Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 43, 1–14.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP773

Existence of invariant probability measures for functional McKean-Vlasov SDEs*

Jianhai Bao† Michael Scheutzow‡ Chenggui Yuan§

Abstract

We show existence of an invariant probability measure for a class of functional McKean-Vlasov SDEs by applying Kakutani's fixed point theorem to a suitable class of probability measures on a space of continuous functions. Unlike some previous works [1, 25], we do not assume a monotonicity condition to hold. Further, our conditions are even weaker than some results in the literature on invariant probability measures for functional SDEs without dependence on the law of the solution [12].

Keywords: functional McKean-Vlasov SDE; invariant probability measure; Kakutani's fixed point theorem.
MSC2020 subject classifications: 60J60; 47D07.
Submitted to EJP on July 29, 2021, final version accepted on March 30, 2022.

1 Introduction and main result

Many classical functional SDEs generate Markov processes on a space of functions and it is natural to ask for the existence and uniqueness of invariant probability measures (IPMs) of these processes. A typical way to establish existence of an IPM of an SDE with or without delay is to show boundedness in probability of the solutions and then to apply the Krylov-Bogoliubov theorem (see e.g. [11, Theorem 3.1.1]). Boundedness in probability is usually guaranteed by constructing a suitable *Lyapunov* function (sometimes called Veretennikov-Khasminskii condition); see, for example, [8, 9, 13, 22]. Some authors establish the existence of IPMs without using the Krylov-Bogoliubov theorem: [5] uses the remote start method and [22] employs non-degeneracy and recurrence properties of the Markov process.

*Supported in part by NNSFC (11831014, 12071340).

†Center for Applied Mathematics, Tianjin University, Tianjin 300072, China. E-mail: jianhaibao@tju.edu.cn

‡Institut für Mathematik, MA 7-5, Fakultät II, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany. E-mail: ms@math.tu-berlin.de

§Department of Mathematics, Swansea University, Bay Campus SA1 8EN, UK. E-mail: C.Yuan@swansea.ac.uk

If the transition kernels starting from different initial distributions are contractive under a Wasserstein metric or the total variation metric, then the classical Banach fixed point theorem yields that the associated transition semigroup has a unique IPM; see, for example, [13]. Such conditions are however too strong if one is only interested in the existence of an IPM.

Over the past few decades, McKean-Vlasov SDEs (with or without delay) have gained considerable attention. They are applied in the social sciences, economics, neuroscience, engineering, and finance; see, for example, [17] or the monograph [10]. They often appear as limiting equations for N weakly interacting particles when $N \to \infty$. A McKean-Vlasov SDE is an SDE whose coefficients depend not only on the state of the process but also on its distribution. It is of the form

$$\mathrm{d}X(t) = b(X(t), \mathscr{L}_{X(t)})\mathrm{d}t + \sigma(X(t), \mathscr{L}_{X(t)})\mathrm{d}W(t), \tag{1.1}$$

where $\mathscr{L}_{X(t)}$ denotes the distribution of $X(t)$ and $(W(t))$ is an m-dimensional Brownian motion. In contrast to classical SDEs, the distribution of the McKean-Vlasov SDE (1.1) solves a *nonlinear* Fokker-Planck equation and the solution process of (1.1) is a *nonlinear Markov process in the sense of McKean* (but not a classical linear Markov process). Further, there is no associated (linear) semigroup but the distribution flow $P_t^*\mu := \mathscr{L}_{X_t^\mu}$ generates a nonlinear semigroup. Therefore, the approaches taken in e.g. [5, 12, 13, 16] to address existence of IPMs for linear Markov processes (or semigroups) cannot be applied to the McKean-Vlasov SDE (1.1).

Let us briefly review some previous approaches to show existence of IPMs for McKean-Vlasov SDEs. By applying the generalized Banach fixed-point theorem (see e.g. [26, Theorem 9.A, p449]), [1, Theorem 3] investigated existence of IPMs for semi-linear McKean-Vlasov stochastic partial differential equations (SPDEs for short). By employing the shift coupling approach and using contractivity under Wasserstein distance, [25, Theorem 3.1] discussed existence of IPMs for the McKean-Vlasov SDE (1.1). For further extensions to functional McKean-Vlasov SPDEs, we refer to [21, Theorem 2.3]. The existence of IPMs for McKean-Vlasov SDEs was investigated in [14, Theorem 4.1] under measure dependent Lyapunov conditions. Further, [27] treated non-uniqueness of IPMs for McKean-Vlasov SDEs with quadratic interaction and non-quadratic interaction.

Some authors assume a monotonicity condition to show existence and uniqueness of an IPM, e.g. [1, 25]. Such a condition is, however, very strong if one is only interested in the existence of an IPM. Motivated by the previous works, in this paper, we shall focus on a functional McKean-Vlasov SDE (see (1.2) below), which is much more general than the McKean-Vlasov SDE (1.1), and investigate existence of IPMs under much weaker conditions than those imposed in [1, 7, 25]. In particular, we will not assume a monotonicity condition.

Let us point out that we will not address the question of uniqueness of IPMs which clearly requires stronger assumptions than those in our main result.

Before we present our set-up and results, let us introduce some notation. For a subinterval $U \subset \mathbb{R}$, $C(U; \mathbb{R}^d)$ stands for the collection of all continuous functions $f : U \to \mathbb{R}^d$. For a fixed finite number $r_0 > 0$, set $\mathscr{C} := C([-r_0, 0]; \mathbb{R}^d)$, which is endowed with the uniform norm

$$\|f\|_\infty := \sup_{-r_0 \le \theta \le 0} |f(\theta)|, \quad f \in \mathscr{C}.$$

For $f \in C([-r_0, \infty); \mathbb{R}^d)$ and $t \ge 0$, let $f_t \in \mathscr{C}$ be defined by $f_t(\theta) = f(t + \theta), \theta \in [-r_0, 0]$. Often, $(f_t)_{t\ge 0}$ is called the *segment* (or *window*) process corresponding to $(f(t))_{t\ge -r_0}$. For each $p > 0$, let $\mathscr{P}_p(\mathscr{C})$ be the set of all probability measures on $\mathscr{C}$, denoted by $\mathscr{P}(\mathscr{C})$, with finite p-th moment, i.e.,

$$\mathscr{P}_p(\mathscr{C}) = \left\{ \mu \in \mathscr{P}(\mathscr{C}) \middle| \mu(\|\cdot\|_\infty^p) := \int_{\mathscr{C}} \|\xi\|_\infty^p \mu(\mathrm{d}\xi) < \infty \right\},$$

which is a Polish space (see e.g. [23, Theorem 6.18]) under the Wasserstein distance

$$\mathbb{W}_p(\mu,\nu) = \inf_{\pi \in \mathcal{C}(\mu,\nu)} \left(\int_{\mathscr{C}\times\mathscr{C}} \|\xi-\eta\|_\infty^p \pi(\mathrm{d}\xi,\mathrm{d}\eta) \right)^{\frac{1}{1\vee p}},$$

where $\mathcal{C}(\mu,\nu)$ is the set of all probability measures on $\mathscr{C}\times\mathscr{C}$ with marginals μ and ν, respectively. For a random variable ξ, we denote its law by $\mathscr{L}_\xi$ and we write $\xi \sim \mu$ if $\mu = \mathscr{L}_\xi$.

We fix $p \geq 2$ and consider the following functional McKean-Vlasov SDE

$$\mathrm{d}X(t) = b(X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma(X_t, \mathscr{L}_{X_t})\mathrm{d}W(t), \quad t \geq 0, \quad X_0 \sim \mu \in \mathscr{P}_p(\mathscr{C}), \tag{1.2}$$

where

$$b : \mathscr{C} \times \mathscr{P}_p(\mathscr{C}) \to \mathbb{R}^d, \quad \sigma : \mathscr{C} \times \mathscr{P}_p(\mathscr{C}) \to \mathbb{R}^d \otimes \mathbb{R}^m,$$

and $(W(t))_{t\geq0}$ is an m-dimensional Brownian motion on the probability space $(\Omega,\mathscr{F},\mathbb{P})$ with filtration $(\mathscr{F}_t)_{t\geq0}$. We will assume the following hypotheses.

($\mathbf{H_1}$) b and σ are continuous and bounded on bounded subsets of $\mathscr{C}\times\mathscr{P}_p(\mathscr{C})$;

($\mathbf{H_2}$) There is a constant $K > 0$ such that

$$2\langle \xi(0)-\eta(0), b(\xi,\mu)-b(\eta,\nu)\rangle^+ + \|\sigma(\xi,\mu)-\sigma(\eta,\nu)\|_{\mathrm{HS}}^2 \leq K\big\{\|\xi-\eta\|_\infty^2 + \mathbb{W}_p(\mu,\nu)^2\big\}$$

for any $\xi,\eta \in \mathscr{C}, \mu,\nu \in \mathscr{P}_p(\mathscr{C})$, where, for a real number a, $a^+ := \max\{a,0\}$ denotes its positive part, and, for a matrix $A = \big(a_{ij}\big)$, $\|A\|_{\mathrm{HS}} := \big(\sum_{i,j} a_{ij}^2\big)^{1/2}$ is its Hilbert-Schmidt norm.

We will be interested in the existence of an IPM $\pi \in \mathscr{P}_p(\mathscr{C})$ of (1.2), i.e. a probability measure π for which the functional solution $(X_t)_{t\geq0}$ of

$$\mathrm{d}X(t) = b(X_t,\pi)\mathrm{d}t + \sigma(X_t,\pi)\mathrm{d}W(t), \quad t\geq 0, \quad X_0 \sim \pi,$$

satisfies $\mathscr{L}_{X_t} = \pi$ for every $t \geq 0$. In this case we say that (1.2) (or $(X_t)_{t\geq0}$) *admits an IPM*.

To investigate existence of an IPM we do not need to know whether (1.2) has a unique (strong) solution for every initial condition but we still mention that it has been shown in [4, Lemma 3.1] that under ($\mathbf{H_1}$) and ($\mathbf{H_2}$), for any $X_0 \in L^p(\Omega \to \mathscr{C}, \mathscr{F}_0, \mathbb{P})$ with $p \geq 2$, (1.2) has a unique functional solution $(X_t)_{t\geq0}$ and that there exists a nondecreasing positive function $T \mapsto C_T$ such that

$$\mathbb{E}\Big(\sup_{0\leq t\leq T} \|X_t\|_\infty^p\Big) \leq C_T\big(1+\mathbb{E}\|X_0\|_\infty^p\big), \quad T > 0.$$

Clearly, hypotheses ($\mathbf{H_1}$) and ($\mathbf{H_2}$) are insufficient to guarantee existence of an IPM for (1.2). Therefore, we impose the following additional condition which guarantees that the drift b pushes solutions towards $\mathbf{0}$ whenever $\|X_t\|_\infty$ is large. The condition is formulated in terms of constants $\lambda_0, ..., \lambda_5$. For our main result, Theorem 1.2, to hold, these constants will have to satisfy additional constraints.

($\mathbf{H_3}$) There exist a constant $\lambda_1 > 0$ and constants $\lambda_0, \lambda_i \geq 0, i = 2, \cdots, 5$, such that

$$2\langle \xi(0), b(\xi,\mu)\rangle \leq \lambda_0 - \lambda_1|\xi(0)|^2 + \lambda_2\|\xi\|_\infty^2 + \lambda_3\mathbb{W}_p(\mu,\delta_{\mathbf{0}})^2, \tag{1.3}$$

$$\|\sigma(\xi,\mu)\|_{\mathrm{HS}}^2 \leq \lambda_0 + \lambda_4\|\xi\|_\infty^2 + \lambda_5\mathbb{W}_p(\mu,\delta_{\mathbf{0}})^2 \tag{1.4}$$

for any $\xi \in \mathscr{C}$ and $\mu \in \mathscr{P}_p(\mathscr{C})$.

Note that (1.4) actually follows from ($\mathbf{H_2}$) (for suitable λ_0, λ_4 and λ_5).

In the sequel, we will frequently use the following one-sided version of the Burkholder-Davis-Gundy inequality which is a special case of [19, Theorem 4.1.(ii)].

Proposition 1.1. *For any continuous martingale M satisfying $M(0)=0$ and for any $t \geq 0$, we have*

$$\mathbb{E}\Big(\sup_{0\le s\le t} M(s)\Big) \le \chi\, \mathbb{E}\big([M,M]_t^{1/2}\big),$$

where χ (called ν_1 in [19, Theorem 4.1.(ii)]) is the smallest positive root of the confluent hypergeometric function with parameter 1.

The numerical value is $\chi \approx 1.30693....$

For the parameters $\lambda_1 > 0$, and $\lambda_2, \lambda_3, \lambda_4, \lambda_5 \ge 0$ in ($\mathbf{H_3}$), set

$$\begin{aligned}\mathcal{A} := \Big\{(\varepsilon,\alpha,\gamma_1,\gamma_2) \in (0,1)\times(0,\infty)^3 \Big| \alpha - 2\,\mathrm{e}^{\alpha r_0}\big(\kappa_1(\varepsilon,\gamma_1)+\kappa_2(\varepsilon,\gamma_2)\big) > 0,\\ \psi(\varepsilon,\alpha,\gamma_1,\gamma_2) < 0\Big\},\end{aligned}$$

where

$$\begin{aligned}\psi(\varepsilon,\alpha,\gamma_1,\gamma_2) &:= \frac{1-\varepsilon}{\varepsilon}\Big\{\alpha - \frac{p\lambda_1}{2} + \frac{p\gamma_1}{2}\big(\lambda_2+\lambda_4(p-1)\big) + \frac{p\gamma_2}{2}\big(\lambda_3+\lambda_5(p-1)\big)\Big\}\\ &\quad + \frac{\chi^2p^2}{2\varepsilon^2}\big(\lambda_4\gamma_1+\lambda_5\gamma_2\big),\\ \kappa_1(\varepsilon,\gamma_1) &:= \frac{1}{\varepsilon}\big(\lambda_2+\lambda_4((1+\chi^2/\varepsilon)p-1)\big)\Big(\frac{p-2}{p\gamma_1}\Big)^{\frac{p-2}{2}},\\ \kappa_2(\varepsilon,\gamma_2) &:= \frac{1}{\varepsilon}\big(\lambda_3+\lambda_5((1+\chi^2/\varepsilon)p-1)\big)\Big(\frac{p-2}{p\gamma_2}\Big)^{\frac{p-2}{2}}.\end{aligned} \tag{1.5}$$

Here and below, we interpret expressions of the form 0^0 as 1 (when $p=2$).

Our main result is the following theorem.

Theorem 1.2. *Fix $p \ge 2$ and assume ($\mathbf{H_1}$), ($\mathbf{H_2}$), and ($\mathbf{H_3}$). Then,* (1.2) *has an IPM $\pi \in \mathscr{P}_p(\mathscr{C})$ if λ_1 is large enough. More precisely, if $\mathcal{A} \neq \emptyset$, then* (1.2) *admits an IPM $\pi \in \mathscr{P}_p(\mathscr{C})$.*

Remark 1.3. Note that, for fixed $\lambda_2,\lambda_3,\lambda_4,\lambda_5 \ge 0$ and $p>2$, the set $\mathcal{A}$ is non-empty if λ_1 is sufficiently large: to see this, fix $(\varepsilon,\alpha)\in(0,1)\times(0,\infty)$. Then choose γ_1,γ_2 so large that $\alpha - 2\,\mathrm{e}^{\alpha r_0}\big(\kappa_1(\varepsilon,\gamma_1)+\kappa_2(\varepsilon,\gamma_2)\big) > 0$. Finally choose $\lambda_1>0$ so large that $\psi(\varepsilon,\alpha,\gamma_1,\gamma_2)<0$.

For $p=2$ this argument fails (in general) but it is still true that for fixed $\lambda_2,\lambda_3,\lambda_4,\lambda_5 \ge 0$ and $p=2$ and λ_1 sufficiently large (1.2) admits an IPM $\pi\in\mathscr{P}_2(\mathscr{C})$. To see this, replace $p=2$ by some $q>p$ and observe that conditions ($\mathbf{H_1}$), ($\mathbf{H_2}$), and ($\mathbf{H_3}$) still hold true for q instead of p with the same constants, provided we restrict the domains of b and σ accordingly. By arguing as above, we see that the set $\mathcal{A}$ defined in terms of q instead of p is non-empty for λ_1 sufficiently large, so Theorem 1.2 even guarantees the existence of an IPM $\pi\in\mathscr{P}_q(\mathscr{C})\big(\subset\mathscr{P}_2(\mathscr{C})\big)$.

Remark 1.4. Let us now see how λ_1 has to grow as a function of p for an IPM to exist. Fix $(\varepsilon,\alpha,\gamma_1,\gamma_2)\in(0,1)\times(0,\infty)\times(1,\infty)^2$ and fix $\lambda_2,\lambda_3,\lambda_4,\lambda_5\ge0$. Then κ_1 and κ_2 converge to 0 as $p\to\infty$ and it is clear that there exists some $\beta>0$ such that for $\lambda_1>\beta p$ we have $\psi(\varepsilon,\alpha,\gamma_1,\gamma_2)<0$ for all $p\ge2$. The assumption that λ_1 has to grow at least linearly in p to obtain an IPM in $\mathscr{P}_p(\mathscr{C})$ cannot be avoided in general (not even in the case of a one-dimensional SDE without delay and without dependence on the law of the process) as the example in Remark 1.6 shows.

Remark 1.5. Theorem 1.2 provides a stronger result than the main result in [12] even in the case when neither b nor σ depends on the law $\mathscr{L}_{X_t}$ (i.e., $\lambda_3=\lambda_5=0$) since the main result in [12] requires that the drift is superlinear while we allow that the drift has linear (negative) growth of sufficient strength.

Remark 1.6. Let $d=m=1$, $\sigma>0$, $\lambda_1>0$ and consider the SDE

$$\mathrm{d}X(t)=-\lambda_1 X(t)\,\mathrm{d}t+\sigma\sqrt{|X(t)|^2+1}\,\mathrm{d}W(t)$$

(without delay and without dependence on the law of the process). The corresponding real-valued Markov process admits an invariant measure, the so-called *speed measure*, with density

$$f(x)=\frac{1}{x^2+1}\exp\Big\{-2\int_0^{|x|}\frac{\lambda_1 u}{\sigma^2(u^2+1)}\,\mathrm{d}u\Big\}=\big(x^2+1\big)^{-\frac{\lambda_1}{\sigma^2}-1},\quad x\in\mathbb{R}$$

(see [15, p.343 & p.353]). This measure is finite and hence, after normalizing, an IPM. It has a finite p-th moment if and only if $p-2\frac{\lambda_1}{\sigma^2}<1$ or, equivalently, $\lambda_1>\frac{1}{2}(p-1)\sigma^2$ showing that, in general, λ_1 has to increase at least linearly as a function of p to obtain an IPM in $\mathscr{P}_p(\mathscr{C})$. Of course, in this example, an IPM exists for every $\sigma>0$ and every $\lambda_1>0$ but it does not have a finite p-th moment when λ_1 is too small.

The remainder of this paper is organized as follows. In Section 2, we prepare several auxiliary lemmas which are crucial for the proof Theorem 1.2. Section 3 is devoted to completing the proof of Theorem 1.2.

2 Preliminary Lemmas

Under $(\mathbf{H_1})$ and $(\mathbf{H_2})$, for fixed $\mu,\nu\in\mathscr{P}_p(\mathscr{C})$, [24, Theorem 2.3] shows that the following frozen SDE with memory

$$\mathrm{d}X^{\mu,\nu}(t)=b(X_t^{\mu,\nu},\nu)\mathrm{d}t+\sigma(X_t^{\mu,\nu},\nu)\mathrm{d}W(t),\ \ t\geq 0,\ \ X_0^{\mu,\nu}\sim\mu \tag{2.1}$$

has a unique functional solution $(X_t^{\mu,\nu})_{t\geq 0}$ provided that $X_0^{\mu,\nu}$ is $\mathscr{F}_0$-measurable. Further, $(X_t^{\cdot,\nu})_{t\geq 0}$ is a Markov process which is even Feller, i.e., the corresponding Markov semigroup $(P_t^\nu)_{t\geq 0}$ on $\mathscr{C}$ maps $C_b(\mathscr{C})$, the set of bounded continuous functions on $\mathscr{C}$, to $C_b(\mathscr{C})$; see e.g. [12, Proposition 3.1].

Lemma 2.1. *Assume $(\mathbf{H_1})$ and $(\mathbf{H_2})$. Then, for any $\mu,\nu_1,\nu_2\in\mathscr{P}_p(\mathscr{C})$,*

$$\mathbb{E}\|X_t^{\mu,\nu_1}-X_t^{\mu,\nu_2}\|_\infty^p\leq C\mathbb{W}_p(\nu_1,\nu_2)^p t\mathrm{e}^{Ct},\quad t\geq 0, \tag{2.2}$$

where $X_0^{\mu,\nu_1}=X_0^{\mu,\nu_2}\sim\mu$ and $C:=2(\chi^2+1)Kp^2$.

Proof. Let $\Psi(t)=X^{\mu,\nu_1}(t)-X^{\mu,\nu_2}(t)$, $t\geq -r_0$, and

$$M(t)=p\int_0^t\big|\Psi(s)\big|^{p-2}\big\langle\Psi(s),\big(\sigma(X_s^{\mu,\nu_1},\nu_1)-\sigma(X_s^{\mu,\nu_2},\nu_2)\big)\,\mathrm{d}W(s)\big\rangle,\quad t\geq 0.$$

For $N\in\mathbb{N}$, define $\tau_N=\inf\{s\geq 0:\ |\Psi(s)|\geq N\}$, $\Psi^N(t)=\Psi(t\wedge\tau_N)$, and $M_N(t)=M(t\wedge\tau_N)$. Note that M_N is a martingale. By Itô's formula, it follows from $(\mathbf{H_2})$ that

$$\begin{aligned}
\mathrm{d}\big|\Psi(t)\big|^p&\leq\frac{p}{2}\big|\Psi(t)\big|^{p-2}\big\{2\big\langle\Psi(t),b(X_t^{\mu,\nu_1},\nu_1)-b(X_t^{\mu,\nu_2},\nu_2)\big\rangle\\
&\qquad+(p-1)\big\|\sigma(X_t^{\mu,\nu_1},\nu_1)-\sigma(X_t^{\mu,\nu_2},\nu_2)\big\|_{\mathrm{HS}}^2\big\}\,\mathrm{d}t+\mathrm{d}M(t)\\
&\leq\frac{p^2K}{2}\big|\Psi(t)\big|^{p-2}\big\{\|\Psi_t\|_\infty^2+\mathbb{W}_p(\nu_1,\nu_2)^2\big\}\,\mathrm{d}t+\mathrm{d}M(t)\\
&\leq pK\big\{(p-1)\|\Psi_t\|_\infty^p+\mathbb{W}_p(\nu_1,\nu_2)^p\big\}\,\mathrm{d}t+\mathrm{d}M(t),
\end{aligned}$$

where we used Young's inequality in the last step. Therefore, due to $X_0^{\mu,\nu_1}=X_0^{\mu,\nu_2}$, we have

$$\mathbb{E}\|\Psi_t^N\|_\infty^p\leq pK\int_0^t\big\{(p-1)\mathbb{E}\|\Psi_s^N\|_\infty^p+\mathbb{W}_p(\nu_1,\nu_2)^p\big\}\,\mathrm{d}s+\mathbb{E}\Big[\sup_{(t-r_0)^+\leq s\leq t}M_N(s)\Big]. \tag{2.3}$$

Next, Proposition 1.1, $(\mathbf{H_2})$, and Young's inequality yield

$$\begin{aligned}
\mathbb{E}\Big[\sup_{(t-r_0)^+\le s\le t} M_N(s)\Big] &= \mathbb{E}\Big[\sup_{(t-r_0)^+\le s\le t}\Big(M_N(s)-M_N\big((t-r_0)^+\big)\Big)\Big]\\
&\le \chi p\,\mathbb{E}\bigg[\bigg(\int_{(t-r_0)^+}^t \mathbb{1}_{[0,\tau_N)}(s)\big|\Psi^N(s)\big|^{2(p-1)}\big\|\sigma(X_s^{\mu,\nu_1},\nu_1)-\sigma(X_s^{\mu,\nu_2},\nu_2)\big\|_{\mathrm{HS}}^2\,\mathrm{d}s\bigg)^{1/2}\bigg]\\
&\le \chi p\sqrt{K}\,\mathbb{E}\bigg[\bigg(\|\Psi_t^N\|_\infty^p\int_{(t-r_0)^+}^t \mathbb{1}_{[0,\tau_N)}(s)\big|\Psi^N(s)\big|^{p-2}\big\{\|\Psi_s^N\|_\infty^2+\mathbb{W}_p(\nu_1,\nu_2)^2\big\}\,\mathrm{d}s\bigg)^{1/2}\bigg]\\
&\le \chi p\sqrt{2K}\mathbb{E}\bigg[\bigg(\|\Psi_t^N\|_\infty^p\int_0^t\big\{\|\Psi_s^N\|_\infty^p+\mathbb{W}_p(\nu_1,\nu_2)^p\big\}\,\mathrm{d}s\bigg)^{1/2}\bigg]\\
&\le \frac12\mathbb{E}\|\Psi_t^N\|_\infty^p+\chi^2Kp^2\int_0^t\big\{\mathbb{E}\|\Psi_s^N\|_\infty^p+\mathbb{W}_p(\nu_1,\nu_2)^p\big\}\,\mathrm{d}s.
\end{aligned}$$

Inserting the estimate above back into (2.3) implies

$$\mathbb{E}\|\Psi_t^N\|_\infty^p\le 2(\chi^2+1)Kp^2\int_0^t\big\{\mathbb{E}\|\Psi_s^N\|_\infty^p+\mathbb{W}_p(\nu_1,\nu_2)^p\big\}\,\mathrm{d}s.$$

Applying Gronwall's inequality and letting $N\to\infty$ implies (2.2). □

Before we move on, let us introduce some additional notation. Let

$$\mathcal{U}=\Big\{(\varepsilon,\alpha,\gamma_1,\gamma_2)\in(0,1)\times(0,\infty)^3\Big|\,2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}<\alpha,\quad \psi(\varepsilon,\alpha,\gamma_1,\gamma_2)<0\Big\},\tag{2.4}$$

where the functions κ_1 and ψ were introduced in (1.5). Note that $\mathcal{A}\subseteq\mathcal{U}$. If $\mathcal{U}\neq\emptyset$, then there exist $(\varepsilon,\alpha,\gamma_1,\gamma_2,\gamma_3)\in(0,1)\times(0,\infty)^4$ such that

$$\phi(\varepsilon,\alpha,\gamma_1,\gamma_2,\gamma_3):=\psi(\varepsilon,\alpha,\gamma_1,\gamma_2)+\frac12\Big(\frac{1-\varepsilon}{\varepsilon}+\frac{\chi^2}{\varepsilon^2}\Big)p^2\gamma_3\lambda_0<0,\tag{2.5}$$

and

$$\alpha>2\mathrm{e}^{\alpha r_0}\kappa_1(\varepsilon,\gamma_1).\tag{2.6}$$

Furthermore, we set

$$\kappa_3(\varepsilon,\gamma_3):=\frac{p\lambda_0}{\varepsilon}\big(1+\chi^2/\varepsilon\big)\Big(\frac{p-2}{p\gamma_3}\Big)^{\frac{p-2}{2}}.\tag{2.7}$$

Lemma 2.2. *Assume* $(\mathbf{H_1})$, $(\mathbf{H_2})$, *and* $(\mathbf{H_3})$. *If* $\mathcal{U}\neq\emptyset$, *then for any* $\mu,\nu\in\mathscr{P}_p(\mathscr{C})$ *and* $t>0$,

$$\begin{aligned}
\mathbb{E}\|X_t^{\mu,\nu}\|_\infty^p\le{}&\frac{2\mathrm{e}^{\alpha r_0}}{\alpha-2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}}\big(\kappa_3(\varepsilon,\gamma_3)+\kappa_2(\varepsilon,\gamma_2)\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^p\big)\\
&+\mathrm{e}^{\alpha r_0}\big(1+4/\varepsilon+4\kappa_1(\varepsilon,\gamma_1)r_0\mathrm{e}^{\alpha r_0}\big)\mu(\|\cdot\|_\infty^p)\mathrm{e}^{-(\alpha-2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0})t},
\end{aligned}\tag{2.8}$$

where $\kappa_1(\cdot)$ *and* $\kappa_2(\cdot)$ *were introduced in* (1.5).

Proof. Since $\mathcal{U}\neq\emptyset$, there exist $(\varepsilon,\alpha,\gamma_1,\gamma_2,\gamma_3)\in(0,1)\times(0,\infty)^4$ such that (2.5) and (2.6) hold. In the sequel, we shall stipulate the parameters $(\varepsilon,\alpha,\gamma_1,\gamma_2,\gamma_3)\in(0,1)\times(0,\infty)^4$ satisfying (2.5) and (2.6) and fix $\mu,\nu\in\mathscr{P}_p(\mathscr{C})$ for $p\ge 2$. Below, for notational simplicity, we shall write $X(t)$ and X_t instead of $X^{\mu,\nu}(t)$ and $X_t^{\mu,\nu}$, respectively. By Itô's formula, it follows from $(\mathbf{H_3})$ that

$$\begin{aligned}
\mathrm{d}\big(\mathrm{e}^{\alpha t}|X(t)|^p\big) &\le \mathrm{e}^{\alpha t}\Big\{\alpha|X(t)|^p + \frac{p}{2}|X(t)|^{p-2}\big(2\langle X(t), b(X_t,\nu)\rangle \\
&\quad + (p-1)\|\sigma(X_t,\nu)\|_{\mathrm{HS}}^2\big)\Big\}\,\mathrm{d}t + \mathrm{d}N(t) \\
&\le \mathrm{e}^{\alpha t}\Big\{\big(\alpha - p\lambda_1/2\big)\big|X(t)\big|^p + \frac{p^2\lambda_0}{2}\big|X(t)\big|^{p-2} \\
&\quad + \frac{p}{2}\big(\lambda_2 + \lambda_4(p-1)\big)\|X_t\|_\infty^2\big|X(t)\big|^{p-2} \\
&\quad + \frac{p}{2}\big(\lambda_3 + \lambda_5(p-1)\big)\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^2\big|X(t)\big|^{p-2}\Big\}\,\mathrm{d}t + \mathrm{d}N(t),
\end{aligned} \tag{2.9}$$

where

$$N(t) := p\int_0^t \mathrm{e}^{\alpha s}|X(s)|^{p-2}\langle X(s), \sigma(X_s,\nu)\,\mathrm{d}W(s)\rangle, \quad t\ge 0.$$

Applying Young's inequality in case $p>2$, we obtain

$$\begin{aligned}
\|X_t\|_\infty^2\big|X(t)\big|^{p-2} &\le \gamma_1\big|X(t)\big|^p + \frac{2}{p}\Big(\frac{p-2}{p\gamma_1}\Big)^{\frac{p-2}{2}}\|X_t\|_\infty^p, \\
\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^2\big|X(t)\big|^{p-2} &\le \gamma_2\big|X(t)\big|^p + \frac{2}{p}\Big(\frac{p-2}{p\gamma_2}\Big)^{\frac{p-2}{2}}\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^p, \\
\big|X(t)\big|^{p-2} &\le \gamma_3\big|X(t)\big|^p + \frac{2}{p}\Big(\frac{p-2}{p\gamma_3}\Big)^{\frac{p-2}{2}}.
\end{aligned} \tag{2.10}$$

Note that these inequalities also hold in case $p=2$ (due to our convention that $0^0=1$). Thus, we derive from (2.9) that

$$\begin{aligned}
\mathrm{d}\big(\mathrm{e}^{\alpha t}|X(t)|^p\big) &\le \mathrm{e}^{\alpha t}\Big\{\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)\big|X(t)\big|^p + \big(\lambda_2+\lambda_4(p-1)\big)\Big(\frac{p-2}{p\gamma_1}\Big)^{\frac{p-2}{2}}\|X_t\|_\infty^p \\
&\quad + p\lambda_0\Big(\frac{p-2}{p\gamma_3}\Big)^{\frac{p-2}{2}} + \big(\lambda_3+\lambda_5(p-1)\big)\Big(\frac{p-2}{p\gamma_2}\Big)^{\frac{p-2}{2}}\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^p\Big\}\,\mathrm{d}t + \mathrm{d}N(t),
\end{aligned}$$

in which

$$\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3) := \alpha - \frac{p\lambda_1}{2} + \frac{p^2\lambda_0\gamma_3}{2} + \frac{p\gamma_1}{2}\big(\lambda_2+\lambda_4(p-1)\big) + \frac{p\gamma_2}{2}\big(\lambda_3+\lambda_5(p-1)\big).$$

Therefore, for $t\ge s\ge (t-r_0)^+$,

$$\begin{aligned}
\mathrm{e}^{\alpha s}|X(s)|^p + \big(-\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)\big)\int_0^s \mathrm{e}^{\alpha u}|X(u)|^p\,\mathrm{d}u &\le |X(0)|^p + \Gamma(s) + N(s) \\
&\le |X(0)|^p + \Gamma(t) + \sup_{(t-r_0)^+\le r\le t} N(r),
\end{aligned}$$

where

$$\begin{aligned}
\Gamma(s) := \int_0^s \mathrm{e}^{\alpha u}\Big(&\big(\lambda_2+\lambda_4(p-1)\big)\Big(\frac{p-2}{p\gamma_1}\Big)^{\frac{p-2}{2}}\|X_u\|_\infty^p + p\lambda_0\Big(\frac{p-2}{p\gamma_3}\Big)^{\frac{p-2}{2}} \\
&+ \big(\lambda_3+\lambda_5(p-1)\big)\Big(\frac{p-2}{p\gamma_2}\Big)^{\frac{p-2}{2}}\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^p\Big)\,\mathrm{d}u.
\end{aligned}$$

Since $\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)<0$ due to (2.5), and since $\varepsilon\in(0,1)$, we have

$$\begin{aligned}
&\sup_{(t-r_0)^+\le r\le t}\big(\mathrm{e}^{\alpha r}|X(r)|^p\big)\le |X(0)|^p+\Gamma(t)+\sup_{(t-r_0)^+\le r\le t}N(r)\\
&=|X(0)|^p+\Gamma(t)+\sup_{(t-r_0)^+\le r\le t}N(r)-\big(-\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)\big)\int_0^t\mathrm{e}^{\alpha u}|X(u)|^p\,\mathrm{d}u\\
&\qquad+\big(-\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)\big)\int_0^t\mathrm{e}^{\alpha u}|X(u)|^p\,\mathrm{d}u\\
&\le\frac{1}{\varepsilon}\Big(|X(0)|^p+\Gamma(t)+\sup_{(t-r_0)^+\le r\le t}N(r)-\big(-\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)\big)\int_0^t\mathrm{e}^{\alpha u}|X(u)|^p\,\mathrm{d}u\Big)\\
&\qquad+\big(-\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)\big)\int_0^t\mathrm{e}^{\alpha u}|X(u)|^p\,\mathrm{d}u\\
&=\frac{1}{\varepsilon}\Big(|X(0)|^p+\Gamma(t)+\sup_{(t-r_0)^+\le r\le t}N(r)\Big)-\Big(\frac{1-\varepsilon}{\varepsilon}\Big)\big(-\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)\big)\int_0^t\mathrm{e}^{\alpha u}|X(u)|^p\,\mathrm{d}u.
\end{aligned}$$

In the following computation, it is at first not clear that the expected values are finite. This can, however, be shown by a stopping argument like in the proof of Lemma 2.1.

Define

$$\begin{aligned}
g(t):=\mathbb{E}\Big(\sup_{(t-r_0)^+\le r\le t}\big(\mathrm{e}^{\alpha r}|X(r)|^p\big)\Big)&\le\frac{1}{\varepsilon}\mathbb{E}|X(0)|^p+\frac{1}{\varepsilon}\mathbb{E}\Gamma(t)+\frac{1}{\varepsilon}\mathbb{E}\Big(\sup_{(t-r_0)^+\le r\le t}N(r)\Big)\\
&\quad+\frac{1-\varepsilon}{\varepsilon}\Theta(\alpha,\gamma_1,\gamma_2,\gamma_3)\int_{(t-r_0)^+}^t\mathrm{e}^{\alpha s}\mathbb{E}|X(s)|^p\,\mathrm{d}s.
\end{aligned}\tag{2.11}$$

Next, applying Proposition 1.1 and Young's inequality, we obtain

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{(t-r_0)^+\le r\le t}N(r)\Big)\\
&\le\chi p\mathbb{E}\bigg(\sup_{(t-r_0)^+\le r\le t}\big(\mathrm{e}^{\alpha r}|X(r)|^p\big)\int_{(t-r_0)^+}^t\mathrm{e}^{\alpha s}|X(s)|^{p-2}\|\sigma(X_s,\nu)\|_{\mathrm{HS}}^2\,\mathrm{d}s\bigg)^{1/2}\\
&\le\chi p\bigg\{\frac{\varepsilon}{2\chi p}g(t)+\frac{\chi p}{2\varepsilon}\mathbb{E}\int_{(t-r_0)^+}^t\mathrm{e}^{\alpha s}|X(s)|^{p-2}\|\sigma(X_s,\nu)\|_{\mathrm{HS}}^2\,\mathrm{d}s\bigg\}\\
&\le\frac{\varepsilon}{2}g(t)+\frac{\chi^2p^2}{2\varepsilon}\mathbb{E}\int_{(t-r_0)^+}^t\mathrm{e}^{\alpha s}|X(s)|^{p-2}\big(\lambda_0+\lambda_4\|X_s\|_\infty^2+\lambda_5\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^2\big)\,\mathrm{d}s\\
&\le\frac{\varepsilon}{2}g(t)+\frac{\chi^2p^2}{2\varepsilon}\int_{(t-r_0)^+}^t\mathrm{e}^{\alpha s}\bigg(\big(\lambda_0\gamma_3+\lambda_4\gamma_1+\lambda_5\gamma_2\big)\mathbb{E}|X(s)|^p+\frac{2\lambda_4}{p}\Big(\frac{p-2}{p\gamma_1}\Big)^{\frac{p-2}{2}}\mathbb{E}\|X_s\|_\infty^p\\
&\quad+\frac{2\lambda_5}{p}\Big(\frac{p-2}{p\gamma_2}\Big)^{\frac{p-2}{2}}\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^p+\frac{2\lambda_0}{p}\Big(\frac{p-2}{p\gamma_3}\Big)^{\frac{p-2}{2}}\bigg)\,\mathrm{d}s,
\end{aligned}$$

where the last inequality is due to (2.10). Whence, we deduce from (2.11) that

$$\begin{aligned}
g(t)\le\frac{1}{\varepsilon}\mathbb{E}|X(0)|^p&+\big(\kappa_3(\varepsilon,\gamma_3)+\kappa_2(\varepsilon,\gamma_2)\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^p\big)\int_0^t\mathrm{e}^{\alpha s}\,\mathrm{d}s\\
&+\phi(\varepsilon,\alpha,\gamma_1,\gamma_2,\gamma_3)\int_{(t-r_0)^+}^t\mathrm{e}^{\alpha s}\mathbb{E}|X(s)|^p\,\mathrm{d}s\\
&+\kappa_1(\varepsilon,\gamma_1)\int_0^t\mathrm{e}^{\alpha s}\mathbb{E}\|X_s\|_\infty^p\,\mathrm{d}s+\frac{1}{2}g(t),
\end{aligned}$$

where $\phi(\varepsilon,\alpha_1,\gamma_1,\gamma_2,\gamma_3)$ was given in (2.5). Since $\phi(\varepsilon,\alpha_1,\gamma_1,\gamma_2,\gamma_3)<0$, we obtain

$$\begin{aligned}g(t)\le{}&\frac{2}{\varepsilon}\mathbb{E}|X(0)|^p+\frac{2}{\alpha}\big(\kappa_3(\varepsilon,\gamma_3)+\kappa_2(\varepsilon,\gamma_2)\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^p\big)(\mathrm{e}^{\alpha t}-1)\\&+2\kappa_1(\varepsilon,\gamma_1)\int_0^t\mathrm{e}^{\alpha s}\mathbb{E}\Big(\sup_{s-r_0\le u\le s}|X(u)|^p\Big)\,\mathrm{d}s.\end{aligned}$$

This, together with

$$\begin{aligned}\int_0^t\mathrm{e}^{\alpha s}\mathbb{E}\Big(\sup_{s-r_0\le u\le s}|X(u)|^p\Big)\,\mathrm{d}s&\le\mathrm{e}^{\alpha r_0}\int_0^t\mathbb{E}\Big(\sup_{s-r_0\le u\le s}(\mathrm{e}^{\alpha u}|X(u)|^p)\Big)\,\mathrm{d}s\\&\le r_0\mathrm{e}^{\alpha r_0}\mathbb{E}\|X_0\|_\infty^p+\mathrm{e}^{\alpha r_0}\int_0^t g(s)\,\mathrm{d}s,\end{aligned}$$

yields

$$g(t)\le\Phi(t)+2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}\int_0^t g(u)\,\mathrm{d}u,$$

where

$$\begin{aligned}\Phi(t):&=\frac{2}{\alpha}\big(\kappa_3(\varepsilon,\gamma_3)+\kappa_2(\varepsilon,\gamma_2)\mathbb{W}_p(\nu,\delta_{\mathbf{0}})^p\big)\mathrm{e}^{\alpha t}\\&\quad+2\Big(\frac{1}{\varepsilon}\mathbb{E}|X(0)|^p+\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}r_0\mathbb{E}\|X_0\|_\infty^p\Big)\\&=:A\mathrm{e}^{\alpha t}+B.\end{aligned}\tag{2.12}$$

Applying Gronwall's inequality, we get

$$g(t)\le\Phi(t)+2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}\int_0^t\Phi(s)\mathrm{e}^{2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}(t-s)}\,\mathrm{d}s.\tag{2.13}$$

Observe that

$$\begin{aligned}\mathbb{E}\|X_t\|_\infty^p&\le\mathrm{e}^{-\alpha(t-r_0)}\mathbb{E}\Big(\sup_{t-r_0\le s\le t}(\mathrm{e}^{\alpha s}|X(s)|^p)\Big)\\&\le\mathrm{e}^{-\alpha(t-r_0)}\mathbb{E}\Big(\sup_{t-r_0\le s\le(t-r_0)^+}(\mathrm{e}^{\alpha s}|X(s)|^p)\vee\sup_{(t-r_0)^+\le s\le t}(\mathrm{e}^{\alpha s}|X(s)|^p)\Big)\\&\le\mathrm{e}^{-\alpha(t-r_0)}\big(\mathbb{E}\|X_0\|_\infty^p+g(t)\big).\end{aligned}$$

This, together with (2.13) as well as $\alpha>2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}$ in view of (2.6), implies

$$\begin{aligned}\mathbb{E}\|X_t\|_\infty^p&\le\mathrm{e}^{-\alpha(t-r_0)}\Big(\mathbb{E}\|X_0\|_\infty^p+A\mathrm{e}^{\alpha t}+B\\&\qquad+2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}\int_0^t\big(A\mathrm{e}^{\alpha s}+B\big)\mathrm{e}^{2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}(t-s)}\,\mathrm{d}s\Big)\\&\le\frac{A\alpha\mathrm{e}^{\alpha r_0}}{\alpha-2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}}+\big(\mathbb{E}\|X_0\|_\infty^p+B\big)\mathrm{e}^{-\alpha(t-r_0)}+B\mathrm{e}^{\alpha r_0}\mathrm{e}^{-(\alpha-2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0})t}\\&\le\frac{A\alpha\mathrm{e}^{\alpha r_0}}{\alpha-2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0}}+\mathrm{e}^{\alpha r_0}\big(\mathbb{E}\|X_0\|_\infty^p+2B\big)\mathrm{e}^{-(\alpha-2\kappa_1(\varepsilon,\gamma_1)\mathrm{e}^{\alpha r_0})t}.\end{aligned}$$

As a result, (2.8) follows by inserting the expressions of A and B, given in (2.12). □

Our next goal is to find a suitable compact subset $\mathcal{K}$ of $\mathscr{P}_p(\mathscr{C})$ to which we can apply Kakutani's fixed point theorem to obtain an IPM of the Markov process generated by (1.2). Since a subset of $\mathscr{P}_p(\mathscr{C})$ which is relatively compact with respect to the weak (sometimes called *narrow*) topology on the space of probability measures on $\mathscr{C}$ is not necessarily relatively compact in $\mathscr{P}_p(\mathscr{C})$ endowed with the $\mathbb{W}_p$-Wasserstein distance, we have to work with different values of p in the following. So far, we fixed the value of $p \geq 2$ and we will continue to regard p as fixed. Observe however, that if $(\mathbf{H_1})$, $(\mathbf{H_2})$, and $(\mathbf{H_3})$ hold for a given $p \geq 2$, then they also hold for p replaced by $q > p$ (with the same constants) by restricting the domain of definition of b and σ accordingly. Now we assume that the set $\mathcal{A}$ (defined in terms of p) is non-empty and we fix $(\varepsilon, \alpha, \gamma_1, \gamma_2) \in \mathcal{A}$; then the same quadruple is also in the set $\mathcal{A}$ defined with respect to q instead of p provided that $q - p > 0$ is sufficiently small. This holds true since the functions ψ, κ_1 and κ_2 depend continuously on p. From now on we fix such a quadruple and $q > p$. Then, in particular, Lemma 2.2 holds for q. To avoid confusion, we write $\widetilde{\kappa}_1$ if p is replaced by q and similarly for other functions.

For

$$M \geq M_0 := \frac{2\widetilde{\kappa}_3(\varepsilon, \gamma_3)\mathrm{e}^{\alpha r_0}}{\alpha - 2\mathrm{e}^{\alpha r_0}(\widetilde{\kappa}_1(\varepsilon, \gamma_1) + \widetilde{\kappa}_2(\varepsilon, \gamma_2))} \vee \mathbb{E}\|X_{r_0}^{\delta_0, \delta_0}\|_\infty^q, \tag{2.14}$$

let

$$\mathcal{M}_M = \big\{\mu \in \mathscr{P}_q(\mathscr{C}) : \mu\big(\|\cdot\|_\infty^q\big) \leq M\big\},$$

and

$$\mathcal{K}_0 := \mathcal{M}_M \cap \big\{\mathscr{L}_{X_{r_0}^{\mu,\nu}}, \mu, \nu \in \mathcal{M}_M\big\}.$$

Note that $\mathcal{M}_M$ is closed in $\mathscr{P}_q(\mathscr{C})$ and hence also in $\mathscr{P}_p(\mathscr{C})$ and that $\mathcal{K}_0 \neq \emptyset$ since $\mathscr{L}_{X_{r_0}^{\delta_0,\delta_0}} \in \mathcal{K}_0$. For $\mu, \nu \in \mathscr{P}_p(\mathscr{C})$, write $\mu_t^\nu = \mathscr{L}_{X_t^{\mu,\nu}}$ for notational brevity and set

$$\Lambda_\nu := \big\{\mu \in \mathscr{P}_p(\mathscr{C}) : \mu_t^\nu = \mu, \quad t \geq 0\big\},$$

which is the collection of all IPMs of $(X_t^{\cdot,\nu})_{t\geq 0}$ solving (2.1). We will see in the next lemma, that we automatically obtain $\Lambda_\nu \subseteq \mathscr{P}_q(\mathscr{C})$ if $\nu \in \mathcal{M}_M$.

Lemma 2.3. *Assume $(\mathbf{H_1})$, $(\mathbf{H_2})$, $(\mathbf{H_3})$ and assume that $\mathcal{A} \neq \emptyset$. Choose $q > p$ and M_0 as above and fix $M \geq M_0$. Then $\Lambda_\nu \subseteq \mathcal{K}_0$ for any $\nu \in \mathcal{M}_M$.*

Proof. Fix $\nu \in \mathcal{M}_M$. We write $(X_t^{\xi,\nu})_{t\geq 0}$ in lieu of $(X_t^{\mu,\nu})_{t\geq 0}$ when the initial distribution is $\mu = \delta_\xi$ for $\xi \in \mathscr{C}$. Due to our assumption on q, we can apply Lemma 2.2 with q and obtain, for each $\rho_\nu \in \Lambda_\nu$, $N > 0$, and a suitable function $\beta(t)$, $t \geq 0$, decreasing to 0, that

$$\begin{aligned}
\int_{\mathscr{C}} \big(\|\xi\|_\infty^q \wedge N\big)\, \rho_\nu(\mathrm{d}\xi) &= \int_{\mathscr{C}} \mathbb{E}\big(\|X_t^{\xi,\nu}\|_\infty^q \wedge N\big)\, \rho_\nu(\mathrm{d}\xi) \\
&\leq \int_{\mathscr{C}} \big(N \wedge \mathbb{E}\|X_t^{\xi,\nu}\|_\infty^q\big)\, \rho_\nu(\mathrm{d}\xi) \\
&\leq \frac{2\mathrm{e}^{\alpha r_0}}{\alpha - 2\widetilde{\kappa}_1(\varepsilon, \gamma_1)\mathrm{e}^{\alpha r_0}} \big(\widetilde{\kappa}_3(\varepsilon, \gamma_3) + \widetilde{\kappa}_2(\varepsilon, \gamma_2)\mathbb{W}_q(\nu, \delta_{\mathbf{0}})^q\big) \\
&\quad + \int_{\mathscr{C}} N \wedge \Big(\beta(t)\|\xi\|_\infty^q\Big)\, \rho_\nu(\mathrm{d}\xi),
\end{aligned}$$

where the identity is due to the invariance of ρ_ν and the first inequality holds true by Jensen's inequality since $x \mapsto N \wedge x$ is a concave function. Then, taking $t \to \infty$ and using the Lebesgue dominated convergence theorem, one has

$$\int_{\mathscr{C}} \big(\|\xi\|_\infty^q \wedge N\big)\, \rho_\nu(\mathrm{d}\xi) \leq \frac{2\mathrm{e}^{\alpha r_0}}{\alpha - 2\widetilde{\kappa}_1(\varepsilon, \gamma_1)\mathrm{e}^{\alpha r_0}} \big(\widetilde{\kappa}_3(\varepsilon, \gamma_3) + \widetilde{\kappa}_2(\varepsilon, \gamma_2)\mathbb{W}_q(\nu, \delta_{\mathbf{0}})^q\big).$$

Applying the monotone convergence theorem, we get

$$\int_{\mathscr{C}} \|\xi\|_\infty^q\, \rho_\nu(\mathrm{d}\xi) \leq \frac{2\mathrm{e}^{\alpha r_0}}{\alpha - 2\widetilde{\kappa}_1(\varepsilon, \gamma_1)\mathrm{e}^{\alpha r_0}} \big(\widetilde{\kappa}_3(\varepsilon, \gamma_3) + \widetilde{\kappa}_2(\varepsilon, \gamma_2)\mathbb{W}_q(\nu, \delta_{\mathbf{0}})^q\big) =: C(\nu).$$

According to (2.14), we infer $C(\nu) \leq M$ so that $\rho_\nu \in \mathcal{M}_M$ for $\nu \in \mathcal{M}_M$ and therefore $\Lambda_\nu \subseteq \mathcal{M}_M$. On the other hand, according to the structure of Λ_ν, one has

$$\Lambda_\nu \subseteq \big\{\mathscr{L}_{X_{r_0}^{\mu,\nu}}, \mu, \nu \in \mathcal{M}_M\big\}.$$

Thus, we arrive at $\Lambda_\nu \subseteq \mathcal{K}_0$ and the desired assertion follows. □

Lemma 2.4. *Under the same assumptions as in the previous lemma the closure $\overline{\mathcal{K}_0}$ of the set $\mathcal{K}_0$ in $\mathscr{P}_p(\mathscr{C})$ is a compact subset of $\mathscr{P}_p(\mathscr{C})$.*

Proof. The set $\overline{\mathcal{K}_0}$ is a closed subset of $\mathscr{P}_p(\mathscr{C})$, so it suffices to show that $\big\{\mathscr{L}_{X_{r_0}^{\mu,\nu}}, \mu, \nu \in \mathcal{M}_M\big\}$ is a relatively compact set in $\mathscr{P}_p(\mathscr{C})$ which follows once we have verified the following two conditions (see [3, Proposition 7.1.5 and (5.1.20)])

(a)
$$\sup_{\mu,\nu\in\mathcal{M}_M} \mathbb{E}\|X_{r_0}^{\mu,\nu}\|_\infty^q < \infty,$$

(b)
$$\big\{\mathscr{L}_{X_{r_0}^{\mu,\nu}}, \mu, \nu \in \mathcal{M}_M\big\} \text{ is a tight subset of } \mathscr{P}(\mathscr{C}).$$

Condition (a) is an immediate consequence of Lemma 2.2 with p replaced by q. Then, (b) follows once we have verified the following property (see e.g. [6, Theorem 7.3, p.82])

$$\lim_{\delta\to 0} \sup_{\mu,\nu\in\mathcal{M}_M} \mathbb{P}\Big(\sup_{0\leq u\leq v\leq r_0, v-u\leq\delta} |X^{\mu,\nu}(v) - X^{\mu,\nu}(u)| \geq \gamma\Big) = 0, \quad \text{for every } \gamma > 0. \tag{2.15}$$

By Chebyshev's inequality, for $\mu, \nu \in \mathcal{M}_M$ and $R > 0$, we obtain from (2.8) that there exists a constant $c_1 > 0$ (dependent on M but independent of R, μ, and ν) such that

$$M_R^{\mu,\nu} := \mathbb{P}\big(\|X_{r_0}^{\mu,\nu}\|_\infty \geq R\big) + \mathbb{P}\big(\|X_0^{\mu,\nu}\|_\infty \geq R\big) \leq \frac{1}{R^p}\big\{\mathbb{E}\|X_{r_0}^{\mu,\nu}\|_\infty^p + \mu(\|\cdot\|_\infty^p)\big\} \leq \frac{c_1}{R^p}.$$

Whence, for any $\varepsilon_0 > 0$, there exists an $R_0 = R_0(\varepsilon_0, c_1, p)$ sufficiently large such that

$$\sup_{\mu,\nu\in\mathcal{M}_M} M_{R_0}^{\mu,\nu} \leq \varepsilon_0. \tag{2.16}$$

Fix $\varepsilon_0 > 0$ and $R_0 > 0$ satisfying (2.16) and define the stopping time

$$\tau^{\mu,\nu} := \inf\big\{t \geq -r_0 : \big|X^{\mu,\nu}(t)\big| > R_0\big\}.$$

Then $\mathbb{P}\big(\tau^{\mu,\nu} \leq r_0\big) \leq \varepsilon_0$ and, for $\gamma > 0$, we get

$$\begin{aligned}
&\mathbb{P}\bigg(\sup_{0\leq u\leq v\leq r_0, v-u\leq\delta} |X^{\mu,\nu}(v) - X^{\mu,\nu}(u)| \geq \gamma\bigg) \\
&\quad\leq \mathbb{P}\big(\tau^{\mu,\nu} \leq r_0\big) + \mathbb{P}\bigg(\sup_{0\leq u\leq v\leq r_0, v-u\leq\delta} \int_u^v \mathbb{1}_{(0,\tau^{\mu,\nu}]}(s)|b(X_s^{\mu,\nu},\nu)|\mathrm{d}s \geq \frac{\gamma}{2}\bigg) \\
&\qquad+ \mathbb{P}\bigg(\sup_{0\leq u\leq v\leq r_0, v-u\leq\delta} \Big|\int_u^v \mathbb{1}_{(0,\tau^{\mu,\nu}]}(s)\sigma(X_s^{\mu,\nu},\nu)\mathrm{d}W(s)\Big| \geq \frac{\gamma}{2}\bigg) \\
&\quad=: \mathbb{P}\big(\tau^{\mu,\nu} \leq r_0\big) + N^{\mu,\nu}(\delta) + J^{\mu,\nu}(\delta).
\end{aligned}$$

Since b is bounded on bounded subsets of $\mathscr{C} \times \mathscr{P}_p(\mathscr{C})$ by $(\mathbf{H_1})$, we obtain

$$\lim_{\delta\to 0} \sup_{\mu,\nu\in\mathcal{M}_M} N^{\mu,\nu}(\delta) = 0. \tag{2.17}$$

If we further have

$$\lim_{\delta\to 0}\sup_{\mu,\nu\in\mathcal{M}_M} J^{\mu,\nu}(\delta)=0. \tag{2.18}$$

then the claim (b) holds by combining (2.16), (2.17), and (2.18) and by invoking the fact that $\varepsilon_0>0$ is arbitrary. It remains to show (2.18). This is in fact a standard argument (see, e.g. [12]) since the integrand of the stochastic integral is bounded uniformly in s, μ, ν, ω, so each component of the stochastic integral is a time-changed Brownian motion where the time change process has a derivative which is uniformly bounded away from 0. Therefore, (2.18) follows and the proof of the lemma is complete. □

3 Proof of Theorem 1.2

The key idea of the proof of Theorem 1.2 is to apply the following Kakutani's fixed point theorem for multivalued maps (see e.g. [18, Lemma 20.1, p20]):

Lemma 3.1. *Let $\mathcal{K}$ be a non-empty convex compact subset of a locally convex topological vector space and assume that $\Gamma:\mathcal{K}\to 2^{\mathcal{K}}$ is a set-valued function (where $2^{\mathcal{K}}$ is the set of subsets of $\mathcal{K}$) satisfying the following conditions:*

(a) *$\Gamma(x)$ is nonempty and convex for every $x\in\mathcal{K}$;*
(b) *The graph $Gr(\Gamma):=\{(x,y)\in\mathcal{K}\times\mathcal{K}: y\in\Gamma(x)\}$ is a closed subset of $\mathcal{K}\times\mathcal{K}$.*

Then, Γ admits a fixed point $\xi^*\in\mathcal{K}$, i.e., $\xi^*\in\Gamma(\xi^*)$.

Let

$$\mathcal{M}^p=\big\{\mu|\,\mu \text{ is a finite signed measure on } \mathscr{C} \text{ such that } |\mu|(\|\cdot\|_\infty^p)<\infty\big\},$$

which is a normed space so it is a Hausdorff locally convex space under the Kantorovich-Rubinstein metric

$$\mathcal{W}_p(\mu,\nu)=\Big(\sup_{(f,g)\in\mathscr{F}_{\mathrm{L}ip}}\big(\mu(f)-\nu(g)\big)\Big)^{1/p},\quad \mu,\nu\in\mathcal{M}^p,$$

where

$$\mathscr{F}_{\mathrm{L}ip}=\big\{(f,g): f,g \text{ are Lipschitz such that } f(\xi)\le g(\eta)+\|\xi-\eta\|_\infty^p,\xi,\eta\in\mathscr{C}\big\}.$$

Observe that

$$\mathcal{W}_p(\mu,\nu)=\mathbb{W}_p(\mu,\nu),\quad \mu,\nu\in\mathscr{P}_p(\mathscr{C}),$$

see, for example, [20]. Therefore, Lemma 3.1 is applicable to the present set-up.

With Lemma 3.1 at hand, we are in the position to complete the proof of Theorem 1.2.

Proof of Theorem 1.2. Let $\mathcal{K}$ be the convex hull of $\overline{\mathcal{K}_0}$ in $\mathscr{P}_p(\mathscr{C})$, so $\mathcal{K}$ is a convex subset of $\mathcal{M}_M$. Further, by Lemma 2.4, $\overline{\mathcal{K}_0}$ is a compact subset of $\mathscr{P}_p(\mathscr{C})$. Thus, [2, Theorem 5.35] enables us to deduce that $\mathcal{K}$ is also compact in $\mathscr{P}_p(\mathscr{C})$. So we conclude that $\mathcal{K}$ is a nonempty convex compact subset of $\mathscr{P}_p(\mathscr{C})$.

By taking advantage of Lemma 2.2, together with $(\mathbf{H_1})$ and $(\mathbf{H_2})$, it follows from [12, Theorem 2.3] (see also [16, Theorem 3.1.1]) that the (Feller) semigroup (P_t^ν) admits at least one IPM in the set $\mathcal{M}_M$, so the set Λ_ν of IPMs of (P_t^ν) in $\mathcal{M}_M$ is non-empty. Define the multivalued map $\Gamma:\mathcal{K}\to 2^{\mathcal{K}}$ by

$$\Gamma(\nu)=\Lambda_\nu,\quad \nu\in\mathcal{K}. \tag{3.1}$$

As long as the map Γ has a fixed point, i.e., there exists a probability measure $\mu\in\mathcal{K}$ such that $\mu\in\Gamma(\mu)$, the definition of Λ_μ enables us to get

$$\mathscr{L}_{X_t^{\mu,\mu}}=\mu_t^\mu=\mu,\quad t\ge 0,$$

so the $\mathscr{C}$-valued Markov process $(X_t)_{t\geq 0}$ solving (1.2) has an IPM. To complete the proof of Theorem 1.2, it is sufficient to show that the multivalued map Γ, defined in (3.1), has a fixed point. To apply Kakutani's fixed point theorem, we need to verify

(i) Λ_ν is convex for each $\nu \in \mathcal{K}$;

(ii) the graph of Γ is closed.

The claim **(i)** is clear since convex combinations of IPMs are invariant and $\mathcal{K}$ is convex. To show **(ii)**, consider sequences $(\nu_n)_{n\geq 1}$ and $(\pi_n)_{n\geq 1}$ in $\mathcal{K}$ such that $\pi_n \in \Lambda_{\nu_n}$ and such that there exist ν and π in $\mathcal{K}$ such that $\lim_{n\to\infty} \mathbb{W}_p(\nu_n, \nu) = 0$ and $\lim_{n\to\infty} \mathbb{W}_p(\pi_n, \pi) = 0$. We have to show that $\pi \in \Lambda_\nu$. To see this, for any $t > 0$, by $\pi_n \in \Lambda_{\nu_n}$, we have

$$\int_{\mathscr{C}} f(\xi)\pi_n(\mathrm{d}\xi) = \int_{\mathscr{C}} P_t^{\nu_n} f(\xi)\pi_n(\mathrm{d}\xi), \quad f \in \mathrm{Lip}_b(\mathscr{C}),$$

where $\mathrm{Lip}_b(\mathscr{C})$ means the set of bounded Lipschitz continuous functions on $\mathscr{C}$. Thus, one has

$$\int_{\mathscr{C}} f(\xi)\pi_n(\mathrm{d}\xi) = \int_{\mathscr{C}} P_t^{\nu} f(\xi)\pi_n(\mathrm{d}\xi) + \Upsilon_n, \tag{3.2}$$

where

$$\Upsilon_n := \int_{\mathscr{C}} \big(P_t^{\nu_n} f(\xi) - P_t^{\nu} f(\xi)\big)\pi_n(\mathrm{d}\xi).$$

Taking advantage of Lemma 2.1, we infer

$$|\Upsilon_n| \leq \|f\|_{\mathrm{Lip}} \int_{\mathscr{C}} \mathbb{E}\|X_t^{\delta_\xi,\nu_n} - X_t^{\delta_\xi,\nu}\|_\infty \pi_n(\mathrm{d}\xi) \leq \|f\|_{\mathrm{Lip}} \big(Ct\mathrm{e}^{Ct}\big)^{1/p} \mathbb{W}_p(\nu_n, \nu).$$

This, together with $\mathbb{W}_p(\nu_n, \nu) \to 0$, $\mathbb{W}_p(\pi_n, \pi) \to 0$, and the Feller property of $(P_t^\nu)_{t\geq 0}$, allows us to conclude from (3.2) that

$$\int_{\mathscr{C}} P_t^{\nu} f(\xi)\pi(\mathrm{d}\xi) = \int_{\mathscr{C}} f(\xi)\pi(\mathrm{d}\xi), \quad f \in \mathrm{Lip}_b(\mathscr{C}).$$

Therefore, $\pi \in \mathscr{P}_p(\mathscr{C})$ is an IPM so that $\pi \in \Lambda_\nu$ and the proof of the theorem is complete. □

References

[1] Ahmed, N. U., Ding, X., On invariant measures of nonlinear Markov processes, *J. Appl. Math. Stochastic Anal.* **6**, (1993), 385–406. MR1255258

[2] Aliprantis, C.D., Border, Kim C., Infinite Dimensional Analysis, A hitchhiker's guide, Third edition, *Springer*, Berlin, 2006. MR2378491

[3] Ambrosio, L., Gigli, N., Savaré, G., Gradient Flows in Metric Spaces and in the Space of Probability Measures, Second edition, *Birkhäuser*, Basel, 2008. MR2401600

[4] Bao, J., Ren, P., Wang, F.-Y., Bismut formula for Lions derivative of distribution-path dependent SDEs, *J. Differential Equations* **282**, (2021), 285–329. MR4219322

[5] Bao, J., Yin, G., Yuan, C., Ergodicity for functional stochastic differential equations and applications, *Nonlinear Anal.* **98**, (2014), 66–82. MR3158446

[6] Billingsley, P., Convergence of Probability Measures. Second edition, *Wiley*, New York, 1999. MR1700749

[7] Butkovsky, O., On ergodic properties of nonlinear Markov chains and stochastic McKean-Vlasov equations, *Theory Probab. Appl.* **58**, (2014), 661–674. MR3403022

[8] Butkovsky, O., Subgeometric rates of convergence of Markov processes in the Wasserstein metric, *Ann. Appl. Probab.* **24**, (2014), 526–552. MR3178490

[9] Butkovsky, O., Scheutzow, M., Invariant measures for stochastic functional differential equations, *Electronic J. Probab.* **22**, (2017), 1–23. MR3724566

[10] Carmona, R., Delarue, F., Probabilistic Theory of Mean Field Games with Applications, I-II, *Springer*, Cham, 2018. MR3752669

[11] Da Prato, G., Zabczyk, J., Ergodicity for Infinite Dimensional Systems, London Mathematical Society Lecture Notes, vol. 229, *Cambridge University Press*, Cambridge, 1996. MR1417491

[12] Es-Sarhir, A., Scheutzow, M., van Gaans, O., Invariant measures for stochastic functional differential equations with superlinear drift term, *Differential Integral Equations* **23**, (2010), 189–200. MR2588808

[13] Hairer, M., Mattingly, J.C., Scheutzow, M., Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations, *Probab. Theory Related Fields* **149**, (2011), 223–259. MR2773030

[14] Hammersley, W. R. P., Šiška, D., Szpruch, L., McKean-Vlasov SDEs under measure dependent Lyapunov conditions, *Ann. Inst. H. Poincaré Probab. Statist.* **57**, (2021), 1032–1057. MR4260494

[15] Karatzas, I., Shreve, S., Brownian Motion and Stochastic Calculus, Graduate Texts in Mathematics, vol. 113, *Springer*, New York, 1988. MR0917065

[16] Kinnally, M.S., Williams, R.J., On existence and uniqueness of stationary distributions for stochastic delay differential equations with positivity constraints, *Electron. J. Probab.* **15**, (2010), 409–451. MR2639731

[17] Mehri, S., Scheutzow, M., Stannat, W., Zangeneh, B., Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions, *Ann. Appl. Probab.* **30**, (2020), 175–207. MR4068309

[18] Osborne, M. J., Rubinstein, A., A Course in Game Theory, *MIT Press*, Cambridge, MA, 1994. MR1301776

[19] Osekowski, A., Sharp maximal inequalities for the martingale square bracket, *Stochastics* **82**, (2010), 589–605. MR2783421

[20] Rachev, S.T., The Monge-Kantonovich mass transference problem and its stochastic applications, *Theory Probab. Appl.* **29**, (1984), 647–676. MR0773434

[21] Ren, P., Wang, F.-Y., Donsker-Varadhan large deviations for path-distribution dependent SPDEs, *J. Math. Anal. Appl.* **499** (2021), 125000, 32 pp. MR4207324

[22] Scheutzow, M., Qualitative behaviour of stochastic delay equations with a bounded memory, *Stochastics* **12**, (1984), 41–80. MR0738934

[23] Villani, C., Optimal Transport. Old and New, *Springer*, Berlin, 2009. MR2459454

[24] von Renesse, M.-K., Scheutzow, M., Existence and uniqueness of solutions of stochastic functional differential equations, *Random Oper. Stoch. Equ.* **18**, (2010), 267–284. MR2718125

[25] Wang, F.-Y.: Distribution dependent SDEs for Landau type equations, *Stochastic Process. Appl.* **128**, (2018), 595–621. MR3739509

[26] Zeidler, E.: Nonlinear Functional Analysis and its Applications I. Fixed-point Theorems, *Springer*, New York, 1986. MR0816732

[27] Zhang, S.-Q.: Existence and non-uniqueness of stationary distributions for distribution dependent SDEs, arXiv:2105.04226.