Existence of weak solutions of mean-field stochastic differential equations

Jesse Koivu

Master's thesis in Mathematics
University of Jyväskylä
Department of Mathematics and Statistics
Fall 2021

Abstract

In this thesis we consider mean-field stochastic differential equations, which are an extension of classical stochastic differential equations, where the coefficients may depend on an additional measure component in the law of the solution. We consider the existence of weak solutions of such equations under the assumption that the coefficients are bounded and continuous, where continuity is understood in the 2-Wasserstein metric in the measure component. We follow the treatment given in the article of Li, J. and Min, H., Weak solutions of mean-field stochastic differential equations (2017). We start by recalling some fundamental notions from stochastic analysis. Then we introduce the path space, along with the classical local martingale problem and functional stochastic differential equations. Furthermore we introduce the Wasserstein spaces of measures and how to differentiate functions depending on a measure variable. Finally we show the existence of weak solutions to mean-field stochastic differential equations under bounded, measurable and continuous coefficients by showing that there exists a solution to the corresponding local martingale problem.

Tiivistelmä

Tässä tutkielmassa käsittelemme odotusarvokentällisiä stokastisia differentiaaliyhtälöitä, mitkä ovat yleistys klassisille stokastisille differentiaaliyhtälöille. Odotusarvokentällisen stokastisen differentiaaliyhtälön kerroinfunktiot saattavat riippua ylimääräisestä mittakomponentista ratkaisun jakauman muodossa. Käsittelemme heikkojen ratkaisujen olemassaoloa tällaisille yhtälöille olettaen, että kerroinfunktiot ovat rajoitettuja ja jatkuvia, missä jatkuvuus mittakomponentin suhteen ymmärretään jatkuvuutena 2-Wasserstein metriikan suhteen. Seuraamme artikkelia Li, J. ja Min, H. *Weak solutions of mean-field stochastic differential equations* (2017). Aloitamme palauttamalla mieliimme joitakin keskeisiä käsitteitä stokastisesta analyysistä. Tämän jälkeen esittelemme polkuavaruuden, klassisen lokaalin martingaaliongelman ja funktionaaliset stokastiset differentiaaliyhtälöt. Lisäksi esittelemme Wassersteinin mittojen avaruudet ja funktioiden differentioituvuuden mittakomponentin suhteen. Lopuksi osoitamme heikkojen ratkaisujen olemassaolon odotusarvokentällisille stokastisille differentiaaliyhtälöille olettaen että kerroinfunktiot ovat rajoitettuja, mitallisia ja jatkuvia. Tämä tehdään näyttämällä, että vastaavalla lokaalilla martingaaliongelmalla on olemassa ratkaisu.

Contents

1 Introduction

Throughout the thesis we make the convention to use the abbreviation SDE for stochastic differential equation.

This thesis deals with a type of SDEs called mean-field SDEs. How the classical SDE

$$dX_t = b(s, X_s)ds + \sigma(s, X_s)dW_s \tag{1}$$

has coefficients depending on a time variable and a spatial variable, in the mean-field case we have

$$dX_t = b(s, X_s, \mathbb{P}_{X_s})ds + \sigma(s, X_s, \mathbb{P}_{X_s})dW_s. \tag{2}$$

This means that in the mean-field case the coefficients are allowed to depend on the law of the solution, which is a measure. This extension brings difficulties, for example defining derivatives with respect to measure components and applying existing theory of SDEs to the mean-field case. Typically one can freeze the measure component and then use the corresponding result in the classical SDE case to obtain results, which we will see later in this thesis to be useful. However there is a degree of complexity that is added and needs to be treated.

The above discussion raises the question why are we interested in mean-field SDEs? To answer this question we will first attempt to give an answer on why are we interested in classical SDEs. One of the first examples of a classical SDE is the Ornstein-Uhlenbeck process (see [23]), given in the differential form by

$$dX_t = -bX_t dt + \sigma dW_t. \tag{3}$$

The equation corresponds to the Langevin (see [17]) equation for the Brownian motion of a particle with friction. Physics overall is a source of many mathematical problems and the fields of SDEs and mean-field SDEs are no exception to such problems. The theory of ordinary differential equations and partial differential equations comes to a stop when trying to treat such an equation, because of the random element involved; the integral $\int_0^t \sigma dW_s$ is a random variable. Examples exist in other disciplines as well, for example in finance and economics the randomness of a particle is replaced with the randomness of a value of a stock. In this approach the value of an asset called a bond can be modeled by an ordinary differential equation, but the value of other assets, called stocks, must be modeled by an SDE (see [16]). Another example from the field of finance is the so-called Black-Scholes (see [3]) option valuation formula, which is of fundamental importance in the field of mathematical finance.
Coming back to the motivation of mean-field SDEs, we have again the roots

in physics. The interest began with the Boltzmann equations and modeling them with mean-field SDEs, as done by Kac [15], McKean [20] and Vlasov [26]. Mean-field SDEs are often called McKean-Vlasov SDEs based on these contributions to the field.

The field of SDEs and stochastic analysis has been extensively studied for almost a century now. The field began largely with Itô's contributions in [13] and [14]. Itô studied SDEs with Lipschitz continuous coefficients, and over time the existence of weak solutions to SDEs began to be more widely known, first under just continuous coefficients, and then merely measurable and bounded coefficients. The field started to be well-investigated until a few decades ago mean-field SDEs brought many new problems. Various existence and uniqueness results for mean-field SDEs were achieved in articles of Funaki [8] and Gärtner [11], mostly relating to various assumptions on the Lipschitz continuity of coefficients. Further interest was brought to the context of mean-field SDEs via the concept of a mean-field game, that were studied for example by Lasry and Lions in [18]. Carmona and Lacker [6] found weak solutions to mean-field SDEs with measurable coefficients where the diffusion coefficient σ does not depend on the law of the solution and the drift coefficient b is sequentially continuous in the measure component.

There are (at least) two main approaches to deal with mean-field SDEs, the Lyapunov-type approach and the martingale problem approach. For the Lyapunov-type approach the reader can consult for example [21], [12]. In this thesis we take the martingale problem approach. The papers of Funaki and Gärtner used the classical local martingale problem (this originates from Stroock and Varadhan, see [16]) to treat mean-field SDEs under their assumptions.

We will look to investigate the existence of a weak solution to the equation

$$dX_t = b(s, X_s, \mathbb{P}_{X_s})ds + \sigma(s, X_s, \mathbb{P}_{X_s})dW_s, \tag{4}$$

under the assumption that the coefficients b and σ are bounded and continuous. For this case it is not sufficient to deal with the classical local martingale problem, so we will look to extend the local martingale problem for our needs. As a motivation to the above equation we consider the following system of interacting diffusion, studied by Chiang in [7]

$$\begin{aligned} dX_t^{n,i} &= b\Big(X_t^{n,i}, \frac{1}{n}\sum_{j=1}^n \delta_{X_t^{n,j}}\Big)dt + \sigma\Big(X_t^{n,i}, \frac{1}{n}\sum_{j=1}^n \delta_{X_t^{n,j}}\Big)dW_t^i \\ X_0^{n,i} &= x_0, x_0 \in \mathbb{R}^d,\ 1 \le i \le n, \end{aligned} \tag{5}$$

where W_i are independent Brownian motions and δ_x denotes the Dirac measure at $x \in \mathbb{R}^d$. This system models the behaviour of n weakly interacting particles, similarly one can think in the context of mean-field games and games about n players. The importance of our mean-field SDE (4) is that it models the asymptotic behaviour of this large system of diffusions as $n \to \infty$, i.e. when the number of particles (or players) becomes large.

Finally we note that the classical Itô formula is of fundamental importance in the study of SDEs and the field of mean-field SDEs is no exception. However in the mean-field case there is the lack of measure derivatives, and thus it has been extended for example the paper of Buckdahn et al. [4], to contain the derivatives with respect to the measure variable.

The thesis is organized as follows. In the first section we recall some classical stochastic analysis, including the Brownian motion, stochastic integration, the classical Itô formula, and then introduce the classical SDEs. In the second section we fix some notations and setting for our future needs. The third section is devoted to the path space

$$(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d))) \tag{6}$$

and the classical local martingale problem. In the fourth section we introduce the Wasserstein spaces. In the fifth section we consider measure derivatives via a lifting to L^2 and then look to extend the Itô formula to the measure dependant case. Finally, in the sixth section we consider mean-field SDEs, extend the local martingale problem to the mean-field case and show that solving the local martingale problem is equivalent to solving the mean-field SDE. Finally the main theorem of the thesis considers the existence of a weak solution to the mean-field SDE (4) under bounded and continuous coefficients. Our treatment follows closely the treatment in [19].

1.1 Table of notation

Here we summarize some notations from the article that are used.

- $\mathcal{A}'f(s,y) = \sum_{i=1}^d b_i(s,y)\frac{\partial}{\partial x_i}f(s,y(s)) + \frac{1}{2}\sum_{i,j,k=1}^d(\sigma_{ik}\sigma_{jk})(s,y)\frac{\partial^2}{\partial x_i \partial x_j}f(s,y(s))$ where $y \in C([0,T];\mathbb{R}^d)$
- $(\widetilde{\mathcal{A}}f)(s,y,\nu) = \sum_{i=1}^d b_i(s,y,\nu)\partial_{y_i}f(s,y) + \frac{1}{2}\sum_{i,j,k=1}^d(\sigma_{ik}\sigma_{jk})(s,y,\nu)\partial_{y_iy_j}f(s,y)$ where $(s,y,\nu) \in ([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$
- $\mathcal{A}f(s,y,\nu) = (\widetilde{\mathcal{A}}f)(s,y,\nu) + \sum_{i=1}^d \int_{\mathbb{R}^d}(\partial_\mu f)_i(s,y,\nu,z)b_i(s,z,\nu)\nu(dz) + \frac{1}{2}\sum_{i,j,k=1}^d \int_{\mathbb{R}^d}\partial_{z_i}(\partial_\mu f)_j(s,y,\nu,z)(\sigma_{ik}\sigma_{jk})(s,z,\nu)\nu(dz)$
- $M_t^f = f(t,y(t)) - f(0,y(0)) - \int_0^t(\partial_s + \mathcal{A}')f(s,y(s))ds$ where $f \in C^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R})$
- $C_t^f = f(t,y(t)) - f(0,y(0)) - \int_0^t(\partial_s + \widetilde{\mathcal{A}})f(s,y(s),\mu(s))ds$ where $f \in C_b^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R})$
- $C^f(t,y,\mu) = f(t,y(t),\mu(0)) - f(0,y(0),\mu(0)) - \int_0^t(\partial_s + \mathcal{A})f(s,y(s),\mu(s))ds$ where $f \in C_b^{1,2,1}([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d);\mathbb{R})$

2 Classical stochastic analysis

In this section we will introduce some parts of the framework and central objects from classical stochastic analysis that are used later or generalized to fit the mean-field case. The main object of study in this section is the stochastic integral $\int_0^T X_t dW_t$. We first define what the process $(W_t)_{t\in[0,T]}$ is and then discuss what it means to integrate a stochastic process against this process. We will then recall the classical Itô's formula, which can be thought of as the Fundamental Theorem of Calculus for stochastic integrals. Finally at the end of the section we introduce the classical SDEs.

2.1 Filtrations and measurability

First we define a filtration and a stochastic basis. Let us fix a time horizon $T > 0$. Here we only consider the time interval $[0, T]$ to fit the setting of the thesis, but all of the following concepts could also be defined for $[0, \infty)$.

Definition 2.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A collection of σ-algebras $(\mathcal{F}_t)_{t\in[0,T]}$ is called a filtration given that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for all $0 \leq s \leq t \leq T$. The quadruple $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ is called a stochastic basis.

Definition 2.2. We say that the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ satisfies the usual conditions provided that the following is satisfied:

1. $(\Omega, \mathcal{F}, \mathbb{P})$ is complete.
2. $\mathcal{F}_0$ contains all $\mathcal{F}$-null sets.
3. The filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is right continuous, that is $\mathcal{F}_t = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}$ for all $t \in [0, T)$.

We will recall also the concepts of progressive measurability and of an adapted process:

Definition 2.3. Assume that $X = (X_t)_{t\in[0,T]}$, $X_t : \Omega \to \mathbb{R}^d$ is a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ be a filtration.

1. We say that X is progressively measurable with respect to the filtration $\mathbb{F}$ given that for all $s \in [0, T]$ the map $(\omega, t) \to X_t(\omega)$ from $\Omega \times [0, s]$ to $\mathbb{R}^d$ is measurable with respect to $\mathcal{F}_s \otimes \mathcal{B}([0, s])$ and $\mathcal{B}(\mathbb{R}^d)$.
2. We say that X is adapted with respect to the filtration $\mathbb{F}$ given that for all $t \in [0, T]$ we have that X_t is $\mathcal{F}_t$-measurable.

These concepts of measurability will be for the most part sufficient for our needs. We will now consider the concept of a local martingale in finite time. Often local martingales are considered in infinite time along with the sequence of stopping times tending to infinity. The modification for a finite time interval is as follows.

Definition 2.4. Assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ satisfying the usual conditions. We say an $\mathbb{F}$-adapted process $X = (X_t)_{t\in[0,T]}$ is a local martingale given that there exists a non-decreasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$, such that $0 \le \tau_n \le T$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} \mathbb{P}(\tau_n = T) = 1$ and that $X^{\tau_n} = (X_{t\wedge\tau_n})_{t\in[0,T]}$ is a martingale for all $n \in \mathbb{N}$.

We will also need to estimate the expectation of a supremum of a process. For this we have the Doob's maximal inequality which we recall here. For more information one can see [16, Theorem 1.3.8].

Proposition 2.5. *Assume $X = (X_t)_{t\in[0,T]}$ is a continuous martingale and $p \in (1, \infty)$. Then for $t \in [0, T]$ it holds*

$$\mathbb{E}\big(\sup_{s\in[0,t]} |X_s|\big)^p \le \Big(\frac{p}{p-1}\Big)^p \mathbb{E}|X_t|^p. \tag{7}$$

2.2 Brownian motion and stochastic integrals

We shall next define the Brownian motion which is of fundamental importance in stochastic analysis and in the study of SDEs.

Definition 2.6. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a stochastic basis. An adapted stochastic process $W = (W_t)_{t\in[0,T]}$, $W_t : \Omega \to \mathbb{R}$ is called a standard $(\mathcal{F}_t)_{t\in[0,T]}$-Brownian motion given the following is satisfied:

1. $W_0 = 0$ holds $\mathbb{P}$-a.s.
2. For all $0 \le s < t \le T$ the increment $W_t - W_s$ is independent from $\mathcal{F}_s$. This means the sets C and $\{W_t - W_s \in A\}$ are independent for all $C \in \mathcal{F}_s$ and $A \in \mathcal{B}(\mathbb{R})$.
3. For all $0 \le s < t \le T$ the increment $W_t - W_s$ is normally distributed with mean 0 and variance $t - s$.
4. The map $t \to W_t(\omega)$ is continuous for all $\omega \in \Omega$.

We call W a standard d-dimensional Brownian motion if instead of 3. we have that for all $0 \le s < t \le T$ the increment $W_t - W_s$ is normally distributed with mean 0 and covariance matrix $(t-s)I_d$, where I_d is the identity matrix.

It is not a trivial question whether such a process even exists. It turns out the existence in the d-dimensional case follows from the 1-dimensional case. Here we have the following:

Proposition 2.7. *The standard Brownian motion exists.*

This can be proven in multiple ways, for example Kolmogorov's extension theorem, see [16, Theorem 2.2.2].

Next we will turn to the problem of defining the stochastic integral $\int_0^T X_t dW_t$, where $(X_t)_{t\in[0,T]}$ is a stochastic process. As with Riemann and Lebesgue integration there is a problem of determining what functions, or in our case processes, can be integrated. We will start by defining the integrals of the so-called simple processes. Assume from now on that we have a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ that satisfies the usual conditions, and that we have a 1-dimensional $(\mathcal{F}_t)_{t\in[0,T]}$-Brownian motion.

Definition 2.8. A stochastic process $X = (X_t)_{t\in[0,T]}$ is called simple, given that there exists a partition $\mathcal{P} = \{0 = t_0, \dots, t_n = T\}$ of the interval $[0, T]$ and $\mathcal{F}_{t_i}$-measurable random variables $v_i : \Omega \to \mathbb{R}$, $i \in \{0, \dots, n-1\}$ that are uniformly bounded over Ω, such that

$$X_t(\omega) = \sum_{i=1}^{n} \chi_{(t_{i-1},t_i]}(t) v_{i-1}(\omega). \tag{8}$$

We denote the class of simple processes by $\mathcal{L}_0^T$.

For the case of stochastic processes on $[0, \infty)$ the simple processes are defined with an increasing sequence that tends to infinity instead of a partition, and countable set of random variables. We now define a stochastic integral for simple functions.

Definition 2.9. For a process X given by Equation (8) and $t \in [0, T]$ we define the stochastic integral of X with respect to W to be

$$\Big(\int_0^t X_s dW_s\Big)(\omega) := \sum_{i=1}^{n} v_{i-1}(\omega)(W_{t\wedge t_i}(\omega) - W_{t\wedge t_{i-1}}(\omega)). \tag{9}$$

Notice that for a fixed t the integral is a random variable. Further on we will drop the ω from the notation of stochastic integrals for notational convenience. In this way we get a linear operator that sends simple processes to continuous, square-integrable martingales that start at zero. One also has the Itô isometry which states that for $0 \le s \le t \le T$ and a simple process X it holds

$$\mathbb{E}\Big(\Big[\int_0^t X_u dW_u - \int_0^s X_u dW_u\Big]^2 \Big| \mathcal{F}_s\Big) = \mathbb{E}\Big(\int_s^t X_u^2 du \Big| \mathcal{F}_s\Big) \quad a.s. \tag{10}$$

Next we define a larger class of processes, which is still reasonably nice. Its main property is that one can approximate processes in this new class by simple processes in an appropriate metric. The definition is as follows.

Definition 2.10. Define $\mathcal{L}_2^T$ to be the space of all progressively measurable processes $X = (X_t)_{t\in[0,T]}$, $X_t : \Omega \to \mathbb{R}$ such that

$$|X|_T = \mathbb{E}\Big(\int_0^T X_s^2 ds\Big)^{\frac{1}{2}} < \infty, \tag{11}$$

equipped with the metric

$$d(X, Y) := |X - Y|_T. \tag{12}$$

One then defines the stochastic integral of a process in $\mathcal{L}_2^T$ by the limit (in L^2)

$$\int_0^t X_s dW_s = \lim_{n\to\infty} \int_0^t X_s^{(n)} dW_s \tag{13}$$

Where $X^{(n)} \in \mathcal{L}_0^T$ is a sequence approximating X in L^2. This sequence exists, but we omit the proof, see [9, Proposition 3.1.11]. Furthermore one can choose the version of $\int_0^t X_s dW_s$ such that $\big(\int_0^t X_s dW_s\big)_{t\in[0,T]}$ is a continuous, L^2-martingale that starts identically at zero, see [9, Proposition 3.1.12].

One can extend the class $\mathcal{L}_2^T$ slightly in a way that integration still makes sense as follows

Definition 2.11. Define $\mathcal{L}_2^{T,loc}$ to be the class of all progressively measurable $X = (X_t)_{t\in[0,T]}$ such that

$$\mathbb{P}\Big(\Big\{\omega \in \Omega : \int_0^T X_s^2(\omega) ds < \infty\Big\}\Big) = 1. \tag{14}$$

The integral is defined by localizing the process in $\mathcal{L}_2^{T,loc}$ with an increasing sequence of stopping times such that the stopped processes are in $\mathcal{L}_2^T$.

Now we perform an extension of the stochastic integral to the case where the integrator is the d-dimensional Brownian motion. To this end we assume we have a d-dimensional $(\mathcal{F}_t)_{t\in[0,T]}$-Brownian motion W.

Definition 2.12. For a process $Y = (Y_t)_{t\in[0,T]}$, $Y_t : \Omega \to \mathbb{R}^{d\times d}$, such that for each $i, j = 1, \ldots, d$ it holds $Y_t^{ji} \in \mathcal{L}_2^T$, we define the stochastic integral of Y with respect to W to be

$$\int_0^t Y_s dW_s := \Big(\sum_{i=1}^d \int_0^t Y_t^{1i} dW_s^i, \ldots, \sum_{i=1}^d \int_0^t Y_t^{di} dW_s^i\Big), \quad t \in [0, T]. \tag{15}$$

2.3 Itô's formula

In this section we recall the classical Itô formula, which corresponds to the fundamental theorem of calculus, but for stochastic integrals. Assume that we have a continuous and adapted process $X = (X_t)_{t\in[0,T]}$ that can be represented as

$$X_t = x_0 + \int_0^t Y_s dW_s + \int_0^t Z_s ds,\ t \in [0,T],\ a.s., \tag{16}$$

for $Y \in \mathcal{L}_2^T$ and Z progressively measurable and integrable for all $\omega \in \Omega$. Furthermore we recall the definition of $C^{1,2}([0,T] \times \mathbb{R}; \mathbb{R})$, [16, Remark 5.4.1, p.312].

Definition 2.13. A continuous function $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ belongs to $C^{1,2}([0,T] \times \mathbb{R}; \mathbb{R})$ provided that the partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$ and $\frac{\partial^2 f}{\partial x^2}$ exist on $(0,T) \times \mathbb{R}$, are continuous and can be continuously extended to $[0,T] \times \mathbb{R}$.

Then we have the following:

Theorem 2.14. *For $f \in C^{1,2}([0,T] \times \mathbb{R}; \mathbb{R})$ and X as above one has*

$$\begin{aligned} f(t, X_t) - f(0, X_0) &= \int_0^t \frac{\partial f}{\partial s}(s, X_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)Y_s dB_s \\ &\quad + \int_0^t \frac{\partial f}{\partial x}(s, X_s)Z_s ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s)Y_s^2 ds, \\ &\textit{for} \quad t \in [0,T]\, a.s. \end{aligned} \tag{17}$$

Itô's formula extends also to $f \in C^{1,2}([0,T] \times \mathbb{R}^d; \mathbb{R})$ and processes X which have a representation with respect to $Y, Z : [0,T] \times \Omega \to \mathbb{R}^{d\times d}, \mathbb{R}^d$ and Y is coordinate-wise in $\mathcal{L}_2^T$ and Z is progressively measurable and integrable for every $\omega \in \Omega$, see for example [16, Theorem 3.3.6].

2.4 Classical SDEs

Next we want to make sense of the formal equation

$$dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt. \tag{18}$$

One has to interpret this as an integral equation because dW_t does not make sense as a derivative in the classical sense. This is because the set of ω for which the standard Brownian motion is nowhere differentiable contains a full measure set. However towards this goal we defined the stochastic integral with respect to the Brownian motion so we can define the following:

Definition 2.15. Assume we have a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ along with an $(\mathcal{F}_t)_{t\in[0,T]}$ Brownian motion $(W_t)_{t\in[0,T]}$. Assume $x_0 \in \mathbb{R}$ and $\sigma, b : [0,T] \times \mathbb{R} \to \mathbb{R}$ are continuous and bounded. A pathwise continuous and adapted stochastic process $X = (X_t)_{t\in[0,T]}$ is a strong solution of the SDE

$$dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt, \quad \text{with} \quad X_0 = x_0, \tag{19}$$

if $X_0 = x_0$ and

$$X_t = x_0 + \int_0^t \sigma(s, X_s)dW_s + \int_0^t b(s, X_s)ds \quad \text{for} \quad t \in [0,T] a.s. \tag{20}$$

Notice that for strong solutions we obtain an adapted solution given a specific stochastic basis. One has the classical existence result for strong solutions, see [16, Theorem 5.2.9]:

Proposition 2.16. *If σ, b are uniformly Lipschitz in the space coordinate, then there exists a strong solution to equation (19).*

It turns out one can also take another approach to finding a solution to equation (19). This is the concept of a weak solution.

Definition 2.17. Assume $\sigma, b : [0,T] \times \mathbb{R} \to \mathbb{R}$ are continuous and bounded. A weak solution of

$$dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt, \quad \text{with} \quad X_0 = x_0 \tag{21}$$

is a pair $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, $(X_t, \widetilde{W}_t)_{t\in[0,T]}$ such that the stochastic basis satisfies the usual conditions, the process X is continuous and $\mathcal{F}_t$-adapted, the process $(\widetilde{W}_t)_{t\in[0,T]}$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-Brownian motion and

$$X_t = x_0 + \int_0^t \sigma(s, X_s)d\widetilde{W}_s + \int_0^t b(s, X_s)ds \quad \text{for} \quad t \in [0,T] \quad a.s. \tag{22}$$

In this approach we are more free with the particular stochastic basis our solution process will be defined on. A straightforward verification shows that strong solutions are weak solutions so the terminology makes sense. These definitions also extend to the d-dimensional case, i.e. where $b, \sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d, \mathbb{R}^{d\times d}$.

Finally to end this section we recall the useful Burkholder-Davis-Gundy inequalities which are used in estimating norms related to integrals of $\mathcal{L}_2^{T,loc}$ processes with respect to the Brownian motion by square functions. See for example [16, Theorem 3.3.28].

Proposition 2.18. *Assume $p \in (0,\infty)$. There exists constants $a_p, b_p > 0$ such that for any $X = (X_t)_{t\in[0,T]} \in \mathcal{L}_2^{T,loc}$, it holds*

$$b_p \left\| \sqrt{\int_0^T X_t^2 dt} \right\|_p \leq \left\| \sup_{s\in[0,T]} \Big| \int_0^t X_s dB_s \Big| \right\|_p \leq a_p \left\| \sqrt{\int_0^T X_t^2 dt} \right\|_p \tag{23}$$

The bound $a_p \leq c\sqrt{p}$ for some absolute $c > 0$ can be obtained for $p \in [2,\infty)$. However we do not need this.

3 The Setting

In this section we will fix some notation.

We denote by $\mathcal{P}(\mathbb{R}^d)$ as the set of probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Further we denote the law of the random variable X by $\mathbb{P}_X$. Fix a time horizon $T > 0$ and consider a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ that satisfies the usual conditions on which we have a d-dimensional $(\mathcal{F}_t)_{t\in[0,T]}$ Brownian motion $W = (W_t)_{t\in[0,T]}$. We assume that W is independent from $\mathcal{F}_0$ such that

$$\mathcal{F}_t = \sigma(W_s : s \in [0,t]) \cup \mathcal{F}_0, \tag{24}$$

so the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is the natural filtration of W, augmented by $\mathcal{F}_0$. Furthermore we assume that for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ there exists a random variable $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ with $\mathbb{P}_\xi = \mu$.

The space $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ is defined in the usual way, equipped with the inner product $(\xi, \eta)_{L^2} = \mathbb{E}(\xi \cdot \eta)$ and the norm given by this inner product. Here $\cdot$ denotes the scalar product on $\mathbb{R}^d$. We identify two random variables in $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ if they are $\mathbb{P}$-a.s. equal.

We assume assumptions made in this section hold for the rest of the thesis.

4 The classical local martingale problem

In this section we will introduce the classical local martingale problem (for more details see [16, Section 5.4, p.311-319]). It will be of fundamental importance to the proof of our main theorem. We start by introducing the path space and then move to functional SDEs and the corresponding local martingale problems.

4.1 The Path space

In this section we define the space of continuous functions $C([0,T];\mathbb{R}^d)$ and equip it with a σ-algebra, and a compatible metric topology. Further we will define the coordinate process on this space that will give us a natural filtration on the space. The coordinate process also explains the name of the space. The space of continuous, $\mathbb{R}^d$-valued functions on $[0,T]$ is denoted by $C([0,T];\mathbb{R}^d)$. We want to give this space a Borel σ-algebra, but this depends on the topology given to the space. To this end we equip the space $C([0,T];\mathbb{R}^d)$ with the norm $\|x\| := \sup_{t\in[0,T]}|x(t)|$. This norm corresponds to the concept of uniform convergence. Now we also have a complete metric space, which yields us the topology given by this metric, that is we have the topology of open sets $\mathcal{T}$ given by the supremum norm. This allows us to define the Borel σ-algebra $\mathcal{B}(C([0,T];\mathbb{R}^d))$ to be the smallest σ-algebra that contains $\mathcal{T}$ i.e. $\mathcal{B}(C([0,T];\mathbb{R}^d)) = \sigma(\mathcal{T})$. This makes the path space a measurable space.

Next we want to define a certain process on the path space, known as the coordinate process. To this end we define the coordinate process $y = (y_t)_{t\in[0,T]}$ on $(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d)))$ by setting $y_t(\omega) := \omega(t)$ for any $\omega \in C([0,T];\mathbb{R}^d)$. We also want to equip this space with a filtration, we do this with the help of the coordinate process. To this end we define on $(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d)))$ the filtration $\widetilde{\mathbb{F}}^y = (\widetilde{\mathcal{F}}^y_t)_{t\in[0,T]}$ by $\widetilde{\mathcal{F}}^y_t := \sigma(y_s : s \le t)$. We say this filtration is the one generated by the coordinate process $y = (y_t)_{t\in[0,T]}$.

4.2 Functional SDEs and local martingale problems

We now turn to consider the equivalence of weak solution of a functional SDE and the solution of the corresponding local martingale problem. For this we assume that $b : [0,T] \times C([0,T];\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0,T] \times C([0,T];\mathbb{R}^d) \to \mathbb{R}^{d\times d}$ are continuous with respect to the product topology and non-anticipating, which is defined to mean that $b(t,f) = b(t, s \to f(s \wedge t))$. The definitions are as follows.

Definition 4.1. A weak solution of the functional SDE

$$X_t = \xi + \int_0^t b(s, X_{\cdot\wedge s})ds + \int_0^t \sigma(s, X_{\cdot\wedge s})d\widetilde{W}_s, \quad t \in [0,T], \tag{25}$$

is a six-tuple $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \widetilde{\mathbb{P}}, \widetilde{W}, X)$ such that the following is satisfied.

1. $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ is a complete probability space and $\widetilde{\mathbb{F}} = (\mathcal{F}_t)_{0\le t\le T}$ is a filtration on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ that satisfies the usual conditions, recall Definitions 2.1 and 2.2.

2. $X = (X_t)_{0\le t\le T}$ is a continuous $\mathbb{R}^d$-valued process that is adapted to $\widetilde{\mathbb{F}}$ and $\widetilde{W} = (\widetilde{W}_t)_{0\le t\le T}$ is a d-dimensional Brownian motion with respect to $(\widetilde{\mathbb{F}}, \widetilde{\mathbb{P}})$.

3. $\widetilde{\mathbb{P}}(\int_0^T (|b(s, X_{\cdot\wedge s})| + |\sigma(s, X_{\cdot\wedge s})|^2)ds < \infty) = 1$ and Equation (25) is satisfied $\widetilde{\mathbb{P}}$-a.s.

Definition 4.2. A solution to the local martingale problem associated with $\mathcal{A}'$ is a probability measure $\widehat{\mathbb{P}}$ on $(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d)))$ if for every $f \in C^{1,2}([0,T] \times \mathbb{R}^d; \mathbb{R})$ the process

$$M_t^f := f(t, y(t)) - f(0, y(0)) - \int_0^t (\partial_s + \mathcal{A}')f(s, y(s))ds, \quad t \in [0,T] \tag{26}$$

is a continuous local martingale with respect to $(\mathbb{F}^y, \widehat{\mathbb{P}})$ where $y = (y(t))_{t\in[0,T]}$ is the coordinate process on $C([0,T];\mathbb{R}^d)$, the filtration $\mathbb{F}^y$ is generated by y and augmented by the $\widehat{\mathbb{P}}$-null sets and made right-continuous, that is $\mathcal{F}_t^y = \bigcap_{s>t} \sigma(\widetilde{\mathcal{F}}_s^y \cup \widehat{\mathcal{N}})$, where $\widehat{\mathcal{N}} = \{A \subset C([0,T];\mathbb{R}^d)$: there exists $B \in \mathcal{B}(C([0,T];\mathbb{R}^d))$, such that $A \subset B$ and $\widehat{\mathbb{P}}(B) = 0\}$. Furthermore the second order differential operator $\mathcal{A}'$ is given by

$$\begin{aligned} \mathcal{A}'f(s,y) := &\sum_{i=1}^d b_i(s,y) f(s, y(s)) \\ &+ \frac{1}{2} \sum_{i,j,k=1}^d (\sigma_{ik}\sigma_{jk})(s,y)\partial^2_{x_i x_j} f(s, y(s)), \quad y \in C([0,T];\mathbb{R}^d). \end{aligned} \tag{27}$$

Now we can state the relevant Lemmas concerning the equivalence of the above concepts. See [16, p. 312-319].

Lemma 4.3. *The existence of a weak solution* $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \widetilde{\mathbb{P}}, \widetilde{W}, X)$ *to the functional SDE (25) with a given initial distribution* μ *on* $\mathcal{B}(C([0,T];\mathbb{R}^d))$ *(i.e. the law of* ξ *is* μ*) is equivalent to the existence of a solution* $\widehat{\mathbb{P}}$ *to the local martingale problem (26) associated with* $\mathcal{A}'$ *given by (27) and with* $\widehat{\mathbb{P}}_{y(0)} = \mu$*. The solutions are related by* $\widehat{\mathbb{P}} = \widetilde{\mathbb{P}} \circ X^{-1}$*.*

Lemma 4.4. *The uniqueness of the solution* $\widehat{\mathbb{P}}$ *of the local martingale problem (26) for a fixed initial distribution* $\widehat{\mathbb{P}}_{y(0)} = \mu$*, where* μ *is a probability measure on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$*, is equivalent to the uniqueness in law for the Equation (25) with* $\widetilde{\mathbb{P}}_{X_0} = \mu$*.*

To end the section, we introduce the concept of tightness and Prohorov's theorem. Recall that relative compactness for a family of probability measures means that every sequence of the elements of the family has a weakly convergent subsequence. For more details see [2, Theorems 5.1, 5.2].

Definition 4.5. A family $\mathcal{M}$ of probability measures on a metric measure space $(S, \mathcal{S})$ is tight if for every $\varepsilon > 0$ there exists a compact $K \subset S$ such that $\mathbb{P}(K) > 1 - \varepsilon$ for every $\mathbb{P} \in \mathcal{M}$.

In our setting the metric measure space will be the path space

$$(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d))). \tag{28}$$

Theorem 4.6. *Assume we have a family of probability measures* $\mathcal{M}$ *on a metric measure space* $(S, \mathcal{S})$*. If* $\mathcal{M}$ *is tight, then it is relatively compact. If the space* S *is separable and complete and* $\mathcal{M}$ *is relatively compact, then it is tight.*

What to note here is that the path space is separable and complete, so we have equivalence of tightness and relative compactness for families of probability measures on it.

5 The Wasserstein spaces

In this section we will consider the spaces of measures with a finite given moment. We make a standing assumption that $p \in [1, \infty)$. Furthermore we denote by $\mathcal{P}_p(\mathbb{R}^d)$ the space of probability measures μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $\int_{\mathbb{R}^d} |x|^p \mu(dx) < \infty$. We endow this space with the p-Wasserstein metric

$$W_p(\mu,\nu) := \inf \Bigg\{ \Big(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p \rho(dxdy) \Big)^{\frac{1}{p}}, \\ \rho \in \mathcal{P}_p(\mathbb{R}^{2d}) \, \text{with} \, \rho(\cdot \times \mathbb{R}^d) = \mu, \rho(\mathbb{R}^d \times \cdot) = \nu \Bigg\}, \tag{29}$$

for $\nu, \mu \in \mathcal{P}_p(\mathbb{R}^d)$.

First of all we should justify the fact that the p-Wasserstein metric actually is a metric as we call it. For this we have the following, see [25, Theorem 7.3]

Proposition 5.1. *W_p defines a metric on $\mathcal{P}_p(\mathbb{R}^d)$.*

We also have the useful monotonicity property for the moments $\widehat{W}_p(\mu) = \big(\int_{\mathbb{R}^d} |x|^p \mu(dx) \big)^{\frac{1}{p}}$:

$$\text{If} \quad 1 \le p \le q, \quad \text{then} \quad \widehat{W}_p \le \widehat{W}_q. \tag{30}$$

Later in this article we will consider the case $p = 2$. We also note the following estimate that will be used later on:

Remark 5.2. $W_2(\mathbb{P}_\xi, \mathbb{P}_\zeta) \le \mathbb{E}(|\xi - \zeta|^2)^{\frac{1}{2}}$ whenever $\xi, \zeta \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$.

This remark follows straightforwardly from the definition of the 2-Wasserstein metric as long as our σ-algebra $\mathcal{F}_0$ is rich enough to support random variables with laws in $\mathcal{P}_2(\mathbb{R}^d)$, which we assumed it to be, see [4, Proof of Lemma 3.1].

For our future interest we will also formulate the following lemma regarding the compactness of a specific set in $\mathcal{P}_2(\mathbb{R}^d)$:

Lemma 5.3. *Consider for a fixed $C > 0$ the set $\mathcal{E} \subset \mathcal{P}_2(\mathbb{R}^d)$ defined as*

$$\mathcal{E} = \{\mu \in \mathcal{P}_2(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^4 \mu(dx) \le C\}. \tag{31}$$

Then $\mathcal{E}$ is compact.

Proof. First we set $B_r^c := \mathbb{R}^d \setminus B(0,r)$. We will use the result that a subset $\mathcal{M} \subset \mathcal{P}_2(\mathbb{R}^d)$ is relatively compact if and only if

$$\limsup_{r\to\infty} \sup_{\mu\in\mathcal{M}} \int_{B_r^c} |x|^2 \mu(dx) = 0. \tag{32}$$

For this result, see [25, Theorem 7.12]. Using this we can estimate using Hölder's inequality for arbitrary $r > 0$ and $\nu \in \mathcal{E}$

$$\int_{B_r^c} |x|^2 \nu(dx) \leq \nu(B_r^c)^{\frac{1}{2}} \Big(\int_{B_r^c} |x|^4 \nu(dx) \Big)^{\frac{1}{2}} \leq C^{\frac{1}{2}} \nu(B_r^c)^{\frac{1}{2}}. \tag{33}$$

Now we estimate $\nu(B_r^c)$.

$$\nu(B_r^c) = \int_{B_r^c} d\nu = \frac{1}{r^4} \int_{B_r^c} r^4 d\nu \leq \frac{1}{r^4} C. \tag{34}$$

Continuing from (33) we get

$$\int_{B_r^c} |x|^2 \nu(dx) \leq C^{\frac{1}{2}} \Big(\frac{1}{r^4} \Big)^{\frac{1}{2}} C^{\frac{1}{2}} = \frac{C}{r^2} \to 0, \tag{35}$$

as $r \to \infty$. We also have

$$\sup_{\nu\in\mathcal{E}} \int_{B_r^c} |x|^2 \nu(dx) \leq \frac{C}{r^2} \to 0, \tag{36}$$

as $r \to \infty$. This proves the relative compactness of $\mathcal{E}$. We get compactness by showing that $\mathcal{E}$ is closed. This follows since the limit μ of a weakly convergent sequence $(\mu_k)_{k\in\mathbb{N}}$ in $\mathcal{E}$ is still an element of $\mathcal{E}$. □

We will later use the 2-Wasserstein space of probability measures on the path space. This is a straightforward extension where one changes the space $\mathbb{R}^d$ to $C([0,T];\mathbb{R}^d)$ and thus one integrates over $C([0,T];\mathbb{R}^d)$ and instead of the Euclidean norm on $\mathbb{R}^d$ we consider the supremum norm on $C([0,T];\mathbb{R}^d)$. Precisely we let $\mathcal{P}_p(C([0,T];\mathbb{R}^d))$ to be the space of probability measures μ on $(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d)))$ such that $\int_{C([0,T];\mathbb{R}^d)} \|y\|^p \mu(dy) < \infty$. This space is further endowed with the p-Wasserstein metric

$$\begin{aligned} W_p(\mu,\nu) := \inf \Big\{ \Big(& \int_{C([0,T];\mathbb{R}^d)\times C([0,T];\mathbb{R}^d)} \|x-y\|^p \rho(dxdy) \Big)^{\frac{1}{p}}, \\ & \rho \in \mathcal{P}_p(C([0,T];\mathbb{R}^d) \times C([0,T];\mathbb{R}^d)) \text{ with } \rho(\cdot \times C([0,T];\mathbb{R}^d)) = \mu, \\ & \rho(C([0,T];\mathbb{R}^d) \times \cdot) = \nu \Big\}. \end{aligned} \tag{37}$$

6 Measure derivatives

In this section we consider what it means to take derivatives with respect to a measure variable. We also introduce a version of Itô's formula for the law dependant case. We consider measure derivatives via a lifting to L^2 and using the Fréchet differentiability structure in L^2. To this end we recall the notion of a Fréchet derivative. In this section we follow the framework of Buckdahn et al., for more details refer to [4, Section 2].

Definition 6.1. We say that a function $\tilde{f} : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \to \mathbb{R}$ is Fréchet differentiable at $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$, if there exists a continuous linear map $D\tilde{f}(\xi) : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \to \mathbb{R}$ (notice ξ is not the input of the function), such that $\tilde{f}(\xi+\eta) - \tilde{f}(\xi) = D\tilde{f}(\xi)(\eta) + o(|\eta|_{L^2})$ where $\eta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ is such that $|\eta|_{L^2} \to 0$.

The notation o means that $|\tilde{f}(\xi+\eta) - \tilde{f}(\xi) - D\tilde{(f)}(\xi)(\eta)|_{L^2} \le \varepsilon|\eta|_{L^2}$ for any $\varepsilon > 0$ as long as $|\eta|_{L^2}$ is sufficiently small. Using Definition 6.1 we define the derivative of a function $f : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$, see [5, Definition 6.1]:

Definition 6.2. We say that a function $f : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is differentiable at a probability measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, if for the function $\tilde{f} : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \to \mathbb{R}$ defined by $\tilde{f}(\xi) := f(\mathbb{P}_\xi)$ there exists a $\zeta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ with $\mathbb{P}_\zeta = \mu$ and such that $\tilde{f}$ is Fréchet differentiable at ζ.

This definition explains what we mean by lifting the map to L^2. Now we can use the Riesz representation theorem in the Hilbert space L^2 to find a $\mathbb{P}$-a.s. unique random variable $\gamma \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ such that $D\tilde{f}(\zeta)(\eta) = \mathbb{E}(\gamma \cdot \eta)$ for all $\eta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$. For this random variable γ it was shown by Lions, see [5, Section 6.1], that there exists a Borel function $g : \mathbb{R}^d \to \mathbb{R}^d$ such that $\gamma = g(\zeta)$ $\mathbb{P}$-a.s. and g only depends on ζ via it's law $\mathbb{P}_\zeta$. Based on the above we have

$$\begin{aligned} f(\mathbb{P}_\xi) - f(\mathbb{P}_\zeta) &= \tilde{f}(\xi) - \tilde{f}(\zeta) \\ &= D\tilde{f}(\zeta)(\xi - \zeta) + o(|\xi - \zeta|_{L^2}) \\ &= \mathbb{E}(g(\zeta) \cdot (\xi - \zeta)) + o(|\xi - \zeta|_{L^2}), \xi \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d). \end{aligned} \tag{38}$$

Definition 6.3. We call the function $\partial_\mu f(\mathbb{P}_\zeta, y) := g(y)$, $y \in \mathbb{R}^d$, the derivative of the function $f : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ at $\mathbb{P}_\zeta$, $\zeta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$.

Remark 6.4. Notice that $\partial_\mu f(\mathbb{P}_\zeta, y)$ is $\mathbb{P}_\zeta(dy)$-a.e. uniquely determined.

With this notion of differentiation now in hand we can define the classes of continuously differentiable functions and related notions that are needed in the following. We will first define the class of C^1-functions on $\mathcal{P}_2(\mathbb{R}^d)$, we will then use this to define the subspaces of higher orders of differentiability.

Definition 6.5. 1. We say that a function $f : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is of class $C^1(\mathcal{P}_2(\mathbb{R}^d))$, which we denote by $f \in C^1(\mathcal{P}_2(\mathbb{R}^d))$, if for all $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ there exists a $\mathbb{P}_\xi$-modification of $\partial_\mu f(\mathbb{P}_\xi, \cdot)$ (which we denote also by $\partial_\mu f(\mathbb{P}_\xi, \cdot)$) such that $\partial_\mu f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ is continuous with respect to the product topology of the 2-Wasserstein topology on $\mathcal{P}_2(\mathbb{R}^d)$ and the standard Euclidean topology on $\mathbb{R}^d$. This modified function is identified as the derivative of f.

2. The function f is said to be of class $C_b^{1,1}(\mathcal{P}_2(\mathbb{R}^d))$, if $f \in C^1(\mathcal{P}_2(\mathbb{R}^d))$ and $\partial_\mu f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ is bounded and Lipschitz continuous (again with respect to the product topology, where we assume this is the one given by the sum of the two metrics).

Comparing with the remark earlier we have that $\partial_\mu f(\mathbb{P}_\xi, \cdot)$ is unique. Further we denote $\partial_\mu f(\mu, x) = ((\partial_\mu f)_i(\mu, x))_{1 \le i \le d}$.

Definition 6.6. 1. We say that a function $f : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is of class $C^2(\mathcal{P}_2(\mathbb{R}^d))$ if $f \in C^1(\mathcal{P}_2(\mathbb{R}^d))$ is such that $\partial_\mu f(\mu, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is differentiable for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and it's derivative $\partial_y \partial_\mu f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is continuous and jointly measurable.

2. The function f is said to be of class $C_b^{2,1}$ given $f \in C^2(\mathcal{P}_2(\mathbb{R}^d)) \cap C_b^{1,1}(\mathcal{P}_2(\mathbb{R}^d))$ and the derivative $\partial_y \partial_\mu f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is bounded and Lipschitz continuous.

Now we can use the above definitions to consider f defined on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, in other words to the case where f depends also on a temporal and a spatial variable.

Remark 6.7. For the following definition we assume that f, along with all its appropriate derivatives are jointly measurable in all three variables.

Definition 6.8. 1. We say that a function $f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is of class $C^2(\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ if the following holds:

- $f(x, \cdot) \in C^2(\mathcal{P}_2(\mathbb{R}^d))$ for all $x \in \mathbb{R}^d$ and $f(\cdot, \mu) \in C^2(\mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.
- All derivatives $\partial_{x_k} f, \partial^2_{x_k x_l} f$ and $\partial_\mu f, \partial_{y_k} \partial_\mu f$, $1 \le k, l \le d$ are continuous over $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ and $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, respectively.

Furthermore we say that f is of class $C_b^{2,1}(\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ if $f \in C^2(\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ and all the derivatives are bounded and Lipschitz continuous.

2. We say that a function $f : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is of class $C^{1,2}([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ if $f(\cdot, x, \mu) \in C^1([0,T])$, for all $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ and $f(t, \cdot, \cdot) \in C^2(\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ for all $t \in [0,T]$.

3. Finally we say that f is of class $C_b^{1,2,1}([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d);\mathbb{R})$ if $f\in C^{1,2}([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d);\mathbb{R})$ and all the derivatives are uniformly bounded over $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ and Lipschitz in (x,μ,y) uniformly with respect to t.

We now finish the section by extending Itô's formula to the measure dependant case. To this end we need to introduce some notations. We denote $(\bar{\Omega},\bar{\mathcal{F}},\bar{\mathbb{P}})\otimes(\Omega,\mathcal{F},\mathbb{P})$ to be the product of $(\Omega,\mathcal{F},\mathbb{P})$ with itself. For a random variable ξ on $(\Omega,\mathcal{F},\mathbb{P})$ we denote by $\bar{\xi}$ it's copy over $(\bar{\Omega},\bar{\mathcal{F}},\bar{\mathbb{P}})$. Furthermore the expectation $\bar{\mathbb{E}}(\cdot)=\int_{\bar{\Omega}}(\cdot)d\bar{\mathbb{P}}$ only acts on random variables with a bar. This formalism extends to stochastic processes with the extension that $(\bar{\xi}_s)_{s\geq 0}$ denotes the copy process on $(\bar{\Omega},\bar{\mathcal{F}},\bar{\mathbb{P}})$. Note that the copy random variable and process share laws with the original random variable and process.

Theorem 6.9. *Assume $\sigma=(\sigma_s)$, $\gamma=(\gamma_s)$ are $\mathbb{R}^{d\times d}$-valued and $b=(b_s)$, $\beta=(\beta_s)$ are $\mathbb{R}^d$-valued progressively measurable stochastic processes such that the following holds:*

1. *There exists a constant $q>6$ such that $\mathbb{E}[(\int_0^T(|\sigma_s|^q+|b_s|^q)ds)^{\frac{3}{q}}]<\infty$.*
2. *$\int_0^T(|\gamma_s|^2+|\beta_s|^2)ds<\infty$ $\mathbb{P}$-a.s.*

Further assume $F\in C_b^{1,2,1}([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$. Then for the Itô-processes

$$X_t=X_0+\int_0^t\sigma_s dW_s+\int_0^t b_s ds,\quad t\in[0,T],\ X_0\in L^2(\Omega,\mathcal{F}_0,\mathbb{P}) \tag{39}$$

$$Y_t=Y_0+\int_0^t\gamma_s dW_s+\int_0^t \beta_s ds,\quad t\in[0,T],\ Y_0\in L^2(\Omega,\mathcal{F}_0,\mathbb{P}), \tag{40}$$

it holds that

$$\begin{aligned}&F(t,Y_t,\mathbb{P}_{X_t})-F(0,Y_0,\mathbb{P}_{X_0})\\&=\int_0^t\Bigg(\partial_r F(r,Y_r,\mathbb{P}_{X_r})+\sum_{i=1}^d\partial_{y_i}F(r,Y_r,\mathbb{P}_{X_r})\beta_r^i+\frac{1}{2}\sum_{i,j,k=1}^d\partial^2_{y_iy_j}F(r,Y_r,\mathbb{P}_{X_r})\gamma_r^{ik}\gamma_r^{jk}\\&+\bar{\mathbb{E}}\left[\sum_{i=1}^d(\partial_\mu F)_i(r,Y_r,\mathbb{P}_{X_r},\bar{X}_r)\bar{b}_r^{\,i}+\frac{1}{2}\sum_{i,j,k=1}^d\partial_{z_i}(\partial_\mu F)_j(r,Y_r,\mathbb{P}_{X_r},\bar{X}_r)\bar{\sigma}_r^{ik}\bar{\sigma}_r^{jk}\right]\Bigg)dr\\&+\int_0^t\sum_{i,j=1}^d\partial_{y_i}F(r,Y_r,\mathbb{P}_{X_r})\gamma_r^{ij}dW_r^j,\quad t\in[0,T]\quad a.s.\end{aligned} \tag{41}$$

We omit the proof (the reader can consult for example [19, Appendix]).

7 Mean-Field SDEs

We assume that $b : [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d}$ are continuous and bounded throughout this section. Formally we are looking for a weak solution of the following mean-field SDE:

$$X_t = \xi + \int_0^t b(s, X_s, \mathbb{Q}_{X_s})ds + \int_0^t \sigma(s, X_s, \mathbb{Q}_{X_s})dW_s, \quad t \in [0,T], \tag{42}$$

where $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ obeys a given law $\mathbb{Q}_\xi = \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $(W_t)_{t\in[0,T]}$ is a d-dimensional Brownian motion with respect to $\mathbb{Q}$.

Two other questions can be asked:

- Does uniqueness hold for the mean-field SDE under the conditions of boundedness and continuity on the coefficients?
- Can we extend the result to coefficients b, σ defined on $[0,T]\times C([0,T];\mathbb{R}^d)\times \mathcal{P}_2(C([0,T];\mathbb{R}^d))$? I.e. the coefficients are path-dependant.

The answer to both of these is yes, but we will not further explore these. Both of these are considered in [19].

Next we define what is meant by a weak solution of Equation (42) and a solution of the corresponding local martingale problem.

Definition 7.1. A six-tuple $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \mathbb{Q}, W, X)$ is called a weak solution of the mean-field SDE (42), given the following conditions are satisfied

1. $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \mathbb{Q})$ is a stochastic basis that satisfies the usual conditions (recall Definitions 2.1 and 2.2).
2. $X = (X_t)_{t\in[0,T]}$ is an $\mathbb{R}^d$-valued continuous process that is adapted to $\widetilde{\mathbb{F}}$ and $W = (W_t)_{t\in[0,T]}$ is a d-dimensional Brownian motion with respect to $(\widetilde{\mathbb{F}}, \mathbb{Q})$.
3. Equation (42) is satisfied $\mathbb{Q}$-almost surely.

Definition 7.2. A probability measure $\widehat{\mathbb{P}}$ is a solution of the local martingale problem associated with the operator $\widetilde{\mathcal{A}}$ if for every $f \in C_b^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R})$ the process

$$C^f(t,y,\mu) := f(t,y(t)) - f(0,y(0)) - \int_0^t ((\partial_s + \widetilde{\mathcal{A}})f)(s,y(s),\mu(s))ds, \quad t\in[0,T] \tag{43}$$

is a continuous local $(\mathbb{F}^y, \widehat{\mathbb{P}})$-martingale, where $\mu(t) = \widehat{\mathbb{P}}_{y(t)}$ is the law of the coordinate process $y(t)$ on $C([0,T];\mathbb{R}^d)$ at time t, the filtration $\mathbb{F}^y$ is generated by the coordinate process y, completed with the $\widehat{\mathbb{P}}$-null sets and made right-continuous, see 4.2. The (second order differential) operator $\widetilde{\mathcal{A}}$ is defined by

$$(\widetilde{\mathcal{A}}f)(s,y,\nu) := \sum_{i=1}^{d} \partial_{y_i} f(s,y) b_i(s,y,\nu) + \frac{1}{2} \sum_{i,j,k=1}^{d} \partial^2_{y_i y_j} f(s,y)(\sigma_{ik}\sigma_{jk})(s,y,\nu). \tag{44}$$

Here $\partial_s + \widetilde{\mathcal{A}}$ denotes the pointwise sum of the operators ∂_s and $\widetilde{\mathcal{A}}$ (notice $\partial_s f$ does not depend on the measure ν).

The local martingale problem and it's solution above extend the case where the coefficients b_i and σ_{ik} only depend on (s,y) to the case where they also depend on the probability measure $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. With the above extension we can also extend Lemma 4.3:

Lemma 7.3. *The existence of a weak solution $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \mathbb{Q}, B, X)$ of Equation (42) with given initial distribution ν on $\mathcal{B}(\mathbb{R}^d)$ is equivalent to the existence of a solution $\widehat{\mathbb{P}}$ of the local martingale problem associated with $\widetilde{\mathcal{A}}$ given by Definition 7.2 with $\widehat{\mathbb{P}}_{y(0)} = \nu$.*

Proof. (a) We start with the sufficiency by assuming that we have a solution $\widehat{\mathbb{P}}$ on $(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d)))$ of the local martingale problem associated with $\widetilde{\mathcal{A}}$. We then define the coefficients $\tilde{b}(s,x) = b(s,x,\widehat{\mathbb{P}}_{y(s)})$ and $\tilde{\sigma}(s,x) = \sigma(s,x,\widehat{\mathbb{P}}_{y(s)})$. For these coefficients and all $f \in C^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R})$ the operator $\widetilde{\mathcal{A}}$ is given by

$$\widetilde{\mathcal{A}}f(s,y) = \sum_{i=1}^{d} \partial_{y_i} f(s,y) \tilde{b}_i(s,x) + \frac{1}{2} \sum_{i,j,k=1}^{d} \partial^2_{y_i y_j} f(s,y)(\tilde{\sigma}_{ik}\tilde{\sigma}_{jk})(s,y) \tag{45}$$

and we have that

$$M_t^f := f(t,y(t)) - f(0,y(0)) - \int_0^t (\partial_s + \widetilde{\mathcal{A}}) f(s,y(s))ds, \quad t \in [0,T], \tag{46}$$

is a continuous local martingale with respect to $(\mathbb{F}^y, \widehat{\mathbb{P}})$. Now we notice that $\widehat{\mathbb{P}}$ is a solution of the classical local martingale problem given in Definition 4.2, and thus we can invoke Lemma 4.3 to get a weak solution $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \mathbb{Q}, W, X)$ of the SDE

$$X_t = X_0 + \int_0^t \tilde{b}(s,X_s)ds + \int_0^t \tilde{\sigma}(s,X_s)dW_s, \quad \in [0,T]. \tag{47}$$

We also notice (see [16, Proposition 4.6, p. 315]) that $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \mathbb{Q})$ can be chosen as an extension of the space

$$(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}) := (C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d)), \widehat{\mathbb{P}}). \tag{48}$$

This is done in the following way: For a suitable probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{F}} = (\bar{\mathcal{F}}_t), \bar{\mathbb{P}})$, on which a d-dimensional $(\bar{\mathbb{F}}, \bar{\mathbb{P}})$-Brownian motion is defined, $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \mathbb{Q})$ is the completed product space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \mathbb{Q}) = (\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}) \bigotimes (\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$, equipped with the smallest right-continuous and augmented filtration $\widetilde{\mathbb{F}} = (\widetilde{\mathcal{F}}_t)_{t\in[0,T]}$, for which $\widehat{\mathcal{F}}_t \bigotimes \bar{\mathcal{F}}_t \subset \widetilde{\mathcal{F}}_t$ for all $t \in [0,T]$ (see the Appendix). Furthermore we extend every process Z adapted to $\widehat{\mathbb{F}}$ and defined on $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ to a process $\widetilde{Z}$ that is defined on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \mathbb{Q})$ and adapted to $\widetilde{\mathbb{F}}$ by setting $\widetilde{Z}_t(\widehat{\omega}, \bar{\omega}) = Z_t(\widehat{\omega})$, where $(\widehat{\omega}, \bar{\omega}) \in \widetilde{\Omega}, t \in [0,T]$. Now making the observation that $\mathbb{Q}(A \times \bar{\Omega}) = \widehat{\mathbb{P}}(A)$ holds for all $A \in \widehat{\mathcal{F}}$ we get for all $\Gamma \in \mathcal{B}(\mathbb{R}^d)$, $s \in [0,T]$,

$$\begin{aligned} \mathbb{Q}_{\widetilde{Z}_s}(\Gamma) &= \mathbb{Q}((\widehat{\omega}, \bar{\omega}) \in \widehat{\Omega} \times \bar{\Omega} : \widetilde{Z}_s(\widehat{\omega}, \bar{\omega}) \in \Gamma) \\ &= \mathbb{Q}(\{\widehat{\omega} \in \widehat{\Omega} : Z_s(\widehat{\omega}) \in \Gamma\} \times \bar{\Omega}) \\ &= \widehat{\mathbb{P}}(\widehat{\omega} \in \widehat{\Omega} : Z_s(\widehat{\omega}) \in \Gamma) = \widehat{\mathbb{P}}_{Z_s}(\Gamma). \end{aligned} \tag{49}$$

Further we note that $X = (X_t)_{t\in[0,T]}$ can be chosen as the extension of the coordinate process $y = (y(t))_{t\in[0,T]}$, in other words $X_t(\widehat{\omega}, \bar{\omega}) = y(t, \widehat{\omega}) = \widehat{\omega}(t)$, $t \in [0,T]$.

Combining Equations (47) and (49) we obtain

$$\begin{aligned} X_t &= X_0 + \int_0^t b(s, x, \widehat{\mathbb{P}}_{y(s)})ds + \int_0^t \sigma(s, X_s, \widehat{\mathbb{P}}_{y(s)})dW_s \\ &= X_0 + \int_0^t b(s, x, \mathbb{Q}_{X_s})ds + \int_0^t \sigma(s, X_s, \mathbb{Q}_{X_s})dW_s, \quad t \in [0,T]. \end{aligned} \tag{50}$$

This gives that $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \mathbb{Q}, W, X)$ is a weak solution of the mean-field SDE (42).

(b) We then proceed to consider the necessity. To this end we assume that $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \mathbb{Q}, W, X)$ is a weak solution of the mean-field SDE

$$X_t = X_0 + \int_0^t b(s, X_s, \mathbb{Q}_{X_s})ds + \int_0^t \sigma(s, X_s, \mathbb{Q}_{X_s})dW_s, \quad t \in [0,T], \tag{51}$$

where for $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$ it holds $\mathbb{Q}_{X_0} = \nu$ and $W = (W_t)_{t\in[0,T]}$ is a d-dimensional $(\widetilde{\mathbb{F}}, \mathbb{Q})$-Brownian motion.

We will show $(C^f(t, y, \mu))_{t\in[0,T]}$ is a continuous local $(\mathbb{F}^y, \mathbb{Q}_X)$-martingale. Here $\mathbb{Q}_X$ denotes the law of X with respect to the probability measure $\mathbb{Q}$.

We first fix the law $\mathbb{Q}_{X_s}$. Further defining $\bar{b}(s,x) = b(s,x,\mathbb{Q}_{X_s})$ and similarly $\bar{\sigma}(s,x) = \sigma(s,x,\mathbb{Q}_{X_s})$ we have from (51) the solution $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \mathbb{Q}, B, X)$ of the classical SDE

$$X_t = X_0 + \int_0^t \bar{b}(s,X_s)ds + \int_0^t \bar{\sigma}(s,X_s)dB_s, \quad t \in [0,T],$$

and then the classical local martingale problem gives us a probability measure $\widehat{\mathbb{P}}$ on $(C([0,T];\mathbb{R}^d), \mathcal{B}(([0,T];\mathbb{R}^d)))$ such that $\widehat{\mathbb{P}}_{y(0)} = \nu$ and $\widehat{\mathbb{P}} = \mathbb{Q}_{X_s}$, and such that $M_t^f = f(t,y(t)) - f(0,y(0)) - \int_0^t (\partial_s + \mathcal{A}')f(s,y(s))ds$ is a continuous local $(\mathbb{F}^y, \widehat{\mathbb{P}})$-martingale. Recalling the definition of the classical differential operator we have $\mathcal{A}'f(s,y) = \widetilde{\mathcal{A}}f(s,y,\mathbb{Q}_y)$; so that $M_t^f = C^f(t,y,\mathbb{Q})$ is a continuous local martingale. This finishes the proof. □

Using the extension given in Lemma 7.3 we can prove the following Lemma:

Lemma 7.4. *Let the probability $\widehat{\mathbb{P}}$ on $(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d)))$ be a solution of the local martingale problem associated with $\widetilde{\mathcal{A}}$ given in Definition 7.2. Then for the operator $\mathcal{A}$ applied to functions $f \in C^{1,2,1}([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d);\mathbb{R})$ which is given by*

$$\begin{aligned}(\mathcal{A}f)(s,y,\nu) := (\widetilde{\mathcal{A}}f)(s,y,\nu) + \sum_{i=1}^d \int_{\mathbb{R}^d} (\partial_\mu f)_i(s,y,\nu,z) b_i(s,z,\nu)\nu(dz) \\ + \frac{1}{2}\sum_{i,j,k=1}^d \int_{\mathbb{R}^d} \partial_{z_i}(\partial_\mu f)_j(s,y,\nu,z)(\sigma_{ik}\sigma_{jk})(s,z,\nu)\nu(dz),\end{aligned} \tag{52}$$

we have that for every $f \in C^{1,2,1}([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d);\mathbb{R})$ the process

$$\begin{aligned}C^f(t,y,\mu) &:= f(t,y(t),\mu(t)) - f(0,y(0),\mu(0)) \\ &- \int_0^t ((\partial_s + \mathcal{A})f)(s,y(s),\mu(s))ds, \quad t \in [0,T],\end{aligned} \tag{53}$$

is a continuous local martingale with respect to $(\mathbb{F}^y, \widehat{\mathbb{P}})$, where $\mu(t) = \widehat{\mathbb{P}}_{y(t)}$ is the law of the coordinate process on $C([0,T];\mathbb{R}^d)$ at time t and the filtration $\mathbb{F}^y$ is generated by y, completed and made right-continuous. Further, if $f \in C_b^{1,2,1}([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d);\mathbb{R})$, then the process C^f is a martingale with respect to $(\mathbb{F}^y, \widehat{\mathbb{P}})$.

Proof. Assume we are given the solution $\widehat{\mathbb{P}}$ of the local martingale problem associated with $\widetilde{\mathcal{A}}$. We then have by Lemma 7.3 a weak solution

$(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \mathbb{Q}, B, X)$ to the SDE

$$X_t = X_0 + \int_0^t b(s, X_s, \mathbb{Q}_{X_s})ds + \int_0^t \sigma(s, X_s, \mathbb{Q}_{X_s})dW_s, \quad t \in [0, T], \quad (54)$$

where $\mathbb{Q}_{X_s} = \widehat{\mathbb{P}}$. Now for an arbitrary $f \in C^{1,2,1}([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d); \mathbb{R})$, the necessity part of the proof of the previous lemma gives precisely the argument that shows C^f in (53) is a continuous local $(\mathbb{F}^y, \widehat{\mathbb{P}})$-martingale.

Notice the same argument works, since the C^f changes according to whether we have a function f that changes in the measure component, Itô's formula Theorem 6.9 gives two extra terms, but they get absorbed by the extension of the operator $\widetilde{\mathcal{A}}$. □

7.1 The Existence Theorem

Now we have the required ingredients from the prerequisites and Lemmas 7.3 and 7.4 to prove the main theorem of the thesis regarding the existence of a weak solution to Equation (42):

Theorem 7.5. *There exists a weak solution $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \widetilde{\mathbb{Q}}, B, X)$ of the mean-field SDE (42).*

Recall that we assume b, σ are continuous and bounded where boundedness and continuity in the measure variable is understood with respect to the 2-Wasserstein metric.

Proof. We prove this by showing that the local martingale problem has a solution and then invoking Lemma 7.3. To this end we partition the interval $[0, T]$ in the following way. For each $n \in \mathbb{N}$ let $t_i^n = iT2^{-n}$, where $0 \leq i \leq 2^n$, then let $P_n := \{t_0^n, \dots, t_{2^n}^n\}$. We then define for any $y \in C([0, T]; \mathbb{R}^d)$ and $\mu \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d))$ the non-anticipating functionals $b^{(n)}(s, y, \mu) = b(s, y(t_i^n), \mu(t_i^n))$, $\sigma^{(n)}(s, y, \mu) = \sigma(s, y(t_i^n), \mu(t_i^n))$, where $s \in (t_i^n, t_{i+1}^n]$ and $0 \leq i \leq 2^n$.

Now let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a stochastic basis and let $W = (W_t)_{t \in [0,T]}$ be a d-dimensional $(\mathbb{F}, \mathbb{P})$-Brownian motion. Furthermore let $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ be a random variable with the law $\mathbb{P}_\xi = \nu$.

We define now, for $n \geq 1$, the continuous, $\mathbb{F}$-adapted and unique process $X^{(n)} = (X_t^{(n)})_{t \in [0,T]}$ by the Euler scheme

$$\begin{aligned} X_t^{(n)} &= X_{t_i^n}^{(n)} + \int_{t_i^n}^t b(s, X_{t_i^n}^{(n)}, \mathbb{P}_{X_{t_i^n}^{(n)}}) ds + \int_{t_i^n}^t \sigma(s, X_{t_i^n}^{(n)}, \mathbb{P}_{X_{t_i^n}^{(n)}}) dW_s \\ X_0^{(n)} &= \xi, \qquad\qquad t \in (t_i^n, t_{i+1}^n],\ 0 \leq i \leq 2^{n-1}. \end{aligned} \tag{55}$$

We note that $X^{(n)}$ then solves the SDE

$$X_t^{(n)} = \xi + \int_0^t b^{(n)}(s, X^{(n)}, \mathbb{P}_{X^{(n)}}) ds + \int_0^t \sigma^{(n)}(s, X^{(n)}, \mathbb{P}_{X^{(n)}}) dW_s, \quad 0 \leq t \leq T. \tag{56}$$

Next we want to prove tightness for the laws of the processes $X^{(n)}$. To this end we will next prove the inequality

$$\sup_{n \geq 1} \mathbb{E}[|X_t^{(n)} - X_s^{(n)}|^{2m}] \leq C_{m,T,b,\sigma} |t - s|^m, \quad m \geq 1, 0 \leq s, t \leq T, \tag{57}$$

where the constant $C_{m,T,b,\sigma}$ does not depend on the process $X^{(n)}$. To this end assume $s < t$, and recall that b, σ are bounded. We now compute, using equation (56), the inequality $(a+b)^p \leq c_p(a^p + b^p)$, valid for $a, b \geq 0, p \geq 1$, and the Burkholder-Davis-Gundy inequality (recall Proposition 2.18), that

$$\begin{aligned}
&\mathbb{E}[|X_t^{(n)} - X_s^{(n)}|^{2m}] \\
&= \mathbb{E}\left[\Big|\int_s^t b^{(n)}(r, X^{(n)}, \mathbb{P}_{X^{(n)}})dr + \int_s^t \sigma^{(n)}(r, X^{(n)}, \mathbb{P}_{X^{(n)}})dW_r\Big|^{2m}\right] \\
&\leq \mathbb{E}\left[c_{2m}\Big(\int_s^t \big|b^{(n)}(r, X^{(n)}, \mathbb{P}_{X^{(n)}})\big|dr\Big)^{2m} + c_{2m}\Big(\int_s^t \big|\sigma^{(n)}(r, X^{(n)}, \mathbb{P}_{X^{(n)}})\big|dW_r\Big)^{2m}\right] \\
&\leq c_{2m}c_b^{2m}|t-s|^{2m} + c_{2m}\mathbb{E}\left[\Big(\int_s^t c_\sigma dW_r\Big)^{2m}\right] \\
&\leq c_{2m}c_b T|t-s|^m + c_{2m}c_\sigma^{2m}\sqrt{2m}|t-s|^m \\
&= C_{m,T,b,\sigma}|t-s|^m,
\end{aligned} \tag{58}$$

which is what was desired.

We will next verify that for any $\varepsilon > 0$ and any $t \in [0,T]$ it holds

$$\frac{1}{\delta}\sup_{n\geq 1}\mathbb{P}\big(\sup_{t\leq s\leq t+\delta}|X_t^{(n)} - X_s^{(n)}| \geq \varepsilon\big) = 0. \tag{59}$$

To this end, fix $\varepsilon > 0$ and let $n \geq 1, \delta > 0$ be arbitrary. We estimate using Markov's inequality

$$\begin{aligned}
&\mathbb{P}(|X_t^{(n)} - X_s^{(n)}| \geq \varepsilon) \leq \frac{\mathbb{E}[|X_t^{(n)} - X_s^{(n)}|^{2m}]}{\varepsilon^{2m}} \\
&\leq \sup_{n\geq 1}\frac{\mathbb{E}[|X_t^{(n)} - X_s^{(n)}|^{2m}]}{\varepsilon^{2m}} \leq \frac{C_m}{\varepsilon^{2m}}|t-s|^m.
\end{aligned} \tag{60}$$

We will now perform a similar estimate on $\mathbb{E}\sup_{t\leq s\leq t+\delta}|X_t^{(n)} - X_s^{(n)}|^{2m}$. First of all we have $\sup_{t\leq s\leq t+\delta}|X_t^{(n)} - X_s^{(n)}| \leq \delta c_b + \sup_{t\leq s\leq t+\delta}\Big|\int_s^t \sigma_r^{(n)}dW_r\Big|$. By linearity we only need to estimate the expectation of the second term (to the $2m$th power). We have using the simple inequality for pth power's of sums and the Burkholder-Davis-Gundy-inequality that

$$\begin{aligned}
&\mathbb{E}\Big|\sup_{t\leq s\leq t+\delta}\int_t^s \sigma_r^{(n)}dW_r\Big|^{2m} \\
&\leq C_{BDG}^{(2m)}\Big(\mathbb{E}\int_t^{t+\delta}|\sigma_r^{(n)}|^2 dr\Big)^{\frac{2m}{2}} \\
&\leq C_{BDG}^{(2m)}c_\sigma^{2m}\delta^m.
\end{aligned} \tag{61}$$

From this it follows that

$$\Big(\mathbb{E}\big|\sup_{t\le s\le t+\delta}|X_t^{(n)}-X_s^{(n)}|^{2m}\big|\Big)^{\frac{1}{2m}}\le \delta c_b+(C_{BDG}^{(2m)})^{\frac{1}{2m}}c_\sigma\delta^{\frac{m}{2}}. \tag{62}$$

Now using the above for $m\ge 2$ we get

$$\sup_{n\ge1}\mathbb{P}\big(\sup_{t\le s\le t+\delta}|X_t^{(n)}-X_s^{(n)}|\ge\varepsilon\big)\le\frac{\delta^m}{\varepsilon^{2m}}c_{b,\sigma,m}. \tag{63}$$

Now dividing both sides of the inequality by δ and letting $\delta\to0$ gives us (59).

Denote $\mathbb{P}^{(n)}$ to be the law of $X^{(n)}$. From this estimate and [2, Theorem 7.3 and Corollary p.82-83] it follows that the sequence $(\mathbb{P}^{(n)})_{n\in\mathbb{N}}$ is tight on $(C([0,T];\mathbb{R}^d),\mathcal{B}(C([0,T];\mathbb{R}^d)))$. Thus by Prohorov's theorem, see Theorem 4.6, there exists a probability measure $\mathbb{Q}$ on

$$(C([0,T];\mathbb{R}^d),\mathcal{B}(C([0,T];\mathbb{R}^d))) \tag{64}$$

and a subsequence $(n_k)_{k=1}^\infty$ such that $\mathbb{P}^{(n_k)}\to\mathbb{Q}$ weakly as $k\to\infty$. The weak convergence along a subsequence implies that $\mathbb{Q}_{y(0)}=\nu$.

Now using the facts that $X^{(n)}$ is a solution to Equation (56) and the coefficients b,σ are bounded along with Lemmas 7.3 and 7.4 we get that for every $f\in C_b^{1,2,1}([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d);\mathbb{R})$ and $n\ge1$,

$$\begin{aligned}C^f_{b^{(n)}\sigma^{(n)}}(t,y,\mathbb{P}^{(n)}_y)&=f(t,y(t),\mathbb{P}^{(n)}_{y(t)})-f(0,y(0),\mathbb{P}^{(n)}_{y(0)})\\&\quad-\int_0^t(\partial_s+\mathcal{A}^{(n)})f(s,y,\mathbb{P}^{(n)}_{y(s)})ds,\quad t\in[0,T],\end{aligned} \tag{65}$$

is an $(\mathbb{F},\mathbb{P}^{(n)})$-martingale such that $\mathbb{P}^{(n)}_{y(0)}=\mathbb{P}_{X_0^{(n)}}=\mathbb{P}_\xi=\nu$. Here

$$\begin{aligned}\mathcal{A}^{(n)}f(s,z,\mu)&=\sum_{i=1}^d\partial_{y_i}f(s,z(s,\mu(s)))b_i(s,z(s_n),\mu(s_n))\\&\quad+\frac12\sum_{i,j,k=1}^d\partial^2_{y_iy_j}f(s,z(s),\mu(s))(\sigma_{ik}\sigma_{jk})(s,z(s_n),\mu(s_n))\\&+\sum_{i=1}^d\int_{C([0,T];\mathbb{R}^d)}(\partial_\mu f)_i(s,z(s),\mu(s),y(s))b_i(s,z(s_n),\mu(s_n))\mu(dy)\\&+\frac12\sum_{i,j,k=1}^d\int_{C([0,T];\mathbb{R}^d)}\partial_{z_i}(\partial_\mu f)_j(s,z(s),\mu(s),y(s))(\sigma_{ik}\sigma_{jk})(s,z(s_n),\mu(s_n))\mu(dy),\end{aligned} \tag{66}$$

$(s, z, \mu) \in [0, T] \times C([0, T]; \mathbb{R}^d) \times \mathcal{P}_2(C([0, T]; \mathbb{R}^d))$ and $s_n = s_n(s) := t_i^n$ for $t_i^n \le s < t_{i+1}^n$ and $s_n = s_n(s) := T$ for $s = T$.

Now $C^f_{b^{(n)}\sigma^{(n)}} : [0, T] \times C([0, T]; \mathbb{R}^d) \times \mathcal{P}_2(C([0, T]; \mathbb{R}^d)) \to \mathbb{R}$ is bounded and continuous, thanks to the boundedness and continuity of b, σ and f. Furthermore $C^f_{b^{(n)}\sigma^{(n)}}(\cdot, y, \mu)$ is an $(\mathbb{F}, \mathbb{P}^{(n)})$-martingale so that for any bounded, continuous and non-anticipating function $\phi : [0, T] \times C([0, T]; \mathbb{R}^d) \to \mathbb{R}$, $0 \le s \le t \le T$, it holds (see Lemma 8.4)

$$0 = \mathbb{E}_{\mathbb{P}^{(n)}}\big[(C^f_{b^{(n)}\sigma^{(n)}}(t, y, \mathbb{P}^{(n)}_y) - C^f_{b^{(n)}\sigma^{(n)}}(s, y, \mathbb{P}^{(n)}_y))\phi(s, y)\big]. \tag{67}$$

Now define $C^f_{b\sigma}(t, y, \mu) := f(t, y(t), \mu(t)) - f(0, y(0), \mu(0)) - \int_0^t (\partial_s + \mathcal{A}) f(s, y, \mu) ds$, where $\mathcal{A} f(s, y, \mu) = \mathcal{A} f(s, y(s), \mu(s))$ is defined by (52). Furthermore let $F_n(t, z) := C^f_{b^{(n)}\sigma^{(n)}}(t, z, \mathbb{P}^{(n)}_y)$ and $F(t, z) := C^f_{b\sigma}(t, z, \mathbb{Q}_y)$. Our main goal for the majority of the remaining proof is to show that $F_n(t, \cdot) \to F(t, \cdot)$ uniformly on compact subsets of $C([0, T]; \mathbb{R}^d)$ along the subsequence for which $\mathbb{P}^{(n)} \to \mathbb{Q}$ weakly. For notational simplicity we identify the subsequence with the sequence itself. Now we note that for all $k \ge 2$, $\sup_{n \ge 1} \int_{C([0,T];\mathbb{R}^d)} \|y\|^k \mathbb{P}^{(n)}(dy) = \sup_{n \ge 1} \mathbb{E}[\sup_{t \in [0,T]} |X^{(n)}_t|^k] < \infty$. Here the second term can be estimated from above using the fact that $X^{(n)}$ solves equation (55), the Burkholder-Davis-Gundy inequality and the fact that b, σ are bounded. We skip this estimation as it is essentially identical to the estimate we did earlier in the proof. Now note that $\mathbb{P}^{(n)} \to \mathbb{Q}$ in the 2-Wasserstein metric on $\mathcal{P}_2(C([0, T]; \mathbb{R}^d))$ (Recall Equation (37) and [25, Theorem 7.12]).

We will now move to the uniform convergence on compact subsets. To this end we take an arbitrary compact $\mathcal{K} \subset C([0, T]; \mathbb{R}^d)$. Since $\mathcal{K}$ is compact, it is pre-compact so for any $\varepsilon > 0$, there exists $\omega_1, \dots, \omega_m \in \mathcal{K}$, $m \ge 1$, such that the closed balls $\bar{B}_\varepsilon(\omega_i) := \{g \in C([0, T]; \mathbb{R}^d) : \|g - \omega_i\| \le \varepsilon\}$, $i = 1, \dots, m$ cover $\mathcal{K}$, i.e. $\mathcal{K} \subset \bigcup_{i=1}^m \bar{B}_\varepsilon(\omega_i)$.

Now we want to show that for the centers of these balls ω_i it holds $F_n(t, \omega_i) \to F(t, \omega_i)$ for $t \in [0, T], 1 \le i \le m$. For this we recall that

$$\begin{aligned} F_n(t, \omega_i) &= C^f_{b^{(n)}\sigma^{(n)}}(t, \omega_i, \mathbb{P}^{(n)}_y) \\ &= f(t, \omega_i(t), \mathbb{P}^{(n)}_{y(t)}) - f(0, \omega_i(0), \mathbb{P}^{(n)}_{y(0)}) \\ &- \int_0^t (\partial_s + \mathcal{A}^{(n)}) f(s, \omega_i(s), \mathbb{P}^{(n)}_{y(s)}) ds, \end{aligned} \tag{68}$$

and similarly

$$
\begin{aligned}
F(t,\omega_i) =& C^f_{b\sigma}(t,\omega_i,\mathbb{Q}_y)\\
=& f(t,\omega_i(t),\mathbb{Q}_{y(t)}) - f(0,\omega_i(0),\mathbb{Q}_{y(0)})\\
&- \int_0^t (\partial_s+\mathcal{A})f(s,\omega_i(s),\mathbb{Q}_{y(s)})ds.
\end{aligned}
\tag{69}
$$

From these definitions, for $t\in[0,T], 1\le i\le m$ after multiple applications of the triangle inequality and by Itô's formula Theorem 6.9 we get (we denote $\mathcal{I}$ to be the indicator function of the set in it's subscript)

$$
\begin{aligned}
&|F_n(t,\omega_i) - F(t,\omega_i)|\\
\le& |f(t,\omega_i(t),\mathbb{P}^{(n)}_{y(t)}) - f(t,\omega_i(t),\mathbb{Q}_{y(t)})| + |f(0,\omega_i(0),\mathbb{P}^{(n)}_{y(0)}) - f(0,\omega_i(0),\mathbb{Q}_{y(0)})|\\
&+ \int_0^t |\partial_s f(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)}) - \partial_s f(s,\omega_i(s),\mathbb{Q}_{y(s)})|ds\\
&+ \sum_{l=0}^{2^n-1}\sum_{j=1}^{d} |\partial_{y_j} f(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)}) b_j(s,\omega_i(t^n_l),\mathbb{P}^{(n)}_{y(t^n_l)})\\
&- \partial_{y_j} f(s,\omega_i(s),\mathbb{Q}_{y(s)}) b_j(s,\omega_i(s),\mathbb{Q}_{y(s)})|\mathcal{I}_{[t^n_l,t^n_{l+1})}(s)\\
&+ \frac{1}{2}\sum_{l=0}^{2^n-1}\sum_{j,k,r=1}^{d} |\partial^2_{y_jy_k} f(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)})(\sigma_{jr}\sigma_{kr})(s,\omega_i(t^n_l),\mathbb{P}^{(n)}_{y(t^n_l)})\\
&- \partial^2_{y_jy_k} f(s,\omega_i(s),\mathbb{Q}_{y(s)})(\sigma_{jr}\sigma_{kr})(s,\omega_i(s),\mathbb{Q}_{y(s)})|\mathcal{I}_{[t^n_l,t^n_{l+1})}(s)\\
&+ \sum_{l=0}^{2^n-1}\sum_{j=1}^{d} \int_{C([0,T];\mathbb{R}^d)} |(\partial_\mu f)_j(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)},z(s)) b_j(s,\omega_i(t^n_l),\mathbb{P}^{(n)}_{y(t^n_l)})\\
&- (\partial_\mu f)_j(s,\omega_i(s),\mathbb{Q}_{y(s)},z(s)) b_j(s,\omega_i(s),\mathbb{Q}_{y(s)})|\mathcal{I}_{[t^n_l,t^n_{l+1})}(s)\mu(dz)\\
&+ \frac{1}{2}\sum_{l=0}^{2^n-1}\sum_{j,k,r=1}^{d} \int_{C([0,T];\mathbb{R}^d)} |\partial_{z_j}(\partial_\mu f)_k(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)},z(s))(\sigma_{jr}\sigma_{kr})(s,\omega_i(t^n_l),\mathbb{P}^{(n)}_{y(t^n_l)})\\
&- \partial_{z_j}(\partial_\mu f)_j(s,\omega_i(s),\mathbb{Q}_{y(s)},z(s))(\sigma_{jr}\sigma_{kr})(s,\omega_i(s),\mathbb{Q}_{y(s)})|\mathcal{I}_{[t^n_l,t^n_{l+1})}(s)\mu(dz).
\end{aligned}
\tag{70}
$$

Denote this as a sum $I_n + II_n + III_n + IV_n + V_n$, where

$$
\begin{aligned}
I_n :=& |f(t,\omega_i(t),\mathbb{P}^{(n)}_{y(t)}) - f(t,\omega_i(t),\mathbb{Q}_{y(t)})|\\
&+ |f(0,\omega_i(0),\mathbb{P}^{(n)}_{y(0)}) - f(0,\omega_i(0),\mathbb{Q}_{y(0)})|\\
&+ \int_0^t |\partial_s f(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)}) - \partial_s f(s,\omega_i(s),\mathbb{Q}_{y(s)})|ds,
\end{aligned}
\tag{71}
$$

$$II_n := \sum_{l=0}^{2^n-1}\sum_{j=1}^{d}|\partial_{y_j} f(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)})b_j(s,\omega_i(t_l^n),\mathbb{P}^{(n)}_{y(t_l^n)}) - \partial_{y_j} f(s,\omega_i(s),\mathbb{Q}_{y(s)})b_j(s,\omega_i(s),\mathbb{Q}_{y(s)})|\mathcal{I}_{[t_l^n,t_{l+1}^n)}(s), \tag{72}$$

$$III_n := \frac{1}{2}\sum_{l=0}^{2^n-1}\sum_{j,k,r=1}^{d}|\partial^2_{y_j y_k} f(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)})(\sigma_{jr}\sigma_{kr})(s,\omega_i(t_l^n),\mathbb{P}^{(n)}_{y(t_l^n)}) - \partial^2_{y_j y_k} f(s,\omega_i(s),\mathbb{Q}_{y(s)})(\sigma_{jr}\sigma_{kr})(s,\omega_i(s),\mathbb{Q}_{y(s)})|\mathcal{I}_{[t_l^n,t_{l+1}^n)}(s), \tag{73}$$

$$IV_n := \sum_{l=0}^{2^n-1}\sum_{j=1}^{d}\int_{C([0,T];\mathbb{R}^d)}|(\partial_\mu f)_j(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)},z(s))b_j(s,\omega_i(t_l^n),\mathbb{P}^{(n)}_{y(t_l^n)}) - (\partial_\mu f)_j(s,\omega_i(s),\mathbb{Q}_{y(s)},z(s))b_j(s,\omega_i(s),\mathbb{Q}_{y(s)})|\mathcal{I}_{[t_l^n,t_{l+1}^n)}(s)\mu(dz), \tag{74}$$

and

$$V_n := \frac{1}{2}\sum_{l=0}^{2^n-1}\sum_{j,k,r=1}^{d}\int_{C([0,T];\mathbb{R}^d)}|\partial_{z_j}(\partial_\mu f)_k(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)},z(s))(\sigma_{jr}\sigma_{kr})(s,\omega_i(t_l^n),\mathbb{P}^{(n)}_{y(t_l^n)}) - \partial_{z_j}(\partial_\mu f)_j(s,\omega_i(s),\mathbb{Q}_{y(s)},z(s))(\sigma_{jr}\sigma_{kr})(s,\omega_i(s),\mathbb{Q}_{y(s)})|\mathcal{I}_{[t_l^n,t_{l+1}^n)}(s)\mu(dz). \tag{75}$$

We now want to show that $I_n + II_n + III_n + IV_n + V_n \to 0$ when $n \to \infty$, which amounts to showing that each of these terms goes to zero separately, since each of them is non-negative. Here we restrict ourselves on the term IV_n, the other terms are treated similarly. We elaborate more on this after we show that $IV_n \to 0$ as $n \to \infty$. We use the following equality $ab - cd = b(a-c) + c(b-d)$ on the integrand in IV_n to get

$$IV_n = \sum_{l=0}^{2^n-1}\sum_{j=1}^{d}\int_{C([0,T];\mathbb{R}^d)}\big|b_j(s,\omega_i(t_l^n),\mathbb{P}^{(n)}_{y(t_l^n)})((\partial_\mu f)_j(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)},z(s)) - (\partial_\mu f)_j(s,\omega_i(s),\mathbb{Q}_{y(s)},z(s))) + (\partial_\mu f)_j(s,\omega_i(s),\mathbb{Q}_{y(s)},z(s)) (b_j(s,\omega_i(t_l^n),\mathbb{P}^{(n)}_{y(t_l^n)}) - b_j(s,\omega_i(s),\mathbb{Q}_{y(s)}))\big|\mathcal{I}_{[t_l^n,t_{l+1}^n)}(s)\mu(dz). \tag{76}$$

Which we can now estimate using the triangle inequality and the fact that b_j are bounded and all the derivatives of f are uniformly bounded to get

$$IV_n \le \sum_{l=0}^{2^n-1}\sum_{j=1}^{d}\int_{C([0,T];\mathbb{R}^d)}\big(C\big|(\partial_\mu f)_j(s,\omega_i(s),\mathbb{P}^{(n)}_{y(s)},z(s)) - (\partial_\mu f)_j(s,\omega_i(s),\mathbb{Q}_{y(s)},z(s))\big| + C\big|b_j(s,\omega_i(t_l^n),\mathbb{P}^{(n)}_{y(t_l^n)}) - b_j(s,\omega_i(s),\mathbb{Q}_{y(s)})\big|\big)\mathcal{I}_{[t_l^n,t_{l+1}^n)}(s)\mu(dz) \tag{77}$$

Furthermore the derivatives $(\partial_\mu f)_j$ are Lipschitz in the measure component so we further estimate

$$\begin{aligned} IV_n \leq C \int_0^T \big(W_2(\mathbb{P}^{(n)}_{y(s)}, \mathbb{Q}_{y(s)}) + \sum_{l=0}^{2^n-1} \sum_{j=1}^{d} |b_j(s, \omega_i(t_l^n), \mathbb{P}^{(n)}_{y(t_l^n)}) \\ - b_j(s, \omega_i(s), \mathbb{Q}_{y(s)})| \mathcal{I}_{[t_l^n, t_{l+1}^n)}(s) \big) ds. \end{aligned} \tag{78}$$

We are now looking to show that the last integral converges to zero as $n \to \infty$. To this end we show the following:

1. $W_2(\mathbb{P}^{(n)}_{y(s)}, \mathbb{Q}_{y(s)}) \leq W_2(\mathbb{P}^{(n)}_{y}, \mathbb{Q}_y)$, $s \in [0,T]$.
2. $W_2(\mathbb{P}^{(n)}_{y(t_l^n)}, \mathbb{Q}_{y(s)}) \leq C 2^{-\frac{n}{2}} + W_2(\mathbb{P}^{(n)}_{y}, \mathbb{Q}_y)$, $s \in (t_l^n, t_{l+1}^n]$, $0 \leq l \leq 2^n - 1$.

The first item shows that the first term in the integral converges to zero as $n \to \infty$ since $\mathbb{P}^{(n)} \to \mathbb{Q}$ in the 2-Wasserstein metric on $\mathcal{P}_2(C([0,T];\mathbb{R}^d))$. The second item on the other hand when paired with the facts that $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ is continuous and $\omega_i \in C([0,T];\mathbb{R}^d)$ shows that the second term in the integral also converges to zero as $n \to \infty$. Then by the boundedness of the terms and the bounded convergence theorem one concludes that $IV_n \to 0$ as $n \to \infty$. Now on to showing items 1 and 2.

For 1 we note that for any $\varepsilon > 0$ and $n \geq 1$, there is a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ and on it continuous stochastic processes $y^P = (y^P(s))_{s \in [0,T]}$ and $y^Q = (y^Q(s))_{s \in [0,T]}$ such that $\mathbb{P}'_{y^P} = \mathbb{P}^{(n)}_y$ and $\mathbb{P}'_{y^Q} = \mathbb{Q}_y$ and

$$\varepsilon + W_2(\mathbb{P}^{(n)}_y, \mathbb{Q}_y) \geq \big(\mathbb{E}\big[\sup_{s \in [0,T]} |y^P(s) - y^Q(s)|^2 \big] \big)^{\frac{1}{2}}. \tag{79}$$

Now using the inequality from Remark 5.2 we obtain for all $s \in [0,T]$ that

$$\begin{aligned} \varepsilon + W_2(\mathbb{P}^{(n)}_y, \mathbb{Q}_y) &\geq \big(\mathbb{E}\big[\sup_{s \in [0,T]} |y^P(s) - y^Q(s)|^2 \big] \big)^{\frac{1}{2}} \\ &\geq \big(\mathbb{E}\big[|y^P(s) - y^Q(s)|^2 \big] \big)^{\frac{1}{2}} \geq W_2(\mathbb{P}^{(n)}_{y(s)}, \mathbb{Q}_{y(s)}). \end{aligned} \tag{80}$$

From which item 1 follows by letting $\varepsilon \to 0$.

For 2 we first use the triangle inequality on the Wasserstein space and then the same inequality from Remark 5.2 along with the item 1 we just proved to obtain

$$\begin{aligned} W_2(\mathbb{P}^{(n)}_{y(t_l^n)}, \mathbb{Q}_{y(s)}) &\leq W_2(\mathbb{P}^{(n)}_{y(t_l^n)}, \mathbb{P}^{(n)}_{y(s)}) + W_2(\mathbb{P}^{(n)}_{y(s)}, \mathbb{Q}_{y(s)}) \\ &\leq \big(\mathbb{E}\big[|X^{(n)}_{t_l^n} - X^{(n)}_s|^2 \big] \big)^{\frac{1}{2}} + W_2(\mathbb{P}^{(n)}_y, \mathbb{Q}_y). \end{aligned} \tag{81}$$

We now recall the equation (55) and use the same method as inequality (58) on the expectation in the inequality above to obtain

$$W_2(\mathbb{P}^{(n)}_{y(t^n_l)}, \mathbb{Q}_{y(s)}) \le C2^{-\frac{n}{2}} + W_2(\mathbb{P}^{(n)}_y, \mathbb{Q}_y). \tag{82}$$

Which proves the second item. Thus as we mentioned before, by the bounded convergence theorem the integral $IV_n \to 0$ as $n \to \infty$. V_n is essentially exactly the same, instead of invoking the properties of b, here we use the properties of σ, which both are assumed to be bounded and continuous. Further we use the fact that $\partial_{z_j}(\partial_\mu f)_k$ is Lipschitz in the measure component, which is the same process. The terms II_n and III_n are in spirit the same to each other, here we only need the boundedness and continuity along with the items 1 and 2, so this also works. For the term I_n we need only the continuity and boundedness properties of f and do not use the properties of b, σ.

Thus we are done with our task of showing that $F_n(t, \omega_i) \to F(t, \omega_i)$ for $t \in [0, T], 1 \le i \le m$ for all centers ω_i of the balls covering $\mathcal{K}$. Therefore also for any $\varepsilon > 0$, there exists $n_\varepsilon \ge 1$ such that for all $n \ge n_\varepsilon$ we have $|F_n(t, \omega_i) - F(t, \omega_i)| \le \varepsilon$ for $1 \le i \le m$.

We next want to show that we can control $|F_n(t, \omega_l) - F_n(t, \omega)|$ whenever ω, ω_l are contained in a small ball.

To this end we first note that since $\mathcal{K}$ is a compact set, and thus bounded so we find an $R > 0$, such that $\mathcal{K} \subset \bar{B}_R(0) = \{\omega \in C([0, T]; \mathbb{R}^d) : \|\omega\| \le R\}$.

Now notice again by the boundedness of b, σ and a similar inequality as inequality (58) that there exists a constant $C_0 > 0$ such that $\mathbb{E}[|X^{(n)}_s|^4] \le C_0$ for $s \in [0, T], n \ge 1$. Now we recall the subspace $\mathcal{E}$ of $\mathcal{P}_2(\mathbb{R}^d)$, defined by

$$\mathcal{E} = \{\mu \in \mathcal{P}_2(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^4 \mu(dx) \le C_0\}, \tag{83}$$

which we showed to be compact in Lemma 5.3. Furthermore for any $t \in [0, T]$, $n \ge 1$ we have that $\mathbb{P}^{(n)}_{y(t)} \in \mathcal{E}$, which follows by the bound on the fourth moment we indicated above.

We now define the continuity modulus $m_{R,\mathcal{E}} : \mathbb{R}^+ \to \mathbb{R}^+$ for any $\delta > 0$ by

$$\begin{aligned} m_{R,\mathcal{E}}(\delta) := \sup\{&|\gamma(s, x, \nu, z) - \gamma(s, x', \nu, z)| : s \in [0, T], \nu \in \mathcal{E} \\ &x, x' \in \mathbb{R}^d, \text{ such that } |x|, |x'| \le R, |x - x'| \le \delta, z \in \mathbb{R}^d \\ &\gamma \in \{b, \sigma, f, \partial_s f, \partial_y f, \partial^2_{yy} f, \partial_\mu f, \partial_z(\partial_\mu f)\}\}. \end{aligned} \tag{84}$$

This is done to control every term after estimating $|F_n(t, \omega_l) - F_n(t, \omega)|$ with the triangle inequality and the bounds for b, σ.

Now we note by the facts that $f \in C_b^{1,2,1}([0,T] \times \mathbb{R}^d, \mathcal{P}_2(\mathbb{R}^d); \mathbb{R})$ and $(b, \sigma) \in C_b([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d); \mathbb{R}^d \times \mathbb{R}^{d \times d})$ and that $[0,T] \times \{x \in \mathbb{R}^d; |x| \leq R\} \times \mathcal{E}$ is a compact set as a product of compact sets we obtain the fact that all γ in the definition of the continuity modulus are uniformly continuous on $[0,T] \times \{x \in \mathbb{R}^d; |x| \leq R\} \times \mathcal{E}$. This implies that $m_{R,\mathcal{E}}(\delta) \to 0$ as $\delta \to 0$.

Now by the definition of F_n for any $\omega, \omega_l \in \mathcal{K}, n \geq 1$ it holds

$$
\begin{aligned}
&|F_n(t,\omega_l) - F_n(t,\omega)| \\
&\leq |f(t,\omega_l(t),\mathbb{P}^{(n)}_{y(t)}) - f(t,\omega(t),\mathbb{P}^{(n)}_{y(t)})| \\
&+ |f(0,\omega_l(0),\mathbb{P}^{(n)}_{y(0)}) - f(0,\omega(0),\mathbb{P}^{(n)}_{y(0)})| \\
&+ \int_0^t |\partial_s f(s,\omega_l(s),\mathbb{P}^{(n)}_{y(s)}) - \partial_s f(s,\omega(s),\mathbb{P}^{(n)}_{y(s)})| ds \\
&+ C_b \int_0^t \sum_{i=1}^d |\partial_{y_i} f(s,\omega_l(s),\mathbb{P}^{(n)}_{y(s)}) - \partial_{y_i} f(s,\omega(s),\mathbb{P}^{(n)}_{y(s)})| ds \\
&+ \int_0^t \sum_{i=1}^d |\partial_{y_i} f(s,\omega(s),\mathbb{P}^{(n)}_{y(s)})||b_i^{(n)}(s,\omega_l(s),\mathbb{P}^{(n)}_y) - b_i^{(n)}(s,\omega(s),\mathbb{P}^{(n)}_y)| ds \\
&+ C_\sigma \int_0^t \sum_{i,j,k=1}^d |\partial^2_{y_i y_j} f(s,\omega_l(s),\mathbb{P}^{(n)}_{y(s)}) - \partial^2_{y_i y_j} f(s,\omega(s),\mathbb{P}^{(n)}_{y(s)})| ds \\
&+ \frac{1}{2} \int_0^t \sum_{i,j,k=1}^d |\partial_{y_i y_j} f(s,\omega(s),\mathbb{P}^{(n)}_{y(s)})||\sigma_{ik}^{(n)}(s,\omega_l(s),\mathbb{P}^{(n)}_y) \\
&- \sigma_{ik}^{(n)}(s,\omega(s),\mathbb{P}^{(n)}_y)| ds + C_b \int_0^t \sum_{i=1}^d \int_{\mathbb{R}^d} |(\partial_\mu f)_i(s,\omega_l(s),\mathbb{P}^{(n)}_{y(s)},z) \\
&- (\partial_\mu f)_i(s,\omega(s),\mathbb{P}^{(n)}_{y(s)},z)| \mathbb{P}^{(n)}_{y(s)}(dz) ds \\
&+ C_\sigma \int_0^t \sum_{i,j,k=1}^d \int_{\mathbb{R}^d} |\partial_{z_i}(\partial_\mu f)_j(s,\omega_l(s),\mathbb{P}^{(n)}_{y(s)},z) \\
&- \partial_{z_i}(\partial_\mu f)_j(s,\omega(s),\mathbb{P}^{(n)}_{y(s)},z)| \mathbb{P}^{(n)}_{y(s)}(dz) ds.
\end{aligned} \tag{85}
$$

Here we did a similar trick as in estimating the integral IV_n earlier and then exploited the boundedness of the coefficients b, σ. However we notice that all of these terms can be crudely estimated from above by the continuity modulus for any $\omega, \omega_l \in \mathcal{K} \subset \bar{B}_R(0)$ such that $|\omega(s) - \omega_l| \leq \varepsilon,\ s \in [0,t]$. Thus it holds for these ω, ω_l that $|F_n(t,\omega_l) - F_n(t,\omega)| \leq C m_{R,\mathcal{E}}, \varepsilon > 0$. Now for all $\omega \in \mathcal{K}$ we find $i, 1 \leq i \leq m$ such that $\omega \in \bar{B}_\varepsilon(\omega_i)$ and thus we obtain by a double application of the triangle inequality

$$\begin{aligned}
&|F_n(t,\omega) - F(t,\omega)| \\
&\le 2Cm_{R,\mathcal{E}}(\varepsilon) + \max_{1\le i\le m} |F_n(t,\omega_i) - F(t,\omega_i)| \\
&\le 2Cm_{R,\mathcal{E}}(\varepsilon) + \varepsilon
\end{aligned} \tag{86}$$

for $n \ge n_\varepsilon, \omega \in \mathcal{K}$. The last term in the inequality goes to zero as $\varepsilon \to 0$. Therefore $F_n(t,\cdot)$ converges uniformly on the compact set $\mathcal{K}$ to $F(t,\cdot)$ for all $t \in [0,T]$.

We are now done with the main part of the proof. To conclude, we fix $s < t$ in $[0,T]$. We have for all non-anticipating $\phi \in C_b([0,T] \times C([0,T];\mathbb{R}^d);\mathbb{R})$ by the uniform convergence

$$\begin{aligned}
0 &= \mathbb{E}_{\mathbb{P}^{(n)}}[(C^f_{b^{(n)}\sigma^{(n)}}(t,y,\mathbb{P}^{(n)}) - C^f_{b^{(n)}\sigma^{(n)}}(s,y,\mathbb{P}^{(n)}))\phi(s,y)] \\
&= \mathbb{E}_{\mathbb{P}^{(n)}}[(F_n(t,y) - F_n(s,y))\phi(s,y)] \\
&\to \mathbb{E}_{\mathbb{Q}}[(F(t,y) - F(s,y))\phi(s,y)] \\
&= \mathbb{E}_{\mathbb{Q}}[(C^f_{b\sigma}(t,y,\mathbb{Q}) - C^f_{b\sigma}(s,y,\mathbb{Q}))\phi(s,y)],
\end{aligned} \tag{87}$$

as $n \to \infty$ along the (sub)sequence of $n \ge 1$ for which $\mathbb{P}^{(n)} \to \mathbb{Q}$ weakly. Now from this equation we can use the characterization for martingales against integrals of continuous functions (Lemma 8.4) and thus we get that for any $f \in C_b^{1,2,1}([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d);\mathbb{R})$,

$$C^f(t) = f(t,y(t),\mathbb{Q}_{y(t)}) - f(0,y(0),\mathbb{Q}_{y(0)}) - \int_0^t (\partial_s + \mathcal{A})f(s,y(s),\mathbb{Q}_{y(s)})ds \tag{88}$$

is an $(\mathbb{F}^y,\mathbb{Q})$-martingale. Now due to the definitions of $\mathcal{A}$ and $\widetilde{\mathcal{A}}$, we can invoke Lemma 7.3 to obtain a d-dimensional Brownian motion $\widetilde{W} = (\widetilde{W}_s)_{s\in[0,T]}$ defined on an extension $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{Q}})$ of $(C([0,T];\mathbb{R}^d),\mathcal{B}((C([0,T];\mathbb{R}^d)),\mathbb{Q})$ such that $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{F}},\widetilde{\mathbb{Q}},\widetilde{W},X)$ is a weak solution to the mean-field SDE (42), where $X = y$ and $\widetilde{\mathbb{F}} \supset \mathbb{F}^y$ is a suitable filtration. □

Thus we have established a weak existence result for the mean-field SDE (42) under the condition of merely bounded and continuous coefficients b and σ.

8 Appendix

Here we mention some of the technical results used before.

Remark 8.1. Recall first in the proof of Lemma 7.3 we claimed that there exists a smallest right-continuous and complete filtration $\widetilde{\mathbb{F}} = (\widetilde{\mathcal{F}}_t)_{t\in[0,T]}$ for which $\widehat{\mathcal{F}}_t \bigotimes \bar{\mathcal{F}}_t \subset \widetilde{\mathcal{F}}_t$. Here recall that the $\widehat{\mathcal{F}}_t$ denotes the Borel σ-algebra on the path space and $\bar{\mathcal{F}}_t$ is a σ-algebra on a probability space on which we have defined a d-dimensional Brownian motion. We elaborate the construction of such a σ-algebra. Denote $\mathcal{G}_t := \widehat{\mathcal{F}}_t \bigotimes \bar{\mathcal{F}}_t$. The problem is that $(\mathcal{G}_t)_{t\in[0,T]}$ might not satisfy the usual conditions. However we can define $\widetilde{\mathcal{F}}_t := \bigcap_{s>t} \sigma(\mathcal{G}_s \cup \mathcal{N})$. Each of the σ-algebras $\sigma(\mathcal{G}_s \cup \mathcal{N})$ is complete, and defining the intersection of these σ-algebras forces it to be right-continuous, this follows since $\bigcap_{r>t} \widetilde{\mathcal{F}}_r = \bigcap_{r>t} \bigcap_{s>r} \sigma(\mathcal{G}_s \cup \mathcal{N}) = \bigcap_{s>t} \sigma(\mathcal{G}_s \cup \mathcal{N}) = \widetilde{\mathcal{F}}_t$. Here $\mathcal{N}$ denotes the null-sets of the product probability measure $\widehat{\mathbb{P}} \times \mathbb{P}$ where the probability with a hat corresponds to the solution to the local martingale problem. Clearly also $\widehat{\mathcal{F}}_t \bigotimes \bar{\mathcal{F}}_t \subset \widetilde{\mathcal{F}}_t$.

Here we will also recall the following result from [16, p. 315] to represent a solution of an SDE as the coordinate process against a certain stochastic basis and a Brownian motion:

Proposition 8.2. *Assume* $b : [0,T] \times C([0,T];\mathbb{R}^d) \to \mathbb{R}^d$ *and* $\sigma : [0,T] \times C([0,T];\mathbb{R}^d) \to \mathbb{R}^{d\times d}$ *are progressively measurable and let*

$$M_t^f = f(y(t)) - f(y(0)) - \int_0^t (\mathcal{A}'f)(y)ds,\ t \in [0,T]. \tag{89}$$

Let $\widehat{\mathbb{P}}$ *be a probability measure on* $(C([0,T];\mathbb{R}^d), \mathcal{B}(C([0,T];\mathbb{R}^d)))$ *such that* M^f *is a continuous local* $(\mathbb{F}^y, \widehat{\mathbb{P}})$*-martingale for the choices* $f(x) = x_i$ *and* $f(x) = x_i x_j$, $1 \le i,j \le d$. *Then there exists a d-dimensional Brownian motion* $W = (W_t)_{t\in[0,T]}$ *defined on an extension* $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ *of*

$$(C([0,T];\mathbb{R}^d), \mathcal{B}((C([0,T];\mathbb{R}^d)), \widehat{\mathbb{P}}) \tag{90}$$

such that $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{F}}, \widehat{\mathbb{P}}, W, X)$ *is a solution to the functional SDE* (25)*, where* $X = y$.

Here $(\mathcal{A}'f)(y)$ is defined as $\mathcal{A}'f(s,y)$ where f does not depend on s. This is helpful as whenever we are investigating something and have a solution to the local martingale problem, we can represent this solution as the coordinate process.

In the proof of Theorem 7.5 we used the following relation for martingales:

$M = (M_t)_{t\in[0,T]}$ on $C([0,T];\mathbb{R}^d)$ is a martingale with respect to $\mathbb{F}^y$ if and only if $\mathbb{E}[(M_t(y) - M_s(y))\phi(s,y)] = 0$ for any non-anticipating $\phi \in C_b([0,T] \times C([0,T] \times \mathbb{R}^d);\mathbb{R})$ and $s < t$. This result is not obvious, we will first prove it when ϕ is measurable, and not continuous. This is a weaker assumption since in our setting continuous functions are measurable.

Lemma 8.3. *$M = (M_t)_{t\in[0,T]}$ is a martingale with respect to $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ if and only if for $0 \le s < t \le T$ and any $\phi : \Omega \to \mathbb{R}$ bounded and $\mathcal{F}_s$-measurable function it holds*

$$\mathbb{E}[(M_t - M_s)\phi] = 0 \tag{91}$$

Proof. Assume that M is a martingale. We get by a simple computation using the tower property of conditional expectation and the fact that ϕ is $\mathcal{F}_s$-measurable ('take out what is known' property)

$$\begin{aligned}\mathbb{E}[(M_t - M_s)\phi] &= \mathbb{E}[\mathbb{E}[(M_t - M_s)\phi|\mathcal{F}_s]] \\ &= \mathbb{E}[\phi\mathbb{E}[(M_t - M_s)|\mathcal{F}_s]] = 0, \quad a.s.\end{aligned} \tag{92}$$

Where the last equality follows from the fact that M is a martingale.

Now assume that $\mathbb{E}[(M_t - M_s)\phi] = 0$ holds for all $\phi : \Omega \to \mathbb{R}$, $\mathcal{F}_s$-measurable and bounded. Especially for $A \in \mathcal{F}_s$ it holds that $\mathbb{E}[(M_t - M_s)\mathcal{I}_A] = 0$. However upon rewriting we obtain

$$\int_A M_t d\mathbb{P} = \int_A M_s d\mathbb{P} \tag{93}$$

and thus $\mathbb{E}(M_t|\mathcal{F}_s) = M_s$ a.s. which finishes the proof. □

We want to extend this characterization to the case where ϕ is assumed to be continuous, this makes one direction of the above implications harder, and the other easier. As every continuous function is measurable in our setting, we obtain the first direction immediately. We now formulate the result:

Lemma 8.4. *$M = (M_t)_{t\in[0,T]}$ is a martingale with respect to $\mathbb{F}^y = (\mathcal{F}^y_t)_{t\in[0,T]}$ if and only if for $0 \le s < t \le T$ and any $\phi : [0,T] \times C([0,T];\mathbb{R}^d) \to \mathbb{R}$, non-anticipating, bounded and continuous it holds*

$$\mathbb{E}[(M_t - M_s)\phi(s,y)] = 0. \tag{94}$$

Proof. We assume there exists $A \in \mathcal{B}(C([0,s];\mathbb{R}^d))$ such that (the expectation is taken against $\mathbb{Q}$ on $C([0,s];\mathbb{R}^d)$)

$$\mathbb{E}[(M_t - M_s)\mathcal{I}_A] > 0. \tag{95}$$

By the regularity of probability measures on metric spaces (see [24, Theorem 1.2]), for any $\varepsilon > 0$ we can find an open set G, such that $A \subset G$ and $\mathbb{Q}(G) \leq \mathbb{Q}(A) + \varepsilon$. Thus we also have

$$\mathbb{E}[(M_t - M_s)\mathcal{I}_G] > 0. \tag{96}$$

Now we can approximate the function $\mathcal{I}_G$ pointwise and in probability (with respect to $\mathbb{Q}$) by continuous functions ϕ_n. This can be done for example by the usual distance function construction. Most importantly we find a continuous function $\phi : [0,T] \times C([0,T]; \mathbb{R}^d) \to \mathbb{R}$ for which it holds

$$\mathbb{E}[(M_t - M_s)\phi] > 0, \tag{97}$$

which is a contradiction. □

References

[1] Bauer, M. & Meyer-Brandis, T. & Proske, F. (2018), *Strong solutions of mean-field stochastic differential equations with irregular drift.* Electronic Journal of Probability, 23, no. 132, 1-35.

[2] Billingsley, P. (1999), *Convergence of Probability Measures (Second edition).* Chicago: Wiley series in Probability and Statistics.

[3] Black, F. & Scholes, M. (1973), *The Pricing of Options and Corporate Liabilities.* University of Chicago Press. The Journal of Political Economy, Vol. 81, No. 3.

[4] Buckdahn, R. & Li, J. & Peng, S. G. & Rainer, C. (2014), *Mean-field stochastic differential equations and associated PDEs.* arXiv:1407.1215.

[5] Cardaliaguet, P. (2013), *Notes on mean field games (from P.-L. Lions' lectures at Collége de France).* https://www.ceremade.dauphine.fr/ cardaliaguet/MFG100629.pdf, 15.1.2012.

[6] Carmona, R. & Lacker, D. (2015), *A probabilistic weak formulation of mean field games and applications.* Annals of Applied Probability 25(3):1189-1231.

[7] Chiang, T. S. (1994), *McKean-Vlasov equations with discontinuous coefficients.* Soochow J. Math. 20:507-526.

[8] Funaki, T. (1984), *A certain class of diffusion processes associated with nonlinear parabolic equations.* Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 67(3):331-348.

[9] Geiss, S. (2020), *Stochastic Differential Equations.* Lecture note, University of Jyväskylä. http://users.jyu.fi/ geiss/lectures/stochasticsde.pdf, 25.11.2020.

[10] Geiss, S. & Ylinen, J. (2021), *Decoupling on the Wiener Space, Related Besov Spaces, and Applications to BSDEs.* Memoirs of the American Mathematical Society. Number 1335.

[11] Gärtner, J. (1988), *On the McKean-Vlasov limit for interacting diffusions.* Mathematische Nachrichten 137(1):197-248.

[12] Hammersley, W. R. P. & Siska, D. & Szpruch, L., (2020), *McKean-Vlasov SDEs under Measure Dependent Lyapunov Conditions.* School of Mathematics, University of Edinburgh.

[13] Itô, K. (1942), *Differential equations determining Markov processes.* Zenkoku shijo Sugaku Danwakai 244(1077):1352-1400.

[14] Itô, K. (1951), *On stochastic differential equations.* Memoirs of the American Mathematical Society, No. 4.

[15] Kac, M. (1956), *Foundations of Kinetic Theory.* Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. 3: Contributions to Astronomy and Physics. University of California Press.

[16] Karatzas, I. & Shreve, S. (1991), *Brownian Motion and Stochastic Calculus.* New York: Springer Verlag.

[17] Langevin, P. (1908), *On the Theory of Brownian Motion.* C. R. Acad. Sci. Paris. 146: 530–533.

[18] Lasry, J.-M. & Lions, P.-L. (2007), *Mean field games.* Japanese Journal of Mathematics 2, no. 1, 229-260.

[19] Li, J. & Min, H. (2017), *Weak solutions of mean-field stochastic differential equations.* Stochastic Analysis and Applications.

[20] McKean, H. P. (1966), *A class of Markov processes associated with nonlinear parabolic equations.* Proc Natl Acad Sci USA 56, no. 6.

[21] Mehri, S & Stannat, W. (2019), *Weak solutions to Vlasov-McKean equations under Lyapunov-type conditions.* Stochastics and Dynamics, Vol. 19, No. 6. World Scientific.

[22] Mishura, S. & Veretennikov, A. (2018), *Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations.* arXiv:1603.02212v8.

[23] Ornstein, L. S. & Uhlenbeck, G. E. (1930), *On the Theory of the Brownian Motion.* Amercian Physical Society.

[24] Parthasarathy, K. R. (1967), *Probability Measures on Metric Spaces.* Academic Press. New York and London.

[25] Villani, C. (2003), *Topics in Optimal Transportation.* Graduate Studies in Mathematics. 58, American Mathematical Society.

[26] Vlasov, A. A. (1968), *The vibrational properties of an electron gas.* Soviet Physics Uspekhi 10, no. 6, 721.