# Declaration of Authorship

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Berlin, den 06.11.2019

Technische Universität Berlin

Masters Thesis

---

# Rough McKean Vlasov Equations

---

*Author:*
Nikolai Bobenko

*Supervisor:*
Dr. Peter Friz

*A thesis submitted in fulfillment of the requirements for the degree of Master of Science*

*in the*

Study Group Probability Theory and Mathematical Finance
Mathematics and Natural Sciences

November 7, 2019

TECHNISCHE UNIVERSITÄT BERLIN

# *Abstract*

Probability Theory and Mathematical Finance
Mathematics and Natural Sciences

Master of Science

**Rough McKean Vlasov Equations**

by Nikolai BOBENKO

McKean-Vlasov equations enable analysis of large particle systems with interactions. We consider such systems with rough drivers and interactions only in the drift component. Following the approach laid out in [7] we extend the results of Cass-Lyons to the case of non-linear interactions in law using the techniques of controlled rough paths. This pathwise controlled approach shows promise and has been proven to be extendable to more general settings for example in [1]. We show existence of solutions, continuity of the solution map and a convergence speed result.

TECHNISCHE UNIVERSITÄT BERLIN

# *Zusammenfassung*

Probability Theory and Mathematical Finance
Mathematics and Natural Sciences

Master of Science

**Rough McKean Vlasov Equations**

von Nikolai Bobenko

McKean-Vlasov Gleichungen ermöglichen die Analyse von dynamischen Systemen mit großen Anzahlen von Teilchen. In dieser Arbeit betrachten wir solche stochastischen Systeme mit rough Path Input und Interaktion nur in der Driftkomponente. Wir folgen der Herangehensweise in [7] und erweitern die darin vorgestellten Resultate auf nicht-lineare Interaktionen in der Verteilung, indem wir die Methoden der Controlled Rough Paths verwenden. Ein solcher pfadweiser Ansatz liefert Resultate, die mithilfe anderer Methoden außer Reichweite zu liegen scheinen und wurde in [1] verallgemeinert um auch Interaktion in der Diffusion zuzulassen. In dieser Arbeit zeigen wir Existenz der Lösung, Stetigkeit der Lösungfunktion und ein Konvergenzgeschwindigkeitsresultat.

# *Acknowledgements*

I would like to thank Prof. Peter Friz and Michele Coghi for guiding me through this process and taking the time to meet with me whenever I was stuck or needed some inspiration.

Furthermore I thank my family and friends for supporting me along the way. It is greatly appreciated.

# Contents

# Chapter 1

# Introduction

Systems of a random set of $N$ particles $(Y^i)_{i\in\{1,\dots,N\}}$ that evolve according to dynamics with pairwise interactions of the type

$$dY_t^i(\omega) = \frac{1}{N}\sum_{j=1}^{N} g\left(Y_t^i(\omega), Y_t^j(\omega)\right) dt + \frac{1}{N}\sum_{j=1}^{N} f\left(Y_t^i(\omega), Y_t^j(\omega)\right) dX_t^i(\omega) \qquad (1.1)$$

with i.i.d. starting conditions and $X^i$ independent input processes have been studied extensively. It has been shown in [21] and others that for $X$ a semi-martingale such a system exhibits propagation of chaos properties. That is for $N \to \infty$ each particle's behaviour can be approximated in law by the solution to a non-linear equation

$$dY_t = g(Y_t, \mathcal{L}(Y_t))dt + f(Y_t, \mathcal{L}(Y_t))dX_t. \qquad (1.2)$$

Here $\mathcal{L}(Y_t)$ denotes the law of $Y_t$ and both $g$ and $f$ are presumed to be Lipschitz in their respective law components with regard to the 2-Wasserstein metric. The significance of such results lies in giving us suitable approximations that allow us to reason about some large particle systems that defy direct analysis.

Trying to extend the type of input processes for which solutions can be found, Cass and Lyons [7] study a rough setup with weak interactions. That is $f(y,\mu) = f(y)$ independent of the measure argument with an input process $\mathbf{X} : \Omega \to \mathscr{C}_g$ where $\mathscr{C}_g$ is a suitable rough path space further defined in section 1.2. Using rough path techniques, existence of solution and propagation of chaos results are proven under specific integrability conditions for $\mathbf{X}$. We apply the controlled rough path techniques introduced by Gubinelli [17] to expand these results allowing $g$ to be Lipschitz in the measure thus expanding the linearity assumption imposed in [7]. The pathwise approach chosen furthermore naturally allows one to drop the previsibility and semi-martingale conditions imposed on the input driver in [21].

In [1] Bailleul, Catelier and Delarue consider a general setup

$$dY_t = f(Y_t, \mathcal{L}(Y_t))d\mathbf{X}_t.$$

for a random rough path valued driver $\mathbf{X}$. Regularity assumptions in the law variable are leveraged to introduce mixed finite-infinite dimensional controlled rough paths and again existence and propagation of chaos results are shown. This work is more advanced than ours and yields more general results but we see value in the simplicity of the theory that comes with not considering interactions in the diffusion.

In all of the mentioned rough paths based approaches [1, 7] as well as in this work the results rely fundamentally on limiting the permitted input noise to variables with finite exponential moments of the local accumulated variation $N_1$ (Definition 1.2). This stems from the stability estimates of rough differential equations (RDEs) carrying an exponential term $e^{CN_1(w_X)}$ which when assumed to be integrable allows one to control the Wasserstein distance of the law of solutions. Luckily [6] provides a large family of processes for which that is the case. While this assumption ensures all the results mentioned above it is not clear whether it is in any way sharp or could be weakened substantially.

In [8] McKean-Vlasov equations with additive noise are considered. This extends the theory since similarly to the case of ordinary differential equations (ODEs) with additive noise there are no regularity assumptions imposed on the input driver other than continuity. The setting is instructive due to its simplicity and permits proofs of large deviations and central limit theorem type results.

McKean-Vlasov equations are encountered as limits of systems of interacting particles in several fields. Some applications arise in fluid dynamics [20, 12, 2], mean field games [5, 4] and economics [18].

In this work we extend the setting of [7] allowing for $g$ to be Lipschitz rather than linear in the measure, remain however in the realm of weak interactions. We show existence, uniqueness and propagation of chaos results seen in a more general form in [1] and show a strong rate of convergence result in Theorem 4.5 for the system of particles.

We structure this thesis as follows. In chapter 1 we introduce relevant concepts relating to rough path theory and show a few helpful lemmas along the way. Chapter 2 is devoted to showing existence and uniqueness theorems for rough differential equations with drift in a setting of controlled rough paths. We use these in chapter 3 to show the existence of solutions for McKean-Vlasov equations as well as continuity of the solution map. Finally, in chapter 4 we show propagation of chaos results using the continuity of the solution map and prove an estimate for the strong rate of convergence.

## 1.1 Hölder and p-variation spaces

Let $(E, \|\cdot\|)$ be a Banach space. For $T > 0$ we denote by $\Delta_T$ the two-simplex $\Delta_T = \{(s,t) \mid 0 \le s < t \le T\}$.

We denote by $\mathcal{C}_2^{p-var}([0,T], E)$ the space of all continuous maps $X : \Delta_T \to E$ for which

$$\|X\|_{p-var,[0,T]} := \left( \sup_{\pi \in \Pi([0,T])} \sum_{(t_i)=\pi} \|X_{t_i,t_{i+1}}\|^p \right)^{\frac{1}{p}} < \infty.$$

Here $\Pi(I)$ denotes the set of all finite partitions of the set $I$.

**Definition 1.1.** A continuous function $w : [0,T]^2 \to \mathbb{R}^+$ is called a control if it is superadditive, i.e.

$$w(s,t) \ge w(s,u) + w(u,t) \quad \forall\, s < u < t \in [0,T].$$

We call a control satisfying $w(s,s) = 0$, $\forall\, s \in [0,T]$ regular.

For a function $X \in \mathcal{C}_2^{p-var}([0,T], E)$ we define $w_X(s,t) := \|X\|^p_{p-var,[s,t]}$. It can be seen that $w_X$ is indeed a control [16, Proposition 5.8]. Furthermore it is known that the existence of a control $w$ with $X_{s,t} \le Cw(s,t)^{\frac{1}{p}}$ for all $s,t$ implies $w_X(s,t) \le Cw(s,t)$ [16, Proposition 5.10]. Therefore $p$-variation can also be equivalently defined as

$$\|X\|_{p-var,[s,t]} = \inf \left\{ w(s,t)^{\frac{1}{p}} \mid w \text{ is a control and } \|X_{u,v}\| \le w(u,v)^{\frac{1}{p}}\ \forall\, s \le u < v \le t \right\}.$$

We will neglect the interval and just write $\|X\|_{p-var}$ when the interval is implied. We will denote by $w_{id}$ the control given by $w_{id}(s,t) = |t-s|$.

We will use the following construction for a greedy partition and local accumulated variation in the later chapters to ensure integrability.

**Definition 1.2.** For a control $w$ for some fixed $\beta > 0$ we define the greedy $\beta$-partition $\{\tau_n\}_n$ by setting

$$\tau_0 = s, \quad \tau_n = \inf\{t \mid w(\tau_{n-1}, t) \geq \beta, \tau_{n-1} < t < T\} \wedge T.$$

Let $N$ be the size of this partition. We then have $w(\tau_{n-1}, \tau_n) = \beta, \ \forall n < N$ and $w(\tau_{N-1}, \tau_N) \leq \beta$. Furthermore we define the local accumulation of $w$ as

$$N_\beta(w, [s,t]) := \sup\{n \geq 0 \mid \tau_n < t\}.$$

Following the discussion in [15, Chapter 11] it is known that for lifted Brownian motion $\mathbf{B}$ with control $w_{\mathbf{B}}$ the amount of greedy partition increments $N_\beta(w_{\mathbf{B}}, [0,T])$ is exponentially integrable. In contrast the random variable $|||\mathbf{B}|||_\alpha^{\frac{1}{\alpha}}$ is not exponentially integrable.

For some $\alpha > 0$ we define $\mathcal{C}_2^\alpha([0,T], E)$ to be the space of all $\alpha$-Hölder functions. That is all functions $X : \Delta_T \to E$ with

$$\|X\|_\alpha := \sup_{t,s \in [0,T]} \frac{\|X_{s,t}\|}{|t-s|^\alpha} < \infty.$$

We associate a path $X : [0,T] \to E$ with its increment function

$$X^c : (s,t) \mapsto X_t - X_s$$

and define $\|X\|_\alpha = \|X^c\|_\alpha$ as well as $\|X\|_{p-var,[s,t]} = \|X^c\|_{p-var,[s,t]}$. We denote the corresponding path spaces as $\mathcal{C}^\alpha([0,T], E)$ and $\mathcal{C}^{p-var}([0,T], E)$ respectively.

The following well known lemma provides a connection between local and global Hölder norms.

**Lemma 1.3.** *Let $X : [0,T] \to E$ be a path, $\alpha \in (0,1]$ and $h > 0$ with*

$$\|X\|_{\alpha,h} := \sup_{\substack{0 \leq s < t \leq T \\ t-s \leq h}} \frac{|X_{s,t}|}{|t-s|^\alpha} \leq \infty.$$

*Then we have*

$$\|X\|_\alpha \leq \|X\|_{\alpha,h} \left(1 \vee 2h^{\alpha-1}\right).$$

For a proof see e.g. [15, Exercise 4.24].

## 1.2 Rough Paths

There are many comprehensive expositions to the topic of Rough Paths that do well introducing and motivating the subject. We refer the reader to [15, 16] and will only present some basic notions here.

Let $p \in [2,3)$. We call a pair

$$\mathbf{X} = (X, \mathbb{X}) \in \mathcal{C}^{p-var}([0,T], E) \times \mathcal{C}_2^{\frac{p}{2}-var}([0,T], E \otimes E)$$

that satisfies Chen's relation

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t} \quad \forall\, s,u,t \tag{1.3}$$

a $p$-rough path. We denote the set of all such pairs by $\mathscr{C}^{p-var}([0,T],E)$ and equip it with the homogeneous $p$-variation metric given by

$$\varrho_{p-var,[0,T]}(\mathbf{X}^1,\mathbf{X}^2) = \left\|X^1 - X^2\right\|_{p-var,[0,T]} + \sqrt{\|\mathbb{X}^1 - \mathbb{X}^2\|_{p-var,[0,T]}}\,.$$

Note that this is not a linear space but the corresponding norm $|||\cdot|||_{p-var} := \varrho_{p-var}(\cdot,0)$ behaves well under the natural dilation operation in $\mathscr{C}^{p-var}$ given by $(X,\mathbb{X}) \mapsto (\lambda X, \lambda^2\mathbb{X})$.

For a path of bounded variation $Z:[0,T]\to E$ there exists a canonical lift given via the well defined Riemann-Stieljes integral $\mathbb{Z}_{s,t} = \int_s^t Z_{s,u}\otimes dZ_u$. With the second level defined this way it is easy to check that $\mathbf{Z} = (Z,\mathbb{Z})$ satisfies Chen's relation (1.3). We denote by $\mathscr{C}_g^{p-var}([0,T],E)$ the set of geometric rough paths, that is the closure of lifted paths of bounded variation under $|||\cdot|||_{p-var}$.

Geometric rough paths on $\mathbb{R}^e$ can be identified with paths taking values in a Lie-Group $G^{(2)}(\mathbb{R}^e)$ that have finite $p$-variation with respect to the corresponding Carnot-Carathéodory metric $d_{CC}$. With this in mind for $\mathbf{X} \in \mathscr{C}_g^{p-var}([0,T],\mathbb{R}^e)$ we have

$$|||\mathbf{X}|||_{p-var,[0,T]} = \sup_{\pi\in\Pi([0,T])} \sum_{(t_i)=\pi} d_{CC}(\mathbf{X}_{t_i},\mathbf{X}_{t_{i+1}})^p.$$

For some $\alpha \in (\frac{1}{3},\frac{1}{2}]$ we call $\mathbf{X} = (X,\mathbb{X})$ an $\alpha$-Hölder rough path if it satisfies Chen's relation and

$$\|X\|_\alpha \vee \|\mathbb{X}\|_{2\alpha} < \infty.$$

We denote the space of such paths by $\mathscr{C}^\alpha$ with the associated homogeneous norm $|||\mathbf{X}|||_\alpha := \|X\|_\alpha + \sqrt{\|\mathbb{X}\|_{2\alpha}}$.

### Controlled Rough Paths

For a path $X \in \mathcal{C}^{p-var}([0,T],\mathbb{R}^e)$ we say that

$$(Y,Y') \in \mathcal{C}^{p-var}([0,T],\mathbb{R}^d) \times \mathcal{C}^{p-var}([0,T],\mathcal{L}(\mathbb{R}^e,\mathbb{R}^d))$$

is controlled by $X$ if for the remainder term given by

$$Y^\#_{s,t} = Y_{s,t} - Y'_s X_{s,t}$$

we have $Y^\# \in \mathcal{C}^{\frac{p}{2}-var}$. This notion was first introduced by Gubinelli [17] in the Hölder setting but also makes sense in the p-variation setting. We will write $(Y,Y') \in \mathscr{D}_X^{\frac{p}{2}}$ for a pair satisfying this condition.

There is a canonical way in which smooth enough functions preserve controlled paths. This is made concrete in the following lemma and in some situations allows one not to provide $Y'$ explicitly as it is implicitly derived from context.

**Lemma 1.4.** *Let $f \in \mathcal{C}_b^2$, $X \in \mathcal{C}^{p-var}$ and $(Y,Y') \in \mathscr{D}_X^{\frac{p}{2}}$ with controls $w_Y, w_{Y^\#}$. Then $(f(Y)_t, f(Y)'_t) = (f(Y_t), Df(Y_t)Y'_t)$ is also an element of $\mathscr{D}_X^{\frac{p}{2}}$ and for all $s,t\in[0,T]$*

*there exists a constant $C$ depending only on $f$ such that*

$$w_{f(Y)}(s,t) \leq C w_Y(s,t)$$
$$w_{f(Y)^{\#}}(s,t)^{\frac{2}{p}} \leq C(w_Y(s,t)^{\frac{2}{p}} + w_{Y^{\#}}(s,t)^{\frac{2}{p}})$$

*Proof.* First we note that $f(Y)$ and $f(Y)'$ are controlled by the following quantities.

$$\begin{aligned} |f(Y)_{s,t}| &\leq \|Df\|_\infty |Y_{s,t}| \leq C w_Y(s,t)^{\frac{1}{p}} \\ |f(Y)'_{s,t}| &\leq |Df(Y_s)Y'_s - Df(Y_s)Y'_t| + |Df(Y_s)Y'_t - Df(Y_t)Y'_t| \\ &\leq \|Df\|_\infty |Y'_{s,t}| + \|Y'_t\|_\infty \|D^2 f\|_\infty |Y_{s,t}|. \end{aligned}$$

Thus both $f(Y)$ and $f(Y)'$ are in $\mathcal{C}^{p-var}$. For the remainder we have

$$\begin{aligned} |f(Y)^{\#}_{s,t}| &= |f(Y)_t - f(Y)_s - Df(Y_s)Y'_s X_{s,t}| \\ &\leq |f(Y)_t - f(Y)_s - Df(Y_s)Y_{s,t}| + |Df(Y_s)Y^{\#}_{s,t}| \\ &\leq \frac{1}{2}\|D^2 f\|_\infty |Y_{s,t}|^2 + \|Df\|_\infty |Y^{\#}_{s,t}| \\ &\leq C(w_Y(s,t)^{\frac{2}{p}} + w_{Y^{\#}}(s,t)^{\frac{2}{p}}). \end{aligned}$$

Hence we get a control $w_{f(Y)^{\#}}(s,t) \leq C(w_Y(s,t) + w_{Y^{\#}}(s,t))$ with $|f(Y)^{\#}_{s,t}| \leq w_{f(Y)^{\#}}(s,t)^{\frac{2}{p}}$ and the result follows. □

## Rough Path Integration

The sewing lemma is the main result used to define and obtain estimates for rough path integration. We could not find this explicit version of the sewing lemma in the literature therefore include it here for completeness. For a Banach space $V$ and a function $\Xi : \Delta_T \to V$ we write $\delta\Xi_{s,u,t} := \Xi_{s,t} - \Xi_{s,u} - \Xi_{u,t}$ which is used to measure how close $\Xi$ is to being additive in time which is clearly a property we want to impose for any integral notion.

**Lemma 1.5.** *Let $V$ be a Banach space and $\beta > 1$. For a function $\Xi : \Delta_T \to V$ let there be a control $w$ such that for all $0 \leq s < u < t \leq T$ we have*

$$\|\delta\Xi_{s,u,t}\| \leq w^\beta(s,t).$$

*Then there exists a unique element $I(\Xi) : [0,T] \to V$ such that there exists a constant $C$ depending only on $\beta$ with*

$$\|\Xi_{s,t} - I(\Xi)_{s,t}\| \leq C w^\beta(s,t)$$

*for all $s < t$ in $[0,T]$. Furthermore we have*

$$I(\Xi)_{s,t} = \lim_{|\pi|\to 0} \sum_{[t_i,t_{i+1}]\in\pi} \Xi_{t_i,t_{i+1}}$$

*where $\pi$ is a partition of $[s,t]$ and $|\pi|$ denotes the size of its largest element.*

*Proof.* This proof is a modification of the classical Young argument.

Let $\pi$ be a partition on $[s,t]$ consisting of a total of $r \geq 1$ intervals. Then for $r \geq 2$ there exists a $u \in [s,t]$ such that $[u_-, u], [u, u_+] \in \pi$ and

$$w(u_-, u_+) \leq \frac{2}{r-1} w(s,t).$$

Otherwise one would get $2w(s,t) \geq \sum_{u \in \pi^o} > \frac{2r}{r-1} w(s,t) > 2w(s,t)$.

We use the notation $\int_\pi \Xi := \sum_{[u,v]\in\pi} \Xi_{u,v}$. For our choice of $u \in \pi$ it follows that

$$\left\| \int_{\pi\setminus\{u\}} \Xi - \int_\pi \Xi \right\| = \|\delta\Xi_{u_-,u,u_+}\| \leq w^\beta(u_-, u_+) \leq \left(\frac{2}{r-1}\right)^\beta w^\beta(s,t).$$

By iterating this procedure down to the trivial partition $\{[s,t]\}$ we arrive at

$$\left\| \Xi_{s,t} - \int_\pi \Xi \right\| \leq \sum_{r=1}^{\infty} \left(\frac{2}{r}\right)^\beta w(s,t)^\beta = 2^\beta \zeta(\beta) w(s,t)^\beta \tag{1.4}$$

where $\zeta$ denotes the Riemann zeta function. Now consider a partition $\pi_\epsilon$ on $[0,T]$ such that for any $[u,v] \in \pi_\epsilon$ we have $|u-v| \vee w(u,v) < \epsilon$. Then for any subpartition $\pi \supset \pi_\epsilon$ we have

$$\begin{aligned} \left\| \int_{\pi_\epsilon} \Xi - \int_\pi \Xi \right\| &= \sum_{[u,v]\in\pi_\epsilon} \left\| \Xi_{u,v} - \int_{\pi\cap[u,v]} \Xi \right\| \\ &\leq 2^\beta \zeta(\beta) \sum_{[u,v]\in\pi_\epsilon} w(u,v)^\beta \\ &\leq 2^\beta \zeta(\beta) \frac{T}{\epsilon} \epsilon^\beta \\ &\in \mathcal{O}(\epsilon^{\beta-1}). \end{aligned}$$

Therefore, there exists a $K \in V$ such that for any $\epsilon > 0$ there exists a partition $\pi_\epsilon$ such that for any subpartition $\pi \supset \pi_\epsilon$ we have $\left\| \int_\pi \Xi - K \right\| \leq \epsilon$.

Let $\pi$ now be some partition on $[0,T]$ with $|\pi| \leq |\pi_\epsilon|$. Then we observe that

$$\begin{aligned} \left\| \int_\pi \Xi - K \right\| &\leq \left\| \int_\pi \Xi - \int_{\pi\vee\pi_\epsilon} \Xi \right\| + \left\| \int_{\pi\vee\pi_\epsilon} \Xi - K \right\| \\ &\leq C\epsilon^{\beta-1} + \epsilon. \end{aligned}$$

Thus we have convergence. Uniqueness of $I(\Xi)$ follows by considering two sequences of partitions $(\pi_n)$, $(\pi'_n)$ for which $\int_{\pi_n} \Xi$ and $\int_{\pi'_n} \Xi$ both converge. Then the sequence of their refinements given by $\tilde{\pi}_n = \pi_n \wedge \pi'_n$ also converges. It is easy to see that by the definition of our convergence the limits agree, hence we get uniqueness. □

We use Lemma 1.5 to define integration of controlled rough paths against the paths they are controlled by. Let $p \in [2,3)$. Given a rough path $\mathbf{X} \in \mathscr{C}^{p-var}$ and a controlled path $(Y, Y') \in \mathscr{D}_X^{\frac{p}{2}}([0,T], V)$ we set the local expansion

$$\Xi_{s,t} = Y_s X_{s,t} + Y'_s \mathbb{X}_{s,t}$$

defined in coordinates by $\Xi^i_{s,t} = (Y_s)^i_j (X_{s,t})^j + (Y'_s)^i_{j,k} (\mathbb{X}_{s,t})^{j,k}$ written in Einstein notation. Using Chen's relation it is now easily checked that there exists a control $w$

with $|\delta\Xi_{s,u,t}| < w(s,t)^{\frac{3}{p}}$ allowing us to apply the sewing lemma to define $\int Y d\mathbf{X} := I(\Xi)$. We think of $Y$ as the first order local expansion of $\int Y d\mathbf{X}$.

In fact this gives us the following corollary.

**Corollary 1.6.**

$$\left|\int_s^t Y_r - Y_s X_{s,t} - Y_s' \mathbb{X}_{s,t}\right| \le C\left(w_X(s,t)^{\frac{1}{p}} w_{Y\#}(s,t)^{\frac{2}{p}} + w_X(s,t)^{\frac{2}{p}} w_Y(s,t)^{\frac{1}{p}}\right)$$

*Proof.* This is a direct consequence of Lemma 1.5 applied to $\Xi_{s,t} = Y_s X_{s,t} + Y_s' \mathbb{X}_{s,t}$. A simple application of Chen's relation then yields

$$\delta\Xi_{s,\theta,t} = -Y^{\#}_{s,u} X_{u,t} - Y'_{s,u} \mathbb{X}_{u,t}\ .$$

By observing that indeed

$$w_1(s,t) := w_{Y\#}(s,t)^{\frac{2}{3}} w_X(s,t)^{\frac{1}{3}} + w_X(s,t)^{\frac{2}{3}} w_Y(s,t)^{\frac{1}{3}}$$

is a control that satisfies $|\delta\Xi_{s,u,t}| \le C w_1(s,t)^{\frac{3}{p}}$ we can apply the sewing lemma and the claim follows. □

The following rough Grönwall lemma is a modified version of [10, Lemma 2.11] which allows us to advance from local to global estimates. The only difference to [10] is that there is no requirement for $\phi$ to be a control - we only assume it to be bounded by $\phi(0,T)$. This proof will appear in an upcoming version of [9].

**Proposition 1.7** (Rough Grönwall)**.** *Let $W : [0,T] \to \mathbb{R}_+$ be a function. If there exist constants $L > 0$ and $\kappa > 0$, a regular control $w$ and a function $\phi : \Delta_T \to \mathbb{R}_+$ with $\phi(s,t) \le \phi(0,T)$ such that for all $0 \le s < t \le T$ with $w(s,t) \le L$ we have*

$$W_{s,t} \le w(s,t)^{\frac{1}{\kappa}} \sup_{0\le r\le t} W_r + \phi(s,t) \tag{1.5}$$

*then for $\alpha := 1 \vee L^{-1}(2e^2)^{-\kappa}$ we have*

$$\sup_{0\le t\le T} W_t \le 2e^{\frac{2w(0,T)}{L\alpha}}(W_0 + \phi(0,T)).$$

*Proof.* We build a partition of size $K$ such that $\alpha L(K-1) \le w(0,T) \le \alpha LK$. Let us define this partition $0 = t_0 < t_1 < t_2 < \ldots < t_K = T$ via

$$w(0,t_k) = \alpha k L, \quad \forall k = 0,1,\ldots,K-1.$$

Since $w$ is continuous and 0 at the diagonal this is well defined. Superadditivity of $w$ then implies $w(t_k,t_{k+1}) \le \alpha L$. For a fixed $t \in [t_k,t_{k+1}]$ by plugging in the assumption (1.5) we then get

$$\begin{aligned} W_{0,t} &= \sum_{i=0}^{k-2} W_{t_i,t_{i+1}} + W_{t_{k-1},t_k} \\ &\le \sum_{i=0}^{k-1} w(t_i,t_{i+1})^{\frac{1}{\kappa}} \|W\|_{\infty,[0,t_{i+1}]} + \sum_{i=0}^{k-2} \phi(t_i,t_{i+1}) + \phi(t_{k-1},t_k) \qquad (1.6) \\ &\le (\alpha L)^{\frac{1}{\kappa}} \sum_{i=0}^{k-1} \|W\|_{\infty,[0,t_{i+1}]} + k\phi(0,T). \end{aligned}$$

By now introducing the function

$$M_t = \|W\|_{\infty,[0,t]} \, e^{-\frac{w(0,t)}{\alpha L}}$$

we get by the definition of our partition $(t_i)$

$$\sum_{i=0}^{k-1} \|W\|_{\infty,t_{i+1}} = \sum_{i=0}^{k-1} M_{t_{i+1}} e^{\frac{w(0,t)}{\alpha L}} \le \|M\|_{\infty,[0,t_k]} \sum_{i=0}^{k-1} e^{i+1} \le \|M\|_{\infty,[0,t_k]} \, e^{k+1}.$$

Since the right hand side of (1.6) is independent of $t \in [t_{k-1}, t_k]$ we can plug this back into it and get

$$\|W\|_{\infty,[t_{k-1},t_k]} \le W_0 + (\alpha L)^{\frac{1}{\kappa}} \|M\|_{\infty,[0,t_k]} \, e^{k+1} + k\phi(0,T) =: B_k \, .$$

Since the right side here is monotone in $k$ this gives us as a consequence that in fact $\|W\|_{\infty,[0,t]} \le B_k$. Multiplying this inequality by $e^{-\frac{w(0,t)}{\alpha L}}$ then implies

$$M_t = \|W\|_{\infty,[0,t]} \, e^{-\frac{w(0,t)}{\alpha L}} \le e^{-\frac{w(0,t)}{\alpha L}} B_k$$

which gives us

$$\|M\|_{\infty,[0,t_k]} \le e^{-\frac{w(0,t)}{\alpha L}} B_k \le W_0 + (\alpha L)^{\frac{1}{\kappa}} \, e^2 \, \|M\|_{\infty,[0,t_k]} + K\phi(0,T).$$

Now the definition of $\alpha$ ensures that $(\alpha L)^{\frac{1}{\kappa}} \, e^2 \le \frac{1}{2}$ and we get

$$\|M\|_{\infty,[0,t]} \le 2(W_0 + K\phi(0,T))$$

which implies

$$\|W\|_{\infty,[0,T]} \le e^{\frac{w(0,T)}{\alpha L}} 2(W_0 + K\phi(0,T)).$$

Since $K$ was chosen to satisfy $K \le \frac{w(0,T)}{\alpha L} + 1 \le e^{\frac{w(0,T)}{\alpha L}}$, the result follows. □

There is the obvious connection between $N_1(w_X, [0,T])$ and $w_X(0,T)$ given by

$$N_1(w_X, [0,T]) \le w_X(0,T) + 1 \tag{1.7}$$

which follows directly from the superadditivity of $w_X$. Estimates going the other way are less obvious.

The following lemma will be used to estime the $p$-variation metric of rough paths by its local accumulation.

**Lemma 1.8.** *For any* $\mathbf{X} \in \mathscr{C}_g^{p-var}(\mathbb{R}^e)$ *we have*

$$\|\!|\mathbf{X}|\!\|_{p-var}^p = w_X(0,T) \le C_p e^{C_p N_1(w_X,[0,T])}.$$

*Proof.* For a geometric rough path $\mathbf{X}$ the homogeneous $p$-variation norm can be identified with

$$\|\!|\mathbf{X}|\!\|_{p-var,[0,T]}^p = \sup_{\pi \in \Pi([0,T])} \sum_{(t_i)=\pi} d_{CC}(\mathbf{X}_{t_i}, \mathbf{X}_{t_{i+1}})^p$$

where $d_{CC}$ denotes the Carnot-Carathéodory metric on $G^{(2)}(\mathbb{R}^e)$. We will write $\left\|\mathbf{X}_{t_i,t_{i+1}}\right\|$ to indicate $d_{CC}(\mathbf{X}_{t_i}, \mathbf{X}_{t_{i+1}})$.

Consider a fixed partition of $[0,T]$ given by $D = (t_i)_{i\in\{0,\dots n\}}$. For a fixed $i$ let $(\tau^i_j)_{j\in\{0,\dots N_i\}}$ be the 1-greedy partition on $[t_{i-1}, t_i]$. That is $\tau_0 = t_{i-1}$ and

$$\tau^i_j = \inf\left\{t > \tau^i_{j-1} \mid w(\tau^i_{j-1}, t) = 1\right\} \wedge t_j \ .$$

Since the greedy partition has $N_i := N_1(w, [t_{i-1}, t_i]) + 1$ elements, i.e. $\tau_{N_i} = t_i$ we have

$$\begin{aligned}\left\|\mathbf{X}_{t_{i-1},t_i}\right\|^p &\le \left(\sum_{j=1}^{N_i}\left\|\mathbf{X}_{\tau^i_{j-1},\tau^i_j}\right\|\right)^p \\ &\le (N_i)^{p-1}\sum_{j=1}^{N_i}\left\|\mathbf{X}_{\tau^i_{j-1},\tau^i_j}\right\|^p \\ &\le (N_1(w,[0,T])+1)^{p-1}\sum_{j=1}^{N_i}\left\|\mathbf{X}_{\tau^i_{j-1},\tau^i_j}\right\|^p .\end{aligned}$$

By summing up over our entire partition we get

$$\begin{aligned}\sum_{i=1}^{n}\left\|\mathbf{X}_{t_{i-1},t_i}\right\|^p &\le (N_1(w,[0,T])+1)^{p-1}\sum_{i=1}^{n}\sum_{j=1}^{N_i}\left\|\mathbf{X}_{\tau^i_{j-1},\tau^i_j}\right\|^p \\ &\le (N_1(w,[0,T])+1)^{p-1}\sup_{\substack{D([0,T])\\ w(t_{i-1},t_i)\le 1}}\sum_{[t_{j-1},t_j]\in D([0,T])} w(t_{j-1},t_j) \\ &\le 2(N_1(w,[0,T])+1)^p\end{aligned}$$

where in the last step we used [6, Proposition 4.11]. Since the right hand side is now independent of the partition $D$ we can take the supremum over all partitions. Then using $x^e \le e^x \ \forall x > 0$ which implies $x^p \le e^{\frac{p}{e}x}$ we get

$$|||\mathbf{X}|||^p_{p-var} \le 2e^{\frac{p}{e}}e^{\frac{p}{e}N_1(w_X,[0,T])}$$

which is the claimed result. □

## 1.3 Notation

Here is a collection of notation we use throughout this work.

- We use $C$ as a constant that is independant of whatever the context ask it to be independent of. When we want to indicate its dependencies explicitly we do so via subscripts. The constant may change from line to line without extra notice. We try to first write out terms clearly before hiding them in the constant.
- By $\delta_x$ we denote the dirac delta on $x$.
- As is customary we use $\omega$ to indicate the random parameter and $w$ when talking about controls. The reader is advised to take care not to confuse the two.
- We use the notation $W^{(q)}_E$ to indicate the $q$-Wasserstein distance on $\mathcal{P}_q(E)$, the space of probability measures on $E$ with finite $q$-th moment. See Appendix A for a summary of relevant properties.

- For a controlled path $(Y, Y') \in \mathscr{D}_X^{\frac{p}{2}}$ we write the remainder of second kind as

$$Y^{\#}_{s,t} := Y_{s,t} - Y'_s X_{s,t}$$

and the remainder of third kind as

$$Y^{\natural}_{s,t} := Y_{s,t} - Y'_s X_{s,t} - Y''_s \mathbb{X}_{s,t}$$

when it is clear how $Y''$ is defined. Note that in the case $Y' = f(Y)$ we have $Y''_s = f(Y)'_s = Df(Y_s)Y'_s$ as explained in Lemma 1.4.

# Chapter 2

# Controlled RDEs With Drift

In this chapter we discuss rough differential equations with a drift component. We will proceed to show a priori estimates for solutions and show existence and uniqueness of solutions. This is a deterministic theory and no randomness will be considered at this point.

## 2.1 Preliminaries

General solution theory for differential equations with rough driver is well understood and can be found in e.g. [15] and [16].

Let again $p \in [2,3)$. We are interested in solutions to equations of the form

$$dY_t = g(Y_t, t)dt + f(Y_t)d\mathbf{X}_t \tag{2.1}$$

for some well behaved functions $g$ and $f$ and $\mathbf{X} \in \mathscr{C}^{p-var}$ a rough path. Note that $g$ depends on both $Y$ and $t$ whereas $f$ only depends on $Y$. This is enough for our application so we restrict our attention to this case. Including a dependence on $t$ for $f$ is possible but requires one to assume $f(y, \cdot)$ to have finite $\frac{p}{2}$-variation. However this is not implied as a consequence of the McKean-Vlasov equations studied in chapter 3 thus is of no interest to us.

Clearly, one can consider $(X_t, t)$ as a rough path by adding the remaining cross-integrals $(\mathbb{X}, \int X\,dt, \int t\,dX, \int t\,dt)$. Here the second and third entry are interpreted as Riemann-Stieltjes integrals. This allows us to apply the standard theory seen e.g. in [15, 16] to gain a solution $\tilde{Y} = (Y_t, t)$ to the modified problem

$$d\tilde{Y}_t = \tilde{f}(Y_t, t)d\tilde{\mathbf{X}}_t$$

with $\tilde{f}_i = (f_i, 0)\,\forall i \in \{1 \ldots d\}$ and $\tilde{f}_{d+1} = (g, 1)$. Taking the projection onto $Y$ then gives us both existence and uniqueness conditions for our problem. The condition $\tilde{f} \in \mathcal{C}_b^3$ ensures global existence of that unique solution. This however forces $\tilde{f}$ to be in $\mathcal{C}^3$ jointly in $y$ and $t$ which is not a reasonable assumption in light of (1.2). For this reason we will consider the theory of rough differential equations with a drift component here.

## 2.2 A Priori Estimates

We are looking for solutions to differential equations of the form (2.1).

**Definition 2.1.** Let $g$ be measurable and bounded, $f \in C_b^2$, $Y_0 \in \mathbb{R}^d$ some initial condition and $\mathbf{X} \in \mathscr{C}^{p-var}([0,T], \mathbb{R}^e)$. Then we say $(Y, Y') \in \mathscr{D}_X^{\frac{p}{2}}$ is a solution to (2.1) if for all $t$ in $[0,T]$ we have

$$Y_t = Y_0 + \int_0^t g(Y_r, r)dr + \int_0^t f(Y_r)d\mathbf{X}_r \ .$$

Here $f(Y)$ is interpreted as an element of $\mathscr{D}_X^{\frac{p}{2}}$ via Lemma 1.4.

In the following we will need a priori estimates for solutions to RDEs with drift. Note that these bounds only require $f \in \mathcal{C}_b^2$ whereas the existence theorem assumes $f \in \mathcal{C}_b^3$.

**Proposition 2.2** (Local A Priori Estimate)**.** *Let $(Y, Y')$ be a solution to the RDE (2.1) in the sense of Definition 2.1 with $f \in \mathcal{C}_b^2$, $g$ measurable and bounded and $\mathbf{X} \in \mathscr{C}^{p-var}$. Let $w_X$ be the smallest control such that $|X_{s,t}| \le w_X(s,t)^{\frac{1}{p}}$ and $|\mathbb{X}_{s,t}| \le w_X(s,t)^{\frac{2}{p}}$ for all $s,t \in [0,T]$. Then for all $s,t$ small enough such that $C(w_X(s,t)^{\frac{1}{p}} + |t-s|) \le \frac{1}{2}$ for some constant $C$ we have*

- $w_Y(s,t)^{\frac{1}{p}} \le C\left(w_X(s,t)^{\frac{1}{p}} + |t-s|\right)$,
- $w_{Y^\#}(s,t)^{\frac{2}{p}} \le C\left(w_X(s,t)^{\frac{2}{p}} + |t-s|\right)$ *and*
- $w_{Y^\natural}(s,t)^{\frac{3}{p}} \le C\left(w_X(s,t)^{\frac{3}{p}} + |t-s|\right)$.

*Proof.* We defined the rough integral via the sewing lemma Lemma 1.5 which combined with Lemma 1.4 gives us

$$\begin{aligned} |Y^\natural_{s,t}| &= \left| \int_s^t f(Y_r)d\mathbf{X}_r - f(Y_s)X_{s,t} - f(Y)'_s\mathbb{X}_{s,t} \right| \\ &\le C\left(w_X(s,t)^{\frac{1}{p}} w_{f(Y)^\#}(s,t)^{\frac{2}{p}} + w_X(s,t)^{\frac{2}{p}} w_{f(Y)}(s,t)^{\frac{1}{p}}\right) \\ &\le C(w_X(s,t)^{\frac{1}{p}}(w_Y(s,t)^{\frac{2}{p}} + w_{Y^\#}(s,t)^{\frac{2}{p}}) + w_X(s,t)^{\frac{2}{p}} w_Y(s,t)^{\frac{1}{p}}). \end{aligned} \tag{2.2}$$

By definition we have

$$|Y^\#_{s,t}| = |f(Y)'_s\mathbb{X}_{s,t} + \int_s^t g(Y_r,r)dr + Y^\natural_{s,t}| \le C\left(w_X(s,t)^{\frac{2}{p}} + |t-s| + w_{Y^\natural}(s,t)^{\frac{3}{p}}\right)$$

thus giving us

$$|Y_{s,t}| = |Y'_s X_{s,t} + Y^\#_{s,t}| \le C\left(w_X(s,t)^{\frac{1}{p}} + |t-s| + w_X(s,t)^{\frac{2}{p}} + w_{Y^\natural}(s,t)^{\frac{3}{p}}\right).$$

Plugging this back into Equation 2.2 we get

$$\begin{aligned} |Y^\natural_{s,t}| \le w_{Y^\natural}(s,t)^{\frac{3}{p}} &\le C\Big(w_X(s,t)^{\frac{2}{p}}(w_X(s,t)^{\frac{1}{p}} + |t-s| + w_{Y^\natural}(s,t)^{\frac{3}{p}}) \\ &\quad + w_X(s,t)^{\frac{1}{p}}(w_X(s,t)^{\frac{2}{p}} + |t-s|^2 + w_{Y^\natural}(s,t)^{\frac{6}{p}}) \\ &\quad + w_X(s,t)^{\frac{1}{p}}(w_X(s,t)^{\frac{2}{p}} + |t-s| + w_{Y^\natural}(s,t)^{\frac{3}{p}})\Big) \\ &\le C\left(w_X(s,t)^{\frac{3}{p}} + |t-s| + w_X(s,t)^{\frac{1}{p}}(w_{Y^\natural}(s,t)^{\frac{3}{p}} + w_{Y^\natural}(s,t)^{\frac{6}{p}})\right). \end{aligned}$$

Thus for $s, t$ such that $Cw_X(s,t)^{\frac{1}{p}} \le \frac{1}{2}$ and multiplying with $Cw_X(s,t)^{\frac{1}{p}}$ we have

$$\begin{aligned} w_{Y^\natural}(s,t)^{\frac{3}{p}} &\le C\left(w_X(s,t)^{\frac{3}{p}} + |t-s| + w_X(s,t)^{\frac{1}{p}} w_{Y^\natural}(s,t)^{\frac{6}{p}}\right) \\ \implies A_{s,t} &\le \lambda_{s,t} + A_{s,t}^2 \end{aligned}$$

for $A_{s,t} := Cw_X(s,t)^{\frac{1}{p}} w_{Y^\natural}(s,t)^{\frac{3}{p}}$ and $\lambda_{s,t} := C(w_X(s,t)^{\frac{4}{p}} + w_X(s,t)^{\frac{1}{p}}|t-s|)$.

Now clearly for $|t-s|$ small enough such that $A_{s,t} < \frac{1}{2}$ we get $A_{s,t} \le 2\lambda_{s,t}$ which yields our result. However we want that choice of $|t-s|$ to be made independently of $Y$. This can be seen to be possible in the following way.

For $s, t$ such that $\lambda_{s,t} < \frac{1}{4}$ we get

$$\begin{aligned} A_{s,t}^2 - A_{s,t} \ge \lambda_{s,t} &\iff \begin{cases} \sqrt{\frac{1}{4} - \lambda_{s,t}} & \le A_{s,t} - \frac{1}{2} \\ \sqrt{\frac{1}{4} - \lambda_{s,t}} & \le -A_{s,t} + \frac{1}{2} \end{cases} \\ &\iff \begin{cases} A_{s,t} & \ge \frac{1}{2} \\ A_{s,t} & \le \frac{1}{2} - \sqrt{\frac{1}{4} - \lambda_{s,t}} \xrightarrow{|t-s| \to 0} 0. \end{cases} \end{aligned}$$

Since $A_{s,t}$ is monotone and we know that it goes to 0 for $|t-s| \to 0$ we have that for all $s, t$ with $\lambda_{s,t} \le \frac{1}{8}$ it holds that $A_{s,t} \le \frac{1}{4}$. This gives us

$$\begin{aligned} A_{s,t} &\le C\lambda_{s,t} \\ \implies w_{Y^\natural}(s,t)^{\frac{3}{p}} &\le C(w_X(s,t)^{\frac{3}{p}} + |t-s|). \end{aligned}$$

Plugging this back into the above estimates and using $w_X(s,t) \le 1$ yields the claimed result.

Along the way we used $\lambda_{s,t} \le \frac{1}{8}$. We see that there exists a $C$ such that $C(w_X(s,t)^{\frac{1}{p}} + |t-s|) \le \frac{1}{2}$ implies $\lambda_{s,t} \le \frac{1}{8}$. □

From this we get a straight forward global a priori estimate in the $\alpha$-Hölder norm.

**Corollary 2.3** (Global A Priori Estimate)**.** *For $\mathbf{X} \in \mathscr{C}^\alpha$ for some $\alpha \in (\frac{1}{3}, \frac{1}{2}]$ and $(Y, Y')$ as above the solution to Equation 2.1 we have for $h > 0$ small enough*

$$\|Y\|_{\alpha,h} \le C\left(|||\mathbf{X}|||_{\alpha,h} + h^{1-\alpha}\right). \tag{2.3}$$

*As a consequence we get the global estimate*

$$\|Y\|_\alpha \le C\left(|||\mathbf{X}|||_\alpha + |||\mathbf{X}|||_\alpha^{1-\frac{1}{\alpha}} \vee |||\mathbf{X}|||_\alpha^{\frac{1}{\alpha}} + 1\right). \tag{2.4}$$

*For both estimates $C$ depends only on $\alpha$.*

*Proof.* For $\mathbf{X}$ we know $w_X(s,t) \le |||\mathbf{X}|||_{\alpha,h}^p |t-s|$ for any $|t-s| \le h$. Thus the assumption from Proposition 2.2 is satisfied for $h$ small enough such that $Cw_X(s,t)^{\frac{1}{p}} \le C|||\mathbf{X}|||_{\alpha,h}|t-s|^\alpha \le \frac{1}{2}$ holds. Pick such an $h$. Then for $|t-s| \le h$ using Proposition 2.2 we have

$$|Y_{s,t}| \le C(w_X(s,t)^{\frac{1}{p}} + |t-s|) \le C(|||\mathbf{X}|||_\alpha |t-s|^\alpha + |t-s|).$$

This gives us Equation 2.3.

By choosing $h \simeq \|\|\mathbf{X}\|\|_\alpha^{-\frac{1}{\alpha}}$ we get $h^{\alpha-1} \simeq \|\|\mathbf{X}\|\|_\alpha^{\frac{1}{\alpha}-1}$ and thus by applying Lemma 1.3 we have

$$\begin{aligned}\|Y\|_\alpha &\leq \|Y\|_{\alpha,h}\left(1 \vee 2h^{\alpha-1}\right)\\ &\leq C\left(\|\|\mathbf{X}\|\|_{\alpha,h} + h^{1-\alpha}\right)\left(1 \vee 2h^{\alpha-1}\right)\\ &\leq C\left(\|\|\mathbf{X}\|\|_\alpha + \|\|\mathbf{X}\|\|_\alpha^{1-\frac{1}{\alpha}} \vee \|\|\mathbf{X}\|\|_\alpha^{\frac{1}{\alpha}} + 1\right).\end{aligned}$$

□

We can bound any solution $Y$ via the control $w_X$ in the following way.

**Corollary 2.4.** *Under the conditions of Proposition 2.2 there exists a universal constant $C$ for which we get*

$$\|Y\|_{\infty,[0,T]} \leq |Y_0| + C\left(w_X(0,T) + T\right).$$

*Proof.* Note that the local a priori estimates in Proposition 2.2 give us for any such $Y$

$$\|Y\|_{\infty,[s,t]} \leq |Y_s| + w_Y(s,t)^{\frac{1}{p}} \leq |Y_s| + C(w_X(s,t)^{\frac{1}{p}} + |t-s|)$$

for $|t-s|$ such that $Cw_X(s,t)^{\frac{1}{p}} \leq \frac{1}{2}$. Thus building the greedy partition $(t_i)$ over $\mathbf{X}$ given inductively by

$$t_{i+1} = \inf\{t_i < t < T \mid Cw_X(t_i,t)^{\frac{1}{p}} \geq \frac{1}{2}\} \wedge T$$

we get

$$\begin{aligned}\|Y\|_{\infty,[0,T]} &\leq |Y_0| + \sum_{i=1}^{N_{\frac{1}{(2C)^p}}(w_X,[0,T])} w_{Y_i}(t_i,t_{t+1})^{\frac{1}{p}}\\ &\leq |Y_0| + C\left(\sum_{i=1}^{N_{\frac{1}{(2C)^p}}(w_X,[0,T])} w_X(t_i,t_{i+1})^{\frac{1}{p}} + T\right)\\ &\leq |Y_0| + C(N_{\frac{1}{(2C)^p}}(w_X,[0,T]))^{\frac{p-1}{p}}\left(\sum_{i=1}^{N_{\frac{1}{(2C)^p}}(w_X,[0,T])} w_X(t_i,t_{i+1})\right)^{\frac{1}{p}} + CT\\ &\leq |Y_0| + C\left(w_X(0,T) + T\right).\end{aligned}$$

Here in the last step we used

$$N_{\frac{1}{(2C)^p}}(w_X,[0,T]) \leq (2C)^p N_1(w_X,[0,T]) \leq (2C)^p w_X(0,T).$$

□

## 2.3 Stability Estimates

After showing our a priori estimates for solutions of RDEs with drift we will now concern ourselves with stability estimates. In the spirit of rough paths we will see that the solution map is locally Lipschitz with respect to input noise and changes in functions $g, f$. We will then later use this to show well-posedness and propagation of chaos for the McKean-Vlasov equation.

**Proposition 2.5.** *Let $\mathbf{X} \in \mathscr{C}^{p-var}$ be a rough path with $w_X$ the smallest control such that $|X_{s,t}| \leq w_X(s,t)^{\frac{1}{p}}$, $|\mathbb{X}_{s,t}| \leq w_X(s,t)^{\frac{2}{p}}$. Let $dY_t = g^1(t, Y_t)dt + f(Y_t)d\mathbf{X}_t$ and $dZ_t = g^2(t, Z_t)dt + f(Z_t)d\mathbf{X}_t$ be two solutions to RDEs with the two drift functions $g^1, g^2$ and initial conditions $\xi^1, \xi^2$. We define $w_1(s,t) = w_X(s,t) + |t-s|$.*

*Assume $f \in \mathcal{C}_b^3$ and that there exists a constant $L$ such that*

1. *$|g^1(t,y) - g^2(t,z)| \leq L(|z-y| + k_t)$ for some measurable function $k$ and*
2. *$|g^i(t,y)| \leq L$ for $i = 1, 2$.*

*Then there exists a constant $C$ such that for $W_t = Y_t - Z_t$ we have for any $[s,t] \subseteq [0,T]$*

$$\|W\|_{\infty,[s,t]} \leq Ce^{CN_1(w_1,[s,t])} \left( |W_s| + \int_s^t k_r dr \right) \tag{2.5}$$

*and for $s, t$ small enough such that $Cw_1(s,t) \leq \frac{1}{2}$ we have*

$$|W_{s,t}| \leq Ce^{CN_1(w_1,[0,T])}(|W_s| + \int_s^t k_r dr). \tag{2.6}$$

*Proof.* We will use the notation $\Delta = f(Y) - f(Z)$.

First, we have

$$W_{s,t} = (f(Y_s) - f(Z_s))X_{s,t} + Y^\#_{s,t} - Z^\#_{s,t}$$

and thus

$$|W_{s,t}| \leq \|f\|_{C^1} |W_s| w_X(s,t)^{\frac{1}{p}} + w_{W^\#}(s,t)^{\frac{2}{p}}. \tag{2.7}$$

For $W^\#$ we have

$$\begin{aligned} W^\#(s,t) &= \left(f(Y)'_s - f(Z)'_s\right)\mathbb{X}_{s,t} + \int_s^t g^1(r, Y_r) - g^2(r, Z_r)dr \\ &\quad + \left( \int_s^t f(Y_r) - f(Z_r)d\mathbf{X}_r - (f(Y_s) - f(Z_s))X_{s,t} - \left(f(Y)'_s - f(Z)'_s\right)\mathbb{X}_{s,t} \right) \\ &= \left(f(Y)'_s - f(Z)'_s\right)\mathbb{X}_{s,t} + \int_s^t g^1(r, Y_r) - g^2(r, Z_r)dr \\ &\quad + \left( \int_s^t \Delta_r d\mathbf{X}_r - \Delta_s X_{s,t} - \Delta'_s \mathbb{X}_{s,t} \right). \end{aligned}$$

By using Corollary 1.6 this gives us

$$\begin{aligned} |W^\#(s,t)| &\leq \|f\|_{C^2} |W_s| w_X(s,t)^{\frac{2}{p}} + K|t-s| \sup_{r\in[s,t]} |W_r| + \int_s^t k_r dr \\ &\quad + w_X(s,t)^{\frac{1}{p}} w_{\Delta^\#}(s,t)^{\frac{2}{p}} + w_X(s,t)^{\frac{2}{p}} w_{\Delta'}(s,t)^{\frac{1}{p}}. \end{aligned} \tag{2.8}$$

By now providing estimates for $|\Delta'(s,t)|, |\Delta^\#(s,t)|$ we can make progress.

We have

$$\begin{aligned}
|\Delta'_{s,t}| &= |f(Y)'_{s,t} - f(Z)'_{s,t}| \\
&= |Df(Y_t)Y'_t - Df(Y_s)Y'_s - \left(Df(Z_t)Z'_t - Df(Z_s)Z'_s\right)| \\
&= |Df(W_t)Y'_t - Df(W_s)Y'_s + Df(Z_t)W'_t - Df(Z_s)W'_s| \\
&\leq \|f\|_{C^2} |W_t||Y'_{s,t}| + \|f\|_{C^2} |W_{s,t}||Y'_s| + \|f\|_{C^2} |W'_{s,t}| + \|f\|_{C^2} |Z_{s,t}||W_s| \\
&\leq C_f \left(|W_s||Z_{s,t}| + |W_t||Y_{s,t}| + |W_{s,t}|\right) \\
&\leq C_f \left( \left(w_X(s,t)^{\frac{1}{p}} + |t-s|\right) \sup_{r\in[s,t]} |W_r| + w_W(s,t)^{\frac{1}{p}} \right).
\end{aligned}$$

Here in the last step we used the a priori estimates from Proposition 2.2. Furthermore

$$\begin{aligned}
|\Delta^{\#}_{s,t}| &= |\Delta_{s,t} - \Delta'_s X_{s,t}| \\
&= |f(Y_t) - f(Y_s) - Df(Y_s)Y_{s,t} - Df(Y_s)Y^{\#}_{s,t} \\
&\quad - \left(f(Z_t) - f(Z_s) - Df(Z_s)Z_{s,t} - Df(Z_s)Z^{\#}_{s,t}\right)| \\
&\leq |f(Y_t) - f(Y_s) - Df(Y_s)Y_{s,t} - \left(f(Z_t) - f(Z_s) - Df(Z_s)Z_{s,t}\right)| \\
&\quad + |Df(Y_s)Y^{\#}_{s,t} - Df(Z_s)Z^{\#}_{s,t}| \\
&=: T_1 + T_2.
\end{aligned}$$

Now for estimating $T_1$ and $T_2$ separately we get

$$\begin{aligned}
T_1 &= |\int_0^1 \left(D^2 f(Y_s + \theta Y_{s,t})(Y_{s,t}, Y_{s,t}) - D^2 f(Z_s + \theta Z_{s,t})(Z_{s,t}, Z_{s,t})\right)(1-\theta)d\theta| \\
&\leq C_f \left(|W_s|(w_X(s,t)^{\frac{2}{p}} + |t-s|) + w_W(s,t)^{\frac{1}{p}}(w_X(s,t)^{\frac{2}{p}} + |t-s|) + w_W(s,t)^{\frac{2}{p}}\right)
\end{aligned}$$

and

$$\begin{aligned}
T_2 &\leq |Df(Y_s)||W^{\#}_{s,t}| + |Df(Y_s) - Df(Z_s)||Z^{\#}_{s,t}| \\
&\leq \|f\|_{C^1} |W^{\#}_{s,t}| + \|f\|_{C^2} |W_s|(w_X(s,t)^{\frac{2}{p}} + |t-s|)C \\
&\leq w_{W^{\#}}(s,t)^{\frac{2}{p}} + |W_s|(w_X(s,t)^{\frac{2}{p}} + |t-s|).
\end{aligned}$$

Collecting these estimates and plugging them into (2.8) we get

$$\begin{aligned}
|W^{\#}_{s,t}| \leq C\Big(&|W_s| w_X(s,t)^{\frac{2}{p}} + |W_s|(w_X(s,t)^{\frac{3}{p}} + |t-s|) + w_X(s,t)^{\frac{1}{p}}|W_{s,t}|^2 \\
&+ w_X(s,t)^{\frac{2}{p}} \left( (w_X(s,t)^{\frac{1}{p}} + |t-s|) \sup_{r\in[s,t]} |W_r| + w_X(s,t)^{\frac{2}{p}} |W_{s,t}| \right) \\
&+ w_X(s,t)^{\frac{1}{p}} |W^{\#}_{s,t}| + |t-s| \sup_{r\in[s,t]} |W_r| + \int_s^t k_r dr\Big).
\end{aligned}$$

Thus for $s, t$ such that $C w_X(s,t) \leq \frac{1}{2}$ we get

$$\begin{aligned}
|W^{\#}_{s,t}| \leq C\Big(&(w_X(s,t) + |t-s|^p)^{\frac{1}{p}} \sup_{r\in[s,t]} |W_r| + w_X(s,t)^{\frac{1}{p}} |W_{s,t}|^2 \\
&+ w_X(s,t)^{\frac{2}{p}} |W_{s,t}| + \int_s^t k_r dr\Big).
\end{aligned}$$

Plugging this back into (2.7) we get for $s, t$ small enough such that $W_{s,t} \leq 1$ via the a priori estimates

$$|W_{s,t}| \leq C\Big((w_X(s,t) + |t-s|)^{\frac{1}{p}} \sup_{r\in[s,t]} |W_r| + w_X(s,t)^{\frac{1}{p}} |W_{s,t}| + \int_s^t k_r dr\Big) \tag{2.9}$$

and thus again for $s, t$ small enough such that $Cw_X(s,t)^{\frac{1}{p}} \leq \frac{1}{2}$ we get

$$|W_{s,t}| \leq C\left(w_1(s,t)^{\frac{1}{p}} \sup_{r\in[s,t]} |W_r| + \int_s^t k_r dr\right).$$

Now applying the rough Grönwall lemma Proposition 1.7 with $w_1(s,t) = C\left(w_X(s,t) + |t-s|^p\right)$ and $w_2(s,t) = C\int_s^t k_r dr$ we get

$$\sup_{r\in[s,t]} W_r \leq C\exp\left(Cw_1(s,t)\right)\left(|W_s| + w_2(s,t)\right)$$

for any interval $[s,t] \subset [0,T]$. Crucially there is no assumption of small $|t-s|$ imposed here.

Now let $\{\tau_n\}_n$ be the greedy partition of $[s,t]$ introduced in Definition 1.2 for $w_1$ and $\beta = 1$. Due to $w_1(s,\tau_1) = 1$ we then have

$$\sup_{r\in[s,\tau_1]} |W_r| \leq Ce^C\left(|W_s| + w_2(s,\tau_1)\right).$$

Equivalently we have for $C \geq 1$ that

$$\begin{aligned}\sup_{r\in[\tau_1,\tau_2]} |W_r| &\leq Ce^C\left(|W_{\tau_1}| + w_2(\tau_1,\tau_2)\right)\\ &\leq Ce^C\left(Ce^C\left(|W_s| + w_2(s,\tau_1)\right) + w_2(\tau_1,\tau_2)\right)\\ &\leq C^2e^{2C}\left(|W_s| + w_2(s,\tau_2)\right).\end{aligned}$$

Thus via an induction we get

$$\begin{aligned}\sup_{r\in[\tau_n,\tau_{n+1}]} |W_r| &\leq C^n e^{nC}\left(|W_s| + w_2(s,\tau_{n+1})\right)\\ &= e^{n(C+ln(C))}\left(|W_s| + w_2(s,\tau_{n+1})\right)\end{aligned}$$

and hence for $\tilde{C} = C + \ln(C)$ we have

$$\begin{aligned}\sup_{r\in[s,t]} |W_r| &\leq \max_{n<N_1(w_1,[s,t])} \sup_{r\in[\tau_n,\tau_{n+1}]} |W_r|\\ &\leq e^{\tilde{C}N_1(w_1,[s,t])}\left(|W_s| + w_2(s,t)\right)\end{aligned}$$

giving us the claimed result (2.5) by setting $[s,t] = [0,T]$.

We can now go back and plug this back into (2.9) to obtain for $s, t$ such that $Cw_1(s,t) \leq 1$

$$\begin{aligned}|W_{s,t}| &\leq C\left(w_1(s,t)^{\frac{1}{p}} e^{CN_1(w_1,[s,t])}\left(|W_s| + w_2(s,t)\right) + w_2(s,t)\right)\\ &\leq Ce^{CN_1(w_1,[s,t])}(|W_s| + w_2(s,t))\\ &\leq Ce^{CN_1(w_1,[0,T])}(|W_s| + \int_s^t k_r dr)\end{aligned}$$

which is exactly (2.6). □

We stress that (2.5) does not carry a smallness assumption for $|t-s|$ whereas (2.6) does.

Note that uniqueness of solution follows directly.

This shows stability under perturbation of the drift function $g$. In our case it will also be important to have stability under change of the driving noise $\mathbf{X}$. The above result can easily be extended to include this in the following way.

**Proposition 2.6.** *Let $\mathbf{X}, \tilde{\mathbf{X}} \in \mathscr{C}^{p-var}$ be two rough paths with $w_X$ the smallest control such that $|X_{s,t}| \leq w_X(s,t)^{\frac{1}{p}}$, $|\mathbb{X}_{s,t}| \leq w_X(s,t)^{\frac{2}{p}}$ likewise $w_{\tilde{X}}$ for $\tilde{\mathbf{X}}$, $f \in \mathcal{C}_b^3$ and $g^1, g^2$ bounded functions Lipschitz in $y$. Let $dY_t = g^1(t, Y_t)dt + f(Y_t)d\mathbf{X}_t$ and $dZ_t = g^2(t, Y_t)dt + f(Y_t)d\tilde{\mathbf{X}}_t$ be two solutions to RDEs with the two drift functions $g^1, g^2$ driven by $\mathbf{X}, \tilde{\mathbf{X}}$ and initial conditions $\xi^1, \xi^2$. We define $w_1(s,t) = w_X(s,t) + w_{\tilde{X}}(s,t) + |t-s|$, $Q = X - \tilde{X}$ and $\mathbb{Q} = \mathbb{X} - \tilde{\mathbb{X}}$ with the corresponding control $w_Q$.*

*Under the same assumptions as in Proposition 2.5 there exists a universal constant $C$ depending only on $p$ such that for $W_t = Y_t - Z_t$ and for any $[s,t] \subseteq [0,T]$ we have*

$$\|W\|_{\infty,[s,t]} \leq Ce^{CN_1(w_1,[s,t])} \left( |W_s| + w_Q(s,t)^{\frac{1}{p}} + \int_s^t k_r dr \right). \tag{2.10}$$

*Furthermore for $s,t$ small enough such that $Cw_1(s,t) \leq \frac{1}{2}$ we have*

$$|W_{s,t}| \leq Ce^{CN_1(w_1,[s,t])} \left( |W_s| + w_Q(s,t)^{\frac{1}{p}} + \int_s^t k_r dr \right). \tag{2.11}$$

*Proof.* This proof works essentially the same way as the proof of Proposition 2.5. Some changes need to be made to get the control estimates for $W_{s,t}$ to then again apply the rough Grönwall lemma. Mostly estimates of the form $|ab - \tilde{a}\tilde{b}| \leq |a||b-\tilde{b}| + |a-\tilde{a}||\tilde{b}|$ are used. We outline the argument without completing every single computation as they are essentially the same as above.

We have

$$\begin{aligned} W_{s,t} &= Y_s' X_{s,t} - Z_s' \tilde{X}_{s,t} + W_{s,t}^{\#} \\ \implies \quad |W_{s,t}| &\leq |Y_s'||X_{s,t} - \tilde{X}_{s,t}| + |Y_s' - Z_s'||\tilde{X}_{s,t}| + |W_{s,t}^{\#}| \\ &\leq C|Q_{s,t}| + C|W_s||\tilde{X}_{s,t}| + |W_{s,t}^{\#}|. \end{aligned}$$

Thus we need to find a control for $W^{\#}$ to proceed.

$$\begin{aligned} W_{s,t}^{\#} &= Y_{s,t} - Y_s' X_{s,t} - \left( Z_{s,t} - Z_s' \tilde{X}_{s,t} \right) \\ &= \int_s^t f(Y_r) d\mathbf{X}_r - Y_s' X_{s,t} - f(Y)_s' \mathbb{X}_{s,t} + f(Y)_s' \mathbb{X}_{s,t} + \int_s^t b^1(r, Y_r) dr \\ &- \Big( \underbrace{\int_s^t f(Z_r) d\tilde{\mathbf{X}}_r - Z_s' \tilde{X}_{s,t} - f(Z)_s' \tilde{\mathbb{X}}_{s,t}}_{:=E_1} + f(Z)_s' \tilde{\mathbb{X}}_{s,t} + \int_s^t b^2(r, Z_r) dr \Big) \end{aligned}$$

Now for $\Xi_{s,t} = Y'_s X_{s,t} + f(Y)'_s \mathbb{X}_{s,t}$, $\tilde{\Xi}_{s,t} = Z'_s \tilde{X}_{s,t} + f(Z)'_s \tilde{\mathbb{X}}_{s,t}$ we define $\Box_{s,t} = \Xi_{s,t} - \tilde{\Xi}_{s,t}$. Observe via Chen's relation that

$$\delta\Box_{s,u,t} = f(Z)^{\#}_{s,u}\tilde{X}_{u,t} - f(Y)^{\#}_{s,u}X_{u,t} + f(Z)'_{s,u}\tilde{\mathbb{X}}_{u,t} - f(Y)'_{s,u}\mathbb{X}_{u,t}$$

$$\implies |\delta\Box_{s,u,t}| \leq w_{f(Y)^{\#}}(s,t)^{\frac{2}{p}} w_Q(s,t)^{\frac{1}{p}} - w_{\Delta^{\#}}(s,t)^{\frac{2}{p}} w_{\tilde{X}}(s,t)^{\frac{1}{p}}$$
$$+ w_{f(Y)'}(s,t)^{\frac{1}{p}} w_Q(s,t)^{\frac{2}{p}} - w_{\Delta'}(s,t)^{\frac{1}{p}} w_{\tilde{X}}(s,t)^{\frac{2}{p}}.$$

Now using the sewing lemma and the discussion in section 1.2 we get

$$\begin{aligned} |E_1| &= \left| \int_s^t f(Y_r)d\mathbf{X}_r - \int_s^t f(Z_r)d\tilde{\mathbf{X}}_r - \Xi_{s,t} + \tilde{\Xi}_{s,t} \right| \\ &= |I(\Box)_{s,t} - (\Box_{s,t})| \\ &\leq C|w_{f(Y)^{\#}}(s,t)^{\frac{2}{p}} w_Q(s,t)^{\frac{1}{p}} - w_{\Delta^{\#}}(s,t)^{\frac{2}{p}} w_{\tilde{X}}(s,t)^{\frac{1}{p}} \\ &+ w_{f(Y)'}(s,t)^{\frac{1}{p}} w_Q(s,t)^{\frac{2}{p}} - w_{\Delta'}(s,t)^{\frac{1}{p}} w_{\tilde{X}}(s,t)^{\frac{2}{p}}|. \end{aligned}$$

Thus we need estimates for $|f(Y)^{\#}_{s,t}|$, $|\Delta^{\#}_{s,t}|$, $|\Delta'_{s,t}|$ and $|f(Y)'_{s,t}|$. Using the same arguments as in Proposition 2.5 for these and combining them we get for $s,t$ small enough such that $Cw_{\tilde{X}}(s,t) \leq \frac{1}{2}$

$$|W_{s,t}| \leq C\left( w_1(s,t)^{\frac{1}{p}} \sup_{r\in[s,t]} |W_r| + w_Q(s,t)^{\frac{1}{p}} + \int_s^t k_r dr \right). \tag{2.12}$$

From here we proceed the same way as above in using the rough Grönwall with $w_1 = w_X(s,t) + w_{\tilde{X}}(s,t) + |t-s|$ and $w_2(s,t) = w_Q(s,t)^{\frac{1}{p}} + \int_s^t k_r dr$ and induction over the greedy partition steps to conclude with (2.10).

Again plugging this back into (2.12) yields (2.6). □

*Remark* 2.7. We note that $w_Q(s,t)^{\frac{1}{p}} + \int_s^t k_r dr$ in (2.12) is not a control which is precisely the reason why we needed a slightly more general statement in the rough Grönwall Lemma Proposition 1.7 than was provided in [10].

## 2.4 Well-Posedness

After having proven these a priori estimates we now show the existence and uniqueness of solutions.

**Theorem 2.8** (Well-posedness). *Let $\mathbf{X} \in \mathscr{C}^{p-var}$ and $g, f$ satisfy the assumptions in Proposition 2.5. Then there exists a unique solution to equation* (2.1) *with initial condition $\xi \in \mathbb{R}^d$.*

*Proof.* Uniqueness is a direct consequence of Proposition 2.6. Given two solutions $Y^1, Y^2$ it follows from (2.10) that $\|Y^1, Y^2\|_\infty = 0$.

It is left to show existence of a solution. We proceed via Picard iteration. Let

$$\left(Y^0_t, (Y^0)'_t\right) = \left(\xi + f(\xi)X_{0,t}, f(\xi)\right).$$

Note that this is indeed a path controlled by $X$ since $Y^{0,\#}_{s,t} = 0 \in \mathcal{C}^{\frac{p}{2}-var}$. Next we define the iteration by

$$\left(Y^{n+1}_t, (Y^{n+1})'_t\right) = \left(\xi + \int_0^t g(r, Y^n_r)dr + \int_0^t f(Y^n_r)d\mathbf{X}_r, f(Y^n_r)\right).$$

Again we note that this is a path controlled by $X$ which can be seen by applying Corollary 1.6. We now show that the sequence $(Y^n)_{n\in\mathbb{N}}$ defined this way is uniformly bounded and equicontinuous to then apply Arzelà-Ascoli and pass to the limit. Specifically we want to show inductively that there exist constants $D, h$ independent of $n$ such that for all $s, t$ satisfying $w_X(s,t)^{\frac{1}{p}} + |t-s| \le h$ we have

$$\begin{aligned} |Y^n(s,t)| &\le 5L|t-s| + 5C_{p,f} w_X(s,t)^{\frac{1}{p}} + D w_X(s,t)^{\frac{1}{p}}, \\ |Y^{n,\#}_{s,t}| &\le 5L|t-s| + 5C_{p,f} w_X(s,t)^{\frac{2}{p}} + D^2 w_X(s,t)^{\frac{2}{p}}. \end{aligned} \tag{2.13}$$

Consider $Y^{n+1}$. Without loss of generality we assume $L \ge 1$ by setting $L = L \vee 1$. Then we have

$$\begin{aligned} |Y^{n+1}_{s,t}| \le{} & |\int_s^t g(r, Y^n_r) dr| + |f(Y^n_s) X_{s,t}| + |Df(Y^n_s) f(Y^{n-1}_s) \mathbb{X}_{s,t}| \\ & + |\int_s^t f(Y^n_r) d\mathbf{X}_r - f(Y^n_s) X_{s,t} - Df(Y^n_s) f(Y^{n-1}_s) \mathbb{X}_{s,t}| \\ \le{} & L|t-s| + C_{p,f} \Big( w_X(s,t)^{\frac{1}{p}} + w_X(s,t)^{\frac{2}{p}} + w_{Y^n}(s,t)^{\frac{1}{p}} w_X(s,t)^{\frac{2}{p}} \\ & + w_{Y^n}(s,t)^{\frac{2}{p}} w_X(s,t)^{\frac{1}{p}} + w_{Y^{n,\#}}(s,t)^{\frac{2}{p}} w_X(s,t)^{\frac{1}{p}} \Big). \end{aligned}$$

Again without loss of generality we assume $C_{p,f} \ge 1$. Let's assume $h \le \frac{1}{5C_{p,f}\sqrt{L}}$. Note that this implies $h \le \frac{1}{5}$. Plugging the induction hypothesis back in and using $w_X(s,t)^{\frac{1}{p}} + |t-s| \le h$ we get

$$\begin{aligned} |Y^{n+1}_{s,t}| \le{} & L|t-s| + 2C_{p,f} w_X(s,t)^{\frac{1}{p}} + C_{p,f} \Big[ \left(5L|t-s|h + 2(5L|t-s|)^2 h^2\right) w_X(s,t)^{\frac{1}{p}} \\ & + w_X(s,t)^{\frac{1}{p}} \Big(5C_{p,f} h + (5C_{p,f} h)^2\Big) + w_X(s,t)^{\frac{1}{p}} \left((D + 2D^2) h\right) \Big] \\ \le{} & 5L|t-s| + 5C_{p,f} w_X(s,t)^{\frac{1}{p}} + w_X(s,t)^{\frac{1}{p}} C_{p,f} h (D + 2D^2). \end{aligned}$$

Now solving the quadratic inequality $Ch(D + 2D^2) \le D$ shows that the induction hypothesis holds for a choice of $D \ge \frac{1}{2hC_{p,f}}$. The argumentation for $|Y^{n+1,\#}_{s,t}|$ follows the same steps which we chose to omit here.

This gives us equicontinuity and we can hence apply Arzelà-Ascoli. We get a subsequence $Y^{n_k}$ that converges uniformly to some $Y$ in $\mathcal{C}([0,T], \mathbb{R}^d)$. Due to the uniform bounds we can pass to the limit and get that $(Y, f(Y))$ is indeed a solution to (2.1). □

# Chapter 3

# Mean Field Equations

## 3.1 Introduction

Let $\mathbf{X}$ be some random rough path on $\mathbb{R}^e$. We call a process described by an SDE of the form

$$dY_t = g(Y_t, \mathcal{L}(Y_t))dt + f(Y_t, \mathcal{L}(Y_t))d\mathbf{X}_t \tag{3.1}$$

a McKean-Vlasov process. Here $\mathcal{L}(Y_t) \in \mathcal{P}(\mathbb{R}^d)$ denotes the law of $Y$ at time $t$. For a fixed law $\mu \in \mathcal{P}(\mathcal{C}([0,T], \mathbb{R}^d))$ replacing $\mathcal{L}(Y_t)$ with $\mu_t$ in (3.1) the dependence on the law can be hidden in a time variable. This gives us

$$dY_t^\mu = g^\mu(t, Y_t^\mu)dt + f^\mu(t, Y_t^\mu)d\mathbf{X}_t \ .$$

For this to have a solution in the sense of chapter 2 it is easy to see that we need $f^\mu$ to have finite $\frac{p}{2}$-variation in $t$. This kind of regularity does not hold in any generality for measures $\mu$. For a way to deal with this problem see [1] where a more sophisticated approach is chosen involving bounding $L$-differentials of $f$ in the law. Here we choose to consider systems with $f$ independent of the law. Since the conditions on $g^\mu$ are only measurability in $t$ and we in fact get even continuity when imposing $g$ to be Lipschitz in the law, we can apply the techniques from chapter 2 to analyse this case.

Note that the exponential term $e^{CN_1(w_1)}$ in Proposition 2.5 and Proposition 2.6 plays a crucial role in limiting the type of input processes that we can handle.

We use the $q$-Wasserstein distance $W^{(q)}$ as a metric on the space of measures. For a definition and a summary of properties we refer to Appendix A.

## 3.2 Classical Theory

Let us first review some classical results about McKean-Vlasov equations in the Itô setting. Here it is possible to include dependence on the measure in the diffusion term without relying on $L$-differentiation and other advanced techniques. We will use the techniques of Itô calculus to derive existence and uniquess of a solution for the McKean-Vlasov equation in this standard setting.

We will provide an overview over the argumentation for existence and uniqueness of solution in the classical case. This is instructive since similar contractive arguments are made in the rough case. For more details see [5].

In the following let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space that supports a $e$-dimensional standard Brownian motion $(W_t)_{t\in[0,T]}$ with $\mathbb{F}$ satisfying the usual conditions. We define the Hilbert space of $\mathbb{F}$-progressively measurable $L^2$-integrable processes as

$$L^2_{\mathbb{F}}(\mathbb{R}^d) = \left\{ X : \Omega \to \mathbb{R}^d \ \mathbb{F}\text{-progressively measurable} \mid \mathbb{E}\left[\int_0^T |X_s|^2 ds\right] < \infty \right\}.$$

Furthermore let $\mathbb{S}^2_{\mathbb{F}}(\mathbb{R}^d)$ be the set of $\mathbb{F}$-progressively measurable continuous processes $X$ such that $\mathbb{E}\left[\|X\|^2_\infty\right] < \infty$.

We will consider the forward McKean-Vlasov SDE given by

$$dY_t = g(t,\omega,Y_t,\mathcal{L}(Y_t))dt + f(t,\omega,Y_t,\mathcal{L}(Y_t))dW_t\ . \tag{3.2}$$

Here we allow even for a dependence on $\omega$ in $g, f$ since this simplifies the argument.

**Condition 1** (McKean-Vlasov SDE). *Under the following assumptions we will see that this system yields a unique solution.*

1. *For any fixed $(y,\mu) \in \mathbb{R}^d \times \mathcal{P}_2(R^d)$ the processes $g(\cdot,\cdot,y,\mu)$ and $f(\cdot,\cdot,y,\mu)$ lie in $L^2_{\mathbb{F}}(\mathbb{R}^d)$ and $L^2_{\mathbb{F}}(\mathbb{R}^{d\times e})$.*
2. *There exists a Lipschitz constant $L$ such that for any $t \in [0,T]$, $\omega \in \Omega$, $y, y' \in \mathbb{R}^d$ and $\mu,\mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have*

$$\begin{aligned}&|g(t,\omega,y,\mu) - g(t,\omega,y',\mu')| + |f(t,\omega,y,\mu) - f(t,\omega,y',\mu')|\\ \le& L\left(|y-y'| + W^{(2)}(\mu,\mu')\right).\end{aligned}$$

**Theorem 3.1.** *Let $g, f$ satisfy Condition 1 and $Y_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};\mathbb{R}^d)$ be a given initial value. Then (3.2) has a unique solution $(Y_t)_{t\in[0,T]} \in \mathbb{S}^d_{\mathbb{F}}(\mathbb{R}^d)$.*

*Proof.* Fix some $\mu = (\mu_t)_{t\in[0,T]} \in \mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}^d))$. By classical SDE solution theory we then have a unique strong solution $Y^\mu$ for

$$dY^\mu_t = g(t,\omega,Y^\mu_t,\mu_t)dt + f(t,\omega,Y^\mu_t,\mu_t)dW_t\ .$$

It is furthermore known that $\mathcal{L}(Y^\mu) \in \mathcal{P}_2(\mathcal{C}_T)$. We introduce the map $\Phi$ associating a law $\mu$ to the corresponding solution law, i.e.

$$\Phi : \mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}^d)) \ni \mu \mapsto (\mathcal{L}(Y^\mu_t))_{t\in[0,T]} \in \mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}^d)).$$

It is easy to see that a process satisfying $\mathbb{E}\left[\|X\|^2_\infty\right] < \infty$ solves Equation 3.2 if and only if its law is a fixed point of $\Phi$. The latter is what we want to show.

Let now $\mu^1,\mu^2 \in \mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}^d))$ be fixed. We can then apply Doob's maximal inequality and Condition 1 to $Y^{\mu^1}, Y^{\mu^2}$ to obtain

$$\begin{aligned}&\mathbb{E}\left[\sup_{r\in[0,t]}\left|Y^{\mu^1} - Y^{\mu^2}\right|^2\right]\\ \le&\mathbb{E}\left[\sup_{s\in[0,t]}\left|\int_0^s \left(g(r,Y^{\mu^1}_r,\mu^1_r) - g(r,Y^{\mu^2}_r,\mu^2_r)\right)dr\right|\right]\\ &+\mathbb{E}\left[\sup_{s\in[0,t]}\left|\int_0^s \left(f(r,Y^{\mu^1}_r,\mu^1_r) - f(r,Y^{\mu^2}_r,\mu^2_r)\right)dW_r\right|\right]\\ \le& C\left(\int_0^t \mathbb{E}\left[\sup_{r\in[0,s]}|Y^{\mu^1}_r - Y^{\mu^2}_r|^2\right]ds + \int_0^t W^{(2)}(\mu^1_s,\mu^2_s)^2 ds\right).\end{aligned}$$

Here $C$ depends only on $T, L$ and is non-decreasing in $T$. Hence by applying Grönwall's inequality we conclude

$$\sup_{r\in[0,t]} W^{(2)}(\Phi(\mu^1)_r, \Phi(\mu^2)_r)^2 \le \mathbb{E}\left[\sup_{r\in[0,t]} \left|Y^{\mu^1} - Y^{\mu^2}\right|^2\right] \le C\int_0^t W^{(2)}(\mu_s^1, \mu_s^2)^2 ds.$$

Therefore for some $k \in \mathbb{N}$ large enough we get

$$\begin{aligned}\sup_{r\in[0,T]} W^{(2)}(\Phi^k(\mu^1)_r, \Phi^k(\mu^2)_r)^2 &\le C^k \int_0^T \int_0^{t_k} \cdots \int_0^{t_2} W^{(2)}(\mu_{t_1}^1, \mu_{t_1}^2)^2 dt_1 \dots dt_k \\ &\le \frac{C^k T^k}{k!} \sup_{r\in[0,T]} W^{(2)}(\mu_r^1, \mu_r^2)^2 \le \frac{1}{2} W^{(2)}(\mu_r^1, \mu_r^2)^2.\end{aligned}$$

Hence for a $k$ large enough $\Phi^k$ is a contraction. This implies that $\Phi$ is a contraction on $\mathcal{C}([0,T], \mathcal{P}_2(\mathbb{R}^d))$ and therefore has a unique fixed point which is the law of our solution. □

We will use this basic contractive argument in Proposition 3.6.

## 3.3 Rough setting

Let us now consider the rough case.

For some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and given initial condition $\xi : \Omega \to \mathbb{R}^d$ and driver $\mathbf{X} : \Omega \to \mathscr{C}^{p-var}([0,T], \mathbb{R}^e)$ we consider equations of the form

$$\begin{aligned} dY_t &= g(t, Y_t, \mathcal{L}(Y_t))dt + f(Y_t)d\mathbf{X}_t \\ Y_0 &= \xi \end{aligned} \tag{3.3}$$

where $\mathcal{L}(Y_t)$ denotes the law of $Y_t$ on $\mathbb{R}^d$. As discussed above including a dependence on $\mathcal{L}(Y_t)$ in $f$ is beyond the scope of this work and introduces many complexities. We will assume the following condition for some $q \ge 1$.

**Condition 2.** ($\mathbf{L^{(q)}}$)

*There exists a constant $L$ such that for any $t \in [0,T]$, $y^1, y^2 \in \mathbb{R}^d$, $\mu^1, \mu^2 \in \mathcal{P}_q(\mathbb{R}^d)$ we have*

1. $|g(t, y^1, \mu^1) - g(t, y^2, \mu^2)| \le L\left(|y^1 - y^2| + W^{(q)}(\mu^1, \mu^2)\right)$,
2. $|g(t, y^1, \mu^1)| \le L$ *and*
3. $f \in \mathcal{C}_b^3$.

**Example 3.2.** Let $G : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ be a function such that for some constant $C$ we have

$$G(t, x, y) \le C, \quad |G(t, x, y) - G(t, x', y')| \le L(|x - x'| + |y - y'|).$$

Then $g(t, x, \mu) := \int_{\mathbb{R}^d} G(t, x, y) d\mu(y)$ satisfies Condition 2 ($\mathbf{L^{(q)}}$).

**Definition 3.3** (Solution)**.** For an initial condition $\xi : \Omega \to \mathbb{R}^d$ and driver $\mathbf{X} : \Omega \to \mathscr{C}^{p-var}([0,T], \mathbb{R}^e)$ we say that a random variable $Y : \Omega \to \mathcal{C}([0,T], \mathbb{R}^d)$ is a solution of (3.3) if $(Y, Y')(\omega) = (Y(\omega), f(Y(\omega)))$ is controlled by $X(\omega)$ for almost all $\omega \in \Omega$, for almost all $t \in [0,T]$ we have $\mathcal{L}(Y_t) \in \mathcal{P}_q(\mathbb{R}^d)$ and the following integral equation is satisfied $\mathbb{P}$-almost surely for all $t \in [0,T]$.

$$Y_t(\omega) = \xi(\omega) + \int_0^t g(r, Y_r(\omega), \mathcal{L}(Y_r))dr + \int_0^t f(Y_r(\omega))d\mathbf{X}_r(\omega).$$

For $g, f$ satisfying Condition 2 ($\mathbf{L^{(q)}}$) we will call a solution $Y$ a $q$-solution.

Note that clearly by replacing $\mathcal{L}(Y_t)$ with a fixed law $\mu \in \mathcal{P}_q(\mathcal{C}_T)$ in (3.3) the equation turns into an RDE with drift as discussed in chapter 2 and thus have a unique pathwise solution given by $(Y, Y') = (Y, f(Y))$. We will denote the first component of that solution by $Y = \Theta_{g,f}(\mu, \xi, \mathbf{X})$.

We introduce the map

$$\begin{aligned} \Psi^{(q)} : \mathcal{P}(\mathbb{R}^d \times \mathscr{C}^{p-var}) \times \mathcal{P}_q(C_T) &\to \mathcal{P}_q(C_T) \\ \mathcal{L}(\xi, \mathbf{X}) \times \mu &\mapsto \mathcal{L}(\Theta_{g,f}(\mu, \xi, \mathbf{X})) = [\Theta_{g,f}(\mu, \cdot, \cdot)]_{\#}\mathcal{L}(\xi, \mathbf{X}). \end{aligned}$$

For $m \in \mathcal{P}_q(\mathbb{R}^d \times \mathscr{C}^{p-var})$ we will write $\Psi_m^{(q)}$ for $\Psi^{(q)}(m, \cdot)$ and omit the upper index $q$ when $q = 1$.

**Lemma 3.4.** *$Y$ is a solution to (2.1) if and only if $\mathcal{L}(Y)$ is a fixed point of $\Psi^{(q)}_{\mathcal{L}(\xi,\mathbf{X})}$.*

*Proof.* If $Y$ solves (2.1) then for $\mu = \mathcal{L}(Y)$ we have $Y = \Theta_{g,f}(\mu, \xi, \mathbf{X})$ and thus $\mathcal{L}(Y)$ is a fixed point of $\Psi_{\mathcal{L}(\xi,\mathbf{X})}$.

On the other hand let $\mu$ be a fixed point of $\Psi_{\mathcal{L}(\xi,\mathbf{X})}$. Then as seen before $Y = \Theta_{g,f}(\mu, \xi, \mathbf{X})$ has finite $q$-th moment and solves (2.1). □

For a measure $\mu$ on $\mathcal{C}([0,T], \mathbb{R}^d)$ we will at times consider its restriction $\mu\big|_{[0,t]} \in \mathcal{P}(\mathcal{C}([0,t], \mathbb{R}^d))$. We will then write

$$W_t^{(q)}(\mu^1, \mu^2) := W^{(q)}_{\mathcal{C}([0,t],\mathbb{R}^d)}(\mu^1\big|_{[0,t]}, \mu^2\big|_{[0,t]}).$$

When there is no confusion we will just write $W_t := W_t^{(1)}$.

Note that clearly by Lemma A.2 we get for any $\mu^1, \mu^2 \in \mathcal{P}_q(\mathcal{C}([0,T], \mathbb{R}^d))$ that $W_t^{(q)}(\mu^1, \mu^2)$ is non-decreasing in $t$ and that for any $t \in [0,T]$ we have $W^{(q)}_{\mathbb{R}^d}(\mu^1_t, \mu^2_t) \leq W_t^{(q)}(\mu^1, \mu^2)$. Furthermore it can be seen that $W_t^{(q)}(\mu^1, \mu^2)$ is continuous in $t$.

## 3.4 Existence of solution

In Proposition 2.5 and Proposition 2.6 we saw contractive bounds with linear factors in the input terms. We will use these to obtain a contractive property for the map $\Psi_m$ and thus the existence of a solution. Since these estimates feature the exponential term $e^{CN_1(w_X,[0,T])}$ we will require this to be integrable for all constants $C$. Hence for the law of the input noise we will need to assume that the accumulated local variation is exponentially integrable in the following sense.

**Condition 3.** *Let $\nu \in \mathcal{P}(\mathscr{C}_g^{p-var}([0,T], \mathbb{R}^e))$. We assume that for the measure defined by $[N_1(w_\cdot, [0,T])]_*(\nu)$ all exponential moments exist. That is for any $\theta \in \mathbb{R}$ we have*

$$\int_{\mathscr{C}^{p-var}} e^{\theta N_1(w_{\mathbf{X}},[0,T])} d\nu(\mathbf{X}) < \infty.$$

We denote by $\mathcal{E}$ the set of all measures $\nu$ satistying Condition 3 and by $\mathcal{E}_0^q$ the set of measures in $\mathcal{P}(\mathbb{R}^d \times \mathscr{C}^{p-var}([0,T],\mathbb{R}^e))$ for which the first marginal is in $\mathcal{P}_q(\mathbb{R}^d)$ and the second marginal in $\mathcal{E}$.

We note that if for a random variable $\mathbf{X} : \Omega \to \mathscr{C}_g^{p-var}$ we have $\mathcal{L}(\mathbf{X}) \in \mathcal{E}$ then by Lemma 1.8 every moment exists for $|||\mathbf{X}|||_{p-var}$. Indeed

$$\mathbb{E}\left[|||\mathbf{X}|||_{p-var}^q\right] \leq C_p^{\frac{q}{p}} \mathbb{E}\left[e^{\frac{q}{p}C_p N_1(w_X,[0,T])}\right] < \infty.$$

*Remark* 3.5. An example of processes with laws in $\mathcal{E}$ is worked out in [6]. Cass, Litterer and Lyons consider continuous centered Gaussian processes $X$ with i.i.d. components $(X^1,\dots,X^e)$. The idea is to examine the rectangular increments of the covariance function of $X$ given by

$$R_{X^i}\begin{pmatrix} s & t \\ s' & t' \end{pmatrix} := \mathbb{E}\left[X^i_{s,t}X^i_{s',t'}\right].$$

Let there exist some $\varrho \in [1,\frac{3}{2})$ and $M < \infty$ such that for all $0 \leq s \leq t \leq T$ and $i \in \{1,\dots e\}$ we have

$$\|R_{X^i}\|_{\varrho-var;[s,t]^2} \leq M|t-s|^{\frac{1}{\varrho}}.$$

Then one can show [15, Theorem 10.4] that $X$ then has a natural geometric rough path lift $\mathbf{X} = (X,\mathbb{X})$ given by the $L^2$-limit

$$\mathbb{X}^{i,j}_{s,t} := \lim_{|\pi|\to 0}\int_\pi (X^i_r - X^i_s)dX^j_r .$$

Furthermore the Gaussian nature of $X$ is leveraged to show that the associated Cameron-Martin space $\mathcal{H}$ can be embedded via

$$\mathcal{H} \hookrightarrow \mathcal{C}^{q-var}([0,T],\mathbb{R}^e).$$

Arguments using a generalized Fernique theorem then provide for any such constructed random rough path $\mathbf{X}$ a concentration inequality showing that for some $C > 0$ we have

$$\mathbb{E}\left[e^{CN_1(w_X,[0,T])^{\frac{2}{\varrho}}}\right] < \infty. \tag{3.4}$$

In particular this ensures $\mathbf{X} \in \mathcal{E}$. This class of processes contains fractional Brownian motion with Hurst parameter $H > \frac{1}{4}$ since there $\varrho$ can be chosen such that $\frac{2}{\varrho} > 1$.

We are now interested in existence of solutions for the McKean-Vlasov equation.

**Proposition 3.6.** *Let $\nu \in \mathcal{E}$, $p \in [2,3)$ and $q \geq 1$. Let $(\xi,\mathbf{X}) : \Omega \to \mathbb{R}^d \times \mathscr{C}_g^{p-var}([0,T],\mathbb{R}^e)$ be a random variable on some probability space $(\Omega,\mathcal{F},\mathbb{P})$ with projected laws $\mathcal{L}(\xi) = u_0$, $\mathcal{L}(\mathbf{X}) = \nu$ and joint law $\mathcal{L}(\xi,\mathbf{X}) = m \in \mathcal{E}_0^q$. Let $f,g$ satisfy $(\mathbf{L}^{(\mathbf{q})})$. Then there exists a unique pathwise solution $Y$ to equation* (3.3).

*Furthermore $\mathcal{L}(Y)$ depends only on $\mathcal{L}(\xi,\mathbf{X})$ and not the random variables themselves.*

*Proof.* Let $\mu^1,\mu^2 \in \mathcal{P}_q(C_T)$ be two measures. We define

$$Y^i = \Theta_{g,f}(\mu_i,\xi,\mathbf{X}) \quad \text{for } i = 1,2.$$

Note that due to Corollary 2.4 for any such $Y^i$ we have

$$\mathbb{E}\left[\left\|Y^i\right\|_\infty^q\right] \le C\left(\mathbb{E}\left[\xi^q\right] + \mathbb{E}\left[\|\mathbf{X}\|_{p-var,[0,T]}^{pq}\right] + T^q\right) < \infty.$$

Namely any such $Y^i$ is integrable, i.e $\mathcal{L}(Y^i) \in \mathcal{P}_q(\mathcal{C}_T)$ and thus $\mathcal{L}((Y^i)_t) \in \mathcal{P}_q(\mathbb{R}^d)$.

Let's now consider the difference between the $Y_i$'s. In the context of Proposition 2.5 we have $k_r = W_{\mathbb{R}^d}^{(q)}(\mu_r^1, \mu_r^2)$ resulting in

$$\left\|Y^1 - Y^2\right\|_\infty \le Ce^{CN_1((w_X + w_{id},[0,T]))} \int_0^T W^{(q)}(\mu_r^1, \mu_r^2)dr.$$

For the purposes of simpler notation we will simply write $\Psi_m$ for $\Psi_m^{(q)}$. By integrating both sides taken to the power of $q$ we get

$$\begin{aligned} W_T^{(q)}(\Psi_m(\mu^1), \Psi_m(\mu^2)))^q &\le \mathbb{E}\left[\left\|Y^1 - Y^2\right\|_\infty^q\right] \\ &\le C\mathbb{E}\left[e^{qCN_1((w_X + w_{id},[0,T]))}\right]\left(\int_0^T W_{\mathbb{R}^d}^{(q)}(\mu_r^1, \mu_r^2)dr\right)^q \\ &\le \tilde{C}\int_0^T W_r^{(q)}(\mu^1, \mu^2)^q dr. \end{aligned}$$

Recall that $\tilde{C} < \infty$ is guaranteed since $\nu \in \mathcal{E}$.

Applying the map $k$ times we use monotonicity of $t \mapsto W_t(\mu^1, \mu^2)$ to get

$$\begin{aligned} W_T^{(q)}(\Psi_m^k(\mu^1), \Psi_m^k(\mu^2)))^q &\le \tilde{C}^k \int_0^T \int_0^{t_k} \dots \int_0^{t_2} W_{t_1}^{(q)}(\mu^1, \mu^2)^q dt_1 \dots dt_k \\ &\le \frac{\tilde{C}^k}{k!} W_T^{(q)}(\mu^1, \mu^2)^q = a_k W_T^{(q)}(\mu^1, \mu^2)^q. \end{aligned}$$

Due to the asymptotics of the factorial compared to polynomial growth we can pick a $k$ large enough such that $a_k \le \frac{1}{2^q}$. Since the metric space $(C_T, \|\cdot\|_\infty)$ is complete and separable, we know that $(\mathcal{P}_q(C_T), W^{(q)})$ is also complete and separable (See e.g. [3]). Thus we can apply Banach's fixed point theorem which gives us a unique fixed point for $\Psi_\nu^k$. Let's denote that fixed point by $\mu$. We then have

$$W_T^{(q)}(\Psi_m(\mu), \mu) = W_T^{(q)}(\Psi_m^k(\Psi_\nu(\mu)), \Psi_\nu^k(\mu)) \le \frac{1}{2} W_T^{(q)}(\Psi_\nu(\mu), \mu).$$

Therefore $W_T^{(q)}(\Psi_m(\mu), \mu) = 0$ and hence $\mu$ is a fixed point of $\Psi_m$. It is trivially unique due to the uniqueness of fixed point in $\Psi_m^k$.

By Lemma 3.4 this corresponds to a unique solution $Y = \Theta_{g,f}(\mu, \xi, \mathbf{X})$. □

We note that any such measure $\mu$ that is a fixed point of $\Psi_m$ has mass 0 outside paths of finite $p$-variation.

*Remark* 3.7. We used Condition 3 to ensure that the estimates in Proposition 2.5 yield that the difference between two solutions $Y_1, Y_2$ is integrable. We remind the reader that originally (see e.g. [16, Theorem 12.10]) contractive estimates of this type featured the term

$$e^{C\|\mathbf{X}\|_{p-var}^p}(K)$$

where $K$ is Lipschitz in the difference of input noise, starting position and vector field changes. This is of limited use for our purposes since $e^{C\|\mathbf{X}\|_{p-var}^p}$ is not integrable for most processes that are of interest. In fact even for a lifted Brownian motion $\mathbf{B}$ we have $\mathbb{E}\left[e^{|||\mathbf{B}|||^p_{p-var}}\right] = \infty$ for any $p > 2$. The trouble here comes from the pathwise approach being forced to make worst-case estimates for every $\omega$ and the global in time estimates not being very strong.

The intuition for improving this via the greedy partition as above comes from observing that for $f \in \mathcal{C}^2_b$ we have in analog to Corollary 2.3 that

$$\left|\int_0^T f(X)d\mathbf{X}\right| \leq C\left||||\mathbf{X}|||_{p-var,[0,T]} \vee |||\mathbf{X}|||^p_{p-var,[0,T]}\right|$$

and thus for the greedy partition constructed for $\mathbf{X}$ given by $\pi = (t_i)$ we have the following

$$\begin{aligned}\left|\int_0^T f(X)d\mathbf{X}\right| &\leq \sum_{[\tau_i,\tau_{i+1}]\in\pi}\left|\int_{\tau_i}^{\tau_{i+1}} f(X)d\mathbf{X}\right| \\ &\leq (N_1(w_X,[0,T])+1)\sup_i\left(|||\mathbf{X}|||_{p-var,[\tau_i,\tau_{i+1}]} \vee |||\mathbf{X}|||^p_{p-var,[\tau_i,\tau_{i+1}]}\right) \\ &\leq (N_1(w_X,[0,T])+1).\end{aligned}$$

This is indeed the motivation for the introduction of $N_1(w_X,[0,T])$.

The superior integrability estimates become relevant due to the authors of [6] providing a large class of processes for which $e^{CN_1(w_X,[0,T])}$ is in fact integrable. These are Gaussian processes that have natural rough lifts. This class includes fractional brownian motion with Hurst parameter $H > \frac{1}{4}$.

## 3.5 Continuity In Input Driver

It is a natural question to ask whether the solution map to the McKean-Vlasov equation is continuous with respect to the Wasserstein distance. We will see this to be the case for a set of measures for which the exponential moments are uniformly bounded.

To this end for any monotone increasing function $K : (0,\infty) \to (0,\infty)$ we define a subset of $\mathcal{P}(\mathscr{C}^{p-var}_g(\mathbb{R}^d))$ by

$$\mathcal{E}(K) = \{\nu \in \mathcal{P}(\mathscr{C}^{p-var}_g(\mathbb{R}^e)) \mid \forall\theta \in \mathbb{R} : \int_{\mathscr{C}^{p-var}_g} e^{\theta N_1(w_\mathbf{X}),[0,T]}d\nu(\mathbf{X}) \leq K(\theta)\}.$$

Furthermore for $q \geq 1$ let $\mathcal{E}^q_0(K)$ be the set of measures $m$ in $\mathcal{P}(\mathbb{R}^d \times \mathscr{C}^{p-var}_g)$ for which the first marginal is in $\mathcal{P}_q(\mathbb{R}^d)$ and the second marginal is in $\mathcal{E}(K)$.

Our input measure $\nu$ being in $\mathcal{E}(K)$ ensures integrability of the exponential term that arises in the stability estimates from section 2.3. Note that $\mathcal{L}(\mathbf{X}) \in \mathcal{E}(K)$ gives us bounds on the moments of $|||\mathbf{X}|||_{p-var}$. Indeed by Lemma 1.8 we obtain

$$\mathbb{E}\left[|||\mathbf{X}|||^q_{p-var}\right] \leq 2^q e^{\frac{pq}{e}}\mathbb{E}\left[e^{\frac{qp}{e}N_1(w_X,[0,T])}\right] \leq 2^q e^{\frac{pq}{e}} K\left(\frac{qp}{e}\right).$$

*Remark* 3.8. It can be seen that $\mathcal{E}(K)$ is a closed set in $\mathcal{P}(\mathscr{C}^{p-var}(\mathbb{R}^e))$ with respect to weak convergence of measures.

Indeed let $(\nu^n) \subset \mathcal{E}(K)$ weakly converge to some measure $\nu$. We will write $f_\theta(x) = e^{\theta N_1(w_x,[0,T])}$. Then for all $M, \theta \in (0,\infty)$

$$\int f_\theta \wedge M d\nu^n \xrightarrow{n\to\infty} \int f_\theta \wedge M d\nu$$

by weak convergence and thus $\int f \wedge M d\nu \leq K(\theta)$. Further by monotone convergence we get

$$\int f_\theta \wedge M d\nu \xrightarrow{M\to\infty} \int f_\theta d\nu.$$

This gives us $\int f_\theta d\nu \leq K(\theta)$ for all $\theta$.

We have shown that given a measure $m \in \mathcal{E}_0^q(K)$ and functions $f, g$ satisfying Condition 2 ($\mathbf{L^{(q)}}$) the function $\Psi_m^{(q)}(\cdot)$ has a unique fixed point. We define the fixed point map

$$\begin{aligned} \Xi^{(q)} : \mathcal{E}_0^q(K) &\to \mathcal{P}_q(C_T) \\ m &\mapsto \text{fixed point of } \Psi_m^{(q)}(\cdot) \end{aligned}$$

and now want to show its continuity.

**Proposition 3.9.** *Let $K$ be a monotone increasing function. Assume that the laws $(\nu^n)_{n\in\mathbb{N}} \subset \mathcal{E}^q(K)$ converge weakly to $\nu^\infty$ and $(u^n)_{n\in\mathbb{N}} \subset \mathcal{P}_q(\mathbb{R}^d)$ converge in $W^{(q)}$ to $u^\infty$. Let $m^n = u^n \times \nu^n$ and $\mu^n := \Xi(m^n)$ for all $n \in \mathbb{N} \cup \{\infty\}$. Then we have*

$$W^{(q)}(\mu^n, \mu^\infty) \xrightarrow{n\to\infty} 0.$$

*Proof.* First note that due to the closedness of $\mathcal{E}_0^q(K)$ we know that $m^\infty$ also lies in $\mathcal{E}_0^q(K)$. Furthermore $\mathscr{C}_g^{p-var}$ and $\mathbb{R}^d$ are separable. Therefore $m^\infty$ has separable support and we can use Skorokhod's representation theorem. This gives us existence of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ hosting random variables $\xi^n, \mathbf{X}^n$ with laws $\mathcal{L}(\xi^n) = u^n, \mathcal{L}(\mathbf{X}^n) = \nu^n$ including $n = \infty$. Furthermore we have independence between $\xi^n$ and $\mathbf{X}^n$ and almost sure convergence

$$\xi^n \xrightarrow{\mathbb{P}-a.s.} \xi^\infty, \quad \mathbf{X}^n \xrightarrow{\mathbb{P}-a.s.} \mathbf{X}^\infty.$$

For $\mu^n = \Xi(m^n)$ we define the random variables $Y^n := \Theta_{g,f}(\mu^n, \xi^n, \mathbf{X}^n)$, again including $n = \infty$.

Since $\mu^n$ is given by the fixed point map, we have $\mathcal{L}(Y^n) = \Psi_{m^n}(\mu^n) = \mu^n$ and $Y^n$ is a solution to the McKean-Vlasov equation with inputs $\xi^n, \mathbf{X}^n$.

Our previous contraction estimates Proposition 2.6 give us

$$\begin{aligned} \|Y^n - Y^\infty\|_\infty \leq &\left( |\xi^n - \xi^\infty| + \|\mathbf{X}^n - \mathbf{X}^\infty\|_{p-var,[0,T]} + \int_0^T W(\mu_r^n, \mu_r^\infty) dr \right) \\ &\times C_1 e^{C_1 N_1(w_{X^\infty} + w_{X^n} + w_{id},[0,T])}. \end{aligned} \tag{3.5}$$

Let us note that the assumption $\nu^n \in \mathcal{E}(K)$ and an application of Hölder's inequality imply

$$\begin{aligned} \mathbb{E}\left[ e^{CN_1(w_{X^\infty} + w_{X^n} + w_{id},[0,T])} \right] &\leq \mathbb{E}\left[ Ce^{CT} e^{CN_1(w_{\mathbf{X}^\infty})} e^{CN_1(w_{\mathbf{X}^n})} \right] \\ &\leq Ce^{CT} K(2C). \end{aligned}$$

Using this and integrating both sides of (3.5) we obtain

$$W_T^{(q)}(\mu^n,\mu^\infty)^q \leq C_2\int_0^T W_r^{(q)}(\mu_n,\mu^\infty)^q dr + C_1 a_n + C_1 b_n \tag{3.6}$$

with

$$\begin{aligned} a_n &= \mathbb{E}\left[A_n\right] = \mathbb{E}\left[e^{qC_1N_1(w_{X^\infty}+w_{X^n}+w_{id},[0,T])}\|\mathbf{X}^n - \mathbf{X}^\infty\|_{p-var}^q\right],\\ b_n &= \mathbb{E}\left[B_n\right] = \mathbb{E}\left[e^{qC_1N_1(w_{X^\infty}+w_{X^n}+w_{id},[0,T])}|\xi^n - \xi^\infty|^q\right].\end{aligned}$$

Due to the independence of $\xi^n$ and $\mathbf{X}^n$ for all $n$ we have

$$b_n \leq C_3\mathbb{E}\left[|\xi^n - \xi^\infty|^q\right] \xrightarrow{n\to\infty} 0.$$

Here we used the fact that $W^{(q)}(u^n,u^\infty)\xrightarrow{n\to\infty} 0$ implies uniform integrability of the $p$-th moment by Proposition A.3. This combined with pointwise convergence gives us convergence in $L_q$.

On the other hand we can use Lemma 1.8 to get

$$\begin{aligned}\|\mathbf{X}^n-\mathbf{X}^\infty\|_{p-var} &\leq 2(w_{X^n}+w_{X^\infty})(0,T)^{\frac{1}{p}}\\ &\leq C_4 e^{C_4N_1(w_{X^n}+w_{X^\infty},[0,T])}\\ &\leq C_4 e^{C_4N_1(w_{X^\infty},[0,T])}e^{C_4N_1(w_{X^n},[0,T])}.\end{aligned}$$

Using this we get the following estimate for $A_n$.

$$\begin{aligned} A_n &\leq C_5 e^{C_1N_1(w_{X^\infty}+w_{X^n}+w_{id},[0,T])}e^{C_4N_1(w_{X^\infty},[0,T])}e^{C_4N_1(w_{X^n},[0,T])}\\ &\leq C_6 e^{C_6N_1(w_{X^\infty},[0,T])}e^{C_6N_1(w_{X^n},[0,T])} =: D_n\end{aligned}$$

$D_n$ is integrable as seen by applying Hölder's inequality and our assumption $\nu^n,\nu^\infty\in\mathcal{E}(K)$. Furthermore $\mathbb{E}\left[D_n\right]\leq C_6K(2C_6)$. This implies that $(D_n)_{n\in\mathbb{N}}$ is uniformly integrable and therefore we get $a_n\xrightarrow{n\to\infty} 0$.

Going back to (3.6) we now apply the Grönwall inequality as before to get

$$W_T^{(q)}(\mu^n,\mu^\infty)^q \leq C(a_n+b_n)e^{CT}\xrightarrow{n\to\infty} 0$$

which implies weak convergence as seen in Proposition A.3. □

By using a higher power Wasserstein distance we even get Lipschitz continuity of the solution map.

**Proposition 3.10.** *Let $q > q' \geq 1$, $K$ be a monotone increasing function and $f,g$ satisfy Condition 2 ($\mathbf{L}^{(\mathbf{q}')}$). Then the solution map*

$$\begin{aligned}\Xi^{(q,q')}:\mathcal{E}_0^q(K)&\to\mathcal{P}_{q'}(\mathcal{C}_T)\\ m\quad&\mapsto \text{fixed point of } \Psi_m^{(q)}(\cdot)\end{aligned}$$

*is Lipschitz continuous. That is there exists a universal constant $L$ depending on $K,p,q,g,f$ such that for any two measures $m^1,m^2\in\mathcal{E}_0^q(K)$ and $\mu^i := \Xi^{(q,q')}(m^i)$, $i=1,2$ we have*

$$W_{\mathcal{C}_T}^{(q')}(\mu^1,\mu^2)\leq LW^{(q)}_{\mathbb{R}^d\times\mathscr{C}^{p-var}}(m^1,m^2).$$

*Proof.* Let $(\xi^1, \mathbf{X}^1), (\xi^2, \mathbf{X}^2)$ be random variables with laws $m^1, m^2$. Again we can define $Y^i = \Theta_{g,f}(\mu^i, \xi^i, \mathbf{X}^i)$, $i = 1, 2$. Let $\pi_o \in \Pi(m^1, m^2)$ be an optimal coupling, i.e.

$$W^{(q)}(m^1, m^2)^q = \mathbb{E}_{\pi_o}\left[|\xi^1 - \xi^2|^q + \left|\left|\left|\mathbf{X}^1 - \mathbf{X}^2\right|\right|\right|_{p-var}^q\right].$$

By definition of $\mu^i$ we then have $\mathcal{L}(Y^i) = \mu^i$. Using again Proposition 2.6 we get

$$\begin{aligned}\left\|Y^1 - Y^2\right\|_\infty \leq &\left(|\xi^1 - \xi^2| + \left|\left|\left|\mathbf{X}^1 - \mathbf{X}^2\right|\right|\right|_{p-var,[0,T]} + \int_0^T W^{(q')}(\mu_r^1, \mu_r^2)dr\right)\\ &\times Ce^{CN_1(w_{X^1}+w_{X^2}+w_{id},[0,T])}.\end{aligned}$$

Taking expectations we again get

$$W_T^{(q')}(\mu^n, \mu^\infty)^{q'} \leq C_2 \int_0^T W_r^{(q')}(\mu_n, \mu^\infty)^{q'} dr + Ca_n \tag{3.7}$$

with

$$a_n = \mathbb{E}_{\pi_o}\left[e^{CN_1(w_{X^1}+w_{X^2}+w_{id},[0,T])}\left(|\xi^1 - \xi^2| + \left|\left|\left|\mathbf{X}^1 - \mathbf{X}^2\right|\right|\right|_{p-var}\right)^{q'}\right].$$

Let now $\tilde{q}$ and $\frac{q}{q'}$ be a Hölder pair, i.e. $\frac{1}{\tilde{q}} + \frac{q'}{q} = 1$. Then we can use Hölder's inequality to get

$$\begin{aligned}a_n &\leq \mathbb{E}_{\pi_o}\left[e^{\tilde{q}CN_1(w_{X^1}+w_{X^2}+w_{id},[0,T])}\right]^{\frac{1}{\tilde{q}}} \mathbb{E}_{\pi_o}\left[\left(|\xi^1 - \xi^2| + \left|\left|\left|\mathbf{X}^1 - \mathbf{X}^2\right|\right|\right|_{p-var}\right)^q\right]^{\frac{q'}{q}}\\ &\leq CK(C)\mathbb{E}_{\pi_o}\left[|\xi^1 - \xi^2|^q + \left|\left|\left|\mathbf{X}^1 - \mathbf{X}^2\right|\right|\right|_{p-var}^q\right]^{\frac{q'}{q}}.\end{aligned}$$

Using Grönwall's inequality in (3.7) gives us

$$\begin{aligned}W_T^{(q')}(\mu^n, \mu^\infty) &\leq C\mathbb{E}_{\pi_o}\left[|\xi^1 - \xi^2|^q + \left|\left|\left|\mathbf{X}^1 - \mathbf{X}^2\right|\right|\right|_{p-var}^q\right]^{\frac{1}{q}}\\ &= CW^{(q)}_{\mathbb{R}^d \times \mathscr{C}^{p-var}}(m^1, m^2)\end{aligned}$$

which is the claimed result. □

The following proposition applies standard arguments to show that progressive measurability from the input driver transfers to the solution.

**Proposition 3.11.** *Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a right-continuous complete filtration on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\xi$ be $\mathcal{F}_0$-measurable and $\mathbf{X} : \Omega \times [0, T] \to G^{(2)}(\mathbb{R}^e)$ an $(\mathcal{F}_t)$-progressively measureable process with their respective laws satisfying the existence conditions in Proposition 3.6 for some $q \geq 1$. Then the q-solution $Y$ to (3.3) is also $(\mathcal{F}_t)$-progressively measurable.*

*Proof.* Let $\mathcal{L}(Y) = \mu$. For a fixed $t \in [0, T]$ it is clear that $Y\big|_{[0,t]}$ satisfies

$$Y\big|_{[0,t]} = \Theta_{g,f,[0,t]}(\mu\big|_{[0,t]}, \xi, \mathbf{X}\big|_{[0,t]}).$$

Since $\Theta_{g,f,[0,t]}(\mu\big|_{[0,t]}, \cdot, \cdot)$ is $\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathscr{C}^{p-var}([0, t], \mathbb{R}^e))$-measurable, we have that $Y\big|_{[0,t]}$ is $\mathcal{F}_t$-measurable because $\xi, \mathbf{X}\big|_{[0,t]}$ are $\mathcal{F}_t$-measurable. Thus $Y_t$ is $\mathcal{F}_t$-measurable.

Hence $Y$ is adapted and therefore progressively measurable due to the continuity of its paths. □

## Non-geometric drivers

So far we have considered input drivers only in the space of geometric rough paths. The space $\mathscr{C}_g^{p-var}([0,T],\mathbb{R}^e)$ is separable and closed under $\varrho_{p-var}$ thus resulting in $(\mathcal{P}_q(\mathscr{C}_g^{p-var}), W^{(q)})$ being separable and closed allowing us to use the fixed point arguments above. Furthermore geometric paths allow us to write $|||\mathbf{X}|||_{p-var}$ as $\sup_{\pi\in\Pi([0,T])}\sum_{(t_i)=\pi} d_{CC}(\mathbf{X}_{t_i},\mathbf{X}_{t_{i+1}})^p$ thus enabling easy moment estimates on $\mathcal{E}$ via Lemma 1.8. This eliminates the need to impose more moment constraints on the input driver. It is however possible to extend our above results to the case of general rough paths by reducing the problem to the case we have studied.

We call any $p$-rough path satisfying the first order chain rule condition

$$\mathrm{Sym}(\mathbb{X}_{s,t}) = \frac{1}{2} X_{s,t}\otimes X_{s,t} \tag{3.8}$$

a weakly geometric rough path and denote the space of all such paths by

$$\mathscr{C}_{\mathrm{wg}}^{p-var}([0,T],\mathbb{R}^e)\subset \mathscr{C}^{p-var}([0,T],\mathbb{R}^e).$$

A simple application of partial integration shows that any lifted smooth path satisfies (3.8) giving us the straight forward inclusion $\mathscr{C}_g^{p-var}([0,T],\mathbb{R}^e)\subset \mathscr{C}_{\mathrm{wg}}^{p-var}([0,T],\mathbb{R}^e)$.

As was shown in [19, Corollary 19], any path $X\in\mathcal{C}^{p-var}$ can be lifted to a weakly geometric path $\bar{\mathbf{X}}$ for $p\in(2,3)$. We denote this lift as $\bar{\mathbf{X}} = L(X)$. Note that this is not guaranteed to be unique. For our case it is enough to just pick some lift $L(X) = (X,\bar{\mathbb{X}})$.

Further it was shown in [14, Theorem 15] that any $\mathbf{X}\in\mathscr{C}_{\mathrm{wg}}^{p-var}([0,T],\mathbb{R}^e)$ can be approximated uniformly by smooth paths and their respective lifts $\mathbf{X}^n$. Via an interpolation argument one can show that this implies convergence

$$\varrho_{q-var}(\mathbf{X},\mathbf{X}^n)\xrightarrow{n\to\infty} 0$$

for any $q$ such that $2\le p<q<3$. (See e.g. [15, Exercise 2.9]). This implies $\mathbf{X}\in\mathscr{C}_g^{q-var}$.

From Chen's relation (1.3) it is apparent that the first level $X$ determines its lift $\mathbb{X}$ only up to the addition of increments of functions $F\in\mathcal{C}^{\frac{p}{2}-var}([0,T],\mathbb{R}^e\times\mathbb{R}^e)$ i.e. all second levels above a path $X$ have the form $\mathbb{X}_{s,t} = \bar{\mathbb{X}}_{s,t} + F_t - F_s$.

Thus we have

$$\begin{aligned}\mathscr{C}^{p-var}([0,T],\mathbb{R}^e) \cong& \mathscr{C}_{\mathrm{wg}}^{p-var}([0,T],\mathbb{R}^e)\oplus\mathcal{C}^{\frac{p}{2}-var}([0,T],\mathbb{R}^e\times\mathbb{R}^e)\\ \hookrightarrow& \mathscr{C}_{\mathrm{g}}^{q-var}([0,T],\mathbb{R}^e)\oplus\mathcal{C}^{\frac{p}{2}-var}([0,T],\mathbb{R}^e\times\mathbb{R}^e)\end{aligned} \tag{3.9}$$

Furthermore for any $(Y,Y')\in\mathscr{D}_X^{\frac{p}{2}}$ it is easy to check that $(Y,Y')\in\mathscr{D}_{\bar{X}}^{\frac{p}{2}}$. Therefore we get the equality

$$\int Y d\mathbf{X} = \lim_{|\pi|\to 0}\sum_{[u,v]\in\pi} Y_u X_{u,v} + Y'_u\mathbb{X}_{u,v} = \int Y d\bar{\mathbf{X}} + \int Y' dF \tag{3.10}$$

where $\bar{\mathbf{X}} = (X,\bar{\mathbb{X}})$ and the second term is a Young integral.

We now apply this to a rough McKean-Vlasov equation of the type

$$dY_t = g(t, Y_t, \mathcal{L}(Y_t))dt + f(Y_t)d\mathbf{X}_t$$

where $\mathbf{X}$ is a random variable with values in $\mathscr{C}^{p-var}$. By the above discussion for almost all $\omega$ there exists a function $F \in \mathcal{C}^{\frac{p}{2}-var}$ such that for the weakly geometric lift $\bar{\mathbf{X}}_{s,t} = \mathbf{X}_{s,t} + F_t - F_s$ we have

$$\begin{aligned} dY_t &= g(t, Y_t, \mathcal{L}(Y_t))dt + f(Y_t)'dF_t + f(Y_t)d\bar{\mathbf{X}}_t \\ &= g(t, Y_t, \mathcal{L}(Y_t))dt + Df(Y_t)f(Y_t)dF_t + f(Y_t)d\bar{\mathbf{X}}_t. \end{aligned}$$

By now interpreting $\bar{\mathbf{X}}$ as a random variable with values in $\mathscr{C}_g^{q-var}$ for some $q > p$ we can again apply the fixed point arguments above to get a solution.

In order to show existence and continuity here we need a more general version of Proposition 2.5 and Proposition 2.6. It is clear how to extend these however to include another drift term interpreted as a Young integral. We can follow the same steps as above using the Young integral inequality

$$\left| \int_s^t Z_r dF_r - Z_s F_{s,t} \right| \leq C \left( w_Z(s,t)^{\frac{1}{q}} w_F(s,t)^{\frac{2}{q}} \right).$$

This is a direct analog to Corollary 1.6 and provides combined with the arguments in Proposition 2.6 estimates of the type

$$\left\| Y^1 - Y^2 \right\|_{\infty,[s,t]} \leq C e^{CN_1(w_1,[s,t])} \left( \left| Y_s^1 - Y_s^2 \right| + w_{F^1-F^2}(s,t)^{\frac{2}{p}} + w_{\bar{\mathbf{X}}^1-\bar{\mathbf{X}}^2}(s,t)^{\frac{1}{p}} \right)$$

for the two processes $Y^i = \Theta_{g,f}(\mu^i, \xi^i, \mathbf{X}^i)$. Existence of a fixed point for the map $\Psi_m^{(q)}(\cdot)$ follows immediately. Furthermore the following continuity result can be shown.

**Proposition 3.12.** *Let $q > q' \geq 1$, $K$ be a monotone increasing function and $f, g$ satisfy Condition 2 ($\mathbf{L}^{(\mathbf{q}')}$). We denote by $\mathcal{E}_2^q$ the space of measures on $(\mathcal{C}^{\frac{p}{2}-var} \times \mathbb{R}^d \times \mathscr{C}^{p-var})$ for which the projection to the last two variables lies in $\mathcal{E}_0^q$ and for the second marginal $\mu_2$*

$$\int_{\mathcal{C}^{\frac{p}{2}-var}} e^{CN_1(w_x,[0,T])} d\mu_2(x) < K(C)$$

*holds for any $C$. Then for the solution map*

$$\begin{aligned} \Xi^{(q,q')} : \mathcal{E}_2^q(K) &\to \mathcal{P}_{q'}(\mathcal{C}_T) \\ m &\mapsto \textit{fixed point of } \Psi_m^{(q)}(\cdot) \end{aligned}$$

*there exists a constant $L$ such that for any $m^1, m^2 \in \mathcal{E}_2^q(K)$ and $\mu^i := \Xi^{(q,q')}(m^i)$ we have*

$$W_{\mathcal{C}_T}^{(q')}(\mu^1, \mu^2) \leq L W^{(q)}_{\mathcal{C}^{\frac{p}{2}-var} \times \mathbb{R}^d \times \mathscr{C}^{p-var}}(m^1, m^2).$$

*Proof.* The proof is an elementary adjustment of the one of Proposition 3.10. □

Our results could potentially also be extended to geometric processes with jumps using the integration theory developed in [13]. It is however not clear if there exists a rich class of processes with jumps that satisfy the appropriate integrability conditions. This remains to be seen in future work.

# Chapter 4

# Applications

The main applications for equations of McKean-Vlasov type come from approximating interacting particle systems. In our case we consider a system of $N$ particles with the following dynamics

$$dY_t^{i,N} = g(t, Y_t^{i,N}, \frac{1}{N}\sum_{j=1}^{N} \delta_{Y_t^{j,N}}) + f(Y_t^{i,N}) d\mathbf{X}_t^i , \quad \forall i \in \{1, \ldots, N\}. \tag{4.1}$$

We will see how to consider these systems as McKean-Vlasov equations themselves and show propagation of chaos results as well as estimates for a strong rate of convergence.

## 4.1 Particle Approximations

The goal of this section is to examine the particle systems associated with McKean-Vlasov equations. We will see how the existence and stability results from chapter 3 can be applied to interpret the particle systems as their corresponding McKean-Vlasov equation and to get propagation of chaos results.

Let now $(\xi, \mathbf{X})$, $(\xi^i, \mathbf{X}^i)_{i\in\mathbb{N}}$ be a family of random variables with values in $\left(\mathbb{R}^d \times \mathscr{C}_g^{p-var}\right)$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and consider the system of particles given by

$$\begin{cases} dY_t^{i,N} &= g(t, Y_t^{i,N}, \frac{1}{N}\sum_{j=1}^{N} \delta_{Y_t^{j,N}}) dt + f(Y_t^{i,N}) d\mathbf{X}_t^i \\ Y_0^{i,N} &= \xi^i. \end{cases} \tag{4.2}$$

This is the natural interacting particle approximation to the McKean-Vlasov equation

$$\begin{cases} dY_t &= g(t, Y_t, \mathcal{L}(Y_t)) dt + f(Y_t) d\mathbf{X}_t \\ Y_0 &= \xi. \end{cases} \tag{4.3}$$

We will now examine how to interpret the interacting particle system (4.2) as a McKean-Vlasov equation.

Let $Y^{(N)} = (Y^1, \ldots, Y^N)$ be an $N$-tuple with entries in some metric space $E$. We define the empirical measure associated with it as

$$L^{(N)}(Y^{(N)}) = \frac{1}{N}\sum_{i=1}^{N} \delta_{Y_i}.$$

For a given $N \in \mathbb{N}$ we introduce the probability space $(\Omega_N, \mathcal{F}_N, \mathbb{P}_N)$ given by $\Omega_N = \{1, \ldots, N\}$, $\mathcal{F}_N = 2^{\Omega_N}$ and $\mathbb{P}_N = \frac{1}{N}\sum_{i=1}^{N} \delta_i$. This is just the uniformly distributed space on $\{1, \ldots, N\}$. Any $N$-tuple $(Y^1, \ldots, Y^N)$ can now be interpreted as the random variable $Y_e^{(N)} : i \mapsto Y^i$. This identification gives us

$$\mathcal{L}(Y_e^{(N)}) = L^{(N)}(Y^{(N)}).$$

Now for our given input family $(\xi^i, \mathbf{X}^i)_{i\in\mathbb{N}}$ and an $N \in \mathbb{N}$ we define for all $\omega \in \Omega$ the random variables

$$\xi^{(N)}(\omega) = \left(\xi^1(\omega), \dots, \xi^N(\omega)\right)$$
$$\mathbf{X}^{(N)}(\omega) = \left(\mathbf{X}^1(\omega), \dots, \mathbf{X}^N(\omega)\right).$$

Then for a fixed $\omega \in \Omega$ the particle system (4.2) is a McKean-Vlasov equation for a random variable $Y^{(N)} : \Omega_N \to \mathcal{C}_T$ with inputs $(\xi^{(N)}(\omega), \mathbf{X}^{(N)}(\omega))$.

Let's first review a simple lemma about convergence of empirical measures.

**Lemma 4.1.** *Let $q \geq 1$ be given. Let $(X^i)_{i\in\mathbb{N}}$ be a family of i.i.d. random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a separable metric space $(E, d)$ with law $\nu$ such that $\int_E d(x_0, x)^q d\nu(x) < \infty$ for some and thus for all $x_0 \in E$. Let $\nu^N(\omega) = \frac{1}{N}\sum_{i=1}^N \delta_{X^i(\omega)}$ for $\omega \in \Omega$ be the empirical law of $X^{(N)} = (X^1, \dots, X^N)$. Then we have*

$$W_{q,E}(\nu^N, \nu) \xrightarrow{N\to\infty} 0 \quad \mathbb{P}\text{-}a.s.$$

*Proof.* We denote by $\mathcal{B}$ the Borel sigma algebra on $(E, d)$. By the strong law of large numbers we know that for any $B \in \mathcal{B}$ there exists a set $\Omega_0(B)$ with $\mathbb{P}(\Omega_0(B)) = 1$ such that

$$\nu^N(\omega)(B) = \frac{1}{N}\sum_{i=1}^N \mathbb{1}_{X^i(\omega)\in B} \xrightarrow{N\to\infty} \nu(B) \quad \forall\, \omega \in \Omega_0(B).$$

To show that the exception set can be chosen independently of $B$ let $D$ be a countable dense subset of $E$. Let $\mathcal{G} = \{B_{\frac{1}{n}}(x) \mid x \in D, n \in \mathbb{N}\}$, where $B_\epsilon(x)$ denotes the open $\epsilon$ ball around $x$. Every open set in $E$ can be represented as a union of elements in $\mathcal{G}$. Now let $\mathcal{G}_f$ be the set of finite intersections of elements in $\mathcal{G}$. Since $\mathcal{G}_f$ is countable, we have

$$\mathbb{P}(\bigcup_{A\in\mathcal{G}_f} \Omega_0^c(A)) \leq \sum_{A\in\mathcal{G}_f} \mathbb{P}(\Omega_0^c(A)) = 0.$$

This gives us

$$\nu^N(w)(A) \xrightarrow{N\to\infty} \nu(A) \quad \forall\, \omega \notin \bigcup_{A\in\mathcal{G}_f} \Omega_0^c(A)$$

Since $\mathcal{G}_f$ generates the borel sigma algebra $\mathcal{B}$ this gives us

$$\nu^N(\omega) \implies \nu \quad a.s.$$

Furthermore for any element $x_0 \in E$ we can apply the strong law of large numbers to the i.i.d. random variables $(d(x_0, X^i))_{i\in\mathbb{N}}$ to get for almost all $\omega$

$$\int_E d(x_0, x)^q d\nu^N(\omega)(x) = \frac{1}{N}\sum_{i=1}^N d(x_0, X^i(\omega))^q \xrightarrow{N\to\infty} \int_E d(x_0, X)^q d\nu < \infty.$$

Thus by Corollary A.4 we get convergence in $W^{(q)}$. □

We now use this to show that the solution to the particle approximation converges in the Wasserstein distance to the associated McKean-Vlasov solution.

**Theorem 4.2.** *Let $q \geq 1$, $K : (0, \infty) \to (0, \infty)$ be given and $g, f$ satisfy Condition 2 $(\mathbf{L^{(q)}})$. Let $(\xi, \mathbf{X}), (\xi^i, \mathbf{X}^i)_{i \in \mathbb{N}}$ be a set of i.i.d. random variables with values in $(\mathbb{R}^d \times \mathscr{C}_g^{p-var}(\mathbb{R}^e))$ and law $m \in \mathcal{E}_0^q(K)$ with marginals $(u_0, \nu)$. Then for the solution $Y^{(N)}$ to the particle system (4.2) we have*

$$W^{(q)}\left(L^N(Y^{(N)}), \mathcal{L}(Y)\right) \xrightarrow{N \to \infty} 0 \quad \mathbb{P}\text{-}a.s.$$

*Proof.* Let us name the laws of the solutions to our systems $\mu^N := L^N(Y^{(N)}) = \frac{1}{N}\sum_{i=1}^N \delta_{Y^{i,N}}$ and $\mu := \mathcal{L}(Y)$. We introduce the random variables given by $Y^i = \Theta_{g,f}(\mu, \xi^i, \mathbf{X}^i)$. Note that since $(\xi^n, \mathbf{X}^n)$ are i.i.d. copies of $(\xi, \mathbf{X})$ and the law of solutions only depends on the law of the input this gives us $\mathcal{L}(Y^i) = \mu$ for all $i$. Hence all the $Y^i$'s are independent copies of solutions to the McKean-Vlasov equation.

We can now apply our contraction estimate Proposition 2.5 with $k_r = W^{(q)}(\mu_r^N, \mu_r)$ to get the $\mathbb{P}$-almost sure estimate

$$\left\|Y^{i,N} - Y^i\right\|_\infty \leq C e^{CN_1(w_{X^i},[0,T])} \int_0^T W_r^{(q)}(\mu^N, \mu) dr.$$

By denoting the empirical law of $(Y^1, \dots, Y^N)$ as $\bar{\mu}^N := \frac{1}{N}\sum_{i=1}^N \delta_{Y^i}$ and noting the general inequality for discrete laws

$$W^{(q)}\left(\sum_{i=1}^N \delta_{a_i}, \sum_{i=1}^N \delta_{b_i}\right) \leq \left(\frac{1}{N}\sum_{i=1}^N \|a_i - b_i\|^q\right)^{\frac{1}{q}}$$

we obtain

$$W_T^{(q)}(\bar{\mu}^N, \mu^N)^q \leq C\left(\int_0^T W_r^{(q)}(\mu^N, \mu) dr\right)^q \frac{1}{N}\sum_{i=1}^N e^{qCN_1(w_{X^i},[0,T])}.$$

Using this we can split our computation in the following way.

$$\begin{aligned} W_T^{(q)}(\mu, \mu^N) &\leq W_T^{(q)}(\mu, \bar{\mu}^N) + W_T^{(q)}(\bar{\mu}^N, \mu^N) \\ &\leq W_T^{(q)}(\mu, \bar{\mu}^N) + C\int_0^T W_r^{(q)}(\mu, \mu^N) dr \left(\frac{1}{N}\sum_{i=1}^N e^{CN_1(w_{X^i},[0,T])}\right)^{\frac{1}{q}}. \end{aligned}$$

Thus by using the Grönwall lemma we get

$$W_T^{(q)}(\mu, \mu^N) \leq C W_T^{(q)}(\mu, \bar{\mu}^N) e^{TC\left(\frac{1}{N}\sum_{i=1}^N e^{CN_1(w_{X^i},[0,T])}\right)^{\frac{1}{q}}}.$$

By Lemma 4.1 we know that $\mathbb{P}$-almost surely $W_T(\mu, \bar{\mu}^N) \xrightarrow{n \to \infty} 0$. Furthermore we have by the law of large numbers that

$$\left(\frac{1}{N}\sum_{i=1}^N e^{CN_1(w_{X^i},[0,T])}\right)^{\frac{1}{q}} \xrightarrow{N \to \infty} \left(\mathbb{E}\left[e^{CN_1(w_X,[0,T])}\right]\right)^{\frac{1}{q}} < \infty \quad \mathbb{P}\text{-a.s.}$$

This implies

$$W_T^{(q)}(\mu, \mu^N) \xrightarrow{N \to \infty} 0 \quad \mathbb{P}\text{-a.s.}$$

□

An alternative way to prove the convergence of law of the particle approximations to the law of the solution of the McKean-Vlasov equation is to use the Lipschitz estimate in Proposition 3.10. We note that this is a slightly weaker result due to the loss of Wasserstein power but include the proof since it is immediate from the interpretation of the particle approximation as a McKean-Vlasov equation.

**Theorem 4.3.** *Let $q > q' \geq 1$ and $K : (0, \infty) \to (0, \infty)$, $q > 0$ be given. Let $(\xi, \mathbf{X}), (\xi^i, \mathbf{X}^i)_{i \in \mathbb{N}}$ be a set of i.i.d. random variables with values in $(\mathbb{R}^e \times \mathscr{C}_g^{p-var}(\mathbb{R}^d))$ and law $m$ with marginals $(u_0, \nu)$ such that $m \in \mathcal{E}_0^q(K)$. Then for the solution to the particle system $Y^{(N)}$ we have*

$$W^{(q')}\left(L^N(Y^{(N)}), \mathcal{L}(Y)\right) \xrightarrow{N\to\infty} 0 \quad \mathbb{P}\text{-a.s.}$$

*Proof.* We first define $\nu^N = \frac{1}{N}\sum_{i=1}^N \delta_{\mathbf{X}^i}$. By the strong law of large numbers there exists a $\Omega_0$ of zero mass such that for all $\omega \notin \Omega_0$ we have

$$\begin{aligned}\int_{\mathscr{C}_g^{p-var}} e^{\theta N_1(w_x,[0,T])} d\nu^N(\omega)(x) &= \frac{1}{N}\sum_{i=1}^N e^{\theta N_1(w_{X^i(\omega)},[0,T])} \\ &\xrightarrow{N\to\infty} \mathbb{E}\left[e^{\theta N_1(w_X,[0,T])}\right] \leq K(\theta) \quad \forall \theta \in (0,\infty).\end{aligned}$$

Thus we have that for any $\omega \in \Omega_0^c$ there exists an $N(\omega)$ such that for all $n > N(\omega)$ we have $\nu^n \in \mathcal{E}_g^q(2K)$.

We now apply Proposition 3.10 to the input random variables given by the interpretation of (4.2) as a McKean-Vlasov equation

$$\begin{aligned}&(\Omega_N, \mathcal{F}_N, \mathbb{P}_N), \quad (\xi^{(N)}, \mathbf{X}^{(N)})(\omega) = (\xi^{(N)}(\omega), \mathbf{X}^{(N)}(\omega)) \\ &(\Omega, \mathcal{F}, \mathbb{P}), \qquad (\xi, \mathbf{X})(\omega) = (\xi(\omega), \mathbf{X}(\omega)).\end{aligned}$$

For all $n > N(\omega)$ we then have $\mathcal{L}((\xi^{(n)}, \mathbf{X}^{(n)})) \in \mathcal{E}_{0,g}^q(2K)$. Let the obtained solutions to the corresponding McKean-Vlasov equations be $Y$ and $Y^{(n)}$.

Then by Proposition 3.10 we get

$$W^{(q')}(\mathcal{L}(Y), L^{(n)}(Y^{(n)})) \leq CW^{(q)}(\mathcal{L}(\xi, \mathbf{X}), L^{(n)}(\xi^{(n)}, \mathbf{X}^{(n)})) \quad \forall n > N(\omega)$$

and therefore by Lemma 4.1 the claim follows. □

## 4.2 Rate Of Convergence

We now consider the rate of convergence of the particle approximation to the solution to the McKean-Vlasov equation as the number of particles increases.

We first state [5, Theorem 5.8]. This gives us a rate of convergence in the 2-Wasserstein metric for empirical measures on $\mathbb{R}^d$. We will then apply this to our setting to recover the same rate of convergence for the particle approximation.

**Theorem 4.4.** *Let $\mu \in \mathcal{P}_q(\mathbb{R}^d)$ for some $q > 4$. Let $\bar{\mu}^N = \sum_{i=1}^N \delta_{X_i}$ be the $N$-th empirical measure given by i.i.d random variables $(X_i)$. Then there exists a constant*

$C = C(d, q, \int_{\mathbb{R}^d} |x|^q d\mu(x))$ *such that for all* $N \geq 2$ *we have*

$$\mathbb{E}\left[W^{(2)}(\mu, \mu^N)^2\right] \leq C\eta_N := C \begin{cases} N^{-\frac{1}{2}} & d < 4 \\ N^{-\frac{1}{2}} \log(N) & d = 4 \\ N^{-\frac{2}{d}} & d > 4. \end{cases}$$

Note that this can be generalized to any $W^{(q)}$ distance under existence of higher moments by including some more terms. Notably for a measure $\mu$ for which high enough moments exist we have

$$\mathbb{E}\left[W^{(q)}(\mu, \mu^N)^q\right] \leq C\eta_N. \tag{4.4}$$

See [11] for more details. We skip the lenghty proof of this theorem and focus on applying it to the setting of the particle approximation.

We recover the same rate of approximation leading us to believe that no better result should be obtainable in this generality.

**Theorem 4.5.** *Let* $(\xi^n, \mathbf{X}^n)_{n\in\mathbb{N}}$ *be i.i.d random variables with law* $m \in \mathcal{E}_0^q(K)$ *for some function* $K : (0, \infty) \to (0, \infty)$ *and* $q > 4$*. Let* $f \in \mathcal{C}_b^3$ *and* $g$ *satisfy*

$$\begin{aligned} |g(t, y, \mu) - g(t, y', \mu')| &\leq L\left(|y - y'| + W^{(2)}(\mu, \mu')\right), \\ |g(t, y, \mu)| &\leq L. \end{aligned}$$

*For* $N \in \mathbb{N}$ *let* $Y^i$ *be the solution to the McKean-Vlasov equation* (3.3) *with input* $(\xi^i, \mathbf{X}^i)$ *and* $Y^{i,N}$ *be the i-th entry of the solution to the particle system with input* $(\xi^{(N)}, \mathbf{X}^{(N)})$*. Then there exists a constant* $C$ *independent of* $N$ *such that*

$$\sup_{1\leq i\leq N} \mathbb{E}\left[\left\|Y^i - Y^{i,N}\right\|_{\infty,[0,T]}^2\right] \leq C \begin{cases} N^{-\frac{1}{2}} & d < 4 \\ N^{-\frac{1}{2}} log(N) & d = 4 \\ N^{-\frac{2}{d}} & d > 4. \end{cases}$$

*Proof.* Let $i, N$ be fixed for now and pick some $q > 4$. As always the constant $C$ can increase from line to line but we try to make clear what factors into it.

We have $Y_i = \Theta_{g,f}(\mu, \xi^i, \mathbf{X}^i)$ and $Y^{i,N} = \Theta_{g,f}(\mu^N, \xi^i, \mathbf{X}^i)$. Thus by Corollary 2.4 and Lemma 1.8 we have

$$\begin{aligned} \left\|Y^i - Y^{i,N}\right\|_{\infty,[0,T]} &\leq \left\|Y^i\right\|_\infty + \left\|Y^{i,N}\right\|_\infty \\ &\leq C\left(|\xi^i| + \left\|\left|\mathbf{X}^i\right|\right\|_{p-var}^p + T\right) \\ &\leq C\left(|\xi^i| + e^{CN_1(w_{X^i},[0,T])} + T\right) \\ &=: CM^i. \end{aligned} \tag{4.5}$$

Furthermore by Proposition 2.5 we have

$$\left\|Y^i - Y^{i,N}\right\|_{\infty,[0,T]} \leq Ce^{CN_1(w_{X^i},[0,T])} \int_0^T W^{(2)}(\mu_r, \mu_r^N) dr.$$

Therefore we can bound the expected value via

$$\begin{aligned}\mathbb{E}\left[\left\|Y^i - Y^{i,N}\right\|_\infty^2\right] &\leq \int_0^T \mathbb{E}\left[C^2 e^{2CN_1(w_{X^i},[0,T])} W^{(2)}(\mu_r, \mu_r^N)^2\right] dr \\ &\leq \int_0^T C\mathbb{E}\left[e^{CN_1(w_{X^i},[0,T])}\left(W^{(2)}(\mu_r, \bar{\mu}_2^N)^2 + W^{(2)}(\bar{\mu}_r^N, \mu_r^N)^2\right)\right] dr \\ &= \int_0^T C(T_1 + T_2) dr.\end{aligned} \tag{4.6}$$

Here, as before we denote $\bar{\mu}^N = \sum_{i=1}^N \delta_{Y^i}$. We now bound the individual terms. Note that since $\mu_r^N$, $\bar{\mu}_r^N$ are both discrete measures we have

$$W^{(2)}(\bar{\mu}_r^N, \mu_r^N) \leq \left(\frac{1}{N}\sum_{i=1}^N |Y_r^i - Y_r^{i,N}|^2\right)^{\frac{1}{2}}.$$

We use this to get

$$\begin{aligned}T_2 &\leq \mathbb{E}\left[e^{CN_1(w_{X^i},[0,T])}\frac{1}{N}\sum_{i=1}^N |Y_r^i - Y_r^{i,N}|^2\right] \\ &\leq \frac{1}{N}\mathbb{E}\left[e^{CN_1(w_{X^i},[0,T])}|Y_r^i - Y_r^{i,N}|^2\right] + \frac{1}{N}\sum_{j\neq i}\mathbb{E}\left[e^{CN_1(w_{X^i},[0,T])}|Y_r^j - Y_r^{j,N}|^2\right] \\ &\leq \frac{1}{N}\mathbb{E}\left[e^{CN_1(w_{X^i},[0,T])}(M^i)^2\right] + K(C)\frac{1}{N}\sum_{j\neq i}\mathbb{E}\left[|Y_r^j - Y_r^{j,N}|^2\right] \\ &\leq \frac{1}{N}\mathbb{E}\left[Ce^{CN_1(w_{X^i},[0,T])}(|\xi^i|^2 + 1 + T)\right] + K(C)\frac{N-1}{N}\mathbb{E}\left[|Y_r^i - Y_r^{i,N}|^2\right] \\ &\leq \frac{1}{N}C\left(K(C) + K(C)\mathbb{E}\left[|\xi^i|^4\right]^{\frac{1}{2}}\right) + K(C)\mathbb{E}\left[|Y_r^i - Y_r^{i,N}|^2\right] \\ &\leq C\left(\frac{1}{N} + \mathbb{E}\left[|Y_r^i - Y_r^{i,N}|^2\right]\right).\end{aligned}$$

Furthermore we define $\bar{\mu}_r^{N,i} = \frac{1}{N}\left(\delta_{Y_r^{N+1}}\sum_{j\neq i}\delta_{Y_r^j}\right)$. Note that $\bar{\mu_r}^{N,i}$ is an empirical measure for $\mu_r$ and thus Theorem 4.4 applies. We then have

$$\begin{aligned}T_1 &\leq C\mathbb{E}\left[e^{CN_1(w_{X^i},[0,T])}\left(W^{(2)}(\mu_r, \bar{\mu}_r^{N,i})^2 + W^{(2)}(\bar{\mu}_r^{N,i}, \bar{\mu}_r^N)^2\right)\right] \\ &\leq CK(C)\mathbb{E}\left[W^{(2)}(\mu_r, \bar{\mu}_r^{N,i})^2\right] + C\mathbb{E}\left[e^{CN_1(w_{X^i},[0,T])}\frac{1}{N}\left(|Y_r^i - Y_r^{N+1}|^2 + \sum_{j\neq i}|Y_r^j - Y_r^j|^2\right)\right] \\ &\leq C\mathbb{E}\left[W^{(2)}(\mu_r, \bar{\mu}_r^{N,i})^2\right] + \frac{1}{N}C\mathbb{E}\left[e^{CN_1(w_{X^i},[0,T])}|Y_r^i - Y_r^{N+1}|^2\right] \\ &\leq C\left(\mathbb{E}\left[W^{(2)}(\mu_r, \bar{\mu}_r^{N,i})^2\right] + \frac{1}{N}\left(K(C) + K(C)\mathbb{E}\left[|\xi^i|^4\right]^{\frac{1}{2}}\right)\right) \\ &\leq C\left(\mathbb{E}\left[W^{(2)}(\mu_r, \bar{\mu}_r^{N,i})^2\right] + \frac{1}{N}\right).\end{aligned}$$

Plugging these estimates back into (4.6) we get

$$\mathbb{E}\left[\left\|Y^i - Y^{i,N}\right\|_{\infty,[0,T]}^2\right] \leq \int_0^T C\left(\frac{1}{N} + \mathbb{E}\left[W^{(2)}(\mu_r, \bar{\mu}_r^{N,i})^2\right] + \mathbb{E}\left[|Y_r^i - Y_r^{i,N}|^2\right]\right) dr.$$

Now applying Grönwall's inequality gives us

$$\mathbb{E}\left[\left\|Y^i - Y^{i,N}\right\|_{\infty}^2\right] \leq C\left(\frac{1}{N} + \sup_{r\in[0,T]} \mathbb{E}\left[W^{(2)}(\mu_r, \bar{\mu}_r^{N,i})^2\right]\right) e^{TC}.$$

Note that since

$$\begin{aligned}\sup_{r\in[0,T]} \int_{\mathbb{R}^d} |x|^q d\mu_r(x) &= \mathbb{E}\left[\left\|Y^1\right\|_{\infty}\right] \\ &\leq C\left(\mathbb{E}\left[|\xi^1|^q\right] + K(C) + 1\right) =: B < \infty\end{aligned}$$

we have a universal bound inpendent of $r$ and can apply Theorem 4.4 with the constant depending on $B$ to obtain

$$\mathbb{E}\left[\left\|Y^i - Y^{i,N}\right\|_{\infty}^2\right] \leq C\left(\frac{1}{N} + \eta_N\right).$$

Since the right side does not depend on $i$ we can take the supremum and the claim follows. □

*Remark* 4.6. This statement and proof can easily be adapted to

$$\sup_{1\leq i\leq N} \mathbb{E}\left[\left\|Y^i - Y^{i,N}\right\|_{\infty,[0,T]}^p\right] \leq C\eta_N$$

if $g$ is Lipschitz in the measure argument with respect to $W^{(p)}$ and high enough moments exist for $|\xi^i|$ and $\left\|\left|\mathbf{X}^i\right|\right\|_{p-var}$. The proof follows the same steps and yields the same result due to (4.4) and the existence of high enough moments.

We stress that Theorem 4.5 does not directly give us a rate of convergence for $W^{(2)}(\mu, \mu^N)$ as $N \to \infty$. It does however imply

$$\begin{aligned}W^{(2)}(\bar{\mu}^N, \mu^N)^2 &\leq \frac{1}{N}\sum_{i=1}^N \left\|Y^i - Y^{i,N}\right\|_{\infty}^2 \\ &\leq \sup_{1\leq i\leq N} \mathbb{E}\left[\left\|Y^i - Y^{i,N}\right\|_{\infty}^2\right] \leq C\eta_N\end{aligned}$$

An approach to extend this to give us a rate of convergence for $W^{(2)}(\mu, \mu^N)$ is to get an approximation rate for $W^{(2)}(\mu, \bar{\mu}^N)$ for measures $\mu \in \mathcal{P}_2(\mathcal{C}([0,T], \mathbb{R}^d))$. We are not sure if the same rate holds. It seems to be an open problem.

For simulation purposes however this result is useful.

*Remark* 4.7. We expect a naive Euler approach of the type

$$\begin{aligned}\bar{Y}_{t_{k+1}}^{i,N,M} = \bar{Y}_{t_k}^{i,N,M} &+ g\left(t_k, \bar{Y}_{t_k}^{i,N,M}, \mu_{t_k}^N\right) h + f\left(\bar{Y}_{t_k}^{i,N,M}\right) \Delta X_{t_k}^i \\ &+ Df\left(\bar{Y}_{t_k}^{i,N,M}\right) f(\bar{Y}_{t_k}^{i,N,M}) \Delta \mathbb{X}_{t_k}^i\end{aligned}$$

for $h = \frac{T}{M}$ to provide approximations with

$$\sup_{1 \leq i \leq N} \mathbb{E}\left[\left\|Y_t^{i,N} - Y_t^{i,N,M}\right\|_\infty^2\right] \leq Ch.$$

Combined with Theorem 4.5 this would directly result in

$$\sup_{1 \leq i \leq N} \mathbb{E}\left[\left\|Y^i - Y^{i,N,M}\right\|_{\infty,[0,T]}^2\right] \leq C(\eta_N + h).$$

This still needs to be verified in future work.

# Appendix A

# Wasserstein Metric

## A.1 Definitions and Properties

There are several metrics on spaces of probability measures. In this work we chiefly use the Wasserstein metric $W$ because of its geometric properties and being particularly well-suited to control functions of empirical measures. We will use this appendix to remind the reader of some definitions, intuition and properties. As this is just a brief overview some proofs will be skipped. For a more complete picture see [5].

Let $(E, d)$ be a separable metric space equipped with the Borel $\sigma$-algebra $\mathcal{E} = \mathcal{B}(E)$. We denote for any $p \geq 1$ by $\mathcal{P}_p(E)$ the space of probability measures on $E$ with finite $p$-th moment. That is

$$\mathcal{P}_p(E) = \left\{ \mu \in \mathcal{P}(E) : \int_E d(x_0, x)^p d\mu(x) < \infty \text{ for some (and thus all) } x_0 \in E \right\}.$$

Let $\mu, \nu \in \mathcal{P}_p(E)$. Then we define their $p$-th Wasserstein distance by

$$W^{(p)}(\mu, \nu)^p := \inf_{\pi \in \Pi(\mu,\nu)} \int_{E \times E} d(x, y)^p d\pi(x, y) \tag{A.1}$$

where $\Pi(\mu, \nu)$ is the set of all couplings of $\mu$ and $\nu$, i.e. probability measures on $E \times E$ whose projections to the first and second coordinates are $\mu$ and $\nu$ respectively. Note that for any $\mu, \nu \in \mathcal{P}_p(E)$ any coupling $\pi$ is in $\mathcal{P}_p(E \times E)$ with the metric being the $p$-product metric. It is well known that the optimum is always achieved, i.e. there exists a $\pi_o \in \Pi(\nu, \mu)$ such that $W^{(p)}(\mu, \nu)^p = \int_{E \times E} d(x, y)^p d\pi_o(x, y)$.

It is known that $W^{(p)}$ is indeed a metric on $P_p(E)$ and furthermore if $(E, d)$ is Polish, so is $(P_p(E), W^{(p)})$. These are classical results.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Note that for any two random variables $X, Y : \Omega \to E$ for which we have $\mathbb{E}\left[d(X, x_0)^p\right] + \mathbb{E}\left[d(Y, x_0)^p\right] < \infty$ for some and thus for all $x_0 \in E$ we have

$$W^{(p)}(\mathcal{L}(X), \mathcal{L}(Y))^p \leq \mathbb{E}\left[d(X, Y)^p\right].$$

Furthermore if $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless then for any law $\mu \in \mathcal{P}_p(E)$ there exists a random variable $X$ on $\Omega$ with law $\mu$. Therefore we can write

$$W^{(p)}(\mu, \nu)^p = \inf_{\mathcal{L}(X)=\mu, \mathcal{L}(Y)=\nu} \mathbb{E}\left[d(X, Y)^p\right].$$

Another well-known property is the Kantorovich duality theorem. It is central in the study of transportation theory. We include it here although we do not use it in our work.

**Proposition A.1.** *Let $(E,d)$ be a Polish space and $p \geq 1$ and $\mu, \nu \in \mathcal{P}_p(E)$. Then*

$$W^{(p)}(\mu,\nu)^p = \sup_{\phi,\psi:\phi(x)+\psi(y)\leq d(x,y)^p} \left( \int_E \phi(x) d\mu(x) + \int_E \psi(x) d\nu(x) \right)$$

An obvious property but one that we use throughout is monotonicity under projection.

**Lemma A.2.** *Let $\mu, \nu \in \mathcal{P}_p(\mathcal{C}([0,T],\mathbb{R}^d))$ for some $p \geq 1$ and $T > 0$. Then for any $0 \leq s \leq t \leq T$ we have $\mu\big|_{[s,t]} \in \mathcal{P}_p(\mathcal{C}([s,t],\mathbb{R}^d))$ and*

$$W^{(p)}(\mu\big|_{[s,t]}, \nu\big|_{[s,t]}) \leq W^{(p)}(\mu,\nu).$$

*Here the projection to $[t,t]$ is associated with just the marginal at time $t$.*

*Proof.* It is clear that for any coupling $\pi \in \Pi(\mu,\nu)$ the projection $\pi\big|_{[s,t]}$ lies in $\Pi(\mu\big|_{[s,t]}, \nu\big|_{[s,t]})$. Let now $\pi_o$ be an optimal coupling in $\Pi(\mu,\nu)$. We then have

$$\begin{aligned} W^{(p)}(\mu\big|_{[s,t]}, \nu\big|_{[s,t]})^p &\leq \int\int_{\mathcal{C}_{[s,t]}\times\mathcal{C}_{[s,t]}} \left\|\gamma - \gamma'\right\|_{\infty,[s,t]} d\pi_o\big|_{[s,t]}(\gamma,\gamma') \\ &\leq \int\int_{\mathcal{C}_{[0,T]}\times\mathcal{C}_{[0,T]}} \left\|\gamma - \gamma'\right\|_{\infty,[0,T]} d\pi_o(\gamma,\gamma') \\ &= W^{(p)}(\mu,\nu)^p \end{aligned}$$

□

## A.2 Weak Convergence

It is well known that weak convergence is metrized by the Lévy-Prokhorov metric. Convergence in the $p$-Wasserstein metric adds convergence of the $p$-th moment as condition in the following way.

**Proposition A.3** (Wasserstein metric metrizes weak convergence)**.** *Let $(E,d)$ be a separable metric space. For some $p \geq 1$ let $(\mu^n), \mu$ be in $\mathcal{P}_p(E)$. Then*

$$\lim_{n\to\infty} W^{(p)}(\mu^n,\mu) = 0 \iff \begin{cases} \mu^n \Rightarrow \mu & \textit{and} \\ \int_E d(x_0,x)^p d\mu^n(x) \xrightarrow{n\to\infty} \int_E d(x_0,x)^p d\mu(x) & \forall x_0 \in E. \end{cases} \tag{A.2}$$

*Proof.* $\boxed{\Rightarrow}$:
Assume $W^{(p)}(\mu,\mu^n) \to 0$ and let $\pi^n \in \Pi_p^{opt}(\mu^n,\mu)$ be optimal $p$-couplings.

We first show weak convergence $\mu^n \implies \mu$. For any bounded uniformly continuous function $f : E \to \mathbb{R}$ and any $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\begin{aligned}A_n &= \int_E f(x)d\mu(x) - \int_E f(y)d\mu^n(y) \\ &= \int_{E\times E} f(x) - f(y)d\pi^n(x,y) \\ &\leq \int_{d(x,y)\leq\delta} |f(x) - f(y)|d\pi^n(x,y) + \int_{d(x,y)>\delta} |f(x) - f(y)|d\pi^n(x,y) \\ &\leq \epsilon + \int_{d(x,y)>\delta} |f(x) - f(y)|d\pi^n(x,y) \quad \xrightarrow{n\to\infty} \epsilon.\end{aligned}$$

Since we can choose $\epsilon$ arbitrarily small this gives us weak convergence $\mu^n \implies \mu$.

For the convergence of the moments we note that using the triangle inequality we have

$$\left|W^{(p)}(\mu^n, \delta_{x_0}) - W^{(p)}(\mu, \delta_{x_0})\right| \leq W^{(p)}(\mu^n, \mu) \xrightarrow{n\to\infty} 0.$$

It follows that

$$\int d(x, x_0)^p d\mu^n(x) = W^{(p)}(\mu^n, \delta_{x_0})^p \xrightarrow{n\to\infty} W^{(p)}(\mu, \delta_{x_0})^p = \int d(x, x_0)^p d\mu(x).$$

$\boxed{\Leftarrow}$ :

Let now the right hand side of (A.2) hold. By Skorokhod's representation theorem we get random variables $(X^n)_{n\in\mathbb{N}}, X$ with corresponding laws $\mu^n, \mu$ converging almost surely $X^n \xrightarrow{\mathbb{P}\text{-a.s.}} X$. Using Fatou's lemma we see that $\mu \in \mathcal{P}_p(E)$. Indeed

$$\begin{aligned}\int_E d(x_0, x)^p d\mu(x) &= \mathbb{E}\left[d(x_0, X)^p\right] \\ &\leq \liminf_{n\to\infty} \mathbb{E}\left[d(x_0, X^n)^p\right] \\ &= \liminf_{n\to\infty} \int_E d(x_0, x)^p d\mu^n(x).\end{aligned}$$

For any $r > 0$ and $A_n := \mathbb{E}\left[d(x_0, X^n)^p \mathbb{1}_{\{d(x_0,X^n)\geq r\}}\right]$ we then get

$$\begin{aligned}\liminf_{n\to\infty} A_n &= \liminf_{n\to\infty} \left(\mathbb{E}\left[d(x_0, X^n)^p\right] - \mathbb{E}\left[(d(x_0, X^n) \wedge r)^p\right] + r^p\mathbb{P}(d(x_0, X^n) \geq r)\right) \\ &\leq \mathbb{E}\left[d(x_0, X)^p\right] - \mathbb{E}\left[(d(x_0, X) \wedge r)^p\right] + r^p\mathbb{P}(d(x_0, X) \geq r) \\ &= \mathbb{E}\left[d(x_0, X)^p \mathbb{1}_{\{d(x_0,X^n)\geq r\}}\right].\end{aligned}$$

For $r$ large enough this is clearly arbitrarily small giving us uniform integrability of $(d(x_0, X^n)^p)_{n\in\mathbb{N}}$. Since we have almost sure convergence from Skorokhod's lemma this is enough to show convergence in $L^p$. We can thus conclude the proof via

$$W^{(p)}(\mu^n, \mu)^p \leq \mathbb{E}\left[d(X^n, X)^p\right] \xrightarrow{n\to\infty} 0.$$

□

From this we get as a consequence a sufficient condition allowing us to pass from weak convergence to convergence in $W^{(p)}$.

**Corollary A.4.** *Let $(E,d)$ be a Polish space and $(\mu^n)_{n\in\mathbb{N}}$, $\mu$ measures in $\mathcal{P}_p(E)$ for some $p \geq 1$. If $\mu^n$ converges weakly to $\mu$ as $n \to \infty$ and for some $x_0 \in E$ we have*

$$\lim_{r\to\infty} \sup_{n\in\mathbb{N}} \int_E d(x_0,x)^p \mathbb{1}_{\{d(x_0,x)\geq r\}} d\mu^n(x) = 0$$

*then $W^{(p)}(\mu^n,\mu) \xrightarrow{n\to\infty} 0$.*

*Proof.* By Skorokhod's representation theorem there exist random variables $X, (X^n)_{n\in\mathbb{N}}$ with laws $\mu, \mu^n$ such that we have almost sure convergence $X^n \xrightarrow{\mathbb{P}-\text{a.s.}} X$. For any $r > 0$ we then have

$$\mathbb{E}\left[d(x_0,X^n)^p\right] \leq \mathbb{E}\left[d(x_0,X^n)^p \mathbb{1}_{\{d(x_0,X_n)\geq r\}}\right] + \mathbb{E}\left[(d(x_0,X^n) \vee r)^p\right].$$

Therefore we obtain

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[d(x_0,X^n)^p\right] \leq \mathbb{E}\left[d(x_0,X)^p\right].$$

Thus we have uniform integrability of $(d(x_0,X^n))_{n\in\mathbb{N}}$ and since it converges almost surely we have

$$\lim_{n\to\infty} \mathbb{E}\left[d(x_0,X^n)^p\right] = \mathbb{E}\left[d(x_0,X)^p\right].$$

Convergence in $W^{(p)}$ follows by Proposition A.3. □

# Bibliography

[1] I. Bailleul, R. Catellier, and F. Delarue. Mean field rough differential equations. *arXiv:1802.05882 [math]*, Feb. 2018. arXiv: 1802.05882.

[2] H. Bessaih, M. Coghi, and F. Flandoli. Mean field limit of interacting filaments for 3D Euler equations. *J. Stat. Phys.*, 174(3):562–578, 2019. ISSN: 0022-4715.

[3] V. I. Bogachev and A. V. Kolesnikov. The monge-kantorovich problem: achievements, connections, and perspectives. *Russian Mathematical Surveys*, 67(5):785–890, 2012.

[4] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. 2: Mean field games with common noise and master equations.* eng, number volume 84 in Probability theory and stochastic modelling. ISBN: 978-3-319-56436-4 978-3-319-56435-7.

[5] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, Cham, 2018.

[6] T. Cass, C. Litterer, and T. Lyons. Integrability and tail estimates for gaussian rough differential equations. *Ann. Probab.*, 41(4):3026–3050, July 2013.

[7] T. Cass and T. Lyons. Evolving communities with individual preferences. *Proc. London Math. Soc.*, 3:83–107.

[8] M. Coghi, J.-D. Deuschel, P. Friz, and M. Maurelli. Pathwise mckean-vlasov theory, 2018. arXiv: `1812.11773 [math.PR]`.

[9] M. Coghi and T. Nilssen. Rough nonlocal diffusions. *arXiv preprint arXiv:1905.07270*, 2019.

[10] A. Deya, M. Gubinelli, M. Hofmanová, and S. Tindel. A priori estimates for rough pdes with application to rough conservation laws. *Journal of Functional Analysis*, 276(12):3577 –3645, 2019.

[11] N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3):707–738, 2015.

[12] N. Fournier, M. Hauray, and S. Mischler. Propagation of chaos for the 2D viscous vortex model. *J. Eur. Math. Soc. (JEMS)*, 16(7):1423–1466, 2014. ISSN: 1435-9855.

[13] P. Friz and A. Shekhar. General rough integration, levy rough paths and a levy–kintchine type formula, 2012. arXiv: `1212.5888 [math.PR]`.

[14] P. Friz and N. Victoir. A note on the notion of geometric rough paths, 2004. arXiv: `math/0403115 [math.FA]`.

[15] P. K. Friz and M. Hairer. *A Course on Rough Paths*. en. Universitext. Springer International Publishing, Cham, 2014. ISBN: 978-3-319-08331-5 978-3-319-08332-2.

[16] P. K. Friz and N. B. Victoir. Multidimensional Stochastic Processes as Rough Paths: Theory and Applications. en:672.

[17] M Gubinelli. Controlling rough paths. *Journal of Functional Analysis*, 216(1):86 –140, 2004.

[18] C. Guhlke, P. Gajewski, M. Maurelli, P. K. Friz, and W. Dreyer. Stochastic many-particle model for LFP electrodes. *Contin. Mech. Thermodyn.*, 30(3):593–628, 2018. ISSN: 0935-1175.

[19] T. Lyons and N. Victoir. An extension theorem to rough paths. en. *Annales de l'I.H.P. Analyse non linéaire*, 24(5):835–847, 2007.

[20] C. Marchioro and M. Pulvirenti. Hydrodynamics in two dimensions and vortex theory. *Comm. Math. Phys.*, 84(4):483–503, 1982. ISSN: 0010-3616.

[21] A.-S. Sznitman. Topics in propagation of chaos. In *Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989*. Volume 1464, pages 165–251. Springer Berlin Heidelberg, Berlin, Heidelberg, 1991. ISBN: 978-3-540-53841-7 978-3-540-46319-1.