholder–Davis–Gundy inequality. But  $|\sigma^\varepsilon(s, x)|^2 \leq C|a^\varepsilon(s, x)| \leq C(1 + |x|^2)$  by Lemma 2.2-(i) again, whence

$$\mathbb{E}[|X_{S'}^\varepsilon - X_S^\varepsilon|] \leq C \mathbb{E}\left[\int_S^{S+\beta} (1 + |X_s^\varepsilon|) ds + \left(\int_S^{S+\beta} (1 + |X_s^\varepsilon|^2) ds\right)^{1/2}\right].$$

Hence  $\mathbb{E}[|X_{S'}^\varepsilon - X_S^\varepsilon|] \leq C(\beta + \sqrt{\beta}) \mathbb{E}[\sup_{[0, T]} (1 + |X_s^\varepsilon|)] \leq C(\beta + \sqrt{\beta})$ , which ends the proof. □

### 4. Conclusion

As Figalli [5], we will need some continuous (in  $x$ ) approximations of  $a$ ,  $b$  and  $h$ .

**Lemma 4.1.** *Let  $(f_t)_{t \in [0, T]} \in L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$  be a weak solution to (2). For all  $\rho > 0$ , we can find  $\tilde{a} : [0, T] \times \mathbb{R}^d \mapsto \mathcal{S}_d^+$  and  $\tilde{b} : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ , both continuous and compactly supported, a set  $A \in \mathcal{E}$  such that  $\mu(A) < \infty$ , and a measurable function  $\tilde{h} : [0, T] \times E \times \mathbb{R}^d \mapsto \mathbb{R}^d$ , continuous on  $[0, T] \times \mathbb{R}^d$  for each  $z \in E$ , such that  $\tilde{h}(t, z, x) = 0$  for all  $(t, z, x) \in [0, T] \times A^c \times \mathbb{R}^d$  and*

$$\int_0^T \int_{\mathbb{R}^d} \left[ \frac{|a(t, x) - \tilde{a}(t, x)|}{1 + |x|} + |b(t, x) - \tilde{b}(t, x)| + \int_E |h(t, z, x) - \tilde{h}(t, z, x)| \mu(dz) \right] f_t(dx) dt < \rho.$$

**Proof.** For  $a$  and  $b$ , this follows from the fact, see Rudin [10, Theorem 3.14], that continuous functions with compact support are dense in  $L^1([0, T] \times \mathbb{R}^d, dt f_t(dx))$ , and that both  $a(t, x)/(1 + |x|)$  and  $b(t, x)$  belong to this space by Assumption 1.1.

Since  $h \in L^1([0, T] \times E \times \mathbb{R}^d, dt \mu(dz) f_t(dx))$  by Assumption 1.1 and since  $\mu$  is  $\sigma$ -finite, we can find  $A \in \mathcal{E}$  such that  $\mu(A) < \infty$  and  $\int_0^T \int_{A^c} \int_{\mathbb{R}^d} |h(t, z, x)| f_t(dx) \mu(dz) dt < \rho/3$ .

Next, can find a simple function  $g = \sum_{n=1}^N \alpha_n \mathbf{1}_{S_n}$ , with  $\alpha_n \in \mathbb{R}_*$ ,  $S_n \in \mathcal{B}([0, T] \times \mathbb{R}^d) \otimes \mathcal{E}$ , such that  $\int_0^T \int_A \int_{\mathbb{R}^d} |g(t, z, x) - h(t, z, x)| f_t(dx) \mu(dz) dt < \rho/3$ .

But for  $S \in \mathcal{B}([0, T] \times \mathbb{R}^d) \otimes \mathcal{E}$  and  $\varepsilon > 0$ , there is  $\varphi_{S, \varepsilon} : [0, T] \times \mathbb{R}^d \times E \mapsto \mathbb{R}$ , measurable, continuous on  $[0, T] \times \mathbb{R}^d$  for each  $z \in E$  and such that  $\int_0^T \int_A \int_{\mathbb{R}^d} |\mathbf{1}_{\{(t, z, x) \in S\}} - \varphi_{S, \varepsilon}(t, z, x)| f_t(dx) \mu(dz) dt < \varepsilon$ . Indeed, when  $S = C \times D$  with  $C \in \mathcal{B}([0, T] \times \mathbb{R}^d)$  and  $D \in \mathcal{E}$ , it suffices to consider  $\psi$  continuous on  $[0, T] \times \mathbb{R}^d$  such that  $\int_0^T \int_{\mathbb{R}^d} |\mathbf{1}_{\{(t, x) \in C\}} - \psi(t, x)| f_t(dx) dt < \varepsilon/\mu(A)$  and to set  $\varphi_{S, \varepsilon}(t, z, x) = \psi(t, x) \mathbf{1}_{\{z \in D\}}$ . The general case follows from the monotone class theorem.

Finally,  $\tilde{h}(t, z, x) = \sum_{n=1}^N \alpha_n \varphi_{S_n, \rho/(3|\alpha_n|2^n)}(t, z, x) \mathbf{1}_{\{z \in A\}}$  is measurable and continuous in  $(t, x)$  for each  $z \in E$ . Writing

$$\begin{aligned} |h(t, z, x) - \tilde{h}(t, z, x)| &\leq |h(t, z, x)| \mathbf{1}_{\{z \in A^c\}} + |g(t, z, x) - h(t, z, x)| \mathbf{1}_{\{z \in A\}} \\ &\quad + \sum_{n=1}^N |\alpha_n| |\varphi_{S_n, \rho/(3|\alpha_n|2^n)}(t, z, x) - \mathbf{1}_{\{(t, z, x) \in S_n\}}| \mathbf{1}_{\{z \in A\}}, \end{aligned}$$

we conclude that  $\int_0^T \int_E \int_{\mathbb{R}^d} |h(t, z, x) - \tilde{h}(t, z, x)| f_t(dx) \mu(dz) dt < \rho$  as desired. □

We now can give the

**Proof of Theorem 1.4.** Let  $(f_t)_{t \in [0, T]} \in L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$  be a weak solution to (2). For each  $\varepsilon \in (0, 1)$ , consider  $(f_t^\varepsilon)_{t \in [0, T]}$  introduced in Lemma 2.1 and the process  $(X_t^\varepsilon)_{t \in [0, T]}$  introduced in Lemma 3.1-(iii). By Lemma 3.2, we can find a sequence  $(X_t^{\varepsilon_n})_{t \in [0, T]}$  converging in law to some process  $(X_t)_{t \in [0, T]}$ . Since we know from Lemma 3.1 that  $\mathcal{L}(X_t^{\varepsilon_n}) = f_t^{\varepsilon_n}$  for each  $t \in [0, T]$ , each  $n \geq 1$  and since  $f_t^{\varepsilon_n}$  goes weakly to  $f_t$  as  $n \rightarrow \infty$  by construction, we deduce that for all  $t \in [0, T]$ ,  $\mathcal{L}(X_t) = f_t$ . It thus only remains to verify that  $X := (X_t)_{t \in [0, T]}$  is a (weak) solution to (1). According to the theory of martingale problems, see Jacod [8, Theorem 13.55], it classically suffices to prove that for any  $\psi \in C_c^2(\mathbb{R}^d)$ , the process

$$\psi(X_t) - \psi(X_0) - \int_0^t [\mathcal{A}_s \psi(X_s) + \mathcal{B}_s \psi(X_s)] ds$$

is a martingale in the filtration  $\mathcal{F}_t = \sigma(X_s, s \leq t)$ . Our goal is thus to check that for any  $0 \leq s_1 \leq \dots \leq s_k \leq s \leq t \leq T$ , any  $\psi_1, \dots, \psi_k \in C_b(\mathbb{R}^d)$  and any  $\psi \in C_c^2(\mathbb{R}^d)$ , we have  $\mathbb{E}[\mathcal{K}(X)] = 0$ , where  $\mathcal{K} : \mathbb{D}([0, T], \mathbb{R}^d) \mapsto \mathbb{R}$  is defined by

$$\mathcal{K}(\lambda) := \left( \prod_{i=1}^k \psi_i(\lambda_{s_i}) \right) \left( \psi(\lambda_t) - \psi(\lambda_s) - \int_s^t [\mathcal{A}_r \psi(\lambda_r) + \mathcal{B}_r \psi(\lambda_r)] dr \right).$$

We fix  $\rho > 0$  and consider  $\tilde{a}, \tilde{b}$  and  $\tilde{h}$  introduced in Lemma 4.1. We introduce  $\tilde{\mathcal{A}}_s$  and  $\tilde{\mathcal{B}}_s$  exactly as in Definition 1.2 with  $\tilde{a}, \tilde{b}$  and  $\tilde{h}$  instead of  $a, b$  and  $h$ . We define  $\tilde{a}^\varepsilon, \tilde{b}^\varepsilon, \tilde{\mathcal{A}}_{s,\varepsilon}$  and  $\tilde{\mathcal{B}}_{s,\varepsilon}$  exactly as in Lemma 2.1, with everywhere  $\tilde{a}, \tilde{b}$  and  $\tilde{h}$  instead of  $a, b$  and  $h$ . Finally, we define  $\tilde{\mathcal{K}}$  (resp.  $\tilde{\mathcal{K}}_\varepsilon$ , resp.  $\mathcal{K}_\varepsilon$ ) exactly as  $\mathcal{K}$  with  $\mathcal{A}_r$  and  $\mathcal{B}_r$  replaced by  $\tilde{\mathcal{A}}_r$  and  $\tilde{\mathcal{B}}_r$  (resp. by  $\tilde{\mathcal{A}}_{r,\varepsilon}$  and  $\tilde{\mathcal{B}}_{r,\varepsilon}$ , resp. by  $\mathcal{A}_{r,\varepsilon}$  and  $\mathcal{B}_{r,\varepsilon}$ ).

First,  $\mathbb{E}[\mathcal{K}_{\varepsilon_n}(X^{\varepsilon_n})] = 0$ . Indeed, since  $X^\varepsilon = (X_r^\varepsilon)_{t \in [0, T]}$  solves (4), by the Itô formula,

$$\begin{aligned} &\psi(X_t^\varepsilon) - \int_0^t [\mathcal{A}_{r,\varepsilon}(X_r^\varepsilon) + \mathcal{B}_{r,\varepsilon}(X_r^\varepsilon)] dr \\ &= \psi(X_t^\varepsilon) - \int_0^t b^\varepsilon(r, X_r^\varepsilon) \cdot \nabla \psi(X_r^\varepsilon) dr - \frac{1}{2} \sum_{i,j=1}^d \int_0^t a_{ij}^\varepsilon(r, X_r^\varepsilon) \partial_{ij} \psi(X_r^\varepsilon) dr \\ &\quad - \int_0^t \int_E \int_{\mathbb{R}^d} [\psi(X_r^\varepsilon + h(s, z, x)) - \psi(X_r^\varepsilon)] F_s^\varepsilon(x, X_r^\varepsilon) f_r(dx) \mu(dz) dr \end{aligned}$$

is a martingale, which implies the claim. We thus may write, for each  $n \geq 1$ ,

$$\begin{aligned} |\mathbb{E}[\mathcal{K}(X)]| &\leq |\mathbb{E}[\mathcal{K}(X)] - \mathbb{E}[\tilde{\mathcal{K}}(X)]| + |\mathbb{E}[\tilde{\mathcal{K}}(X)] - \mathbb{E}[\tilde{\mathcal{K}}(X^{\varepsilon_n})]| \\ &\quad + |\mathbb{E}[\tilde{\mathcal{K}}(X^{\varepsilon_n})] - \mathbb{E}[\tilde{\mathcal{K}}_{\varepsilon_n}(X^{\varepsilon_n})]| + |\mathbb{E}[\tilde{\mathcal{K}}_{\varepsilon_n}(X^{\varepsilon_n})] - \mathbb{E}[\mathcal{K}_{\varepsilon_n}(X^{\varepsilon_n})]|. \end{aligned}$$

We now study the four terms. We denote by  $M$  a constant such that  $\text{Supp } \psi \subset B(0, M)$ . We also define  $\phi(z) = (2\pi)^{-d/2} e^{-|z|^2/2}$ , so that  $\phi_\varepsilon(z) = \varepsilon^{-d/2} \phi(\varepsilon^{-1/2} z)$ .

*Step 1.* Here we prove that  $\lim_{n \rightarrow \infty} \mathbb{E}[\tilde{\mathcal{K}}(X^{\varepsilon_n})] = \mathbb{E}[\tilde{\mathcal{K}}(X)]$ . Since  $X^{\varepsilon_n}$  goes in law to  $X$  by construction, it suffices to verify that  $\tilde{\mathcal{K}}$  is bounded and a.s. continuous at  $X$ .

Since  $\tilde{a}, \tilde{b}$  and  $\tilde{h}$  are continuous in space and time, we easily deduce that  $(r, x) \mapsto \tilde{\mathcal{A}}_r \psi(x)$  and  $(r, x) \mapsto \tilde{\mathcal{B}}_r \psi(x)$  are continuous and bounded on  $[0, T] \times \mathbb{R}^d$ . For  $\tilde{\mathcal{A}}_r \psi(x) = \tilde{b}(r, x) \cdot \nabla \psi(x) + \frac{1}{2} \sum_{i,j} \tilde{a}_{ij}(r, x) \partial_{ij} \psi(x)$  this is obvious, and for  $\tilde{\mathcal{B}}_r \psi(x) = \int_E [\psi(x + \tilde{h}(r, z, x)) - \psi(x)] \mu(dz) = \int_A [\psi(x + \tilde{h}(r, z, x)) - \psi(x)] \mu(dz)$ , this follows from the Lebesgue theorem, because  $\psi$  is bounded and  $\mu(A) < \infty$ .

We easily deduce that  $\tilde{\mathcal{K}}$  is bounded, and that it is continuous at each  $\lambda \in \mathbb{D}([0, T], \mathbb{R}^d)$  which does not jump at  $s_1, \dots, s_k, s, t$ . This is a.s. the case of  $X$ , see Lemma 3.2.

*Step 2.* Here we check that  $\Delta_1 := |\mathbb{E}[\mathcal{K}(X)] - \mathbb{E}[\tilde{\mathcal{K}}(X)]| \leq C\rho$  for some constant  $C$ . We have, since  $\text{Supp } \psi \subset B(0, M)$ ,

$$\begin{aligned} |\mathcal{K}(\lambda) - \tilde{\mathcal{K}}(\lambda)| &\leq C \int_0^t [|\mathcal{A}_r \psi(\lambda_r) - \tilde{\mathcal{A}}_r \psi(\lambda_r)| + |\mathcal{B}_r \psi(\lambda_r) - \tilde{\mathcal{B}}_r \psi(\lambda_r)|] dr \\ &\leq C \int_0^t (|a(r, \lambda_r) - \tilde{a}(r, \lambda_r)| + |b(r, \lambda_r) - \tilde{b}(r, \lambda_r)|) \mathbf{1}_{\{|\lambda_r| < M\}} dr \\ &\quad + C \int_0^t \int_E |h(r, z, \lambda_r) - \tilde{h}(r, z, \lambda_r)| \mu(dz) dr. \end{aligned}$$

Using now that  $\mathbf{1}_{\{|\lambda| < M\}} \leq C(1 + |\lambda|)^{-1}$  and that  $\mathcal{L}(X_r) = f_r$  for each  $r \in [0, T]$ , we conclude that

$$\begin{aligned} \Delta_1 &\leq C \int_0^t \int_{\mathbb{R}^d} \left( \frac{|a(r, x) - \tilde{a}(r, x)|}{1 + |x|} + |b(r, x) - \tilde{b}(r, x)| \right) f_r(dx) dr \\ &\quad + C \int_0^t \int_E \int_{\mathbb{R}^d} |h(r, z, x) - \tilde{h}(r, z, x)| f_r(dx) \mu(dz) dr. \end{aligned}$$

This is smaller than  $C\rho$  by Lemma 4.1.

Step 3. Now we verify that for all  $n \geq 1$ ,  $\Delta_2^n = |\mathbb{E}[\tilde{\mathcal{K}}_{\varepsilon_n}(X^{\varepsilon_n})] - \mathbb{E}[\mathcal{K}_{\varepsilon_n}(X^{\varepsilon_n})]| \leq C\rho$ . As in Step 2,

$$\begin{aligned} \Delta_2^n &\leq C \int_0^t \int_{\mathbb{R}^d} \left( \frac{|a^{\varepsilon_n}(r, y) - \tilde{a}^{\varepsilon_n}(r, y)|}{1 + |y|} + |b^{\varepsilon_n}(r, y) - \tilde{b}^{\varepsilon_n}(r, y)| \right) f_r^{\varepsilon_n}(y) dy dr \\ &\quad + C \int_0^t \int_E \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |h(r, z, x) - \tilde{h}(r, z, x)| \frac{\phi_{\varepsilon_n}(x - y)}{f_r^{\varepsilon_n}(y)} f_r(dx) f_r^{\varepsilon_n}(y) dy \mu(dz) dr. \end{aligned}$$

Recalling (see Lemma 2.1) that  $a^{\varepsilon_n}(r, y) f_r^{\varepsilon_n}(y) = \int_{\mathbb{R}^d} \phi_{\varepsilon_n}(x - y) a(r, x) f_r(dx)$ , that  $\tilde{a}^{\varepsilon_n}(r, y) f_r^{\varepsilon_n}(y) = \int_{\mathbb{R}^d} \phi_{\varepsilon_n}(x - y) \tilde{a}(r, x) f_r(dx)$  and similar formulas for  $b^{\varepsilon_n}(r, y) f_r^{\varepsilon_n}(y)$  and  $\tilde{b}^{\varepsilon_n}(r, y) f_r^{\varepsilon_n}(y)$ , we find

$$\begin{aligned} \Delta_2^n &\leq C \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left( \frac{|a(r, x) - \tilde{a}(r, x)|}{1 + |y|} + |b(r, x) - \tilde{b}(r, x)| \right) \phi_{\varepsilon_n}(x - y) f_r(dx) dy dr \\ &\quad + C \int_0^t \int_E \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |h(r, z, x) - \tilde{h}(r, z, x)| \phi_{\varepsilon_n}(x - y) f_r(dx) dy \mu(dz) dr. \end{aligned}$$

But  $\int_{\mathbb{R}^d} \phi_{\varepsilon_n}(x - y) dy = 1$  and, since  $\frac{1+|x|}{1+|y|} = 1 + \frac{|x|-|y|}{1+|y|} \leq 1 + |x - y| \leq 2 + |x - y|^2$ ,

$$\int_{\mathbb{R}^d} \frac{(1 + |x|)\phi_{\varepsilon_n}(x - y) dy}{1 + |y|} \leq \int_{\mathbb{R}^d} (2 + |x - y|^2)\phi_{\varepsilon_n}(x - y) dy = 2 + d\varepsilon_n \leq 2 + d.$$

Consequently,

$$\begin{aligned} \Delta_2^n &\leq C \int_0^t \int_{\mathbb{R}^d} \left( \frac{|a(r, x) - \tilde{a}(r, x)|}{1 + |x|} + |b(r, x) - \tilde{b}(r, x)| \right) f_r(dx) dr \\ &\quad + C \int_0^t \int_E \int_{\mathbb{R}^d} |h(r, z, x) - \tilde{h}(r, z, x)| f_r(dx) \mu(dz) dr, \end{aligned}$$

which is smaller than  $C\rho$  by Lemma 4.1.

Step 4. Finally, we check that  $\lim_{n \rightarrow \infty} |\mathbb{E}[\tilde{\mathcal{K}}(X^{\varepsilon_n})] - \mathbb{E}[\tilde{\mathcal{K}}_{\varepsilon_n}(X^{\varepsilon_n})]| = 0$ . We first observe that  $|\mathbb{E}[\tilde{\mathcal{K}}(X^{\varepsilon_n})] - \mathbb{E}[\tilde{\mathcal{K}}_{\varepsilon_n}(X^{\varepsilon_n})]| \leq C(I_n + J_n)$ , where

$$I_n := \mathbb{E} \left[ \int_0^t |\tilde{\mathcal{A}}_{r, \varepsilon_n} \psi(X_r^{\varepsilon_n}) - \tilde{\mathcal{A}}_r \psi(X_r^{\varepsilon_n})| dr \right] \quad \text{and} \quad J_n := \mathbb{E} \left[ \int_0^t |\tilde{\mathcal{B}}_{r, \varepsilon_n} \psi(X_r^{\varepsilon_n}) - \tilde{\mathcal{B}}_r \psi(X_r^{\varepsilon_n})| dr \right].$$

Since  $\psi \in C_c^2(\mathbb{R}^d)$  and since  $\mathcal{L}(X_r^{\varepsilon_n}) = f_r^{\varepsilon_n}$ , we have

$$\begin{aligned} I_n &\leq C \int_0^t \int_{\mathbb{R}^d} (|\tilde{b}^{\varepsilon_n}(r, y) - \tilde{b}(r, y)| + |\tilde{a}^{\varepsilon_n}(r, y) - \tilde{a}(r, y)|) f_r^{\varepsilon_n}(y) dy dr \\ &\leq C \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|\tilde{b}(r, x) - \tilde{b}(r, y)| + |\tilde{a}(r, x) - \tilde{a}(r, y)|) \phi_{\varepsilon_n}(x - y) f_r(dx) dy dr, \end{aligned}$$

because  $[\tilde{b}^{\varepsilon_n}(r, y) - \tilde{b}(r, y)] f_r^{\varepsilon_n}(y) = \int_{\mathbb{R}^d} \phi_{\varepsilon_n}(x - y) \tilde{b}(r, x) f_r(dx) - \int_{\mathbb{R}^d} \phi_{\varepsilon_n}(x - y) \tilde{b}(r, y) f_r(dx)$ , with a similar formula concerning  $\tilde{a}$ . Using finally the substitution  $y = x + \sqrt{\varepsilon_n}u$ , we find

$$I_n \leq C \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|\tilde{b}(r, x) - \tilde{b}(r, x + \sqrt{\varepsilon_n}u)| + |\tilde{a}(r, x) - \tilde{a}(r, x + \sqrt{\varepsilon_n}u)|) \phi(u) f_r(dx) du dr.$$

Hence  $\lim_n I_n = 0$  by dominated convergence, since  $\tilde{a}$  and  $\tilde{b}$  are continuous and bounded.

By the same way, since  $f_r^{\varepsilon_n}(y) = \int_{\mathbb{R}^d} \phi_{\varepsilon_n}(x - y) f_r(dx)$ ,

$$\begin{aligned} J_n &= \mathbb{E} \left[ \int_0^t \left| \int_E \int_{\mathbb{R}^d} [\psi(X_r^{\varepsilon_n} + \tilde{h}(r, z, x)) - \psi(X_r^{\varepsilon_n})] \frac{\phi_{\varepsilon_n}(x - X_r^{\varepsilon_n})}{f_r^{\varepsilon_n}(X_r^{\varepsilon_n})} f_r(dx) \mu(dz) \right. \right. \\ &\quad \left. \left. - \int_E [\psi(X_r^{\varepsilon_n} + \tilde{h}(r, z, X_r^{\varepsilon_n})) - \psi(X_r^{\varepsilon_n})] \mu(dz) \right| dr \right] \\ &= \mathbb{E} \left[ \int_0^t \left| \int_E \int_{\mathbb{R}^d} [\psi(X_r^{\varepsilon_n} + \tilde{h}(r, z, x)) - \psi(X_r^{\varepsilon_n} + \tilde{h}(r, z, X_r^{\varepsilon_n}))] \frac{\phi_{\varepsilon_n}(x - X_r^{\varepsilon_n})}{f_r^{\varepsilon_n}(X_r^{\varepsilon_n})} f_r(dx) \mu(dz) \right| dr \right] \\ &\leq C \mathbb{E} \left[ \int_0^t \int_E \int_{\mathbb{R}^d} [1 \wedge |\tilde{h}(r, z, x) - \tilde{h}(r, z, X_r^{\varepsilon_n})|] \frac{\phi_{\varepsilon_n}(x - X_r^{\varepsilon_n})}{f_r^{\varepsilon_n}(X_r^{\varepsilon_n})} f_r(dx) \mu(dz) dr \right] \end{aligned}$$

because  $\psi$  and  $\nabla \psi$  are bounded. Using that  $\mathcal{L}(X_r^{\varepsilon_n}) = f_r^{\varepsilon_n}$ , the substitution  $y = x + \sqrt{\varepsilon_n}u$  and the fact that  $\tilde{h}(r, z, x) = 0$  if  $z \notin A$ ,

$$\begin{aligned} J_n &\leq C \int_0^t \int_A \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [1 \wedge |\tilde{h}(r, z, x) - \tilde{h}(r, z, y)|] \phi_{\varepsilon_n}(x - y) f_r(dx) dy \mu(dz) dr \\ &= C \int_0^t \int_A \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [1 \wedge |\tilde{h}(r, z, x) - \tilde{h}(r, z, x + \sqrt{\varepsilon_n}u)|] \phi(u) f_r(dx) dy \mu(dz) dr. \end{aligned}$$

Hence  $\lim_n J_n = 0$  by dominated convergence, since  $h$  is continuous in  $x$  and since  $\mu(A) < \infty$ .

*Conclusion.* Gathering Steps 1, 2, 3 and 4, we find that  $|\mathbb{E}[\mathcal{K}(X)]| \leq C\rho$ . Since  $\rho$  can be chosen arbitrarily small, we conclude that  $\mathbb{E}[\mathcal{K}(X)] = 0$ , which completes the proof.  $\square$

### Appendix

**Proof of Remark 1.3.** First, it is very easy, using only that  $a$  and  $b$  are locally bounded on  $[0, T] \times \mathbb{R}^d$ , to show that  $\mathcal{A}_t \varphi(x)$  is uniformly bounded as soon as  $\varphi \in C_c^2(\mathbb{R}^d)$ . The case of  $\mathcal{B}_t \varphi$  is more complicated. We consider  $\varphi \in C_c^2(\mathbb{R}^d)$  and  $M > 0$  such that  $\text{Supp } \varphi \subset B(0, M)$  and we write

$$|\mathcal{B}_t \varphi(x)| \leq \mathbf{1}_{\{|x| \leq 2M\}} \|\nabla \varphi\|_{\infty} \int_E |h(t, z, x)| \mu(dz) + \mathbf{1}_{\{|x| \geq 2M\}} \int_E |\varphi(x + h(t, z, x))| \mu(dz).$$

We observe that  $|\varphi(x + h(t, z, x))| \leq \|\varphi\|_{\infty} \mathbf{1}_{\{|x+h(t,z,x)| \leq M\}}$  and that

$$\mathbf{1}_{\{|x| \geq 2M, |x+h(t,z,x)| \leq M\}} \leq \mathbf{1}_{\{|x| \geq 2M, |h(t,z,x)| \geq |x|/2\}} \leq \mathbf{1}_{\{|x| \geq 2M\}} \frac{2|h(t, z, x)|}{|x|}.$$

Since  $\int_E |h(t, z, x)| \mu(dz) \leq C(1 + |x|)$  by assumption, we conclude that

$$|\mathcal{B}_t \varphi(x)| \leq \mathbf{1}_{\{|x| \leq 2M\}} C \|\nabla \varphi\|_{\infty} (1 + |x|) + \mathbf{1}_{\{|x| \geq 2M\}} \frac{C \|\varphi\|_{\infty} (1 + |x|)}{|x|},$$

which is bounded. We have proved point (i).

We next prove (ii). We put  $\varphi(x) = (1 + |x|^2)^{1/2}$ , which satisfies

$$\frac{1 + |x|}{2} \leq \varphi(x) \leq 1 + |x|, \quad |\nabla \varphi| \leq 1 \quad \text{and} \quad |D^2 \varphi| \leq \frac{C}{\varphi}.$$

We also introduce an increasing  $C^2$  function  $\chi : \mathbb{R}_+ \mapsto \mathbb{R}_+$  such that  $\chi(r) = r$  for  $r \in [0, 1]$  and  $\chi(r) = 2$  for  $r \geq 2$ . We thus have

$$r \wedge 1 \leq \chi(r) \leq 2(r \wedge 1), \quad |\chi'(r)| \leq C \mathbf{1}_{\{r \leq 2\}} \quad \text{and} \quad |\chi''(r)| \leq C \mathbf{1}_{\{1 \leq r \leq 2\}}.$$

We then set, for  $n \geq 1$  and  $x \in \mathbb{R}^d$ ,  $\psi_n(x) = n\chi(\varphi(x)/n)$ , which satisfies

$$\varphi \wedge n \leq \psi_n \leq 2(\varphi \wedge n), \quad |\nabla \psi_n| \leq C \mathbf{1}_{\{\varphi \leq 2n\}} \quad \text{and} \quad |D^2 \psi_n| \leq \frac{C}{\varphi} \mathbf{1}_{\{\varphi \leq 2n\}}.$$

Consequently, for all  $s \in [0, T]$ , since  $|b(s, \cdot)| \leq C\varphi$  and  $|a(s, \cdot)| \leq C\varphi^2$  by Assumption 1.1,

$$|\mathcal{A}_s \psi_n| \leq |b(s, \cdot)| |\nabla \psi_n| + |a(s, \cdot)| |D^2 \psi_n| \leq C\varphi \mathbf{1}_{\{\varphi \leq 2n\}} \leq C[\varphi \wedge (2n)] \leq C\psi_n.$$

We next claim that

$$\Delta_n(s, z, x) = |\psi_n(x + h(s, z, x)) - \psi_n(x)| \leq C |h(s, z, x)| \frac{\psi_n(x)}{\varphi(x)}. \quad (\text{A.1})$$

First, if  $\varphi(x) \leq 4n$ , then we only use that  $\nabla \psi_n$  is uniformly bounded to write  $\Delta_n(s, z, x) \leq C|h(s, z, x)|$ , whence the result because  $\psi_n(x) \geq \varphi(x) \wedge n \geq \varphi(x)/4$ . Second, if  $\varphi(x) \geq 4n$  (whence  $|x| \geq 4n - 1 \geq 3n$ ), since  $\psi_n$  is constant (with value  $2n$ ) on  $B(0, 2n)^c$  and bounded on  $\mathbb{R}^d$  by  $2n$ , we can write  $\Delta_n(s, z, x) \leq 4n \mathbf{1}_{\{|x+h(s,z,x)| \leq 2n\}} \leq 4n \mathbf{1}_{\{|h(s,z,x)| \geq |x|/3\}} \leq 12n|h(s, z, x)|/|x|$ . But  $12n = 6\psi_n(x)$  and  $|x| \geq \varphi(x) - 1 \geq \varphi(x)/2$ , whence the result.

We deduce from (A.1), using Assumption 1.1, that

$$|\mathcal{B}_s \psi_n(x)| \leq C \frac{\psi_n(x)}{\varphi(x)} \int_E |h(s, z, x)| \mu(dz) \leq C \frac{\psi_n(x)}{\varphi(x)} (1 + |x|) \leq C\psi_n(x).$$

Applying (3) with the test function  $\psi_n - 2n \in C_c^2(\mathbb{R}^d)$ , for which of course  $(\mathcal{A}_s + \mathcal{B}_s)(\psi_n - 2n) = (\mathcal{A}_s + \mathcal{B}_s)\psi_n$ , and using that  $f_0$  and  $f_t$  are probability measures, we find

$$\begin{aligned} \int_{\mathbb{R}^d} \psi_n(x) f_t(dx) &= \int_{\mathbb{R}^d} \psi_n(x) f_0(dx) + \int_0^t \int_{\mathbb{R}^d} (\mathcal{A}_s \psi_n(x) + \mathcal{B}_s \psi_n(x)) f_s(dx) ds \\ &\leq \int_{\mathbb{R}^d} \psi_n(x) f_0(dx) + C \int_0^t \int_{\mathbb{R}^d} \psi_n(x) f_s(dx) ds. \end{aligned}$$

Since  $f_0 \in \mathcal{P}_1(\mathbb{R}^d)$  by assumption and since  $0 \leq \psi_n(x) \leq 2|x| + 2$ ,  $\sup_{n \geq 1} \int_{\mathbb{R}^d} \psi_n(x) f_0(dx) < \infty$ . We thus conclude, by the Gronwall Lemma, that  $\sup_{n \geq 1} \sup_{t \in [0, T]} \int_{\mathbb{R}^d} \psi_n(x) f_t(dx) < \infty$ , which clearly implies that  $(f_t)_{t \in [0, T]} \in L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$ , because  $\lim_{n \rightarrow 0} \psi_n(x) = \varphi(x) \geq |x|$ .

For point (iii), we introduce a family of functions  $\chi_n \in C_c^2(\mathbb{R}^d)$ , for  $n \geq 1$ , such that  $\mathbf{1}_{\{|x| \leq n\}} \leq \chi_n(x) \leq \mathbf{1}_{\{|x| \leq n+1\}}$  and such that  $|D\chi_n(x)| + |D^2\chi_n(x)| \leq C \mathbf{1}_{\{|x| \in [n, n+1]\}}$ . We then consider  $\varphi \in C^2(\mathbb{R}^d)$  as in the statement, i.e. such that  $(1 + |x|)[|\varphi(x)| + |\nabla\varphi(x)| + |D^2\varphi(x)|]$  is bounded. Of course,  $\varphi\chi_n \in C_c^2(\mathbb{R}^d)$  for each  $n \geq 1$ , so that we can apply (3). We then let  $n \rightarrow \infty$ . Since  $\varphi$  is bounded, we obviously have  $\lim_n \int_{\mathbb{R}^d} \varphi(x) \chi_n(x) f_t(dx) = \int_{\mathbb{R}^d} \varphi(x) f_t(dx)$ . Next, we want to prove that  $\lim_n \int_0^t \int_{\mathbb{R}^d} [\mathcal{A}_s(\varphi\chi_n)(x) + \mathcal{B}_s(\varphi\chi_n)(x)] f_s(dx) ds = \int_0^t \int_{\mathbb{R}^d} [\mathcal{A}_s\varphi(x) + \mathcal{B}_s\varphi(x)] f_s(dx) ds$ . By dominated convergence and since  $(f_t)_{t \in [0, T]} \in L^\infty([0, T], \mathcal{P}_1(\mathbb{R}^d))$  by (ii), it suffices to prove that for all  $s \in [0, T]$ ,  $x \in \mathbb{R}^d$ ,

- (a)  $\sup_n |\mathcal{A}_s(\varphi\chi_n)(x)| \leq C(1 + |x|)$ , (b)  $\lim_n \mathcal{A}_s(\varphi\chi_n)(x) = \mathcal{A}_s\varphi(x)$ ,
- (c)  $\sup_n |\mathcal{B}_s(\varphi\chi_n)(x)| \leq C(1 + |x|)$ , (d)  $\lim_n \mathcal{B}_s(\varphi\chi_n)(x) = \mathcal{B}_s\varphi(x)$ .

Point (a) is easy: since  $|a(s, x)| + |b(s, x)| \leq C(1 + |x|^2)$  by Assumption 1.1 and since  $\chi_n, D\chi_n, D^2\chi_n$  are uniformly bounded,

$$|\mathcal{A}_s(\varphi\chi_n)(x)| \leq C(1 + |x|^2)(|D(\varphi\chi_n)(x)| + |D^2(\varphi\chi_n)(x)|) \leq C(1 + |x|^2)(|\varphi(x)| + |D\varphi(x)| + |D^2\varphi(x)|),$$

which is bounded by  $C(1 + |x|)$  by assumption. Point (b) is not hard, using that  $\lim_n \nabla(\varphi\chi_n)(x) = \nabla\varphi(x)$  and  $\lim_n \partial_{ij}(\varphi\chi_n)(x) = \partial_{ij}\varphi(x)$  for each  $x \in \mathbb{R}^d$ .

Next,  $\nabla(\varphi\chi_n)$  is uniformly bounded, so that  $|(\varphi\chi_n)(x + h(s, z, x)) - (\varphi\chi_n)(x)| \leq C|h(s, z, x)|$  and thus  $|\mathcal{B}_s(\varphi\chi_n)(x)| \leq C \int_E |h(s, z, x)|\mu(dz) \leq C(1 + |x|)$  by Assumption 1.1, whence (c). Also, by dominated convergence, since  $\lim_n \chi_n(y) = 1$  for all  $y \in \mathbb{R}^d$ ,

$$\lim_n \mathcal{B}_s(\varphi\chi_n)(x) = \lim_n \int_E [(\varphi\chi_n)(x + h(s, z, x)) - (\varphi\chi_n)(x)]\mu(dz) = \int_E [\varphi(x + h(s, z, x)) - \varphi(x)]\mu(dz),$$

which is nothing but  $\mathcal{B}_s\varphi(x)$  as desired.  $\square$

## Acknowledgements

We warmly thank Maxime Hauray for his help. The second author was partly supported by China Scholarship Council under Grant No. 201306040132.

## References

- [1] D. Aldous. Stopping times and tightness. *Ann. Probab.* **6** (2) (1978) 335–340. [MR0474446](#)
- [2] R. F. Bass. Stochastic differential equations with jumps. *Probab. Surv.* **1** (2004) 1–19. [MR2095564](#)
- [3] A. G. Bhatt and R. L. Karandikar. Invariant measures and evolution equations for Markov processes characterized via martingale problems. *Ann. Probab.* **21** (4) (1993) 2246–2268. [MR1245309](#)
- [4] S. N. Ethier and T. G. Kurtz. *Markov Processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons, Inc., New York, 1986. [MR0838085](#)
- [5] A. Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.* **254** (1) (2008) 109–153. [MR2375067](#)
- [6] N. Fournier and M. Hauray. Propagation of chaos for the Landau equation with moderately soft potentials. *Ann. Probab.* **44** (6) (2016) 3581–3660. [MR3572320](#)
- [7] J. Horowitz and R. L. Karandikar. Martingale problems associated with the Boltzmann equation. In *Seminar on Stochastic Processes, 1989 (San Diego, CA, 1989)* 75–122. *Progr. Probab.* **18**. Birkhäuser Boston, Boston, MA, 1990. [MR0900154](#)
- [8] J. Jacod. *Calcul stochastique et problèmes de martingales. Lecture Notes in Mathematics* **714**. Springer, Berlin, 1979. [MR0542115](#)
- [9] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**, 2nd edition. Springer-Verlag, Berlin, 2003. [MR1943877](#)
- [10] W. Rudin. *Real and Complex Analysis*. McGraw-Hill Book Co., New York–Toronto, Ont.–London, 1966. [MR0210528](#)
- [11] R. Situ. *Theory of Stochastic Differential Equations with Jumps and Applications. Mathematical and Analytical Techniques with Applications to Engineering*. Springer, New York, 2005. [MR2160585](#)
- [12] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes. Grundlehren der Mathematischen Wissenschaften* **233**. Springer-Verlag, Berlin–New York, 1979. [MR0532498](#)
- [13] H. Tanaka. Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete* **46** (1) (1978/79) 67–105. [MR0595483](#)
- [14] L. Xu. Uniqueness and propagation of chaos for the Boltzmann equation with moderately soft potentials. *Ann. Appl. Probab.* **28** (2) (2018) 1136–1189.