Interacting Measure-Valued Diffusions and their Long-Term Behavior

by

Hardeep Gill

B.Sc. Pure Mathematics and Statistics, Calgary, 2005

M.Sc. Mathematics, University of British Columbia, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES
(Mathematics)

The University Of British Columbia
(Vancouver)

August 2011

Abstract

The focus of this dissertation is a class of random processes known as interacting measure-valued stochastic processes. These processes are related to another class of stochastic processes known as superprocesses. Both superprocesses and interacting measure-valued stochastic processes arise naturally from branching particle systems as scaling limits. A branching particle system is a collection of particles that propagate randomly through space, and that upon death give birth to a random number of particles (children). Therefore when the populations of the particle system and branching rate are large one can often use a superprocess to approximate it and carry out calculations that would be very difficult otherwise.

There are many branching particle systems which do not satisfy the strong independence assumptions underlying superprocesses and thus are more difficult to study mathematically. This dissertation attempts to address two measure-valued processes with different types of dependencies (interactions) that the associated particles exhibit. In both cases, the method used to carry out this work is called Perkins' historical stochastic calculus, and has never before been used to investigate interacting measure-valued processes of these types. That is, we construct the measure-valued stochastic process associated with an interacting branching particle system directly *without* taking a scaling limit.

The first type of interaction we consider is when all particles share a common chaotic drift from being immersed in the same medium, as well as having other types of individual interactions. The second interaction involves particles that attract to or repel from the center of mass of the entire population. For measure-valued processes with this latter interaction, we study the long-term behavior of the process and show that it displays some types of equilibria.

Preface

The work provided below was conducted solely by myself, under the supervision of my advisor Professor Edwin A. Perkins. The problems investigated were originally conceived by Professor Perkins. I proved the results (solutions of those problems), wrote the research articles and subsequently submitted them to journals of my choosing.

This dissertation is based on two publications that resulted from my time as a doctoral student:

(1) **Gill, H.S.** A super Ornstein-Uhlenbeck process interacting with its center of mass. To appear.

(2) **Gill, H.S.** Superprocesses with spatial interactions in a random medium. *Stochastic processes and their applications*. December, 2009. 119 (12): 3981-4003.

Publication (1) forms the basis of Chapter 3, and (2) forms the basis for Chapter 2.

Table of Contents

Acknowledgments

There were many people who made this dissertation possible. Some of them impacted it directly, and some of them kept me in "Chahrdi Kala," which in Punjabi means in high spirits.

My advisor, Professor Edwin A. Perkins, has my deepest gratitude for his patient and thoughtful guidance. I am especially grateful for the research problems he suggested, which provided me with a great deal of enjoyment and an invaluable source for learning. I quickly found that Ed is a veritable fount of ideas for those difficult times I was stuck while working on a problem. I also appreciate his careful readings of this manuscript and the papers which resulted from it.

I would like to thank the members of my supervisory committee, particularly Martin Barlow and John Walsh, for their insightful comments and suggestions. I am grateful to Jie Xiong for suggesting additional avenues of research.

My career path might have been completely different without Ms. Williams, my high school math teacher. Her creative and challenging class fostered my interest in mathematics.

Kiran has my gratitude for her encouragement and support during the writing of this manuscript. This was nowhere more evident than by her willingness to undergo an allergic reaction rather than leave my doctoral defense room. Kiran, Lisa, Sidd, and my brothers Gurbeer and Harbir have my thanks for interspersing my studies with pleasant diversions which helped to keep me going.

Lastly, I would like to thank the many other people with whom I have interacted during the course of this work, for making my stay at UBC vibrant and enjoyable.

*Dedicated to my parents, for their unconditional love and support.
Their perseverance and hard work allowed me the opportunity
to pursue my aspirations.*

Chapter 1

Introduction

Branching processes are a well studied branch of probability and have been of interest since Galton and Watson first considered the problem of extinction in family names, in the late part of the nineteenth century. The process that Galton and Watson first studied consisted of particles that die after each generation leaving behind an independent and identically distributed number of offspring. If one considers a large Galton-Watson population, then its size can be approximated by what is known as Feller's continuous state branching process. This in turn makes it much simpler to determine various properties of the original Galton-Watson process. The goal in what follows is to construct and study processes that act as analogues of Feller's continuous state branching process for more complicated branching particle systems which have interactions between particles.

A branching particle system (BPS) can essentially be thought of as a Galton-Watson branching process where each particle alive at a given time is also assigned a randomly changing position in space. There are many examples of such systems in biology: consider a single celled bacteria which has the ability to move using its flagellum and which reproduces through binary fission. If each individual bacterium is not affected by the location of the other bacteria, this could be considered a non-interacting BPS. A second example is a larger species called phytoplankton which live in the ocean. Each plankton moves on some random path determined by its internal motor and reproduces (branches) randomly. The plankton spatial position is affected both by individual interactions (e.g. reproductive) with each other as well as by a common drift due to the water current and therefore this system is an example of an interacting BPS.

It can be computationally difficult to simulate a BPS when the population is large since this involves simulating the random paths for each of the particles. Fortunately, if for a BPS the number of particles is large and we observe the system over a large period of time, then it can be approximated by a measure-valued stochastic process. Equivalently, the behavior of a large population, rapidly branching, BPS can be inferred from the properties of the corresponding measure-valued process. A mathematically rigorous result of this type was established independently by Watanabe [25] and later by Dawson [1] for non-interacting BPS, where the associated measure-valued process is a superprocess. The situation for interacting BPS is much more difficult to ascertain since different interactions require different methods of examination. Hence the results in this direction tend to be proven on a case-by-case basis.

Chapters 2 and 3 study branching particle systems with two different types of interactions. We construct the scaling limits of these systems explicitly using something that is known as "Perkins historical stochastic calculus," which is a theory pioneered in the early nineties by Edwin Perkins [13] to study many types of interacting measure-valued processes.

1.1 Branching particle systems and superprocesses

Before mathematically describing an interacting particle system, we define a non-interacting branching particle system. Suppose the spatial motion of a typical particle follows the stochastic process $\{Y_t : t \geq 0\}$, which takes values in a space E (usually E is $\mathbb{R}^d$). Also suppose that when the particle dies it branches into k particles with probability $\nu(k)$. Let Λ be the set of all particles in the particle system. We assume that a particle, labeled by α, born at a (possibly random) time $\tau(\alpha)$, lives until time $d(\alpha)$ whereupon it dies after giving birth to M^α offspring, where $\{M^\alpha : \alpha \in \Lambda\}$ are independent and identically distributed random variables each with law ν. Typically, the lifetime of a particle, $\ell(\alpha) \equiv d(\alpha) - \tau(\alpha)$ is taken to be an independent exponential random variable, or to be constant.

We give ourselves the family of stochastic processes

$$\{Y^{\alpha}(t), \tau(\alpha) \leq t < d(\alpha) : \alpha \in \Lambda\}$$

which will describe the paths of the particles. In particular, for particle α we will insist that $Y^{\alpha}(\tau(\alpha)) = Y^{\alpha'}(d(\alpha')^{-})$, where α' is the parent of α. This simply says that where the parent α' died is where the particle α was born. We will also assume that the particles alive at a given time t move independently of each other.

Then one can define a process Z which counts the number of particles in a region of space at a given time: For A a subset of E, define for $t > 0$,

$$Z_t(A) = \sum_{\alpha \sim t} \delta_{Y^{\alpha}(t) \in A},$$

where $\alpha \sim t$ denotes that particle α is alive at time t. Z_0 is defined by the particles alive at time 0, which we will take to be given.

Assume in the above that Y is a nice process (say a continuous Markov process). Suppose the branching law ν^N has mean $1 + \beta/N$ and variance $\gamma > 0$, where β and γ are fixed numbers. Assume that each labeled particle α is born at an integral multiple of N^{-1} and lives for a time $\ell(\alpha) = N^{-1}$. Suppose also that we are given a finite measure $m(\cdot)$ on E. Then for $A \subset E$ and $t > 0$, let

$$X_t^N(A) = \frac{1}{N} \sum_{\alpha \sim t} \delta_{Y_t^{\alpha} \in A}, \tag{1.1}$$

where $X_0^N(\cdot)$ is given by a Poisson point process with intensity $Nm(\cdot)$. This gives rise to the aforementioned result:

Proposition 1.1.1 (Watanabe, 1968; Dawson, 1975)**.** *There is a measure-valued, Markov process X such that $X_t^N \longrightarrow X_t$, for all t, in distribution with $X_0 = m$.*

The process X is called a (Y, β, γ)-Dawson-Watanabe superprocess. This proposition says that to estimate the probability of most types of events for a large population BPS, one can simply calculate the probability of such events for the corresponding superprocess (which is typically easier).

The value of β determines whether the process is subcritical or supercritical. If $\beta > 0$ then it is supercritical and there is a positive probability that the process will exhibit exponential mass growth and survive indefinitely. If $\beta \leq 0$, the process will die out in finite time almost surely. The case $\beta = 0$ is called critical and $\beta < 0$ is subcritical. One immediate property of X is that it is not an atomic measure, as the X^N were. Indeed, if Y is one dimensional Brownian motion, X_t is almost surely absolutely continuous with Lebesgue measure. In higher dimensions, the support of X_t tends to be fractal-like. For these results and much more, Perkins [14] is a wonderful reference.

The superprocess X_t is deficient in the sense that it only keeps track of the mass alive at time t and ignores any other genealogical information present in the associated branching particle system. To remedy this, we can keep track of not just the position of a particle at time t, but instead the entire path of the particle and those of its ancestors until time t: for $B \subset C([0,\infty), E)$ measurable, define

$$H_t^N(B) = \frac{1}{N} \sum_{\alpha \sim t} \delta_{Y^{\alpha}_{\cdot \wedge t} \in B}.$$

Here, for $t < \tau(\alpha)$, we let Y_t^α be the location of the unique ancestor of α alive at time t. Note: we say $y_{\cdot \wedge t}$ is the path y until time t and fixed thereafter. Then one can adapt Proposition 1.1.1 for a special case to show that

$$H_t^N \longrightarrow H_t \text{ for all } t,$$

in distribution. We call H the (Y, β, γ)-historical process. It is then easy to project H_t down at time t to recover the process X_t (where X is the (Y, β, γ)-superprocess): For $A \subset E$,

$$X_t(A) = \int 1(y_t \in A) H_t(dy).$$

One of the special features of superprocesses that results from the independence of particles alive at a given time is a type of log-laplace equation. This is an equation that can be used to infer properties of the superprocess (and therefore the associated BPS) through the solution of a particular non-linear partial differential equation. More precisely, for certain positive functions $\phi : E \to \mathbb{R}$, one can show that

$$\mathbb{E}\left(e^{-\int \phi dX_t}\right) = e^{-\int V_t \phi dX_0} \tag{1.2}$$

where $V_s \equiv V_s\phi$ is the solution to

$$\frac{\partial V_s}{\partial s} = AV_s + \beta V_s - \frac{\gamma}{2}V_s^2, \tag{1.3}$$

with initial condition $V_0 = \phi$, where A is a second order differential operator called the infinitesimal generator of the process Y.

This connection allows one to calculate probabilities and infer properties of X from the behavior of V, and conversely use properties of X to establish results for the non-linear PDE (1.3) (see for example, Le Gall [11]).

1.2 Superprocesses in a random medium

In the first chapter below, which was published in [8], we construct a family of interacting measure-valued diffusions that live in a random medium by solving a stochastic equation driven by historical Brownian motion. This work is intended to provide oceanographers with a probabilistic model for phytoplankton populations during an algal bloom (colloquially known as a "red tide"). During a bloom, the populations of plankton in a region increase rapidly to the point where one can actually observe them in the water by its reddish tinge. This can have negative consequences higher in the food chain as some of these plankton, such as dinoflagellates, release neurotoxins in the water which can then become concentrated in filter-feeding marine animals. When these animals are consumed by humans it can cause possibly serious illnesses like paralytic shellfish poisoning. As such, it is vital to predict the likelihood of algal blooms taking place at a given location.

Plankton are by definition ocean drifters and range in size from less than a micrometer in width to possibly bigger than a centimeter. The position of a typical organism is affected by the current of the ocean on a large scale. On a smaller scale the position of a plankton is determined by its internal motor and by eddies and ripples of water. As the size of the plankton is typically very tiny, we may think of it as Brown originally thought of the motion of a pollen grain in a glass of water - that it is constantly being bombarded by millions of water molecules which affect its position in a random manner. Thus the motion of a small plankton may be modeled by a diffusion.

The first probabilistic work that could be used to model large plankton populations was found in the Ph.D. dissertation of Wang, published in [23] and [24], though it was not motivated directly by such applications. Dawson, Li and Wang in [2] carry this work further to include a variable branching mechanism. The connection to plankton dynamics was made in the work of Skoulakis and Adler [19]. Young in his survey paper [27] for the oceanographic community, discusses similar models under the guise 'Brownian bugs.' Young et. al. in [28] provide more support for the models such as the one above by showing that the observed 'patchy' distribution of organisms in the water can be explained by the effect of the branching (reproduction) rather than other physical or chemical cues.

The technique used in each of [23], [24], [2] and [19] consists of defining a sequence of interacting branching particle systems and taking a high density limit, as in Proposition 1.1.1. Under this method the particle α follows a path Y_t^α on its life time, where Y^α solves the following stochastic differential equation:

$$Y_t^\alpha = Y_\tau^\alpha + \int_\tau^t \sigma_1(Y_s^\alpha) dB_s^\alpha + \int_\tau^t b(Y_s^\alpha) ds + \int_\tau^t \int \sigma_2(Y_s^\alpha, \xi) dW(s,\xi) \qquad (1.4)$$

where σ_i and b are coefficients satisfying some reasonable conditions and $\tau \equiv \tau(\alpha)$. Here the $\{B^\alpha\}_{\alpha\in\Lambda}$ are an independent family of Brownian motions and W is a space-time white noise term independent of that family. The white noise term is shared across all the particles and therefore acts as a random medium which affects all particles simultaneously (i.e. as the ocean does for plankton) . This also means that the spatial motion of the different particles will depend on each other since they all have a common component driven by the same white noise.

The white noise W is an object arrived at by taking a scaling limit of a space-time lattice of random variables. We divide up space-time in to cubes of side-length N^{-1} and to each of these, we assign an appropriately scaled independent random (spatial) direction. Then for a set A in Euclidean space, $W_t^N(A)$ denotes the sum of all the random directions of all the space-time cubes located inside the set $[0,t] \times A$. We then send N to infinity to get the random direction $W_t(A)$. Hence one can think of white noise as a random field generated by the movement of a very large number of infinitesimally small particles moving about according to Brownian motion (such as what happens in a liquid).

Skoulakis and Adler in [19] assume the white noise W in (1.4) is really a Brownian motion (which is a very specific white noise) whereas Wang and his coauthors all assume that the drift term $b = 0$. Also note that even in the full setting of (1.4), the only interactions between particles are as a result of the common drift due to the random medium. That is, one would expect that for a plankton population there should be some interactions of the plankton with each other, which are not present here.

More is known for the setting where W is a Brownian motion in (1.4) than in the generalized case since it is easier to work with a Brownian motion in place of a white noise. In fact Xiong in [26], by utilizing random duals, shows the existence of a log-laplace equation similar to that of (1.2) where the $V_t\phi$ in that equation is replaced by a solution of a non-linear *stochastic* partial differential equation similar to Equation 1.3. This allows one to perform calculations for the limiting measure-valued process in a manner similar to that used for superprocesses, which is remarkable since as mentioned earlier, superprocesses arise as limits of *non-interacting* BPS's, whereas the BPS presented here has interactions.

The approach employed in Chapter 2 takes an existing historical process K and an independent white noise W and uses them to construct another measure-valued diffusion much in the same way one takes Brownian motion and constructs diffusions by solving stochastic differential equations. In particular, we solve the following stochastic equation: Let K be a critical historical Brownian motion (the process constructed above with Y a Brownian motion and $\beta = 0$), and A measurable. Then define (SE)

$$(a)\ Z_t = Z_0 + \int_0^t \sigma_1(X_s, Z_s)dy(s) + \int_0^t b(X_s, Z_s)ds + \int_0^t \int_{\mathbb{R}^m} \sigma_2(X_s, Z_s, \xi)dW(s,\xi)$$

$$(b)\ X_t(A) = \int 1(Z_t(y) \in A)K_t(dy)$$

One can think of (SE) (a) as a stochastic differential equation describing the motion of a given particle driven by a certain path y and a white noise, and as the natural analogue to (1.4). One major difference is that (SE)(a) includes interactions between individual particles (for example reproductive interactions between plankton) whereas the models considering (1.4) do not. (SE)(b) gives a measure-valued process by integrating over all paths y with respect to K. Since K_t typically puts mass on those paths y that resemble Brownian sample-paths, (SE)(a) is akin to an SDE where y is replaced by a Brownian motion. Note that the white noise W does not depend on y.

The main theorem in this work shows that there exists a unique (strong) solution to (SE):

Theorem 1.2.1. *Under some Lipschitz conditions for the coefficients σ_i and b, there exists a pathwise unique solution (X,Z) to (SE).*

With this theorem in hand, additional properties of X such as the compact support property (which shows that a particle cannot travel too far in finite time), strong Markov property and stability of the solutions are established. One can also show without too much difficulty that a sequence of interacting branching particle systems determined by the formula (1.1), with spatial position for particle α given by (1.4), converge in distribution to the solution X of (SE). In the case where there is no spatial dependence of the coefficients, we recover the results of Wang, Skoulakis and Adler.

1.3 An interacting super Ornstein-Uhlenbeck process

The second chapter of the thesis (forming the basis for the article [9]) is devoted to the construction and examination of a particular class of interacting measure-valued processes with a singular drift called the super Ornstein-Uhlenbeck (SOU) process interacting with its center of mass (COM). What this means is that the BPS associated with this measure-valued process has particles whose spatial motion is described by a modified Ornstein-Uhlenbeck process.

The Ornstein-Uhlenbeck process is one of the simplest stochastic processes – originally introduced in 1930 by Ornstein and Uhlenbeck in [21] to model the velocity V of a particle in a medium with a certain friction coefficient. The process satisfies

$$dV = -\gamma V ds + dB_s,$$

where B is a Brownian motion and $\gamma > 0$. The thought was that since Brownian motion is not differentiable, one cannot talk about the velocity of a Brownian particle and hence it is better to model the velocity by a diffusion process directly. Two of the key properties are that this process is mean-reverting and has a stationary distribution that is a Gaussian random variable.

The interacting particle system corresponding to this process is made up of particles that move in a Brownian fashion, but are attracted to or repelled from the COM of the system. That is, for a particle α in the corresponding particle system, its motion during its lifetime is described by

$$Y_t^\alpha = Y_{\tau(\alpha)}^\alpha + \gamma \int_{\tau(\alpha)}^t (\bar{Y}(s) - Y_s^\alpha)\, ds + \int_{\tau(\alpha)}^t dB_i^\alpha(s), \tag{1.5}$$

where

$$\bar{Y}_t = \frac{1}{\#\{\alpha \sim t\}} \sum_{\alpha \sim t} Y_t^\alpha$$

is the center of mass and $\{B^\alpha : \alpha \in \Lambda\}$ is an independent family of Brownian motions. This constructs a branching modified-OU system Z as follows:

$$Z_t(A) = \sum_{\alpha \sim t} \delta_{Y_t^\alpha \in A}. \tag{1.6}$$

The particles in the system Z exhibit attraction to its COM when $\gamma > 0$ and repulsion when $\gamma < 0$.

The motivation for the model came from a paper of Engländer [3] where he constructs a branching particle system Z' where each particle moves according to (1.5). The system starts with a single particle, and all particles live for one unit of time and die at the integral times leaving behind exactly two offspring particles. It is clear that the particle system will have exactly $2^{\lfloor t \rfloor}$ particles alive at time t. Engländer then proves various theorems about the long term behavior of the particle system and shows that it achieves a natural equilibrium. He proceeds by first showing that the center of mass $\bar{Y}_t$ converges as t approaches infinity and then that for $\phi : \mathbb{R}^d \to \mathbb{R}$ nice,

$$2^{-n} \int \phi(x) Z_n'(dx) \longrightarrow \int \phi(x) P^{\bar{Y}_\infty}(dx), \tag{1.7}$$

as n goes to infinity, almost surely, where $P^{\bar{Y}_\infty}$ is a Gaussian random variable centered at the limiting value of the COM, $\bar{Y}_\infty$.

This particle system is particularly simple in that one does not have to worry about the population of particles going extinct at a given time and the singularity that results in the definition of the center of mass, as is the case for the more general BPS, Z_t. Nevertheless, one can make an analogy between this system and a supercritical BPS on its survival set, as in both situations the systems exhibit exponential population growth. Therefore one expects that on the survival set a general supercritical BPS should exhibit similar long-term behavior as in this binary branching case. Given this insight, it is not surprising that if one can construct the scaling limit of such a sequence of BPS's then it should exhibit some form of equilibrium behavior.

To make this mathematically sound we define (similar to (SE) above) the following stochastic equation: Let K be a supercritical historical Brownian motion, (i.e. $\beta > 0$ in the definition). Then define the process (X,Y) by the solution of the following equation:

$$\text{(SE)}^*(a)\; dY_s = \gamma(\bar{Y}_s - Y_s)ds + dy_s$$
$$(b)\; X_t(A) = \int 1(Y_t(y) \in A) K_t(dy)$$

where $A \subset \mathbb{R}^d$ and the COM is defined as

$$\bar{Y}_s = \frac{\int x X_s(dx)}{\int 1 X_s(dx)}.$$

We proceed to show that $\bar{Y}_t$ converges almost surely and use this fact to prove the main theorem:

Theorem 1.3.1. *If P^x_∞ is the stationary distribution of an Ornstein-Uhlenbeck process with attraction to x at rate $\gamma > 0$, then on the survival set*

$$\frac{X_t(\cdot)}{X_t(1)} \xrightarrow{a.s.} P^{\bar{Y}_\infty}(\cdot),$$

as $t \to \infty$.

The convergence obtained here is in fact even stronger than the one Engländer showed for (1.7) since here we have convergence in measure. The methods used to get this result differ considerably from those of Engländer due to the generalized setting being considered.

The novelty of this work is that it is one of few cases where we can explicitly determine the long-term equilibrium behavior of an interacting measure-valued process. Unlike previous applications of the historical calculus, we must use it here to prove uniqueness in a non-Lipschitz setting.

Chapter 2

Superprocesses with spatial interactions in a random media

2.1 Introduction

We will consider the following stochastic equation:

$$
\begin{aligned}
(SE)\ (a)\ Z_t = Z_0 + \int_0^t \sigma_1(X_s, Z_s) dy(s) \\
+ \int_0^t \int_{\mathbb{R}^m} \sigma_2(X_s, Z_s, \xi) dW(s, \xi) + \int_0^t b(X_s, Z_s) ds \\
(b)\ X_t(A) = \int 1(Z_t(y) \in A) K_t(dy)\ \forall A \in \mathcal{B}(\mathbb{R}^d).
\end{aligned}
$$

The coefficients are such that $\sigma_1 : M_F(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$, $\sigma_2 : M_F(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^{d \times d}$ and $b : M_F(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$, where $M_F(\mathbb{R}^d)$ denotes the space of all non-negative finite measures on $\mathbb{R}^d$. K is a historical super-Brownian motion on some space, y is a path in $\mathbb{R}^d$, and W is a d-dimensional white noise (on $\mathbb{R}_+ \times \mathbb{R}^m$) that is independent of K. The precise definition of a historical process can be found in [14] and the exact meaning of a stochastic integral with respect to y is given below in Section 2.2. Later we impose some Lipschitz-like conditions on b, σ_1 and σ_2 to use a Picard iteration argument. Note that Z takes values in the space of paths in $\mathbb{R}^d$, while X takes values in the space of paths in $M_F(\mathbb{R}^d)$.

One can think of (SE) (a) as a stochastic differential equation describing the motion of a given particle driven by a certain path y and a white noise, and (SE)(b) as the measure-valued process obtained by integrating over all paths y with respect to a historical super-Brownian motion. As mentioned earlier, since K_t typically puts mass on those paths y that resemble Brownian sample-paths, (SE)(a) is like a stochastic differential equation where y is replaced by a Brownian motion. Note that the white noise W is independent of y. Although K_t is not normally atomic, if it were, K_t would keep track of the number of particles alive at time t, as well as the history of each particle's driving Brownian motions until time t.

Dawson et al [2] use a high density limit of a sequence of interacting particle systems, where each of the particles move according to an SDE similar to (SE) (a), to construct a measure-valued diffusion. Earlier, Wang in [24] and [23] constructed special cases of the model of [2]. Using a similar method, the special case where the white noise W is replaced by a Brownian motion B (i.e. setting $m = 0$ here) was studied initially by Skoulakis and Adler [19] in relation to plankton dynamics. Their model also did not incorporate the spatial interactions between particles that are built into (SE).

Before we discuss solutions for (SE), we will need some background material, as well as a framework in which to work. This is provided in Section 2.2. In Section 2.3, we will set up (SE) in a more rigorous manner and then prove the existence and pathwise uniqueness of solutions for a broad range of coefficients σ_1, σ_2, and b. In Section 2.4, we show that the solutions of (SE) satisfy a martingale problem– an extension of the martingale problems that appears in [2] and [13]. We conclude with a section about various additional properties of the solutions, including the strong Markov and compact support properties.

2.2 Preliminaries

For a good reference for much of what follows in this section, consult Chapters II and V of [14].

Let K be an $(\mathcal{F}_t)$-adapted, critical, historical super-Brownian motion on the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

Let $C = C(\mathbb{R}_+, \mathbb{R}^d)$ and $\mathcal{C}$ be its Borel σ-field. Also define $C^t = \{y^t : y \in C\}$

and $\mathcal{C}_t = \sigma(y^s, s \le t, y \in C)$ where $y^t = y_{\cdot \wedge t}$. Define $M_F(C)$ to be the space of finite measures on C. For metric spaces E and E', take $C_c(E,E')$ and $C_b(E,E')$ to be the space of compactly supported continuous functions from E to E' and the space of bounded continuous functions from E to E', respectively.

Notation. For a measure μ on a space E and a measurable function $f : E \to \mathbb{R}$, let $\mu(f) = \int f d\mu$.

Note that K_t takes values in $M_F(C)$, and so $K_t(\cdot)$ will typically mean integration over the y variable below.

We define a martingale problem for K. Let $\hat{B}_t = B(\cdot \wedge t)$ be the path-valued process associated with B, taking values in C^t. Then for $\phi \in b\mathcal{C}$ (bounded and $\mathcal{C}$ measurable), if $s \le t$ let $P_{s,t}\phi(y) = \mathbb{E}^{s,y}(\phi(\hat{B}_t))$, where the right hand side denotes expectation at time t given that until time s, $\hat{B}$ follows the path y. We can now introduce the weak generator, $\hat{A}$, of $\hat{B}$. If $\phi : \mathbb{R}_+ \times C \to \mathbb{R}$ we say $\phi \in \mathcal{D}(\hat{A})$ if and only if ϕ is bounded, continuous and $(\mathcal{C}_t)$-predictable, and for some $\hat{A}_s\phi(y)$ with the same properties as ϕ,

$$\phi(t,\hat{B}) - \phi(s,\hat{B}) - \int_s^t \hat{A}_r\phi(\hat{B})dr, t \ge s,$$

is a $(\mathcal{C}_t)$-martingale under $P_{s,y}$ for all $s \ge 0, y \in C^s$.

Then if $m \in M_F(\mathbb{R}^d)$, we say K satisfies the historical martingale problem, $(HMP)_m$, if and only if $K_0 = m$ a.s. and

$$\begin{aligned}
&\forall \phi \in \mathcal{D}(\hat{A}),\ M_t(\phi) \equiv K_t(\phi_t) - K_0(\phi_0) - \int_0^t K_s(\hat{A}_s\phi)ds \\
&\text{is a continuous } (\mathcal{F}_t)\text{-martingale with } \langle M(\phi)\rangle_t = \int_0^t K_s(\phi_s^2)ds\ \forall t \ge 0, \text{ a.s.}
\end{aligned}$$

The definition of $M_t(\cdot)$ can be extended to form an orthogonal martingale measure using the method of Walsh [22]. Denote by $\mathcal{P}$, the σ-field of $(\mathcal{F}_t)$-predictable sets in $\mathbb{R}_+ \times \Omega$. If $\psi : \mathbb{R}_+ \times \Omega \times C \to \mathbb{R}$ is $\mathcal{P} \times \mathcal{C}$-measurable and

$$\int_0^t K_s(\psi_s^2)ds < \infty\ \forall\ t \ge 0, \tag{2.1}$$

then there exists a continuous local martingale $M_t(\psi)$ with quadratic variation given by $\langle M(\psi)\rangle_t = \int_0^t K_s(\psi_s^2)ds$. If the expectation of the term in (2.1) is finite, then $M_t(\psi)$ is an L^2 martingale.

In this paper, we will require $m(\cdot) \equiv \mathbb{P}(K_0(\cdot))$ to be a finite, positive measure on $\mathbb{R}^d$.

Recall that for a measure space $(E, \mathcal{E})$, the universal completion of $\mathcal{E}$, denoted $\mathcal{E}^*$, is given by

$$\mathcal{E}^* = \cap_\mu \bar{\mathcal{E}}_\mu,$$

where $\bar{\mathcal{E}}_\mu$ denotes the completion of $\mathcal{E}$ with respect to the measure μ and the intersection is over all probability measures μ on E.

Definition. Let $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathcal{F}}_t) = (\Omega \times C, \mathcal{F} \times \mathcal{C}, \mathcal{F}_t \times \mathcal{C}_t)$. Let $\hat{\mathcal{F}}_t^*$ denote the universal completion of $\hat{\mathcal{F}}_t$. If T is a bounded $(\mathcal{F}_t)$-stopping time (denote the set of all such stopping times by $\mathcal{T}_b$), the normalized Campbell measure associated with T is the measure $\hat{\mathbb{P}}_T$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ given by

$$\hat{\mathbb{P}}_T(A \times B) = \mathbb{P}(\mathbf{1}_A K_T(B))m(1)^{-1} \text{ for } A \in \mathcal{F}, B \in \mathcal{C}.$$

We denote sample points in $\hat{\Omega}$ by (ω, y). Hence, under $\hat{\mathbb{P}}_T$, ω has law $K_T(1)d\mathbb{P}$ and conditionally on ω, y has law $K_T(\cdot)/K_T(1)$.

Note that here $\hat{\mathbb{P}}_T$ is a probability measure, since

$$\hat{\mathbb{P}}_T(1) = \mathbb{P}(K_T(1))m(1)^{-1} = \mathbb{P}(K_0(1))m(1)^{-1} = 1$$

as $K_t(1)$ is Feller's critical branching diffusion (a uniformly integrable martingale). To avoid carrying constants of $m(1)^{-1}$ from line to line, we will without loss of generality assume $m(1) = 1$.

The following results will be useful later. The first is Proposition V.2.4 and the second, Proposition V.3.1 of [14].

Proposition 2.2.1. *Assume that $T \in \mathcal{T}_b$ and $\psi \in \hat{\mathcal{F}}_T$, bounded. Then*

$$K_t(\psi) = K_T(\psi) + \int_T^t \int \psi(y)dM(s,y) \ \forall t \geq T \ \mathbb{P}\text{-a.s.}$$

Proposition 2.2.2. *If $T \in \mathfrak{T}_b$ then under $\hat{\mathbb{P}}_T$, y is an $(\hat{\mathcal{F}}_t)$-adapted Brownian motion stopped at T.*

Suppose that $\{W(t,\xi), \xi \in \mathbb{R}^m, t \geq 0\}$ is an $(\mathcal{F}_t)$-adapted, d-dimensional Brownian sheet on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ that is independent of K. For a function $\phi \in C_c(\mathbb{R}^m, \mathbb{R}^{n\times d})$, let $W_t(\phi) = \int_0^t \int_{\mathbb{R}^m} \phi(\xi) dW(s,\xi)$. This defines the associated white noise process of W, and hence we will identify the two in what follows. The next theorem will be useful in proving that W is also an $(\hat{\mathcal{F}}_t)$-adapted $\hat{\mathbb{P}}_T$ Brownian sheet for each $T \in \mathfrak{T}_b$.

Theorem 2.2.3. *Suppose N is a real-valued, square-integrable, $(\mathcal{F}_t)$-adapted martingale that is independent of K. Then for each $T \in \mathfrak{T}_b$, N is an $(\hat{\mathcal{F}}_t)$-adapted martingale under $\hat{\mathbb{P}}_T$ and $\langle N, y^j\rangle_t = 0$ $\forall t \geq 0$ $\hat{\mathbb{P}}_T$-a.s $\forall$ $j \leq d$.*

Proof. Let $A \in \mathcal{C}_s$ and $B \in \mathcal{F}_s$. Then

$$\begin{aligned}
&\hat{\mathbb{P}}_T[(N_t - N_s)\mathbf{1}_A \mathbf{1}_B] \\
&= \mathbb{P}[(N_t - N_s)K_T(A)\mathbf{1}_B] \\
&= \mathbb{P}[\mathbf{1}_B \mathbf{1}_{T\leq s} K_T(A)\mathbb{E}(N_t - N_s|\mathcal{F}_s)] + \mathbb{P}[(N_t - N_s)\mathbf{1}_B \mathbf{1}_{T>s} K_T(A)] \\
&= \mathbb{P}\left[(N_t - N_s)\mathbf{1}_B \mathbf{1}_{T>s}\left(K_s(A) + \int_s^T \int \mathbf{1}_A dM(r,y)\right)\right] \\
&= \mathbb{P}\left[(N_t - N_s)\mathbf{1}_B \mathbf{1}_{T>s}\left(K_s(A) + \int_s^{T\wedge t} \int \mathbf{1}_A dM(r,y)\right)\right]
\end{aligned}$$

where in the first equality we use the definition of the Campbell measure and in the second that $K_T(A)1_{\{T\leq s\}} \in \mathcal{F}_s$. For the third equality, we use the fact that N_t is an $(\mathcal{F}_t)$-martingale and Proposition 2.2.1. Finally we condition on $\mathcal{F}_t$ and appeal to the Optional Sampling Theorem for the last expression. By conditioning first on $\mathcal{F}_s$ and then on $\mathcal{F}_T$ we can reduce the above as

$$\begin{aligned}
\hat{\mathbb{P}}_T[(N_t - N_s)\mathbf{1}_A \mathbf{1}_B] &= \mathbb{P}\left[(N_t - N_s)\mathbf{1}_{B'} \int_s^{T\wedge t} \int \mathbf{1}_A dM(r,y)\right] \\
&= \mathbb{P}\left[(N_{T\wedge t} - N_s)\mathbf{1}_{B'} \int_s^{T\wedge t} \int \mathbf{1}_A dM(r,y)\right]
\end{aligned}$$

where $B' = B \cap \{T > s\} \in \mathcal{F}_s$. Now, if we show that $N_t M_t(\psi)$ is an $(\mathcal{F}_t)$-adapted martingale for any bounded function $\psi : \mathbb{R}_+ \times C \to \mathbb{R}$, then we may apply the

Optional Stopping Theorem to show that the above expression reduces to zero. Note that,

$$
\begin{aligned}
N_t M_t(\psi) &= \int_0^t N_s dM_s(\psi) + \int_0^t M_s(\psi) dN_s + \langle N, M(\psi) \rangle_t \\
&= \int_0^t N_s dM_s(\psi) + \int_0^t M_s(\psi) dN_s
\end{aligned}
$$

since $\langle N, M(\psi) \rangle_t = 0$ by independence of N and M. It also follows from the independence that the first term on the right side is a martingale. For the second term we use independence and the square integrability of N to conclude that it is a martingale as well. Hence $N_t M_t(\psi)$ is a martingale and $\hat{\mathbb{P}}_T[(N_t - N_s)\mathbf{1}_A \mathbf{1}_B] = 0$. This proves the first assertion.

From the independence of N and K we have for constant T, a Borel subset A of $C(\mathbb{R}_+, \mathbb{R})$ and $B \in \mathcal{C}$ that,

$$
\begin{aligned}
\hat{\mathbb{P}}_T(N \in A, y \in B) &= \mathbb{P}(\mathbf{1}_{N \in A} K_T(\mathbf{1}_{y \in B})) \\
&= \mathbb{P}(\mathbf{1}_{N \in A}) \mathbb{P}(K_T(\mathbf{1}_{y \in B})) \\
&= \hat{\mathbb{P}}_T(N \in A) \hat{\mathbb{P}}_T(y \in B).
\end{aligned}
$$

Hence, N and y are independent under $\hat{\mathbb{P}}_T$. This implies that $\langle N, y^j \rangle = 0, \hat{\mathbb{P}}_T$-a.s. In the general case where $T \in \mathcal{T}_b$, independence of N and y under $\hat{\mathbb{P}}_T$ no longer necessarily holds. Let $Z \equiv \langle N, y^j \rangle$. Then under $\hat{\mathbb{P}}_T$, for $u > T$ fixed,

$$
\begin{aligned}
\hat{\mathbb{P}}_T\left(\sup_{s \le T} |Z_s| \wedge 1\right) &= \mathbb{P}\left(K_T\left(\sup_{s \le T} |Z_s| \wedge 1\right)\right) \\
&= \mathbb{P}\left(K_u\left(\sup_{s \le T} |Z_s| \wedge 1\right) - \int_T^u \int \sup_{s \le T} |Z_s| \wedge 1 dM(r, y)\right) \\
&= 0
\end{aligned}
$$

In the second line, we have noted that $\sup_{s \le T} |Z_s| \wedge 1$ is bounded, $\hat{\mathfrak{F}}_T$-measurable and applied Proposition 2.2.1. To get the third line, we use the fact that $Z = 0$, $\hat{\mathbb{P}}_u$-a.s. by the special case considered above and that the second term is a martingale. Hence $\langle N, y^j \rangle_t = 0 \ \forall t \ge 0 \ \hat{\mathbb{P}}_T$-a.s. $\forall T \in \mathcal{T}_b$ (note that we have used the property that under $\hat{\mathbb{P}}_T$, $y = y^T$ a.s. implicitly here). $\square$

Let $\mathcal{R}^m$ denote the Borel σ-algebra of $\mathbb{R}^m$. The following proposition is needed to complete the proof.

Proposition 2.2.4. *Let M be an orthogonal martingale measure on $(\mathbb{R}^m, \mathcal{R}^m)$ as defined in Chapter 2 of Walsh [22]. Suppose further that for each $A \in \mathcal{R}$, $t \mapsto M_t(A)$ is continuous. Then M is a white noise if and only if its covariance measure is given by Lebesgue measure.*

For the proof, see Proposition 2.10 of Walsh [22].

Corollary 2.2.5. *For each $T \in \mathcal{T}_b$, $\{W(t,\xi), \xi \in \mathbb{R}^m, t \geq 0\}$ is an $(\hat{\mathcal{F}}_t)$-adapted Brownian sheet under $\hat{\mathbb{P}}_T$. Furthermore, W and y are orthogonal under $\hat{\mathbb{P}}_T$; that is, $\langle W(\phi), y^j \rangle_t = 0 \; \forall \phi \in C_c(\mathbb{R}^m, \mathbb{R}^d), t \geq 0$ $\hat{\mathbb{P}}_T$-a.s. $\forall T \in \mathcal{T}_b$ for all $j \leq d$.*

Proof. Let $W_t(A) = (W_t^1(A), \dots, W_t^d(A))$. We simply need to verify that under $\hat{\mathbb{P}}_T$, $\{W_t^i\}$ are independent white noises. Each W_t^i is an orthogonal martingale measure under $\hat{\mathbb{P}}_T$ since for disjoint bounded sets $A, B \in \mathcal{R}^m$, $W_t^i(A)$ and $W_t^i(A)W_t^i(B)$ are square-integrable $(\mathcal{F}_t)$-martingales with respect to $\mathbb{P}$, hence also $(\hat{\mathcal{F}}_t)$-martingales with respect to $\hat{\mathbb{P}}_T$ by Theorem 2.2.3. The map $t \mapsto W_t^i(A)$ is continuous since $W_t^i(A)(\omega, y)$ defined on $\hat{\Omega}$ is equal to $W_t^i(A)(\omega)$ on Ω, for which the analogous map is continuous and $\hat{\mathbb{P}}_T|_\Omega$ is absolutely continuous with respect to $\mathbb{P}$.

By noting that Theorem 2.2.3 implies $W_t^i(A)^2 - t\nu(A)$ is an $(\hat{\mathcal{F}}_t)$-martingale, where ν is the m-dimensional Lebesgue measure, we see that the covariance measure is deterministic. Hence each W_t^i is a white noise process by Proposition 2.2.4. Then by a final use of Theorem 2.2.3, we see that $W_t^i(A)W_t^j(A) - t\nu(A)\delta_{i,j}$ is an $(\hat{\mathcal{F}}_t)$-martingale where $\delta_{i,j}$ is Kronecker's δ. Therefore $W^i(A)$ and $W^j(A)$ are independent Brownian motions (for any A with $\nu(A) < \infty$). Hence W_t is a d-dimensional white noise process.

The orthogonality of y, W follows from the previous theorem.

□

Note that this corollary will still hold if W is a white noise on a space $E \times \mathbb{R}_+$ with covariance measure $\nu \times \ell$ where ν is an arbitrary σ-finite measure on E and ℓ is the one dimensional Lebesgue measure. Indeed the results below all extend to this generality as well.

For a matrix $A \in \mathbb{R}^{n\times d}$, let $||A||^2 = \sum_{i,j} A_{ij}^2$. Note for a vector $x \in \mathbb{R}^n$, we will use $|x|^2$ instead to denote the same quantity.

Definition. We say $f \in D_1(n,d)$ iff $f : \mathbb{R}_+ \times \hat{\Omega} \to \mathbb{R}^{n\times d}$ is $(\hat{\mathfrak{F}}_t^*)$-predictable and

$$\int_0^t \|f(s,\omega,y)\|^2 ds < \infty \ K_t - a.a.\ y\ \forall t \geq 0, \mathbb{P}\ a.s.$$

We say $g \in D_2(m,n,d)$ iff $g : \mathbb{R}_+ \times \mathbb{R}^m \times \hat{\Omega} \to \mathbb{R}^{n\times d}$ is $(\hat{\mathfrak{F}}_t^*)$-predictable and

$$\int_0^t \int_{\mathbb{R}^m} \|g(s,\xi,\omega,y)\|^2 d\xi ds < \infty \ K_t - a.a.\ y\ \forall t \geq 0, \mathbb{P}\ a.s.$$

Definition. If $X, Y : \mathbb{R}_+ \times \hat{\Omega} \to E$, we say $X = Y$ K-a.e. iff $X(s,\omega,y) = Y(s,\omega,y)$ for all $s \leq t$, K_t-a.e. for all $t \geq 0$ $\mathbb{P}$-a.s. If E is a metric space we say that X is continuous K-a.e. iff $s \mapsto X(s,\omega,y)$ is continuous on $[0,t]$ for K_t-a.a. y for all $t \geq 0$ $\mathbb{P}$-a.s.

Since with respect to each measure $\hat{\mathbb{P}}_T, T \in \mathfrak{T}_b$, $\{W(s,\xi), 0 \leq s\}$ is an $(\hat{\mathfrak{F}}_s)$-adapted Brownian sheet, we may define the stochastic integral of a function f with $\int_0^t \int_{\mathbb{R}^d} \|f(\xi,s,\omega,y)\|^2 d\xi ds < \infty$, $\hat{\mathbb{P}}_T$-a.s. as in Walsh [22]. However as this integral depends on T, we denote it by $\hat{\mathbb{P}}_T$-$\int_0^t \int f(s,\xi,\omega,y) dW(s,\xi)$.

Similarly, define $\hat{\mathbb{P}}_T$-$\int_0^t g(s,\omega,y)dy(s)$ to be the stochastic integral g with respect to y under $\hat{\mathbb{P}}_T$. The following part of Proposition V.3.2 of [14] is useful in relating these different stochastic integrals.

Proposition 2.2.6. *For $g \in D_1(n,d)$, there exists an $(\hat{\mathfrak{F}}_t)$-predictable, K-a.e. continuous process $I(g,t,\omega,y)$ such that*

$$I(g, t\wedge T, \omega, y) = \hat{\mathbb{P}}_T\text{-}\int_0^t g(s,\omega,y)dy(s) \text{ for all } t \geq 0, \hat{\mathbb{P}}_T\text{-a.s. for all } T \in \mathfrak{T}_b.$$

Moreover, I is unique. That is, if I' is an $(\hat{\mathfrak{F}}_t^*)$-predictable process satisfying the above, then $I(g,s,\omega,y) = I'(g,s,\omega,y)$ K-a.e.

We can prove a similar theorem for $\hat{\mathbb{P}}_T$-$\int_0^t \int f(s,\xi,\omega,y) dW(s,\xi)$, but we will need to recognize that it only holds for $t \leq T$ as the following example demonstrates.

Example 2.2.7. Consider $f(s,\xi,\omega,y) = y^u(s)\eta(\xi)\mathbf{1}_{(u,v]}(s)$ with $\eta \in C_c^2(\mathbb{R}^m)$. Note that under $\hat{\mathbb{P}}_T$, $y = y^T$ a.s., and hence

$$\hat{\mathbb{P}}_T\text{-}\int_0^t \int f(s,\xi,\omega,y)dW(s,\xi) = y^T(u)\left(W_{v\wedge t}(\eta) - W_{u\wedge t}(\eta)\right).$$

This expression clearly depends on the stopping time T. If $T = T_1 < u$ is fixed then our integral is $y(T_1)\left(W_{v\wedge t}(\eta) - W_{u\wedge t}(\eta)\right)$, else if $T = T_2 > u$ is fixed, then the integral is $y(u)\left(W_{v\wedge t}(\eta) - W_{u\wedge t}(\eta)\right)$. If $t < T_1$, then the integrals in both cases agree (and are equal to zero).

This shows that even simple functions do not necessarily have unique integrals when integrated with respect to W under the different measures $\hat{\mathbb{P}}_T$ when $t > T$. The next theorem is the analogue of Proposition 2.2.6 and shows how the stochastic integrals of W are related under different measures $\hat{\mathbb{P}}_T$.

Theorem 2.2.8. **(a)** *If $f \in D_2(m,n,d)$, there is an $\mathbb{R}^n$-valued $(\hat{\mathfrak{F}}_t)$-predictable process $J(f,t,\omega,y)$ such that for all $T \in \mathfrak{T}_b$*

$$J(f,t\wedge T,\omega,y) = \hat{\mathbb{P}}_T\text{-}\int_0^{t\wedge T}\int_{\mathbb{R}^m} f(s,\xi,\omega,y)dW(s,\xi)\ \forall t \geq 0, \hat{\mathbb{P}}_T\text{-a.s.}$$

(b) *If $J'(f)$ is an $(\hat{\mathfrak{F}}_t^*)$-predictable process satisfying (a), then $J(f,s) = J'(f,s)$ K-a.e.*

(c) *$J(f,t)$ is continuous in t, K-a.e.*

(d) *(Dominated Convergence) For any $N > 0$, if $f_k, f \in D_2(m,n,d)$ satisfy*

$$\lim_{k\to\infty}\mathbb{P}\left(K_N\left(\int_0^N\int_{\mathbb{R}^m}\|f_k(s,\xi) - f(s,\xi)\|^2 d\xi ds > \varepsilon\right)\right) = 0, \forall\varepsilon > 0,$$

then,

$$\lim_{k\to\infty}\mathbb{P}\left(\sup_{t\leq N}K_t\left(\sup_{s\leq t}\|J(f_k,s) - J(f,s)\|^2 > \varepsilon\right)\right) = 0\ \forall\varepsilon > 0.$$

(e) *For any $S \in \mathfrak{T}_b$ if $f_k, f \in D_2(m,n,d)$ satisfy*

$$\lim_{k\to\infty} \mathbb{P}\left(K_S\left(\int_0^S \int_{\mathbb{R}^m} \|f_k(s,\xi) - f(s,\xi)\|^2 d\xi ds\right)\right) = 0,$$

then,

$$\sup_{t\le S} K_t\left(\sup_{s\le t} \|J(f_k,s) - J(f,s)\|^2\right) \xrightarrow{\mathbb{P}} 0 \text{ as } k \to \infty.$$

The proof of Proposition 2.2.6 can be easily adapted to prove this theorem. We will at times write $\int_0^t \int f(s,\xi,\omega,y) dW(s,\xi)$ or $W_t(f)$ in place of $J(f,t,\omega,y)$ where the latter is well defined.

Corollary 2.2.9. *Let $T \in \mathfrak{T}_b$. If $f \in D_1(n,d)$, $g \in D_2(m,n,d)$ and there exists an $(\hat{\mathfrak{F}}_t^*)$-predictable process $S(f,g)$ satisfying*

$$S(f,g,t\wedge T,\omega,y) = \hat{\mathbb{P}}_T\text{-}\int_0^{t\wedge T} f(s,y)dy(s) + \hat{\mathbb{P}}_T\text{-}\int_0^{t\wedge T}\int g(s,\xi)dW(s,\xi) \qquad (2.2)$$

$\forall t \ge 0, \hat{\mathbb{P}}_T$-a.s., then $S(f,g) = I(f) + J(g), K$-a.e.

Proof. Define

$$L(t,\omega) = \int \sup_{s\le t} \|I(f,s) + J(f,s) - S(f,g,s)\| \wedge 1 K_t(dy). \qquad (2.3)$$

Note that L is $(\mathfrak{F}_t)$-predictable. The proof of this is essentially given in the last half of the proof for Proposition V.3.2(b) of [14]. Assume T is a bounded, predictable stopping time. Then

$$\begin{aligned}\mathbb{P}(L(T,\omega)) &= \hat{\mathbb{P}}_T\left(\sup_{s\le T} \|I(f,s) + J(f,s) - S(f,g,s)\| \wedge 1\right) \\ &= 0\end{aligned}$$

since $S(f,g,T\wedge s) = \hat{\mathbb{P}}_T\text{-}\int_0^{s\wedge T} f(s,y)dy(s) + \hat{\mathbb{P}}_T\text{-}\int_0^{s\wedge T}\int g(s,\xi)dW(s,\xi)$, $I(f,s\wedge T) = \hat{\mathbb{P}}_T\text{-}\int_0^{s\wedge T} f(s,\omega,y)dy(s)$ by Proposition 2.2.6 and since $J(g,s\wedge T) = \hat{\mathbb{P}}_T\text{-}\int_0^{s\wedge T}\int g(s,\xi,\omega,y)dW(s,\xi)$ by Theorem 2.2.8(a) above. Then by the Section Theorem, we have that $L(t,\omega) = 0\ \forall t \ge 0$ a.s.

□

Notation. If $X(t) = (X_1(t), \dots, X_n(t))$ is an $\mathbb{R}^n$-valued process on $(\hat{\Omega}, \hat{\mathcal{F}})$ and $\mu \in M_F(C)$, let $\mu(X_t) = (\mu(X_1(t)), \dots, \mu(X_n(t)))$ where $\mu(X_i(t)) = \int X_i(t, \omega, y)\mu(dy)$. Also let

$$\int_0^t \int X(s)dM(s,y) = \left(\int_0^t \int X_1(s)dM(s,y), \dots, \int_0^t \int X_n(s)dM(s,y) \right)$$

whenever these integrals are defined. We do the same for stochastic integrals with respect to W.

The next theorem is needed in order to prove a version of Itô's Lemma.

Theorem 2.2.10. *If $f \in D_2(m,n,d)$ and $\sup_{s,\omega,y} \int \|f(s,\xi,\omega,y)\|^2 d\xi < \infty$ then*

$$\begin{aligned} &K_t(J(f,t)) \\ &= \int_0^t \int W_s(f(s))dM(s,y) + \int_0^t \int K_s(f(s))dW(s,\xi) \ \forall t \geq 0 \ \mathbb{P}\text{-}a.s. \end{aligned} \tag{2.4}$$

and each integral on the right is a continuous L^2 $(\mathcal{F}_t)$-martingale.

Proof. To simplify the notation, assume that $n = d = 1$. We first show the result for simple functions. Let

$$f(s,\xi,\omega,y) = \phi_1(\omega)\phi_2(y)\phi_3(\xi)\mathbf{1}_{(u,v]}(s) \tag{2.5}$$

where $\phi_1 \in b\mathcal{F}_u, \phi_2 \in b\mathcal{C}_u, \phi_3 \in C_c^\infty(\mathbb{R}^m)$ and $0 \leq u < v$. Then,

$$\begin{aligned} K_t\left(\int_0^t \int f(s,\xi,\omega,y)dW(s,\xi) \right) &= K_t\left(\phi_1(\omega)\phi_2(y)(W_{v\wedge t}(\phi_3) - W_{u\wedge t}(\phi_3))\right) \\ &= \phi_1(\omega)(W_{v\wedge t}(\phi_3) - W_{u\wedge t}(\phi_3))K_t(\phi_2) \\ &= \int_0^t \int \phi_1(\omega)\phi_3(\xi)\mathbf{1}_{(u,v]}(s)dW(s,\xi)K_t(\phi_2) \end{aligned}$$

where in the second line we make use of the fact that W_t and ϕ_1 depend only on ω. We can now apply the integration by parts formula, while noting that

$\langle W(\phi_3), K(\phi_2)\rangle_t = 0$ (by independence), to get

$$K_t\left(\int_0^t \int f(s,\xi)dW(s,\xi)\right) = \int_0^t \int \phi_1(\omega)\phi_3(\xi)\mathbf{1}_{(u,v]}(s)K_s(\phi_2)dW(s,\xi)$$
$$+\int_u^t \int_u^s \int \phi_1(\omega)\phi_3(\xi)\mathbf{1}_{(u,v]}(r)dW(r,\xi)dK_s(\phi_2)$$
$$= \int_0^t \int K_s(f(s,\xi))dW(s,\xi)$$
$$+\int_0^t \left[\int_0^s \int \phi_1(\omega)\phi_3(\xi)\mathbf{1}_{(u,v]}(r)dW(r,\xi)\right]\int \phi_2(y)dM(s,y)$$
$$= \int_0^t \int K_s(f(s))dW(s,\xi) + \int_0^t \int \left[\int_0^s \int f(r,\xi,\omega,y)dW(r,\xi)\right]dM(s,y).$$

Here in the second line we have used Proposition 2.2.1 with $T=u$ and then the fact that the function $\phi_1(\omega)\phi_3(\xi)\mathbf{1}_{(u,v]}(s)$ disappears for $s<u$. The differential $\int \phi_2(y)dM(s,y)$ is just $dM_s(\phi_2)$.

Note that for these simple functions f, we need only look at the square function of $\int_0^t \int K_s(f(s,\xi))dW(s,\xi)$ to show that it is an L^2 martingale. That is, since $\phi_1\phi_2$ is bounded and $\phi_3 \in C_c^\infty(\mathbb{R}^m)$,

$$\mathbb{P}\left(\int_0^t \int K_s(f(s,\xi))^2 d\xi ds\right) \leq c\mathbb{P}\left(\int_0^t \int \phi_3(\xi)^2 K_s(1)^2 d\xi ds\right).$$

This expectation is finite because $K_t(1)$ is Feller's branching diffusion.

To show that $\int_0^t \int J(f,s)dM(s,y)$ is an L^2 martingale consider

$$\mathbb{P}\left(\int_0^t K_s\left(J(f,s)^2\right)ds\right) = \int_0^t \mathbb{P}\left(K_s\left(\phi_2^2\right)\left[\int_0^s \int \phi_1\phi_3\mathbf{1}_{(u,v]}dW(r,\xi)\right]^2\right)ds$$
$$\leq c\int_0^t \mathbb{P}\left(K_s(1)\,\mathbb{P}\left[\int_0^s \int (\phi_1\phi_3\mathbf{1}_{(u,v]})^2 d\xi dr\right]\right)ds$$
$$< \infty.$$

Hence the sum of these two quantities is also an L^2 martingale. Since $M_t(\phi_2)$ is continuous, $\int_0^t \int J(f,s)dM(s,y)$ is continuous. As $\int_0^t \int K_s(f(s,\xi))dW(s,\xi)$ is continuous so is $K_t\left(\int_0^t \int f(s,\xi,\omega,y)dW(s,\xi)\right)$.

Suppose the statement of the theorem holds for a sequence of $(\hat{\mathfrak{F}}_t)$-predictable processes f_k and f is an $(\hat{\mathfrak{F}}_t^*)$-predictable process such that $f_k \to f$ pointwise, $\sup_k \int \|f_k(\xi)\|^2 d\xi \vee \int \|f(\xi)\|^2 d\xi < \infty$ and

$$\lim_{k\to\infty} \mathbb{P}\left(\int\int_0^N \int (f_k(s,\xi)-f(s,\xi))^2 d\xi ds K_N(dy)\right) = 0 \;\forall N \in \mathbb{N}. \tag{2.6}$$

Since $J(f,s)$ is a White noise integral with respect to $\hat{\mathbb{P}}_s$, consider for each $N \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}\left(\langle M(J(f,s)-J(f_k,s))\rangle_N\right) &= \mathbb{P}\left(\int_0^N \int (J(f_k,s)-J(f,s))^2 K_s(dy)ds\right) \\
&= \int_0^N \hat{\mathbb{P}}_s\left((J(f_k,s)-J(f,s))^2\right) ds \\
&= \int_0^N \hat{\mathbb{P}}_s\left(\int_0^s \int (f_k(r,\xi)-f(r,\xi))^2 d\xi dr\right) ds \\
&= \int_0^N \hat{\mathbb{P}}_N\left(\int_0^s \int (f_k(r,\xi)-f(r,\xi))^2 d\xi dr\right) ds
\end{aligned}$$

where in the second line we have used Fubini's theorem and the definition of the Campbell measure $\hat{\mathbb{P}}_s$. In the third line we make use of the Itô isometry and then finally we use Remark V.2.5 (d) of [14] in the last line. The last expression now goes to zero as $k \to \infty$ by (2.6) (and Fubini's theorem). This shows that $\int_0^t \int J(f,s,y) dM(s,y)$ is a continuous L^2 martingale. Now by the Dominated Convergence Theorem, $\forall t$ as $k \to \infty$,

$$\begin{aligned}
\mathbb{P}\big(\langle J(K_s(f(s))) - J(K_s(f_k(s)))\rangle_t\big) &= \mathbb{P}\left(\int_0^t \int K_s(f(s,\xi)-f_k(s,\xi))^2 d\xi ds\right) \\
&\to 0,
\end{aligned}$$

and so $\int_0^t \int K_s(f(s,\xi)) dW(s,\xi)$ is a continuous L^2 martingale as well.

By using Theorem 2.2.8(e) with $S = N$ and (2.6), we get

$$\sup_{t\le N} K_t(|J(f_k,t)-J(f,t)|) \xrightarrow{\mathbb{P}} 0 \; as \; k \to \infty, \forall N \in \mathbb{N}.$$

Sending $k \to \infty$ now shows that (2.4) holds for f as well.

The rest of the proof proceeds by appealing to a Monotone Class Theorem to pass to the bounded pointwise closure of functions satisfying (2.4). This is exactly what is done in the last part of the proof of Proposition V.3.4 of [14].

□

Theorem 2.2.11 (Itô's Lemma). *Let Z_0 be $\hat{\mathfrak{F}}_0$-measurable and take values in $\mathbb{R}^n$. Let $f \in D_1(n,d)$, $g \in D_2(m,n,d)$, h an $\mathbb{R}^n$-valued $(\hat{\mathfrak{F}}^*_t)$-predictable process and $\psi \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$. Assume*

$$\int_0^t K_s(\|f_s\|^2 + |h_s|)ds + \int_0^t K_s\left(\int \|g(s,\xi)\|^2 d\xi\right) ds < \infty, \ \forall t \ a.s., \tag{2.7}$$

and let

$$\begin{aligned} Z_t(\omega,y) &= Z_0(\omega,y) \\ &\quad + \int_0^t f(s,\omega,y)dy(s) + \int_0^t \int g(s,\xi,\omega,y)dW(s,\xi) + \int_0^t h(s,\omega,y)ds, \end{aligned} \tag{2.8}$$

then

$$\begin{aligned} &\int \psi(t,Z_t)K_t(dy) = \int \psi(0,Z_0)dK_0(y) \\ &+ \int_0^t \int \psi(s,Z_s)dM(s,y) + \int_0^t \int K_s(\nabla\psi(s,Z_s)g(s,\xi))dW(s,\xi) \\ &\quad + \int_0^t K_s\left(\frac{\partial\psi}{\partial s}(s,Z_s) + \nabla\psi(s,Z_s)\cdot h_s + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \psi_{i,j}(s,Z_s)a_{ij}(s)\right)ds, \end{aligned} \tag{2.9}$$

where $\nabla\psi$ and ψ_{ij} are the gradient and second order partial derivatives in the spatial variables and $a \equiv ff^* + \int gg^*(\xi)d\xi$. The second term on the right is an L^2 martingale, the third is a local martingale and the last term on the right has continuous paths with finite variation over compact intervals a.s. Therefore $\int \psi(t,Z_t)K_t(dy)$ is a continuous $(\mathfrak{F}_t)$-semimartingale.

Proof. For now assume that $\|f\|$ and $|h|$ are bounded and that $\sup_{s,y}\int \|g(s,\xi,\omega,y)\|^2 d\xi < \infty$. Let $T \in \mathfrak{T}_b$ and $\bar{Z}_t = (t,Z_t)$. Using the classical Itô's Lemma we have, $\hat{\mathbb{P}}_T$-a.s., $\forall t \geq 0$

$$
\begin{aligned}
&\psi(\bar{Z}_{T\wedge t}) - \psi(\bar{Z}_0) = \\
&\int_0^{T\wedge t} \nabla\psi(\bar{Z}_s)\cdot dZ_s + \int_0^{T\wedge t} \frac{\partial\psi}{\partial s}(\bar{Z}_s)ds + \frac{1}{2}\sum_{i,j\le n}\int_0^{T\wedge t}\psi_{ij}(\bar{Z}_s)d\langle Z^i, Z^j\rangle_s \\
&= \hat{\mathbb{P}}_T\text{-}\int_0^{T\wedge t}\nabla\psi(\bar{Z}_s)f(s)\cdot dy(s) + \hat{\mathbb{P}}_T\text{-}\int_0^{T\wedge t}\int \nabla\psi(\bar{Z}_s)g(s,\xi)\cdot dW(s,\xi) \\
&+ \int_0^{T\wedge t}\left(\nabla\psi(\bar{Z}_s)\cdot h(s) + \frac{\partial\psi}{\partial s}(\bar{Z}_s) + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \psi_{i,j}(\bar{Z}_s)a_{ij}(s)\right)ds.
\end{aligned}
$$

Here in the second line we have simply plugged in (2.8), used Corollary 2.2.5 to manage the square function and done some linear algebra. Letting $\tilde{b}(s)$ denote the term inside the last integral above, we see that

$$S(t) = \psi(\bar{Z}_t) - \psi(\bar{Z}_0) - \int_0^t \tilde{b}(s)ds$$

is an $(\hat{\mathfrak{F}}_t^*)$-predictable process satisfying (2.2). Hence we apply Corollary 2.2.9 (with f and g replaced by $\nabla\psi(\bar{Z})f$ and $\nabla\psi(\bar{Z})g$ respectively) to get

$$\psi(\bar{Z}_t) = \psi(\bar{Z}_0) + I(\nabla\psi(\bar{Z})f, t) + J(\nabla\psi(\bar{Z})g, t) + \int_0^t \tilde{b}(s)ds. \tag{2.10}$$

Since $|\tilde{b}|, \|\nabla\psi(\bar{Z})f\|$ and $\int\|\nabla\psi(\bar{Z})g(\xi)\|^2 d\xi$ are all bounded and $(\hat{\mathfrak{F}}_t^*)$-predictable, we can apply Theorem 2.2.10 above, Proposition V.3.4 and Remarks V.2.5 of [14] to get $\mathbb{P}$-a.s. for all $t \geq 0$,

$$\int \psi(\bar{Z}_t)K_t(dy) = \int \psi(\bar{Z}_0)K_0(dy) + \int_0^t \int \psi(\bar{Z}_0)dM(s,y)$$
$$+ \int_0^t \int I(\nabla\psi(\bar{Z})f,s)dM(s,y) + \int_0^t \int J(\nabla\psi(\bar{Z})g,s)dM(s,y)$$
$$+ \int_0^t \int \left[\int_0^s \tilde{b}(r)dr\right] dM(s,y) + \int_0^t \int K_s(\nabla\psi(\bar{Z}_s)g(s,\xi))dW(s,\xi)$$
$$+ \int_0^t K_s(\tilde{b}(s))ds$$
$$= \int \psi(\bar{Z}_0)K_0(dy) + \int_0^t \int \psi(\bar{Z}_s)dM(s,y)$$
$$+ \int_0^t \int K_s(\nabla\psi(\bar{Z}_s)g(s,\xi))dW(s,\xi) + \int_0^t K_s(\tilde{b}(s))ds.$$

In the second equality, we used (2.10) to simplify. This gives the result in this case.

Now assume that f, g and h satisfy (2.7). We may choose f^k, g^k and h^k bounded and $(\hat{\mathcal{F}}_t^*)$-predictable and such that $\int \|g^k(\xi)\|^2 d\xi$ is bounded and $f^k \to f, g^k \to g$ and $h^k \to h$ pointwise (by, for example, truncating f, g and h) with $\|f^k\| \leq \|f\|, \|g^k\| \leq \|g\|$ and $|h^k| \leq |h|$. Therefore, by the Dominated Convergence Theorem,

$$\int_0^t K_s \left(\|f_s^k - f_s\|^2 + |h_s^k - h_s| + \int \|g_s^k - g_s\|^2 d\xi \right) ds \to 0 \ \forall t \geq 0 \text{ a.s.} \qquad (2.11)$$

Using (2.7) choose $S_n \in \mathcal{T}_b$, $S_n \uparrow \infty$, a.s. such that

$$\int_0^{S_n} K_s \left(\|f_s\|^2 + |h_s|\right) ds + \int_0^{S_n} K_s \left(\int \|g_s\|^2 d\xi \right) ds \leq n \text{ a.s.}$$

Let Z^k be as in (2.8) except with f, g, h replaced by f^k, g^k, h^k respectively. Using Lemma V.3.3(b) of [14] we get that $\sup_{t \leq S_n} K_t(\int_0^t |h_s| ds) < \infty$ for all n, a.s. and hence $Z_t(\omega, y)$ is well-defined K_t-a.a. y for all $t \geq 0$ a.s. The same lemma shows that

$$\sup_{t \leq S_n} K_t \left(\int_0^t |h_s^k - h_s| ds \right) \xrightarrow{\mathbb{P}} 0 \text{ as } k \to \infty \ \forall n. \qquad (2.12)$$

Using Proposition V.3.2(e) and Lemma V.3.3(b) of [14] gives

$$\sup_{t \le S_n} K_t \left(\sup_{s \le t} \|I(f^k, s) - I(f, s)\|^2 \right) \xrightarrow{\mathbb{P}} 0 \text{ as } k \to \infty \ \forall n. \tag{2.13}$$

Similarly, using Theorem 2.2.8 above and Lemma V.3.3(b) of [14] gives

$$\sup_{t \le S_n} K_t \left(\sup_{s \le t} \|J(g^k, s) - J(g, s)\|^2 \right) \xrightarrow{\mathbb{P}} 0 \text{ as } k \to \infty \ \forall n. \tag{2.14}$$

Together, (2.12), (2.13) and (2.14) give

$$\sup_{t \le T} K_t \left(\sup_{s \le t} |Z^k(s) - Z(s)| \right) \xrightarrow{\mathbb{P}} 0 \text{ as } k \to \infty \ \forall T > 0. \tag{2.15}$$

Since we have that (2.9) holds for (Z^k, f^k, g^k, h^k) in place of (Z, f, g, h), the boundedness of ψ and its derivatives together with (2.11) and (2.15) lets us send $k \to \infty$ and derive the result using the Dominated Convergence Theorem.

Note that the proofs of Theorem 2.2.10 above and Proposition V.3.4 of [14] imply that the second term of (2.9) is an L^2 martingale and the third is a local martingale. The condition (2.7) can be used to show that the last term of (2.9) has finite variation on bounded intervals a.s. ☐

Corollary 2.2.12. *If f, g, h and Z are as in Theorem 2.2.11, then*

$$X_t(A) = \int \mathbf{1}(Z_t(\omega, y) \in A) K_t(dy)$$

defines an a.s. continuous $(\mathcal{F}_t)$-predictable $M_F(\mathbb{R}^n)$-valued process.

The proof for this corollary is the same as for Corollary V.3.6 of [14].

2.3 The strong equation

In this section, let K be as in Section 2.2 and let ν be the law of K_0. Let $\sigma_1 : M_F(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$, $\sigma_2 : M_F(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^{d \times d}$, $b : M_F(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ and $Z_0 : \mathbb{R}^d \to \mathbb{R}^d$ be Borel maps. If $\int_0^t \sigma_1(X_s, Z_s) dy(s)$ is the integral I and $\int_0^t \int \sigma_2(X_s, Z_s, \xi) dW(s, \xi)$ is the integral J discussed in Section 2.2 above, then the precise interpretation of (SE) is as follows:

$$
(\mathrm{SE})_{Z_0,K} \quad
\begin{aligned}
&(a)\ Z_t(\omega,y) = Z_0(y_0) + \int_0^t \sigma_1(X_s,Z_s)dy(s) \\
&\qquad + \int_0^t \int \sigma_2(X_s,Z_s,\xi)dW(s,\xi) + \int_0^t b(X_s,Z_s)ds\ K\text{-a.e.} \\
&(b)\ X_t(\omega)(A) = \int 1(Z_t(\omega,y)\in A)K_t(dy)\ \forall A\in\mathcal{B}(\mathbb{R}^d), \forall t\ge 0\ a.s.
\end{aligned}
$$

(X,Z) is a solution of $(\mathrm{SE})_{Z_0,K}$ iff Z is an $(\hat{\mathcal{F}}_t^*)$-predictable $\mathbb{R}^d$-valued process and X is an $(\mathcal{F}_t)$-predictable $M_F(\mathbb{R}^d)$-valued process such that $(\mathrm{SE})_{Z_0,K}$ holds.

Before we start discussing the existence of a solution of $(\mathrm{SE})_{Z_0,K}$, we must impose some conditions on the problem to make it more tractable. Let $\mathrm{Lip}_1 = \{\phi : \mathbb{R}^d \to \mathbb{R} : \|\phi\|_\infty \le 1, |\phi(x)-\phi(z)| \le |x-z|\ \forall x,z\in\mathbb{R}^d\}$ and for $\mu,\nu\in M_F(\mathbb{R}^d)$, denote the Vasserstein metric on $M_F(\mathbb{R}^d)$ by

$$
d(\mu,\nu) = \sup_{\phi\in\mathrm{Lip}_1} |\mu(\phi)-\nu(\phi)|.
$$

Let σ_1,σ_2 and b be such that there is an increasing function $L:\mathbb{R}_+\to\mathbb{R}_+$ with

$$
\begin{aligned}
\text{Lip (a)}\ & \|\sigma_1(\mu,z)-\sigma_1(\mu',z')\| + |b(\mu,z)-b(\mu',z')| \\
&\le L(\mu(1)\vee\mu'(1))\left[d(\mu,\mu')+|z-z'|\right]\ \forall \mu,\mu'\in M_F(\mathbb{R}^d), z,z'\in\mathbb{R}^d \\
\text{(b)}\ & \int \|\sigma_2(\mu,z,\xi)-\sigma_2(\mu',z',\xi)\|^2 d\xi \\
&\qquad \le L(\mu(1)\vee\mu'(1))[d(\mu,\mu')^2+|z-z'|^2] \\
\text{(c)}\ & \sup_z(\|\sigma_1(0,z)\|+|b(0,z)|)<\infty \text{ and } \sup_z \int \|\sigma_2(0,z,\xi)\|^2 d\xi < \infty.
\end{aligned}
$$

The following Lemma follows easily from the Lipschitz type conditions above.

Lemma 2.3.1. *There exists a non-decreasing function $\alpha:\mathbb{R}_+\to\mathbb{R}_+$ such that*

$$
\sup_z \left[\|\sigma_1(\mu,z)\| + \int \|\sigma_2(\mu,z,\xi)\|^2 d\xi + |b(\mu,z)|\right] \le \alpha(\mu(1)) \tag{2.16}
$$

Define $T_N = \inf\{t : K_t(1)\ge N\}\wedge N$ and

$$
\begin{aligned}
S_1 &= \{X : \mathbb{R}_+ \times \Omega \to M_F(\mathbb{R}^d) : X \text{ is } (\mathcal{F}_t)\text{-predictable, a.s. continuous and} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad X_t(1) \le N \ \forall t < T_N \ \forall N \in \mathbb{N}\} \\
S_2 &= \{Z : \mathbb{R}_+ \times \hat{\Omega} \to \mathbb{R}^d : Z \text{ is } (\hat{\mathcal{F}}_t^*)\text{-predictable and continuous } K\text{-a.e.}\} \\
S \ &= S_1 \times S_2.
\end{aligned}
$$

Let

$$
\begin{aligned}
&\Phi_2(X,Z)(t,y) \equiv \tilde{Z}_t(y) \\
&= Z_0(y_0) + \int_0^t \sigma_1(X_s,Z_s)dy_s + \int_0^t \int \sigma_2(X_s,Z_s,\xi)dW(s,\xi) + \int_0^t b(X_s,Z_s)ds \\
&\Phi_1(X,Z)(t)(\cdot) \equiv \tilde{X}_t(\cdot) = \int \mathbf{1}(\tilde{Z}_t(\omega,y) \in \cdot)K_t(dy)
\end{aligned}
$$

and let $\Phi = (\Phi_1, \Phi_2)$. Then note that $\Phi : S \to S$ by Corollary 2.2.12.

Lemma 2.3.2. *Let $(\tilde{X}^i, \tilde{Z}^i) = \Phi(X^i, Z^i)$ for $i = 1,2$. If there exist universal constants c_N such that for any stopping time $T \le T_N$,*

$$
\begin{aligned}
&\hat{\mathbb{P}}_T\left(\sup_{s \le t\wedge T} |\tilde{Z}^1(s) - \tilde{Z}^2(s)|^2 \wedge 1\right) \le c_N \int |\tilde{Z}_0^1 - \tilde{Z}_0^2|^2 \wedge 1 dm \qquad (2.17)\\
&+ c_N\left[\mathbb{P}\left(\int_0^{t\wedge T} d(X_s^1, X_s^2)^2 \wedge 1 ds\right) + \hat{\mathbb{P}}_T\left(\int_0^{t\wedge T} |\tilde{Z}^1(s) - \tilde{Z}^2(s)|^2 \wedge 1 ds\right)\right],
\end{aligned}
$$

$\forall N \in \mathbb{N}$, then there is a pathwise unique solution (X,Z) to $(SE)_{Z_0,K}$. In particular, X is unique up to $\mathbb{P}$-null sets and Z is unique K-a.e. Also, the map $t \mapsto X_t$ is a.s. continuous in t.

Proof. The proof of this lemma is essentially provided in the proof of Theorem V.4.1(a) of [14]. To see this, first set $K^1 = K^2$ in that proof, and then note that inequality (2.17) above is the same as inequality V.4.9 there, and from there on follow the proof substituting (2.17) for V.4.9 where it appears.

The idea is to perform a contraction argument on the complete metric space (S, d_0), where d_0 is an appropriately chosen metric that depends on the sequence $\{c_N\}$. One has to prove that Φ is a contraction on S. □

Theorem 2.3.3. *If (Lip) holds, then there exists a pathwise unique solution (X,Z)*

to $(SE)_{Z_0,K}$. In particular, X is unique up to $\mathbb{P}$-null sets and Z is unique K-a.e. Also, the map $t \mapsto X_t$ is a.s. continuous in t.

Proof. We will simply verify that the hypotheses in Lemma 2.3.2 hold. Let T_N be as above and $T \in \mathfrak{T}_b$ with $T \le T_N$. Let $(\tilde{X}^i, \tilde{Z}^i)$ be as above, for $i = 1, 2$. Using first Doob's strong L^2 inequality and then successive applications of the Cauchy-Schwartz inequality and the fact that $T_N \le N$, we have

$$\begin{aligned}
&\hat{\mathbb{P}}_T\left(\sup_{s\le t\wedge T} |\tilde{Z}^1(s) - \tilde{Z}^2(s)|^2 \wedge 1\right) \\
&\le c\hat{\mathbb{P}}_T\left(|Z_0^1 - Z_0^2|^2 \wedge 1 + \int_0^{t\wedge T}\int \|\sigma_2(X_s^1, Z_s^1, \xi) - \sigma_2(X_s^2, Z_s^2, \xi)\|^2 d\xi ds\right) \\
&+ c\hat{\mathbb{P}}_T\left[\int_0^{t\wedge T} \|\sigma_1(X_s^1, Z_s^1) - \sigma_1(X_s^2, Z_s^2)\|^2 + N|b(X_s^1, Z_s^1) - b(X_s^2, Z_s^2)|^2 ds\right] \\
&\le c\hat{\mathbb{P}}_T(|Z_0^1 - Z_0^2|^2 \wedge 1) \\
&+ cN\hat{\mathbb{P}}_T\Bigg(\int_0^{t\wedge T} L(X_s^1(1) \vee X_s^2(1))^2 (d(X_s^1, X_s^2) + |Z_s^1 - Z_s^2|)^2 \\
&\qquad\qquad \wedge \left(\alpha(X_s^1(1))^2 + \alpha(X_s^2(1))^2\right) ds\Bigg) \\
&\le c\hat{\mathbb{P}}_T\Bigg[|Z_0^1 - Z_0^2|^2 \wedge 1 \\
&\qquad + NL(N)^2 \int_0^{t\wedge T} (d(X_s^1, X_s^2) + |Z_s^1 - Z_s^2|)^2 \wedge \alpha(N)^2 ds\Bigg],
\end{aligned} \tag{2.18}$$

where we have used (Lip) (a), (b) and Lemma 2.3.1 for the third inequality. For the fourth inequality we use the fact that $T \le N$ and Corollary 2.2.12 (to conclude that $X_s^i(1) \le N$ for $s \le T$) and that α and L are non-decreasing. Note that in the above, all the constants (not depending on N) are collected under the term c, and this term then varies from line to line. Also, if $T = 0$, then $K_T(1)$ may exceed N, but the integral in (2.18) above is then zero. If $T > 0$ then $T_N > 0$ and so

$K_T(1) \le \sup_{t\le T} K_t(1) \le N$. Hence we have

$$\hat{\mathbb{P}}_T\left(\sup_{s\le t\wedge T} |\tilde{Z}^1(s)-\tilde{Z}^2(s)|^2\wedge 1\right)$$
$$\le c\int |Z_0^1 - Z_0^2|^2 dm + cN^2L(N)^2\alpha(N)^2\mathbb{P}\left(\int_0^{t\wedge T} d(X_s^1,X_s^2)^2\wedge 1 ds\right)$$
$$+ cNL(N)^2\alpha(N)^2\hat{\mathbb{P}}_T\left(\int_0^{t\wedge T} |Z_s^1 - Z_s^2|)^2\wedge 1 ds\right),$$

where we have used Proposition 2.2.1 to handle the first term and then factored out the $\alpha(N)^2$ term inside the integral in (2.18). Letting $c_N = N^2L(N)^2\alpha(N)^2$ and invoking Lemma 2.3.2 completes the proof. □

Examples of functions σ_1, b satisfying (Lip) (a) can be found in Section V.1 of [14]. The following remark gives some broad classes of coefficients σ_2 for which (Lip)(b) holds.

Remark 2.3.4. Suppose there exists a finite constant c_0 such that for all ξ,

$$\|\sigma_2(\mu,z,\xi)-\sigma_2(\mu',z',\xi)\| \le c_0(d(\mu,\mu')+|z-z'|)$$

and

$$\sup_z \int \|\sigma_2(0,z,\xi)\|^2 d\xi < \infty.$$

Additionally, suppose one of the following properties hold:

(i) $\sigma_2(\mu,z,\cdot)$ has compact support for each μ,z and satisfying $|\operatorname{supp}\sigma_2(\mu,z,\cdot)| \le L(\mu(1))$ for all z for some increasing function $L:\mathbb{R}_+\to\mathbb{R}_+$ (where $|A|$ represents the m-dimensional volume of $A\subset\mathbb{R}^m$).

(ii) $\sigma_2(\mu,z,\xi)$ has no measure dependence and has first and second derivatives in z. Also, there exists $M>0$ such that $\sup_z \|\partial^\alpha(\sigma_2)(z,\cdot)\|_2 < M$ for all multi-indices α up to order 2.

(iii) $\sigma_2(\mu,z,\xi)$ has no measure dependence, is continuous, integrable in ξ for each z, and has uniformly bounded first and second derivatives in z. Also assume that there exists $M>0$ such that for all $i\le d$, $\sup_z \|\frac{\partial}{\partial z_i}\sigma_2(z,\cdot)\|_1 < M$.

(iv) $\sigma_2(\mu,z,\xi) = \int f(z,x,\xi)d\mu^n(x)$ where $x \in \mathbb{R}^{nd}$ and $f : \mathbb{R}^{d+nd+m} \to \mathbb{R}^{d\times d}$ such that f is bounded, twice differentiable in x and z and there exists $M > 0$ with $\sup_{z,x} \|\partial^\alpha f(z,x,\cdot)\|_2 < M$ for all multi-indices α with $|\alpha| \le 2$. Additionally assume that there exists an L^2 function $c : \mathbb{R}^m \to \mathbb{R}_+$ such that for all z, the function $f(z,\cdot,\xi)$ is Lipschitz with constant $c(\xi)$.

(v) $\sigma_2(\mu,z,\xi) = f(\mu(\phi_1),\ldots,\mu(\phi_n),z,\xi)$ where $f : \mathbb{R}^{n+d+m} \to \mathbb{R}^{d\times d}$ and such that $f((x,z),\cdot)$ satisfies the conditions on σ_2 in either (ii) or (iii) above and each ϕ_i bounded and measurable.

Then (Lip) (b) holds.

Proof. The proof of (Lip) (b) under case (i) is not hard (just use the Lipschitz property followed by the compact support property). Assume, for case (ii) that $g^\xi(z) = \|\sigma_2(z,\xi) - \sigma_2(z_0,\xi)\|^2$. Since σ_2 is twice differentiable, so is g^ξ. Note that for each y, g^ξ has a minimum when $z = z_0$ and that $g^\xi(z_0) = 0$. Hence, by looking at the Taylor expansion we get

$$g^\xi(z) = \sum_{|\alpha|=2} \frac{2}{\alpha!}(z-z_0)^\alpha \int_0^1 (1-t)\partial^\alpha(g^\xi)(z_0 + t(z-z_0))dt$$

and, after letting $z(t) = z_0 + t(z-z_0)$ we have

$$\begin{aligned}
\int g^\xi(z)d\xi &\le c|z-z_0|^2 \sum_{\alpha=2} \int\int_0^1 (1-t)\|\partial^\alpha(g^\xi)(z(t))\|dtd\xi \\
&\le c_1|z-z_0|^2 \sum_{\alpha=2}\int_0^1\int \sum_{i,j\le d} |\partial^\alpha \left(\sigma_2^{ij}(z(t),\xi) - \sigma_2^{ij}(z_0,\xi)\right)^2 |d\xi dt \\
&\le c_2 M^2 |z-z_0|^2
\end{aligned}$$

by the uniform L^2 bound on the derivatives up to order 2 of σ_2. This proves the lemma under this case.

For (iii) first define the function $h(z) = \int \|\sigma_2(z,\xi) - \sigma_2(z_0,\xi)\|^2 d\xi$. Then note that $h(z_0) = 0$ and that this is a minimum for h. Therefore, since the first partials of h vanish at z_0

$$\begin{aligned} h(z_1) &\leq |\nabla_z h(z_0)\cdot(z_1-z_0)| + c\sup_z \|\nabla_z^2 h(z)\| |z_1-z_0|^2 \\ &\leq C|z_1-z_0|^2 \end{aligned}$$

by the boundedness of the second derivatives of h (which follows from the assumptions on σ_2). This gives us the result in this case.

For case (iv), consider

$$\begin{aligned} &\sigma_2(\mu_1,z_1,\xi)-\sigma_2(\mu_2,z_2,\xi) \\ &= \int_{\mathbb{R}^{nd}} f(z_1,x,\xi)d(\mu_1^n-\mu_2^n)(x) - \int_{\mathbb{R}^{nd}} (f(z_2,x,\xi)-f(z_1,x,\xi))\,d\mu_2^n(x) \end{aligned}$$

and hence through applications of Cauchy-Schwartz and Fubini, we get

$$\begin{aligned} \int \|\sigma_2(\mu_1,z_1,\xi)-\sigma_2(\mu_2,z_2,\xi)\|^2 d\xi & \\ &\leq 2\int c(\xi)^2 d(\mu_1^n,\mu_2^n)^2 d\xi \\ &+ 2\mu_2(1)^{2n}\int\int \|f(z_2,x,\xi)-f(z_1,x,\xi)\|^2 d\xi\, d\mu_2^n(x) \end{aligned}$$

since $f(z,\cdot,\xi)/c(\xi)\in \mathrm{Lip}_1(nd)$ for each pair z,ξ by the assumptions on f. Then by using the same proof as in part (iii) to handle the second term, the fact that $c(\xi)$ is L^2, and the definition of $\mathrm{Lip}_1(nd)$ we get

$$\begin{aligned} \int \|\sigma_2(\mu_1,z_1,\xi)-\sigma_2(\mu_2,z_2,\xi)\|^2 d\xi &\leq Cd(\mu_1,\mu_2)^n + C(\mu_1(1)\mu_2(1))^{2n}|z_1-z_2|^2 \\ &\leq L(\mu_1(1)\vee\mu_2(1))[d(\mu_1,\mu_2)^2+|z_1-z_2|^2] \end{aligned}$$

by choosing $L(x)=Cx^{n-1}+Cx^{2n}$. This gives (Lip) (b).

The proof for case (v) follows easily from the proofs of either (ii) or (iii) and the definition of the Vasserstein metric. □

2.4 A martingale problem

We will now show that X satisfies a certain martingale problem; a generalized version of the martingale problem satisfied by the process constructed in [2] and an extension to the one considered in [13]. First we need to show that any solution (X,Z) to $(\mathrm{SE})_{Z_0,K}$ satisfies the following martingale problem (MP): $X_0 = \mu \in M_F(\mathbb{R}^d)$ and for any $\phi \in C_b^2(\mathbb{R}^d)$,

$$M_t^X(\phi) \equiv X_t(\phi) - \mu(\phi) - \int_0^t X_s(L_{X_s}\phi)ds \tag{2.19}$$

is a continuous martingale with square function

$$\begin{aligned}\langle M^X(\phi)\rangle_t & \\ = \int_0^t X_s(\phi^2)ds + \int_0^t \int\int \nabla\phi(z_1)\rho_{X_s}(z_1,z_2)\nabla\phi^*(z_2)dX_s(z_1)dX_s(z_2)ds &\end{aligned} \tag{2.20}$$

where for $\nu \in M_F(\mathbb{R}^d)$,

$$L_\nu\phi(z) = \sum_{i=1}^d b^i(\nu,z)\phi_i(z) + \frac{1}{2}\sum_{i,j=1}^d a^{ij}(\nu,z)\phi_{ij}(z) \text{ with } a = \sigma_1\sigma_1^* + \int \sigma_2\sigma_2^*(\xi)d\xi \tag{2.21}$$

and ρ_ν is the $d\times d$ matrix given by

$$\rho_\nu(z_1,z_2) = \int_{\mathbb{R}^m} \sigma_2(\nu,z_1,\xi)\sigma_2^*(\nu,z_2,\xi)d\xi.$$

Lemma 2.4.1. *Any solution (X,Z) of $(\mathrm{SE})_{Z_0,K}$ satisfies (MP) with $X_0(\cdot) = \int \mathbf{1}(Z_0 \in \cdot)K_0(dy)$.*

Proof. The proof follows directly from Theorem 2.2.11 and $(\mathrm{SE})_{Z_0,K}$. First we note that for $\phi : \mathbb{R}^d \to \mathbb{R}$ bounded and measurable, $(\mathrm{SE})_{Z_0,K}$ implies that $X_t(\phi) = \int \phi(Z_t(y))K_t(dy)$. Then from Itô's Lemma (Theorem 2.2.11), for $\psi \in C_b^2(\mathbb{R}^d)$ we get

$$\int \psi(Z_t)K_t(dy) = \int \psi(Z_0)dK_0(y) + N_t$$
$$+ \int_0^t K_s \left(\nabla\psi(Z_s) \cdot b(X_s, Z_s) + \frac{1}{2} \sum_{i,j=1}^{n} \psi_{i,j}(Z_s) a^{ij}(X_s, Z_s) \right) ds,$$

where N_t is a martingale. This becomes, by the boundedness of ψ and its partials,

$$X_t(\psi) = X_0(\psi) + N_t + \int_0^t \int L_{X_s}\psi(z)X_s(dz)ds.$$

Since

$$N_t = \int_0^t \int \psi(Z_s)dM(s,y) + \int_0^t \int K_s(\nabla\psi(Z_s)\sigma_2(X_s, Z_s, \xi))dW(s,\xi),$$

by the orthogonality of M and W we have

$$\langle N \rangle_t = \int_0^t X_s(\psi^2)ds$$
$$+ \int_0^t \int \left(\int \nabla\psi\sigma_2(X_s, z, \xi)dX_s(z) \right) \left(\int \nabla\psi\sigma_2(X_s, z, \xi)dX_s(z) \right)^* d\xi ds,$$

which can be rewritten in the form of (2.20). □

Definition. For a function $f : \mathbb{R}^{nd} \to \mathbb{R}$, define $F_{n,f} : M_F(\mathbb{R}^d) \to \mathbb{R}$ by $F_{n,f}(\mu) = \int f d\mu^n$.

It can be shown, by induction and using Lemma 2.4.1, that for $f = \bigotimes_{i=1}^n \phi_i$, where each $\phi_i \in C_b^2(\mathbb{R}^d)$, that

$$N_t(n,f) = F_{n,f}(X_t) - F_{n,f}(X_0) - \int_0^t \hat{L}F_{n,f}(X_s)ds \tag{2.22}$$

is a martingale, where

$$\hat{L}F_{n,f}(\mu) = \sum_{i=1}^{n} \mu(L\phi_i)\left(\prod_{j\neq i}\mu(\phi_j)\right) \tag{2.23}$$
$$+\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{n}\left(\prod_{k\neq i,j}\mu(\phi_k)\right)\left[\mu(\phi_i\phi_j)+\int_{\mathbb{R}^m}\mu(\nabla\phi_i\sigma_2(\xi))\cdot\mu(\nabla\phi_j\sigma_2(\xi))d\xi\right].$$

We have suppressed the measure dependence of the generator L here and will do so below as well.

Definition. If $f \in \mathcal{B}(\mathbb{R}^{nd})$ (i.e. f a real valued, bounded Borel measurable function on $\mathbb{R}^{nd}$) then we can define the operator $\Phi^n_{i,j} : \mathcal{B}(\mathbb{R}^{nd}) \to \mathcal{B}(\mathbb{R}^{(n-1)d})$ acting on f as:

$$\Phi^n_{i,j} f(z_1,\dots,z_{n-1}) = f(z_1,\dots,z_{n-1},\dots,z_{n-1},\dots,z_{n-2})$$

where z_{n-1} is inserted into the vector $(z_1,\dots,z_{n-2})$ such that it appears at the ith and jth spots.

We can now rewrite (2.23) in terms of f as follows:

$$\hat{L}F_{n,f}(\mu) = \sum_{i=1}^{n}\int L^i f d\mu^n + \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{n}\int \Phi^n_{i,j}(f)d\mu^{n-1} \tag{2.24}$$
$$+\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{n}\sum_{k,l=1}^{d}\int \rho^{k,l}_{\mu}(z_i,z_j)\frac{\partial^2}{\partial z_{i_k}\partial z_{j_l}} f(\mathbf{z})d\mu^n(\mathbf{z}).$$

Here $\mathbf{z} = (z_1,\dots,z_n)$ and $L^i f(z_1,\dots,z_n) = Lf_{z_1,\dots,z_{i-1},z_{i+1},\dots,z_n}(z_i)$ (ie, applying L to a function of just z_i and fixing the other coordinates) where L is as in (2.21). Note that we have suppressed the dependence of L^i on ν, as with L. By Lemma 2.4.1

and an inductive argument, one can see that

$$N_t(n,f) = \int_0^t \int \sum_{i=1}^n \left(\prod_{j\neq i} X_s(\phi_j)\right) X_s(\nabla\phi_i\sigma_2(\xi))dW(s,\xi) \qquad (2.25)$$

$$+ \int_0^t \int \sum_{i=1}^n \left(\prod_{j\neq i} X_s(\phi_j)\right) \phi_i(Z_s)dM(s,y)$$

$$= \sum_{i=1}^n \int_0^t \int \left[\int \nabla_{z_i} f(\mathbf{z})\sigma_2(z_i,\xi)X_s^n(d\mathbf{z})\right] dW(s,\xi) \qquad (2.26)$$

$$+ \sum_{i=1}^n \int_0^t \int \left[\int f(Z_s(y_1),\ldots,Z_s(y_n))\, K_s^{n-1}(d\mathbf{y}^i)\right] dM(s,y_i)$$

where $\mathbf{y}^i = (y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_n)$. Then a calculation gives

$$\langle N(n,f)\rangle_t = \sum_{i,j=1}^n \int_0^t \left[\int \nabla_i f(\mathbf{z}_1)\rho_{X_s}(z_i,z_{n+j})(\nabla_{n+j} f(\mathbf{z}_2))^* X_s^{2n}(d\mathbf{z})\right] ds \qquad (2.27)$$

$$+ \sum_{i,j=1}^n \int_0^t \int \left[\int \Phi_{i,j+n}^{2n}(f\otimes f)X_s^{2n-1}(d\mathbf{z}^{2n})\right] ds$$

where $\mathbf{z} = (\mathbf{z}_1,\mathbf{z}_2)$ with $\mathbf{z}_1 = (z_1,\ldots,z_n), \mathbf{z}_2 = (z_{n+1},\ldots,z_{2n})$ and where $\mathbf{z}^{2n} = (z_1,\ldots,z_{2n-1})$.

Theorem 2.4.2. *For any function $f \in C_b^2(\mathbb{R}^{nd})$,*

$$N_t(n,f) = F_{n,f}(X_t) - F_{n,f}(X_0) - \int_0^t \hat{L}F_{n,f}(X_s)ds \qquad (2.28)$$

is a martingale, where $\hat{L}$ is given by (2.24).

Proof. Let

$$\mathcal{H}_n = \left\{ \sum_{i=1}^M a_i \bigotimes_{j=1}^{nd} \phi_{i_j} : \phi_{i_j} \in C_c^\infty(\mathbb{R}), a_i \in \mathbb{R}, M \in \mathbb{N} \right\}.$$

We will first use functions in $\mathcal{H}_n$ to prove that for $f \in C_c^\infty(\mathbb{R}^{nd})$, (2.28) is a martingale and then boost that up to include functions in $C_c^2(\mathbb{R}^{nd})$ and finally those in $C_b^2(\mathbb{R}^{nd})$. By the Stone-Weierstrass theorem, for any $f \in C_c^\infty(\mathbb{R}^{nd})$, there is a sequence $g_k \in \mathcal{H}_n$ such that $g_k \to \partial^\alpha f$ uniformly, for $\alpha = (2,\ldots,2)$. Then, by integrating in the various components, we obtain a sequence f_k such that $\partial^\alpha f_k \to \partial^\alpha f$ uniformly for $|\alpha| \leq 2$.

This in turn gives point-wise convergence of F_{n,f_k} to $F_{n,f}$ and $\hat{L}F_{n,f_k}$ to $\hat{L}F_{n,f}$ on $M_F(\mathbb{R}^d)$ (using the Dominated Convergence Theorem). Now let $T_N = \inf\{t : K_t(1) \geq N\} \wedge N$ as above. Then for each k, l,

$$\mathbb{P}\left(\sup_{s \leq t \wedge T_N} |N_s(n, f_k) - N_s(n, f_l)|^2\right) \leq c\mathbb{P}\left(\langle N(n, f_k) - N(n, f_l)\rangle_{t \wedge T_N}\right)$$

and hence by sending $l \to \infty$ and using the Dominated Convergence Theorem, we conclude that

$$\begin{aligned}\mathbb{P}\left(\sup_{s \leq t \wedge T_N} |N_s(n, f_k) - N_s(n, f)|^2\right) &\leq c\mathbb{P}\left(\langle N(n, f_k) - N(n, f)\rangle_{t \wedge T_N}\right) \\ &\to 0\end{aligned}$$

as $k \to \infty$, by the Dominated Convergence Theorem and Lemma 2.3.1. Hence $N_t(n, f)$ is a continuous local martingale. Since f and its derivatives up to order 2 are bounded and (Lip) (c) holds, we have by (2.27) that

$$\mathbb{P}(\langle N(n, f)\rangle_t) \leq c\mathbb{P}\left(\int_0^t X_s(1)^{2n} + X_s(1)^{2n-1} ds\right).$$

The right side is finite by the boundedness of moments of the mass process of super-Brownian motion. Therefore $N_t(n, f)$ is a martingale.

Now let $f \in C_c^2(\mathbb{R}^{nd})$. Using approximate identities we can construct functions $f_k \in C_c^\infty(\mathbb{R}^{nd})$ such that $\partial^\alpha f_k \to \partial^\alpha f$ uniformly for all α with $|\alpha| \leq 2$. Repeating the above argument again and using the fact that each $N_t(f_k)$ is a martingale shows that $N_t(f)$ is a continuous martingale.

Finally, suppose $f \in C_b^2(\mathbb{R}^{nd})$. Let $\eta_k \in C_c^\infty(\mathbb{R}^{nd})$ such that $\eta_k = 1$ on $B(0, k)$, $\eta_k = 0$ on $B(0, k+1)^c$ and $\|\partial^\alpha \eta_k\| < 2$ for each multi-index with $|\alpha| \leq 2$. Letting $f_k = f\eta_k$, we see that $\partial^\alpha f_k \to \partial^\alpha f$ pointwise for each $|\alpha| \leq 2$. Here $B(0, k)$ denotes the open ball of radius k about the origin.

This gives $\hat{L}F_{n,f_k} \to \hat{L}F_{n,f}$ pointwise (by the Dominated Convergence Theorem). Then by noting that each $N_t(f_k)$ is a continuous martingale and using the method in the first part of this proof shows that $N_t(f)$ is a continuous martingale. $\square$

To see that the above martingale problem reduces to the one found in [2] in their setting, let $m = d = 1$, b = 0, $\sigma_2(\mu, z, \xi) = h(\xi - z)$ where h is square-integrable and $\sigma_1(\mu, z) = \sigma_1(z)$ is Lipschitz.. Additionally, suppose that $\eta(z) \equiv \int_{\mathbb{R}} h(\xi - z)h(\xi)d\xi$ is twice continuously differentiable with η' and η'' bounded. Then one can easily check that these coefficients satisfy (Lip) (use Taylor's theorem on η to show (Lip)(b)). To get the form of the infinitesimal generator, we use (2.24) :

$$\hat{L}F_{n,f}(\mu) = \frac{1}{2}\sum_{i=1}^{n}\int(\sigma_1^2(z_i) + \eta(0))f_{ii}(\mathbf{z})d\mu^n(\mathbf{z}) + \frac{1}{2}\sum_{\substack{i,j=1 \\ i\neq j}}^{n}\int\rho(z_i, z_j)f_{ij}(\mathbf{z})d\mu^n(\mathbf{z})$$
$$+\frac{1}{2}\sum_{\substack{i,j=1 \\ i\neq j}}^{n}\int\Phi_{i,j}^n(f)d\mu^{n-1}$$

which is equal to the sum of the generators (2.1) and (2.2) of [2], since $\rho(z_i, z_j) = \eta(z_i - z_j)$, if we allow their branching rate σ to be constant at 1.

In [2], uniqueness in law of the martingale problem is established. Hence in this particular setting, our solution is both unique in law and pathwise unique and therefore is a version of the process constructed in [2].

2.5 Additional properties of solutions

We will start off with a stability result for solutions of $(SE)_{Z_0,K}$, and then use it to prove the strong Markov property. We conclude with a proof of the compact support property for the solutions of $(\mathrm{SE})_{Z_0,K}$. Recall that d denotes the Vasserstein metric on $M_F(\mathbb{R}^d)$.

Theorem 2.5.1. *Assume (Lip) holds. Let $K^1 \leq K^2 \equiv K$ be $(\mathcal{F}_t)$-historical super-Brownian motions with $m_i(\cdot) \equiv \mathbb{E}(K_0^i(\cdot)) \in M_F(\mathbb{R}^d)$, and let $Z_0^i : \mathbb{R}^d \to \mathbb{R}^d$ be Borel maps, for $i = 1, 2$. Let $T_N = \inf\{t : K_t(1) \geq N\} \wedge N$. There are universal constants*

$\{c_N, N \in \mathbb{N}\}$ *so that if* (X^i, Z^i) *is the unique solution of* $(SE)_{Z_0^i, K^i}$*, then*

$$\mathbb{P}\left(\int_0^{T_N} \sup_{r \le s} d(X^1(r), X^2(r))^2 ds\right) \le c_N \left(\int |Z_0^1 - Z_0^2|^2 \wedge 1 dm_2 + m_2(1) - m_1(1)\right)$$

The proof of this theorem essentially follows from the proof of Theorem V.4.2 of [14]. Note that before we can follow that proof, we will need to show that the condition (2.17) of Lemma 2.3.2 above holds under (Lip), where $(\tilde{X}^i, \tilde{Z}^i) = \Phi^i(X^i, Z^i)$ where each Φ^i depends on different historical super-Brownian motions $K^i, i = 1,2$ and $K^1 \le K^2$ (ie, the Radon-Nikodym derivative $\frac{dK^1}{dK^2} \le 1$). The computation showing this is almost exactly like the one in the proof of Theorem 2.3.3.

We now construct a canonical space upon which we can define a white noise and an independent historical super-Brownian motion. Let K be an $(\mathcal{F}_t)$-adapted historical super-Brownian motion on $\bar{\Omega} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \ge 0}, \mathbb{Q}_\nu)$. Let (X, Z) be the unique solution to $(SE)_{Z_0, K}$. Recall that X is $(\mathcal{F}_t)$-predictable, and Z is $(\hat{\mathcal{F}}_t^*)$-predictable. Denote by Ω_X the space of continuous paths in $M_F(\mathbb{R}^d)$ and $\mathcal{F}_X$ its Borel σ-field.

Let $\Omega_K = \{H_\cdot \in C([0,\infty), M_F(C)) : H_t \in M_F^t(C) \forall t \ge 0\}$ where $M_F^t(C) = \{\mu \in M_F(C) : \mu(A) = \mu(A \cap \{y \in C : y = y^t\}) \forall A \in \mathcal{C}\}$. The historical super-Brownian motion K has sample paths in Ω_K. Let $\mathcal{F}_K$ be the Borel σ-field of Ω_K and let

$$\mathcal{F}_t^K = \bigcap_{n=1}^{\infty} \sigma(H_r : 0 \le r \le t + 1/n).$$

To treat W we will follow the method of [12]. Let $\mathbb{Q}_W$ be the law of the white noise W. For functions $f \in C(\mathbb{R}^d)$, define

$$\|f\|_\lambda = \sup_{x \in \mathbb{R}^d} |f(x)| e^{-\lambda |x|},$$

and let $C_{tem} = \{f \in C(\mathbb{R}^m) : \|f\|_\lambda < \infty \text{ for all } \lambda > 0\}$. We let C_{tem} be endowed with the topology generated by the family of norms $\|\cdot\|_\lambda, \lambda > 0$. Hence, C_{tem} can be thought of as a metric space with metric d given by $d(f,g) = \sum_{k=1}^\infty 2^{-k}(\|f - g\|_{1/k} \wedge 1)$. It turns out that (C_{tem}, d) is a Polish space. Let $W = (W^1, \ldots, W^d)$ be a d-dimensional white noise, as in Section 2.2. We can now identify each white noise W^i with its associated Brownian sheet and say each W^i is a process with sample paths in $C([0,\infty),C_{tem})$. Hence we will say that W has sample paths in $\Omega_W \equiv C([0,\infty),C_{tem}^d)$. Endow Ω_W with the topology of uniform convergence on compacts.

Let $\widetilde{W} : [0,\infty)\times\Omega_W \to C_{tem}^d$ be the coordinate projection map, and let

$$\mathcal{F}_t^W = \bigcap_{n=1}^{\infty} \sigma(\widetilde{W}_s : s \le t + 1/n).$$

Then $(H,\widetilde{W})$ is a Borel strong Markov process on $(\Omega_{K,W},\mathcal{H},\mathcal{H}_t,\mathbb{Q}_{\nu,W})$ where $\Omega_{K,W} = \Omega_K \times \Omega_W, \mathbb{Q}_{\nu,W} = \mathbb{Q}_\nu \times \mathbb{Q}_W$ and for $t \ge 0$, $\mathcal{H}_t$ is given by $\sigma((H,\widetilde{W})(s) : s \le t)$ with the $\mathbb{Q}_{\nu,W}$-null sets thrown in.

We will now treat H and $\widetilde{W}$ as random processes on the space $\Omega_{K,W}$ with $H(t,\omega,\alpha) = H_t(\omega)$ and $\widetilde{W}(t,\omega,\alpha) = \alpha_t$. Under $\mathbb{Q}_{\nu,W}$, H and $\widetilde{W}$ are independent by definition and have laws $\mathbb{Q}_\nu$ and $\mathbb{Q}_W$ respectively.

Recall from the definition in Section 2.2 that $\widehat{\Omega_{K,W}} = \Omega_{K,W} \times C$ and similarly $\hat{\mathcal{H}}_t = \mathcal{H}_t \times \mathcal{C}_t$ with universal completion $\hat{\mathcal{H}}_t^*$.

Statement (a) below shows that $(\mathrm{SE})_{Z_0,K}$ has a strong solution, whereas (b) shows that there is continuity of the laws of the solutions in the initial condition.

Theorem 2.5.2. *Let (X,Z) be the unique solution of $(SE)_{Z_0,K}$ on $\bar{\Omega}$. Then the following holds:*

(a) *There are $(\mathcal{H}_t)$-predictable and $(\hat{\mathcal{H}}_t^*)$-predictable maps $\tilde{X} : \mathbb{R}_+ \times \Omega_{K,W} \to M_F(\mathbb{R}^d)$ and $\tilde{Z} : \mathbb{R}_+ \times \widehat{\Omega_{K,W}} \to \mathbb{R}^d$, respectively, depending only on (Z_0,ν), and such that*

$$\Big(X(t,\omega),Z(t,\omega,y)\Big) = \Big(\tilde{X}\big(t,K(\omega),W(\omega)\big),\tilde{Z}\big(t,K(\omega),W(\omega),y\big)\Big) \quad (2.29)$$

gives the unique solution of $(SE)_{Z_0,K}$.

(b) *There is a continuous map from $X_0 \mapsto \mathbb{P}'_{X_0}$ from $M_F(\mathbb{R}^d)$ to $M_1(\Omega_X)$ such that if (X,Z) is a solution of $(SE)_{Z_0,K}$ on some filtered space $\bar{\Omega}$ then*

$$\mathbb{P}(X \in \cdot) = \int \mathbb{P}'_{X_0(\omega)}(\cdot)\, d\mathbb{P}(\omega). \quad (2.30)$$

(c) *If T is an a.s. finite $(\mathcal{F}_t)$-stopping time, then for all $A \in \mathcal{F}_X$*

$$\mathbb{P}(X(T+\cdot) \in A|\mathcal{F}_T)(\omega) = \mathbb{P}'_{X_T(\omega)}(A), \mathbb{P}\text{-}a.s.$$

Proof. The proof of (a) is an analogue of the proof of Theorem V.4.1 (b) of [14].

For (a), let $(\tilde{X}, \tilde{Z})$ be the unique solution of $(\text{SE})_{Z_0,K}$ where $(H, \widetilde{W})$ is the canonical process on $\bar{\Omega}' \equiv (\Omega_{K,W}, \mathcal{H}, \mathcal{H}_t, \mathbb{Q}_{\nu,W})$. It is clear from Theorem 2.3.3 that $(\tilde{X}, \tilde{Z})$ depends only on Z_0 and ν, and satisfies the desired predictability conditions. To show (2.29), we need only show that the latter process solves $(\text{SE})_{Z_0,K}$ on $\bar{\Omega}$. Let $I'(f,t,H,W,y), J'(g,t,H,W,y)$ and $f \in D'_1(d,d), g \in D'_2(m,d,d)$ denote the sample path stochastic integral and the white noise integral respectively from Section 2.2 on $\bar{\Omega}'$. Similarly define $I(f,t), J(g,t), f \in D_1, g \in D_2$ to be the stochastic integrals in Section 2.2 on $\bar{\Omega}$. We claim that if $f \in D'_1$, then $f \circ (K,W)(t,\omega,y) \equiv f(t,K(\omega),W(\omega),y) \in D_1$ and

$$I(f \circ (K,W)) = I'(f) \circ (K,W), K\text{-a.e.} \tag{2.31}$$

By the definition of D'_1 and D_1 and since K and H have the same laws, we get the first implication. Equation (2.31) clearly holds for simple functions. Using Proposition V.3.2 (d) of [14] allows us to show that it holds for all $f \in D'_1$. The situation with the white noise integral J is similar in that if $g \in D'_2$, then $J \circ (K,W) \in D_2$ and

$$J(g \circ (K,W)) = J'(g) \circ (K,W), K\text{-a.e.}.$$

This can again be shown by first starting at simple functions and then bootstrapping up using Theorem 2.2.8. To show the result, we can now simply replace $(H, \widetilde{W})$ with $(K(\omega), W(\omega))$ in $(\text{SE})_{Z_0,H}$ and get that $(\tilde{X} \circ (K,W), \tilde{Z} \circ (K,W))$ solves $(\text{SE})_{Z_0,K}$ on $\bar{\Omega}$.

The proof of (b) is also a straight-forward analogue of the proof of Theorem V.4.1 (c) in [14]. The continuity condition in (b) depends on the stability result above (which is the analogue of the stability result in [14]).

The proof of (c) is more involved, but follows exactly the same route as the proof of Theorem V.4.1 (d) in [14]. It uses the continuity proven in (b) to reduce the problem to proving the Markov property (via finite-valued stopping times) which is then established by using the uniqueness of solutions of the strong equation. Again, the only additional detail to worry about is the interpretation of the white noise integral J.

□

We now prove the compact support property for solutions of $(SE)_{Z_0,K}$.

Definition. Let $f : [0,\infty) \times \hat{\Omega} \to (E, \|\cdot\|)$ (a normed linear space). A bounded $(\mathcal{F}_t)$-stopping time T is called a reducing time for f if and only if $\mathbf{1}(0 < t \leq T)\|f(t,\omega,y)\|$ is uniformly bounded. We say $\{T_n\}$ reduces f if and only if each T_n reduces f and $T_n \uparrow \infty$ $\mathbb{P}$-a.s. If such a sequence exists, we say f is locally bounded.

Theorem 2.5.3. *Assume that $f \in D_1(n,d)$, $g \in D_2(m,n,d)$ and $b : [0,\infty) \times \hat{\Omega} \to \mathbb{R}^n$ is $(\hat{\mathcal{F}}_t^*)$-predictable, and all three are locally bounded. Let $Z(t) = Z_0 + I(f,t) + J(g,t) + V(b,t)$ where $V(b,t) = \int_0^t b(s)ds$. Define $\hat{K}(\cdot) = K_t(\{y : Z^t(\omega,y) \in \cdot\})$. Then for a.a. ω for each $k \in \mathbb{N}$ there is a compact set $S_k(\omega) \subset C$ such that $supp(\hat{K}_t) \subset S_k$ $\forall t \in [k^{-1},k]$, and furthermore, $supp(\hat{K}_t) \subset S_k$ $\forall t \in [0,k]$ on $\{\omega : supp(\hat{K}_0)$ is compact$\}$.*

The proof of this theorem is given by a small modification of the proofs of Corollaries 3.3 and 3.4 of [13]

Corollary 2.5.4. *Assume (Lip). Let (X,Z) denote the solution of $(SE)_{Z_0,K}$. Then for a.a. ω for each $k \in \mathbb{N}$ there is a compact set $S'_k(\omega) \subset \mathbb{R}^d$ such that $supp(X_t) \subset S'_k$ $\forall t \in [k^{-1},k]$, and on $\{\omega : supp(X_0)$ is compact$\}$, $supp(X_t) \subset S'_k$ $\forall t \in [0,k]$.*

Proof. Note that under (Lip), the coefficients σ_1, σ_2 and b are locally bounded and that in the above, $X_t(A) = \int \mathbf{1}(Z(t) \in A) K_t(dy) = \hat{K}_t(\pi_t^{-1}A)$ where $\pi_t^{-1}A = \{y : Z_t(\omega,y) \in A\}$ for A bounded measurable. Apply Theorem 2.5.3 and use the fact that $\pi_t : C \to \mathbb{R}^d$ is a continuous mapping.

Let S_k be as in Theorem 2.5.3. Then for $t \in [k^{-1},k]$,

$$\begin{aligned} X_t((\pi_t S_k)^c) &= \hat{K}_t(\pi_t^{-1}((\pi_t S_k)^c)) \\ &\leq \hat{K}_t(S_k^c) = 0. \end{aligned}$$

Hence setting $S'_k = \pi_k S_k$ gives the result in this case. Proceed similarly for the extension for $\{\omega : S(X_0)$ is compact$\}$. □

Chapter 3

A super Ornstein-Uhlenbeck process interacting with its center of mass

3.1 Introduction

The existence and uniqueness of a self-interacting measure-valued diffusion that is either attracted or repelled from its centre of mass is shown below. It seems natural to consider a super Ornstein-Uhlenbeck (SOU) process with attractor (repeller) given by the centre of mass of the process initially as it is the simplest diffusion of this type.

This type of model first appeared in a recent paper of Engländer [3] where a d-dimensional binary Brownian motion, with each parent giving birth to exactly two offspring and branching occurring at integral times, is used to construct a binary branching Ornstein-Uhlenbeck process where each particle is attracted (repelled) by the center of mass (COM). This is done by solving the appropriate SDE along each branch of the particle system and then stitching these solutions together.

This model can be generalized such that the underlying process is a branching Brownian motion (BBM), $\mathcal{T}$, (i.e. with a general offspring distribution). We might then solve an SDE on each branch of $\mathcal{T}$:

$$Y_i^n(t) = Y_{p(i)}^{n-1}(n-1) + \gamma \int_{n-1}^{t} \bar{Y}_n(s) - Y_i^n(s) ds + \int_{n-1}^{t} dB_i^n(s), \tag{3.1}$$

for $n-1 < t \leq n$, where B_i^n labels the ith particle of $\mathcal{T}$ alive from time $n-1$ to n, $p(i)$ is the parent of i and

$$\bar{Y}_n(s) = \frac{1}{\tau_n} \sum_{i=1}^{\tau_n} Y_i^n(s)$$

is the center of mass. Here τ_n is the population of particles alive from time $n-1$ to n. This constructs a branching OU system with attraction to the COM when $\gamma > 0$ and repulsion when $\gamma < 0$.

It seems reasonable then, to take a scaling limit of branching particle systems of this form and expect it to converge in distribution to a measure-valued process where the representative particles behave like an OU process attracting to (repelling from) the COM of the process. Though viable, this approach will be avoided in lieu of a second method utilizing the historical stochastic calculus of Perkins [14] which is more convenient for both constructing the SOU process interacting with its COM and for proving various properties. The idea is to use a supercritical historical Brownian motion to construct the interactive SOU process by solving a certain stochastic equation. This approach for constructing interacting measure-valued diffusions was pioneered in [13] and utilized in, for example, [8].

A supercritical historical Brownian motion, K, is a stochastic process taking values in the space of measures over the space of paths in $\mathbb{R}^d$. One can think of K as a supercritical superprocess which has a path-valued Brownian motion as the underlying process. That is, if B_t is a d-dimensional Brownian motion, then $\hat{B}_t = B_{\cdot \wedge t}$ is the underlying process of K. More information about K is provided in Section 3.2.

It can be shown that if a path $y : [0,\infty) \to \mathbb{R}^d$ is chosen according to K_t (loosely speaking – this is made rigorous below), then $y(s)$ is a Brownian motion stopped at t. Then projecting down gives

$$X_t^K(\cdot) = \int \mathbf{1}(y_t \in \cdot) K_t(dy),$$

a (supercritical) super Brownian motion. One can sensibly define the Ornstein-Uhlenbeck SDE driven by y according to K_t:

$$dZ_s = -\gamma Z_s ds + dy_s$$

where $\gamma > 0$ implies Z_t is attracted to the origin, and $\gamma < 0$ implies repulsion. Defining

$$X_t(\cdot) = \int \mathbf{1}(Z_t(y) \in \cdot) K_t(dy)$$

will give an ordinary super Ornstein-Uhlenbeck process. The intuition here is that K keeps track of the underlying branching structure, and Z_t is a function transforming a typical Brownian path into a typical Ornstein-Uhlenbeck path.

Similarly, we can define a function of the path y that is an OU process with attraction to (repelling from) the COM, according to K_t and the SOU process with attraction (repulsion) to COM as follows:

$$\begin{aligned} dY_s &= \gamma(\bar{Y}_s - Y_s)ds + dy_s \\ X'_t(\cdot) &= \int \mathbf{1}(Y_t(y) \in \cdot) K_t(dy) \end{aligned}$$

where the COM is defined as

$$\bar{Y}_s = \frac{\int x X'_s(dx)}{\int 1 X'_s(dx)}.$$

The details are given in Section 3.3 and it is shown that such a process X' can be constructed by a correspondence with the ordinary SOU process, X.

In Section 3.4 we prove various properties of the ordinary SOU process, X, and the interacting SOU process, X'. It is shown in Theorem 3.4.1 that the COM converges in the case where X' goes extinct in finite time, regardless of the value of γ. Theorem 3.4.2 shows that on the set where X' survives indefinitely, the COM $\bar{Y}_t$ converges in the attractive case (where $\gamma > 0$). Theorem 3.4.3 then shows that in the extinction case the mass normalized process $\frac{X'_t}{X'_t(1)}$ converges a.s. to a point as t approaches the extinction time, which is a version of a result of Tribe [20].

With the convergence of the COM on the survival set, Theorem 3.4.9 gives that in the attractive case,

$$\frac{X_t(\cdot)}{X_t(1)} \xrightarrow{\text{a.s.}} P_\infty(\cdot) \text{ as } t \to \infty$$

where $P_\infty(\cdot)$ is the stationary distribution of the OU process, which is an extension of Proposition 20 of Engländer and Turaev [4] and Theorem 1 of Engländer and Winter [5] for the SOU process. The correspondence between X and X' given in Section 3.3 is used in Theorem 3.4.11 to prove in the attractive case,

$$\frac{X_t'(\cdot)}{X_t'(1)} \xrightarrow{\text{a.s.}} P_\infty^{\bar{Y}_\infty}(\cdot) \text{ as } t \to \infty,$$

where the distribution $P_\infty^{\bar{Y}_\infty}(\cdot)$ is equal to $P_\infty(\cdot)$ shifted to the limiting value of the COM $\bar{Y}_t$.

In final section, in Theorem 3.5.1, we prove that the COM of the SOU process repelling from its COM converges (i.e. when $\gamma < 0$), provided the repulsion is not too strong. It is shown that the corresponding result for the SOU process with repulsion from the origin fails. This is one of the most interesting distinctions between the interactive SOU process and the non-interactive SOU process. We conclude with some conjectures, as well as a comparison with what is known for the SOU process repelling from the origin.

3.2 Background

We will take K to be a supercritical historical Brownian motion. Specifically, we let K be a $(\Delta/2, \beta, 1)$-historical superprocess (here Δ is the d-dimensional Laplacian), where $\beta > 0$ constant, on the probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$. The β in the definition corresponds to the branching bias in the offspring distribution, and the 1 to the variance of the offspring distribution. A martingale problem characterizing K is given below. For a more thorough explanation of historical Brownian motion than found here, see Section V.2 of [14].

Let $C = C(\mathbb{R}_+, \mathbb{R}^d)$ with $\mathcal{C}$ its Borel σ-field and define $C^t = \{y^t : y \in C\}$ where $y^t = y_{\cdot\wedge t}$ and $\mathcal{C}_t = \sigma(y^s, s \leq t, y \in C)$.

For metric spaces E and E', take $C_c(E, E')$ and $C_b(E, E')$ to be the space of compactly supported continuous functions from E to E' and the space of bounded continuous functions from E to E', respectively. For a measure space $(E,\mathcal{E})$, let $b\mathcal{E}$ denote the set of bounded $\mathcal{E}$-measurable real valued functions.

Notation. For a measure μ on a space E and a measurable function $f : E \to \mathbb{R}$, let $\mu(f) = \int f d\mu$. Furthermore, if $f : E \to \mathbb{R}^n$, define

$$\mu(f) = (\mu(f_1), \cdots, \mu(f_n)).$$

For a point $p \in \mathbb{R}^d$, let $|p|$ denote Euclidean norm of p, and for a bounded function $f : E \to \mathbb{R}^d$, denote $\|f\| = \sup_{x\in E} \sum_{i=1}^d |f_i(x)|$.

It turns out that K_t is supported on $C^t \subset C$ a.s. and typically, K_t puts mass on those paths that are "Brownian" (until time t and fixed thereafter). As K takes values in $M_F(C)$, $K_t(\cdot)$ will denote integration over the y variable.

Let $\hat{B}_t = B(\cdot \wedge t)$ be the path-valued process associated with B, taking values in C^t. Then for $\phi \in b\mathcal{C}$, if $s \le t$ let $P_{s,t}\phi(y) = \mathbb{E}^{s,y}(\phi(\hat{B}_t))$, where the right hand side denotes expectation at time t given that until time s, $\hat{B}$ follows the path y.

The weak generator, $\hat{A}$, of $\hat{B}$ is as follows. If $\phi : \mathbb{R}_+ \times C \to \mathbb{R}$ we say $\phi \in \mathcal{D}(\hat{A})$ if and only if ϕ is bounded, continuous and $(\mathcal{C}_t)$-predictable, and for some $\hat{A}_s\phi(y)$ with the same properties as ϕ,

$$\phi(t,\hat{B}) - \phi(s,\hat{B}) - \int_s^t \hat{A}_r\phi(\hat{B})dr, t \ge s,$$

is a $(\mathcal{C}_t)$-martingale under $P_{s,y}$ for all $s \ge 0, y \in C^s$.

If $m \in M_F\left(\mathbb{R}^d\right)$, we will say K satisfies the historical martingale problem, $(HMP)_m$, if and only if $K_0 = m$ a.s. and

$$\forall\phi \in \mathcal{D}(\hat{A}),\ M_t(\phi) \equiv K_t(\phi_t) - K_0(\phi_0) - \int_0^t K_s(\hat{A}_s\phi)ds - \beta\int_0^t K_s(\phi_s)ds$$

$(HMP)_m$ is a continuous $(\mathcal{F}_t)$-martingale with

$$\langle M(\phi)\rangle_t = \int_0^t K_s(\phi_s^2)ds\ \forall t \ge 0,\ \text{a.s.}$$

Using the martingale problem $(HMP)_m$, one can construct an orthogonal martingale measure $M_t(\cdot)$ with the method of Walsh [22]. Denote by $\mathcal{P}$, the σ-field of

$(\mathcal{F}_t)$-predictable sets in $\mathbb{R}_+ \times \Omega$. If $\psi : \mathbb{R}_+ \times \Omega \times C \to \mathbb{R}$ is $\mathcal{P} \times \mathcal{C}$-measurable and

$$\int_0^t K_s(\psi_s^2) ds < \infty \; \forall \; t \geq 0, \tag{3.2}$$

then there exists a continuous local martingale $M_t(\psi)$ with quadratic variation $\langle M(\psi) \rangle_t = \int_0^t K_s(\psi_s^2) ds$. If the expectation of the term in (3.2) is finite, then $M_t(\psi)$ is an L^2-martingale.

Definition. Let $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathcal{F}}_t) = (\Omega \times C, \mathcal{F} \times \mathcal{C}, \mathcal{F}_t \times \mathcal{C}_t)$. Let $\hat{\mathcal{F}}_t^*$ denote the universal completion of $\hat{\mathcal{F}}_t$. If T is a bounded $(\mathcal{F}_t)$-stopping time, then the normalized Campbell measure associated with T is the measure $\hat{\mathbb{P}}_T$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ given by

$$\hat{\mathbb{P}}_T(A \times B) = \frac{\mathbb{P}(\mathbf{1}_A K_T(B))}{m_T(1)} \text{ for } A \in \mathcal{F}, B \in \mathcal{C}.$$

where $m_T(1) = \mathbb{P}(K_T(1))$. We denote sample points in $\hat{\Omega}$ by (ω, y). Therefore, under $\hat{\mathbb{P}}_T$, ω has law $K_T(1) d\mathbb{P} \cdot m_T^{-1}(1)$ and conditional on ω, y has law $K_T(\cdot)/K_T(1)$.

Notation. Henceforth we denote the super Ornstein-Uhlenbeck process by X and the interacting super Ornstein-Uhlenbeck process by X'. The symbol $\mathring{X}$ will be used as a catch-all and will vary according to the setting, but will be a measure-valued process taking values in $M_F\left(\mathbb{R}^d\right)$. Occasionally the symbol $\mathring{K}$ will appear and will denote a historical process different from K.

3.3 Existence and preliminary results

Definition. For two $(\hat{\mathcal{F}}_t^*)$-measurable processes X^1 and X^2, we will say that $X^1 = X^2, K$-a.e. if

$$X^1(s, \omega, y) = X^2(s, \omega, y), \forall s \leq t, K_t\text{-a.a. } y$$

for all fixed times $t \geq 0$.

We say that (X, Z) is a solution to the strong equation $(\text{SE})^1_{Z_0,K}$ if it satisfies

$$(\text{SE})^1_{Z_0,K} \quad \text{(a)} \;\; Z_t(\omega, y) = Z_0(\omega, y_0) + y_t - y_0 - \gamma \int_0^t Z_s(\omega, y) ds \; K - \text{a.e.}$$

$$\text{(b)} \;\; X_t(A) = \int 1(Z_t \in A) K_t(dy) \; \forall A \in \mathcal{B}(\mathbb{R}^d), \forall t \geq 0,$$

where Z_t is an $(\hat{\mathcal{F}}_t^*)$-predictable process and X_t is an $(\mathcal{F}_t)$-predictable process. That this equation has a pathwise unique solution does not follow immediately from Theorem 4.10 of Perkins [13]. There it is shown that equations of the above type (but with more general, interactive, drift and diffusion terms in (a)) have solutions if K is a *critical* historical Brownian motion.

Note that in the definitions of $(\mathrm{SE})^1_{Z_0,K}$ and $(\mathrm{SE})^2_{Y_0,K}$ given below, part (b) is unnecessary to solve the equations. It has been included to provide an easy comparison to the strong equation of chapter V.1 of [14].

Theorem 3.3.1. *There is a pathwise unique solution (X,Z) to $(SE)^1_{Z_0,K}$. That is, X is unique $\mathbb{P}$-a.s. and Z K-a.e. unique. Furthermore, the map $t \to X_t$ is continuous and X is a β-super-critical super Ornstein-Uhlenbeck process.*

Proof. Note that $\mathring{K}_t = e^{-\beta t} K_t$ defines a $(\Delta/2, 0, e^{-\beta t})$-Historical superprocess. Let $\hat{\mathbb{P}}^1_T$ be the Campbell measure associated with $\mathring{K}$ (note that if T is taken to be a random stopping time, this measure differs from $\hat{\mathbb{P}}_T$). The proof of Theorem V.4.1 of [14] with minor modifications shows that $(\mathrm{SE})^1_{Z_0,\mathring{K}}$ has a pathwise unique solution. This is because (K3) of Theorem 2.6 of [13] shows that under $\hat{\mathbb{P}}^1_T$, y_t is a Brownian motion stopped at time T and Proposition 2.7 of the same memoir can be used to replace Proposition 2.4 and Remark 2.5 (c) of [14] for the setting where the branching variance depends on time.

Once this is established, it is simple to deduce that if $(\mathring{X}, Z)$ is the solution of $(\mathrm{SE})^1_{Z_0,\mathring{K}}$, and we let

$$X_t(\cdot) \equiv e^{\beta t} \mathring{X}_t(\cdot) = \int 1(Z_t \in \cdot) K_t(dy)$$

then (X,Z) is the pathwise unique solution of $(\mathrm{SE})^1_{Z_0,K}$. The only thing to check this is that $Z_t(\omega,y) = Z_0(\omega,y_0) + y_t - y_0 - \gamma \int_0^t Z_s(\omega,y)ds$ K-a.e., but this follows from the fact that $\mathring{K}_t \ll K_t, \forall t$.

It can be shown by using by Theorem 2.14 of [13] that X satisfies the following martingale problem: For $\phi \in C_b^2(\mathbb{R}^d)$,

$$M_t(\phi) \equiv X_t(\phi) - X_0(\phi) - \int_0^t \int -\gamma x \cdot \nabla\phi(x) + \frac{\Delta}{2}\phi(x) X_s(dx)ds - \beta \int_0^t X_s(\phi)ds$$

is a martingale where $\langle M(\phi)\rangle_t = \int_0^t X_s(\phi^2)ds$. Then by Theorem II.5.1 of [14] this implies that X is a version of a super Ornstein-Uhlenbeck process, with initial distribution given by $K_0\left(Z_0^{-1}(\cdot)\right)$. $\square$

Remark 3.3.2. (a) Under the Lipschitz assumptions of Section V.1 of [14], one can in fact uniquely solve $(\mathrm{SE})^1_{Z_0,K}$ where K is a supercritical historical Brownian motion. The proof above can be extended with minor modifications.

(b) The proof of Theorem 3.3.1 essentially shows that under $\hat{\mathbb{P}}_T$, T fixed, the path process $y : \mathbb{R}^+ \times \hat{\Omega} \to \mathbb{R}^d$ such that $(t,(\omega,y)) \mapsto y_t$ is a d-dimensional Brownian motion stopped at T.

(c) Under $\hat{\mathbb{P}}_T$, T fixed, Z_t can be written explicitly as a function of the driving path: For $t \le T$,

$$\begin{aligned} e^{\gamma t}Z_t &= Z_0 + \int_0^t e^{\gamma s}dZ_s + \int_0^t Z_s\left(\gamma e^{\gamma s}\right)ds \\ &= Z_0 + \int_0^t e^{\gamma s}dy_s + \int_0^t e^{\gamma s}(-\gamma Z_s)ds + \int_0^t Z_s(\gamma e^{\gamma} s)ds. \end{aligned} \tag{3.3}$$

Hence,

$$Z_t(y) = e^{-\gamma t}Z_0 + \int_0^t e^{-\gamma(t-s)}dy_s.$$

We say (X',Y) is the solution to the stochastic equation $(\mathrm{SE})^2_{Y_0,K}$ if:

$$(\mathrm{SE})^2_{Y_0,K} \quad (a)\ \ Y_t(\omega,y) = Y_0(\omega,y) + y_t - y_0 + \gamma\int_0^t \bar{Y}_s(\omega) - Y_s(\omega,y)ds, K\text{-a.a. } y$$

$$(b)\ \ X'_t(A) = \int 1(Y_t \in A)K_t(dy), \forall A \in \mathcal{B}(\mathbb{R}^d), \forall t \ge 0.$$

where Y_t is an $(\hat{\mathcal{F}}^*_t)$-predictable process, X'_t is an $(\mathcal{F}_t)$-predictable process and

$$\bar{Y}_t \equiv \frac{\int xX'_t(dx)}{X'_t(1)} = \frac{K_t(Y_t)}{K_t(1)},$$

is the center of mass of X'.

Intuitively, X' is an interactive measure-valued diffusion where the representative particles are attracted or repelled by the centre of mass of the process (depending on the sign of γ). We will call this the SOU process with attraction (or repulsion) to the COM. It will be shown below that questions regarding (X', Y) can be naturally reformulated as questions about the ordinary SOU process.

Notation. Unless stated otherwise, (X, Z) will refer to a solution of $(\text{SE})^1_{Z_0,K}$ and (X', Y) to the solution of $(\text{SE})^2_{Y_0,K}$.

Definition. For an arbitrary $\hat{\mathfrak{F}}^*_t$-adapted, $\mathbb{R}^d$-valued process z_t, define the centre of mass (COM), $\bar{z}_t$, with respect to K_t as follows:

$$\bar{z}_t = \frac{K_t(z_t)}{K_t(1)}.$$

We also let

$$\tilde{X}_t(\cdot) \equiv \frac{X_t(\cdot)}{X_t(1)} = \frac{K_t(Z_t \in \cdot)}{K_t(1)},\ \tilde{X}'_t(\cdot) \equiv \frac{X'_t(\cdot)}{X'_t(1)} = \frac{K_t(Y_t \in \cdot)}{K_t(1)} \text{ and } \tilde{K}_t(\cdot) = \frac{K_t(\cdot)}{K_t(1)}.$$

Note that as K is supercritical, K_t survives indefinitely on a set of positive probability S, and goes extinct on the set of positive probability S^c. Hence we can only make sense of $\bar{z}_t$ for $t < \eta$ where η is the extinction time.

Next we show that there exists a solution to $(\text{SE})^2_{Y_0,K}$. We cannot adapt the proof of Theorem V.4.1 of [14] to accomplish this as the Lipschitz assumptions on the drift coefficients fail to be satisfied by our model, due to the normalizing factor $K_t(1)$ in the definition of $\bar{Y}_t$. Instead we will utilize the solution of $(\text{SE})^1_{Y_0,K}$ to define a unique solution for $(\text{SE})^2_{Y_0,K}$.

Theorem 3.3.3. *There is a pathwise unique solution to $(SE)^2_{Y_0,K}$.*

Proof. Suppose there exists a solution Y satisfying $(\text{SE})^2_{Y_0,K}$. Then under $\hat{\mathbb{P}}_T$, Y_t can be written as a function of the driving path y and $\bar{Y}$. Using integration by parts gives

$$\begin{aligned} e^{\gamma t} Y_t &= Y_0 + \int_0^t e^{\gamma s} dY_s + \int_0^t \gamma e^{\gamma s} Y_s ds \\ &= Y_0 + \int_0^t e^{\gamma s} dy_s + \int_0^t \gamma e^{\gamma s} \bar{Y}_s ds \end{aligned}$$

and hence

$$Y_t = e^{-\gamma t} Y_0 + \int_0^t e^{\gamma(s-t)} dy_s + \int_0^t \gamma e^{\gamma(s-t)} \bar{Y}_s ds. \tag{3.4}$$

If (X,Z) is the solution to $(\text{SE})^1_{Z_0,K}$ where $Z_0 = Y_0$, then note that by Remark 3.3.2 (c),

$$Y_t = Z_t + \int_0^t \gamma e^{\gamma(s-t)} \bar{Y}_s ds.$$

By taking the normalized measure $\tilde{K}_t$ of both sides of the above equation, we get

$$\bar{Y}_t = \bar{Z}_t + \gamma \int_0^t e^{-\gamma(t-s)} \bar{Y}_s ds \tag{3.5}$$

Hence $\bar{Y}_t$ is seen to satisfy an integral equation. This is a Volterra Integral Equation of the second kind (see Equation 2.2.1 of [16]) and therefore can be solved pathwise to give

$$\bar{Y}_t = \bar{Z}_t + \gamma \int_0^t \bar{Z}_s ds,$$

which is easily verified using integration by parts. Also, if $\bar{Y}_t^1$ is a second process which solves (3.5), then

$$\begin{aligned} |\bar{Y}_t - \bar{Y}_t^1| &= \left| \gamma \int_0^t e^{-\gamma(t-s)} (\bar{Y}_s - \bar{Y}_s^1) ds \right| \\ &\leq |\gamma| \int_0^t e^{-\gamma(t-s)} |\bar{Y}_s - \bar{Y}_s^1| ds. \end{aligned}$$

By Gronwall's inequality, this implies $\bar{Y}_t = \bar{Y}_t^1$, for all t and ω. Pathwise uniqueness of X' follows from the uniqueness of the solution to $(\text{SE})^1_{Z_0,K}$ and the uniqueness of the process $\bar{Y}_t$ solving (3.5).

We have shown that if there exists a solution to $(\text{SE})^2_{Y_0,K}$, then it is necessarily pathwise unique. Turning now to existence to complete the proof, we work in the opposite order and define Y and X' as functions of the pathwise unique solution to

$(\text{SE})^1_{Z_0,K}$ where $Z_0 = Y_0$:

$$Y_t = Z_t + \gamma \int_0^t \bar{Z}_s ds$$
$$X_t'(\cdot) = K_t(Y_t \in \cdot).$$

Then $\bar{Y}_t$ satisfies the integral equation (3.5), and hence

$$\int_0^t \bar{Z}_s ds = \int_0^t e^{-\gamma(t-s)} \bar{Y}_s ds.$$

Therefore

$$\begin{aligned} Y_t &= Z_t + \gamma \int_0^t e^{-\gamma(t-s)} \bar{Y}_s ds \\ &= e^{-\gamma t} Y_0 + \int_0^t e^{-\gamma(t-s)} dy_s + \gamma \int_0^t e^{-\gamma(t-s)} \bar{Y}_s ds, \end{aligned}$$

and so

$$e^{\gamma t} Y_t = Y_0 + \int_0^t e^{\gamma s} dy_s + \gamma \int_0^t e^{\gamma s} \bar{Y}_s ds.$$

Multiplying by $e^{-\gamma t}$ and using integration by parts shows

$$Y_t = Y_0 + y_t - y_0 + \gamma \int_0^t (\bar{Y}_s - Y_s) ds$$

which holds for K-a.e. y, thereby showing (X', Y) satisfies $(\text{SE})^2_{Y_0,K}$.

□

Remark 3.3.4. Some useful equivalences in the above proof are collected below. If $Y_0 = Z_0$, then for $t < \eta$,

(a) $Y_t = Z_t + \gamma \int_0^t \bar{Z}_s ds$

(b) $\bar{Y}_t = \bar{Z}_t + \gamma \int_0^t \bar{Z}_s ds$

(c) $Y_t - \bar{Y}_t = Z_t - \bar{Z}_t$

(d) $\int_0^t e^{-\gamma(t-s)} \bar{Y}_s ds = \int_0^t \bar{Z}_s ds.$

The significance of these equations is that they intimately tie the behaviour of the SOU process with attraction to its center of mass to the SOU process with attraction to the origin. Part (a) says that the SOU process with attraction to the center of mass is the same as the ordinary SOU process pushed by the position of its center of mass.

Definition. Let the process M_t be defined for $t \in [0,\zeta)$ where $\zeta \leq \infty$ possibly random. M is called a local martingale on its lifetime if there exist stopping times $T_N \uparrow \zeta$ such that $M_{T_N \wedge \cdot}$ is a martingale for all N. The interval $[0,\zeta)$ is called the lifetime of M.

We now consider the martingale problem for X'. For $\phi : \mathbb{R}^d \to \mathbb{R}$, let $\bar{\phi}_t \equiv K_t(\phi(Y_t))/K_t(1)$. Note that the lifetime of the process $\bar{\phi}$ is $[0,\eta)$. Then the following theorem holds:

Theorem 3.3.5. *For $\phi \in C_b^2(\mathbb{R}^d,\mathbb{R})$, and $t < \eta$,*

$$\bar{\phi}_t = \bar{\phi}_0 + M_t + \int_0^t \bar{b}_s ds$$

where

$$b_s = \gamma \nabla \phi(Y_s) \cdot (\bar{Y}_s - Y_s) + \frac{1}{2} \Delta \phi(Y_s)$$

and M_t is a continuous local martingale on its lifetime such that

$$M_t = \int_0^t \int \frac{\phi(Y_s) - \bar{\phi}_s}{K_s(1)} dM(s,y)$$

and hence has quadratic variation given by

$$[M]_t = \int_0^t \frac{\overline{\phi_s^2} - (\bar{\phi}_s)^2}{K_s(1)} ds.$$

Proof. The proof is not very difficult; one need only use Itô's Lemma followed by some slight modifications of theorems in Chapter V of [14] to deal with the drift introduced in the historical martingale problem due to the supercritical branching.

Let T be a fixed time and $t \leq T$. Recall that under $\hat{\mathbb{P}}_T$, y is a stopped Brownian motion by Remark 3.3.2 (b), and hence $Y_t(y)$ is a stopped Ornstein-Uhlenbeck process (attracting to $\bar{Y}_t$). Hence, by the classical Itô's Lemma we have, under $\hat{\mathbb{P}}_T$,

$$\begin{aligned}\phi(Y_t) &= \phi(Y_0) + \int_0^t \nabla\phi(Y_s)\cdot dY_s + \frac{1}{2}\sum_{i,j\leq d}\int_0^t \phi_{ij}(Y_s)d[Y^i,Y^j]_s \\ &= \phi(Y_0) + \int_0^t \nabla\phi(Y_s)\cdot dy_s + \int_0^t \gamma\nabla\phi(Y_s)\cdot(\bar{Y}_s - Y_s) + \frac{1}{2}\Delta\phi(Y_s)ds.\end{aligned}$$

Then

$$\begin{aligned}K_t(\phi(Y_t)) &= K_t(\phi(Y_0)) + K_t\left(\int_0^t \nabla\phi(Y_s)\cdot dy_s\right) + K_t\left(\int_0^t b_s ds\right) \\ &= K_0(\phi(Y_0)) + \int_0^t\int \phi(Y_0)dM(s,y) + \beta\int_0^t K_s(\phi(Y_0))\,ds \\ &+ \int_0^t\int\left[\int_0^s \nabla\phi(Y_r)\cdot dy_r\right]dM(s,y) + \beta\int_0^t K_s\left[\int_0^s \nabla\phi(Y_r)\cdot dy_r\right]ds \\ &+ \int_0^t\int b_s dM(s,y) + \beta\int_0^t K_s(b_s)ds + \int_0^t K_s(b_s)ds \\ &= K_0(\phi(Y_0)) + \int_0^t\int \phi(Y_s)dM(s,y) \\ &\qquad + \beta\int_0^t K_s(\phi(Y_s))\,ds + \int_0^t K_s(b_s)ds.\end{aligned}$$

The equality in the third line to $K_t\left(\int_0^t \nabla\phi(Y_s)\cdot dy_s\right)$ follows from Proposition 2.13 of [13]. The equality of the fourth line to the last term in the first line follows from a generalization of Proposition V.2.4 (b) of [14]. The last equality then follows from collecting like terms and using the definition of $\bar{Y}$.

Supposing $t<\eta$, Itô's formula implies (with properties of $K_t(1)$ and $K_t(\phi(Y_t))$)

$$\begin{aligned}\bar{\phi}_t &= \frac{K_t(\phi(Y_t))}{K_t(1)} \\ &= \bar{\phi}_0 + \int_0^t\int\left[\frac{\phi(Y_s)}{K_s(1)} - \frac{K_s(\phi(Y_s))}{K_s(1)^2}\right]dM(s,y) + \int_0^t \frac{K_s(b_s)}{K_s(1)}ds\end{aligned}$$

Since ϕ is bounded, the stochastic integral term can be localized using the stopping times $T_N \equiv \min\{t : K_t(1)\geq N \text{ or } K_t(1)\leq 1/N\}\wedge N$ and hence it is a local martingale on $[0,\eta)$. It is easy to check that it has the appropriate quadratic variation. □

Remark 3.3.6. The method of Theorem 3.3.5 can be used to show that for the

β-supercritical SOU process, X, for $\phi \in C_b^2(\mathbb{R}^d, \mathbb{R})$ and $t < \eta$,

$$\tilde{X}_t(\phi) = \tilde{K}_t(\phi(Z_t)) = \tilde{K}_0(\phi(Z_0)) + N_t + \int_0^t \tilde{K}_s(L\phi(Z_s))ds$$

where

$$L\phi(x) = -\gamma x \cdot \nabla\phi(x) + \frac{1}{2}\Delta\phi(x)$$

and N_t is a continuous local martingale on its lifetime such that

$$N_t = \int_0^t \int \frac{\phi(Z_s) - \tilde{K}_s(\phi(Z_s))}{K_s(1)} dM(s,y)$$

and hence has quadratic variation given by

$$[N]_t = \int_0^t \frac{\tilde{K}_s\left(\phi^2(Z_s)\right) - \tilde{K}_s(\phi(Z_s))^2}{K_s(1)} ds.$$

Lemma 3.3.7. *Let X be a $(\Delta/2, \beta, 1)$ superprocess with $\beta > 0$ constant. Then there is a non-negative random variable W such that*

$$e^{-\beta t} X_t(1) \to W \text{ a.s.}$$

and $\{\eta < \infty\} = \{W = 0\}$ almost surely.

Proof. Note that $\mathring{X}_t = e^{-\beta t} X_t$ is a $(\Delta/2, 0, e^{-\beta t})$-superprocess. The martingale problem then shows that $\mathring{X}_t(1)$ is a non-negative $(\mathcal{F}_t)$-martingale and therefore converges almost surely by the Martingale Convergence Theorem to a random variable W. It follows that $\{\eta < \infty\} \subset \{W = 0\}$, since 0 is an absorbing state for $X_t(1)$. Exercise II.5.3 in [14] shows that

$$\mathbb{P}(\eta < \infty) = e^{-2\beta X_0(1)}.$$

The same exercise also shows

$$\mathbb{P}_{0,X_0}(\exp(-\lambda \mathring{X}_t(1))) = \exp\left(-\frac{2\lambda\beta\mathring{X}_0(1)}{2\beta + \lambda\left(1 - e^{-\beta t}\right)} \right).$$

Now sending $t \to \infty$ gives

$$\mathbb{P}_{0,X_0}(e^{-\lambda W}) = \exp\left(-\frac{2\beta\lambda \mathring{X}_0(1)}{2\beta+\lambda}\right),$$

and sending $\lambda \to \infty$ gives

$$\mathbb{P}(W=0) = e^{-2\beta \mathring{X}_0(1)} = \mathbb{P}(\eta < \infty)$$

since $X_0 = \mathring{X}_0$. Therefore $\{\eta < \infty\} = \{W = 0\}$ almost surely. □

Define $h(\delta) = (\delta \ln^+(1/\delta))^{1/2}$ where $\ln^+(x) = \ln x \vee 1$. Let $S(\delta, c) = \{y : |y_t - y_s| < ch(|t-s|), \forall t, s \text{ with } |t-s| \leq \delta\}$.

Lemma 3.3.8. *Let K be a supercritical historical Brownian motion, with drift β, branching variance* 1*, and initial measure X_0. For $c_0 > 6$ fixed, $c(t) = \sqrt{t+c_0}$, there exists a.s. $\delta(\omega) > 0$ such that $\operatorname{supp}(K_t(\omega)) \subset S(\delta(\omega), c(t))$ for all t. Further, given c_0, $\mathbb{P}(\delta < \lambda) < p_{c_0}(\lambda)$ where $p_{c_0}(\lambda) \downarrow 0$ as $\lambda \downarrow 0$ and for any $\alpha > 0$, c_0 can be chosen large enough so that $p_{c_0}(\lambda) = C(d, c_0)\lambda^\alpha$ for $\lambda \in [0,1]$.*

Proof. We follow the proof of Perkins [14], Theorem III.1.3 (a). First note that if H is another supercritical historical Brownian motion starting at time τ with initial measure m under $\mathbb{Q}_{\tau,m}$ and A a Borel subset of $C(\mathbb{R}^n)$, then the process defined by

$$H'_t(\cdot) = H_t(\cdot \cap \{y : y_\tau \in A\})$$

is also a supercritical historical Brownian motion starting at time τ with initial measure m' given by $m'(\cdot) = m(\cdot \cap A)$ under $\mathbb{Q}_{\tau,m}$. Then using the extinction probabilities for H', we have

$$\mathbb{Q}_{\tau,m}(H_t(\{y : y_\tau \in A\}) = 0 \; \forall t \geq s) = \exp\left\{-\frac{2\beta m(A)}{1-e^{-\beta(s-\tau)}}\right\}. \tag{3.6}$$

Using the Markov property for K at time $\frac{j}{2^n}$ and (3.6) gives

$$\mathbb{P}\left[\exists t > \frac{j+1}{2^n} \text{ s.t. } K_t\left(\left\{y : \left|y\left(\frac{j}{2^n}\right) - y\left(\frac{j-1}{2^n}\right)\right| > c\left(\frac{j}{2^n}\right) h\left(2^{-n}\right)\right\}\right) > 0\right] \tag{3.7}$$

$$= \mathbb{P}\left[1 - \exp\left\{-\frac{2\beta K_{\frac{j}{2^n}}\left(\left\{y : \left|y\left(\frac{j}{2^n}\right) - y\left(\frac{j-1}{2^n}\right)\right| > c\left(\frac{j}{2^n}\right) h\left(2^{-n}\right)\right\}\right)}{1 - e^{-\beta\frac{1}{2^n}}}\right\}\right]$$

$$\leq \mathbb{P}\left[\frac{2\beta K_{\frac{j}{2^n}}\left(\left\{y : \left|y\left(\frac{j}{2^n}\right) - y\left(\frac{j-1}{2^n}\right)\right| > c\left(\frac{j}{2^n}\right) h\left(2^{-n}\right)\right\}\right)}{1 - e^{-\beta\frac{1}{2^n}}}\right]$$

$$\leq \mathbb{P}\left[\frac{2\beta K_{\frac{j}{2^n}}\left(\left\{y : \left|y\left(\frac{j}{2^n}\right) - y\left(\frac{j-1}{2^n}\right)\right| > c\left(\frac{j}{2^n}\right) h\left(2^{-n}\right)\right\}\right)}{\frac{\beta}{2^n} - \frac{\beta^2}{2^{2n+1}}}\right]$$

$$\leq \frac{2^{n+1}}{1 - \frac{\beta}{2^{n+1}}}\mathbb{P}\left(K_{\frac{j}{2^n}}(1)\right)\hat{\mathbb{P}}_{\frac{j}{2^n}}\left(\left|y\left(\frac{j}{2^n}\right) - y\left(\frac{j-1}{2^n}\right)\right| > c\left(\frac{j}{2^n}\right) h\left(2^{-n}\right)\right).$$

Since under the normalized mean measure, y is a stopped Brownian motion by Remark 3.3.2(b), we use tail estimates to get the following bound:

$$\begin{aligned}(3.7) &\leq \frac{2^{n+1}}{1 - \frac{\beta}{2^{n+1}}}\mathbb{P}\left(K_{\frac{j}{2^n}}(1)\right) c_d n^{d/2-1} 2^{-nc\left(\frac{j}{2^n}\right)^2/2} \\ &\leq 2^{n+2} 2^{\frac{\beta j \ln 2}{2^n}} \mathbb{P}\left(X_0(1)\right) c_d n^{d/2-1} 2^{-n\left(\frac{j}{2^n}+c_0\right)/2}.\end{aligned}$$

Hence, summing over j from 1 to $n2^n$ gives

$$\mathbb{P}\Bigg[\text{There exists } 1 \leq j \leq n2^n \text{ s.t. } \exists t > \frac{j+1}{2^n} \text{ s.t.}$$
$$K_t\left(\left\{y : \left|y\left(\frac{j}{2^n}\right) - y\left(\frac{j-1}{2^n}\right)\right| > c\left(\frac{j}{2^n}\right)h\left(2^{-n}\right)\right\}\right) > 0\Bigg]$$
$$\leq \mathbb{P}\left(X_0(1)\right)c_d n^{d/2-1}2^{n+2-nc_0/2}\sum_{j=1}^{n2^n} 2^{\frac{\beta j \ln 2}{2^n} - n\left(\frac{j}{2^n}\right)}$$
$$\leq \mathbb{P}\left(X_0(1)\right)c_d n^{d/2-1}2^{-2n+2}\sum_{j=1}^{n2^n} 2^{-\frac{j}{2^n}}$$
$$\leq \mathbb{P}\left(X_0(1)\right)c_d n^{d/2-1}2^{-n+2},$$

where we have used the fact that $c_0 > 6$. Hence the sum over n of the above shows by Borel-Cantelli that there exists for almost sure ω, $N(\omega)$ such that for all $n > N$, for all $1 \leq j \leq n2^n$, for all $t \geq \frac{j+1}{2^n}$, for K_t-a.a. y,

$$\left|y\left(\frac{j-1}{2^n}\right) - y\left(\frac{j}{2^n}\right)\right| < c\left(\frac{j}{2^n}\right)h\left(2^{-n}\right).$$

Letting $\delta(\omega) = 2^{-N(\omega)}$, note that on the dyadics, by above, we have that

$$\mathbb{P}(\delta < \lambda) = \mathbb{P}\left(N > -\frac{\ln\lambda}{\ln 2}\right)$$
$$= \mathbb{P}\Bigg[\exists n > -\frac{\ln\lambda}{\ln 2}, \exists j \leq n2^n, \exists t > \frac{j+1}{2^n}, \text{ s.t.}$$
$$K_t\left(\left\{y : \left|y\left(\frac{j}{2^n}\right) - y\left(\frac{j-1}{2^n}\right)\right| > c\left(\frac{j}{2^n}\right)h\left(2^{-n}\right)\right\}\right) > 0\Bigg]$$
$$\leq \sum_{n=\lfloor -\frac{\ln\lambda}{\ln 2}\rfloor} C'(d,c_0)n^{d/2-1}2^{2n-nc_0/2}$$
$$\leq C(d,c_0,\varepsilon)\lambda^{\frac{c_0}{2}-\varepsilon}$$

where ε can be chosen to be arbitrarily small (though the constant C will increase as it decreases). The rest of the proof follows as in Theorem III.1.3(a) of [14], via an argument similar to Levy's proof for the modulus of continuity for Brownian motion. □

The following moment estimates are useful in establishing the convergence of $\bar{Y}_t$. Recall that η is the extinction time of K.

Lemma 3.3.9. *Assume* $\mathbb{P}\left(\tilde{K}_0\left(|Z_0|^2+|y_0|^2\right)\right)<\infty$. *Then,*

$$\mathbb{P}\left(\overline{|Y_t|^2};t<\eta\right)<A(\gamma,t),$$

where

$$A(\gamma,t)=\begin{cases} O\left(1+t^6e^{-2\gamma t}\right) \text{ if } \gamma<0 \\ O\left(1+t^5\right) \text{ if } \gamma\geq 0. \end{cases}$$

Proof. Assume that $t<\eta$. Recall

$$Z_t(\omega,y)\equiv e^{-\gamma t}Z_0+\int_0^t e^{\gamma(s-t)}dy_s.$$

Note that below '$\lesssim$' denotes less than up to multiplicative constants independent of t and y. Suppose that $y\in S(\delta,c(t))$, where $S(\delta,c(t))$ is the same as in the previous lemma. Then, as $Y_t=Z_t+\gamma\int_0^t\bar{Z}_s ds$,

$$\begin{aligned}|Y_t|^2 &\lesssim |Z_t|^2+\gamma^2 t\int_0^t|\bar{Z}_s|^2\,ds\\ &\lesssim |Z_t|^2+\gamma^2 t\int_0^t\overline{|Z_s|^2}ds\end{aligned}$$

by Cauchy-Schwartz and Jensen's inequality. Therefore integrating with respect to the normalized measure gives

$$\overline{|Y_t|^2}\leq\overline{|Z_t|^2}+\gamma^2 t\int_0^t\overline{|Z_s|^2}ds \tag{3.8}$$

and therefore we need only find the appropriate bounds for expectation of $\overline{|Z_t|^2}$ to get the result.

After another few applications of Cauchy-Schwartz and integrating by parts,

$$
\begin{aligned}
|Z_t|^2 &\lesssim e^{-2\gamma t}|Z_0|^2 + \left| e^{-\gamma t}\int_0^t e^{\gamma s} dy_s \right|^2 \\
&\lesssim e^{-2\gamma t}|Z_0|^2 + e^{-2\gamma t}\left| e^{\gamma t} y_t - y_0 - \gamma \int_0^t y_s e^{\gamma s} ds \right|^2 \\
&\lesssim e^{-2\gamma t}\left(|Z_0|^2 + |y_0|^2\right) + |y_t|^2 + \gamma^2 t \int_0^t |y_s|^2 e^{-2\gamma(t-s)} ds.
\end{aligned}
$$

As $y \in S(\delta, c(t))$,

$$
\begin{aligned}
|Z_t|^2 \lesssim\ & e^{-2\gamma t}\left(|Z_0|^2 + |y_0|^2\right) + |y_0|^2 + \left(\frac{tc(t)h(\delta)}{\delta}\right)^2 \\
& + \gamma^2 t \int_0^t \left[|y_0|^2 + \left(\frac{sc(s)h(\delta)}{\delta}\right)^2 \right] e^{-2\gamma(t-s)} ds \\
\lesssim\ & e^{-2\gamma t}\left(|Z_0|^2 + |y_0|^2\right) + |y_0|^2 + |y_0|^2 \gamma t \left(1 - e^{-2\gamma t}\right)/2 \\
& + c(t)^2 \left(\frac{h(\delta)}{\delta}\right)^2 \left(t^2 + \gamma t^3 \left(1 - e^{-2\gamma t}\right)\right) \\
\lesssim\ & (1 + |\gamma| t)\left(1 + e^{-2\gamma t}\right)\left(|Z_0|^2 + |y_0|^2\right) \\
& + c(t)^2 \left(\frac{h(\delta)}{\delta}\right)^2 \left(t^2 + \gamma t^3 \left(1 - e^{-2\gamma t}\right)\right)
\end{aligned}
$$

Integrating by the normalized measure $\tilde{K}_t$,

$$
\begin{aligned}
\overline{|Z_t|^2} \lesssim\ & (1 + |\gamma| t)\left(1 + e^{-2\gamma t}\right)\tilde{K}_t\left(|Z_0|^2 + |y_0|^2\right) \\
& + c(t)^2 \left(\frac{h(\delta)}{\delta}\right)^2 \left(t^2 + \gamma t^3 \left(1 - e^{-2\gamma t}\right)\right).
\end{aligned}
$$

Then using (3.8) and using the above bound on $\overline{|Z_t|^2}$ gives

$$
\begin{aligned}
\overline{|Y_t|^2} &\le \overline{|Z_t|^2} + \gamma^2 t \int_0^t \overline{|Z_s|^2} ds \\
&\lesssim \left(1 + t^2\right)\left(1 + e^{-2\gamma t}\right)\left(\tilde{K}_t\left(|Z_0|^2 + |y_0|^2\right) + \int_0^t \tilde{K}_s\left(|Z_0|^2 + |y_0|^2\right) ds\right) \qquad (3.9) \\
&\quad + (1 + t)\, c(t)^2 \left(\frac{h(\delta)}{\delta}\right)^2 \left(t^2 + \gamma t^3 \left(1 - e^{-2\gamma t}\right)\right).
\end{aligned}
$$

Note that $\phi(y) \equiv |Z_0(y)|^2 + |y_0|^2$ and $\phi_n(y) \equiv \phi(y)1(|y| \leq n)$ are $\hat{\mathfrak{F}}_0$ measurable. By applying Itô's formula to $K_t(\phi_n)K_t^{-1}(1)$ and using the decomposition $K_t(\phi_n) = K_0(\phi_n) + \int_0^t \phi_n(y)dM(s,y) + \beta \int_0^t K_s(\phi_n)ds$ (which follows from Proposition 2.7 of [13]) we get

$$\tilde{K}_t(\phi_n) = \tilde{K}_0(\phi_n) + N_t(\phi_n),$$

where $N_t(\phi_n)$ is a local martingale until time η, for each n. In fact the sequence of stopping times $\{T_N\}$ appearing in Theorem 3.3.5 can be used to localize each $N_t(\phi_n)$. Applying first the Monotone Convergence Theorem and then localizing gives,

$$\begin{aligned}
\mathbb{P}\left(\tilde{K}_t(\phi); t < \eta\right) &= \lim_{n\to\infty} \mathbb{P}\left(\tilde{K}_t(\phi_n); t < \eta\right) \\
&= \lim_{n\to\infty} \lim_{N\to\infty} \mathbb{P}\left(\tilde{K}_t(\phi_n)\mathbf{1}\left(t < T_N\right)\right) \\
&= \lim_{n\to\infty} \lim_{N\to\infty} \mathbb{P}\left(\tilde{K}_{t\wedge T_N}(\phi_n) - \tilde{K}_{T_N}(\phi_n)\mathbf{1}\left(t \geq T_N\right)\right) \\
&\leq \lim_{n\to\infty} \lim_{N\to\infty} \mathbb{P}\left(\tilde{K}_{t\wedge T_N}(\phi_n)\right) \\
&= \lim_{n\to\infty} \lim_{N\to\infty} \mathbb{P}\left(\tilde{K}_0(\phi_n) + N_{t\wedge T_N}(\phi_n)\right) \\
&= \lim_{n\to\infty} \mathbb{P}\left(\tilde{K}_0(\phi_n)\right) \\
&= \mathbb{P}\left(\tilde{K}_0(\phi)\right),
\end{aligned}$$

where we have used the positivity of ϕ_n to get the fourth line and the Monotone Convergence Theorem in the last line. Further, note that

$$\begin{aligned}
\mathbb{P}\left(\int_0^t \tilde{K}_s(\phi)ds; t < \eta\right) &\leq \mathbb{P}\left(\int_0^t \tilde{K}_s(\phi)\mathbf{1}(s < \eta)ds\right) \\
&= \int_0^t \mathbb{P}\left(\tilde{K}_s(\phi); s < \eta\right) ds \\
&\leq t\mathbb{P}\left(\tilde{K}_0(\phi)\right),
\end{aligned}$$

by the calculation immediately above. Therefore, taking expectations in (3.9) and plugging in $c(t)=\sqrt{c_0+t}$ gives

$$\mathbb{P}\left(\overline{|Y_t|^2};t<\eta\right)\lesssim\left(1+t^3\right)\left(1+e^{-2\gamma t}\right)\mathbb{P}\left(\tilde{K}_0\left(\phi\right)\right)+t^5\left(1+te^{-2\gamma t}\right)\mathbb{P}\left(\frac{h(\delta)^2}{\delta^2}\right).$$

Now let c_0 be chosen so that $\operatorname{supp}(K_t)\subset S(\delta,c(t))$ and $p_{c_0}(\lambda)=C\lambda^{\alpha}$ for $\lambda\in[0,1]$. Note that

$$\begin{aligned}\mathbb{P}\left(\frac{h(\delta)^2}{\delta^2}\right)=\mathbb{P}\left(\frac{\ln^+(1/\delta)}{\delta}\right)&=\int_0^\infty\left(\frac{\ln^+(1/\lambda)}{\lambda}\right)d\mathbb{P}(\delta<\lambda)\\ &\leq\int_0^1\left(\frac{\ln^+(1/\lambda)}{\lambda}\right)d\mathbb{P}(\delta<\lambda)+\mathbb{P}(\delta>1)\\ &=\lim_{\lambda\downarrow 0}\left(\frac{\ln^+(1/\lambda)}{\lambda}\mathbb{P}(\delta<\lambda)\right)-\mathbb{P}(\delta<1)\\ &\qquad\qquad-\int_0^1\mathbb{P}(\delta<\lambda)d\left(\frac{\ln^+(1/\lambda)}{\lambda}\right)+\mathbb{P}(\delta>1)\\ &<\infty,\end{aligned}$$

by choosing the original constant c_0 so that α is large ($\alpha\geq 2$ is enough).

□

Remark 3.3.10. **(a)** The proof of Lemma 3.3.9, under the same hypotheses yields

$$\mathbb{P}\left(\overline{|Z_t|^2};t<\eta\right)<A(\gamma,t).$$

(b) Lemma 3.3.9 and its proof can be extended to show that for any positive integer k if $\mathbb{P}\left(\tilde{K}_0\left(|Z_0|^k+|y_0|^k\right)\right)<\infty$, then there exists a function $B(\gamma,t,k)$ polynomial in t if $\gamma\geq 0$, exponential if $\gamma<0$ such that

$$\mathbb{P}\left(\overline{|Y_t|^k};t<\eta\right)<B(\gamma,t,k)\text{ and }\mathbb{P}\left(\overline{|Z_t|^k};t<\eta\right)<B(\gamma,t,k).$$

We only need to note that the exponent α in Lemma 3.3.8 can be made arbitrarily large by choosing a sufficiently large constant c_0. Hence by choosing

α appropriately, we can show that

$$\mathbb{P}\left[\left(\frac{h(\delta)}{\delta}\right)^k\right] < \infty,$$

which can then be used to adapt the proof above.

3.4 Convergence

For this section we will assume that $Z_0(y) = Y_0(y) = y_0$. Then by the definition of X and X', $\int \phi(y_0)K_0(dy) = \int \phi(x)X_0(dx) = \int \phi(x)X_0'(dx)$.

Assume that $X_0'(1) = X_0(1) < \infty$ for Theorems 3.4.1, 3.4.2 and 3.4.3 below.

Theorem 3.4.1. *On S^c, $\bar{Y}_t$ and $\bar{Z}_t$ converge as $t \uparrow \eta < \infty$, $\mathbb{P}$-a.s., for any $\gamma \in \mathbb{R}$.*

Proof. Assume for now that $\mathbb{P}\left(\tilde{K}_0\left(|y_0|^2\right)\right) < \infty$ (the case where $K_0 = 0$ can be ignored without loss of generality). By Theorem 3.3.5,

$$\bar{Y}_t = \bar{Y}_0 + \int_0^t \int \frac{Y_s - \bar{Y}_s}{K_s(1)} dM(s,y)$$

and therefore is a local martingale on its lifetime with reducing sequence $\{T_N\}$ as defined in the proof of the same theorem. Using Doob's weak inequality and Lemma 3.3.9,

$$\begin{aligned}
\mathbb{P}\left(\sup_{s<t\wedge\eta} |\bar{Y}_s| > n\right) &= \lim_{N\to\infty} \mathbb{P}\left(\sup_{s<t\wedge T_N} |\bar{Y}_s| > n\right) \\
&\le \lim_{N\to\infty} \frac{1}{n^2}\mathbb{E}(|\bar{Y}_{t\wedge T_N}|^2) \\
&< \frac{A(\gamma,t)}{n^2}.
\end{aligned}$$

By the first Borel-Cantelli lemma,

$$\mathbb{P}\left(\sup_{s<t\wedge\eta} |\bar{Y}_t| > n \ i.o.\right) = 0.$$

It follows that

$$\liminf_{s\to t\wedge\eta} \bar{Y}_s > -\infty \text{ and } \limsup_{s\to t\wedge\eta} \bar{Y}_s < \infty$$

which implies that on the set $\{\eta < t\}$, $\bar{Y}_s$ converges, by Theorem IV.34.12 of Rogers and Williams [18]. This shows convergence on the extinction set as $S^c = \cup_t \{\eta < t\}$.

Note that if $\nu(\cdot) = \mathbb{P}(K_0 \in \cdot)$. Theorem II.8.3 of [14] gives

$$\mathbb{P}(K \in \cdot) = \int \mathbb{P}_{K_0}(K \in \cdot)\, d\nu(K_0),$$

where $\mathbb{P}_{K_0}$ is the law of a historical Brownian motion with initial distribution δ_{K_0}. Hence, the a.s. convergence of $\bar{Y}_t$ in the case where $\tilde{K}_0\left(|y_0|^2\right)$ is finite in mean imply

$$\begin{aligned}\mathbb{P}\left(\lim_{t\uparrow\eta} \bar{Y}_t \text{ exists}; S^c\right) &= \int \mathbb{P}_{K_0}\left(\lim_{t\uparrow\eta} \bar{Y}_t \text{ exists}; S^c\right) d\nu(K_0) \\ &= \int \mathbb{P}_{K_0}(S^c)\, d\nu(K_0) \\ &= \mathbb{P}(S^c)\end{aligned} \tag{3.10}$$

if $K_0\left(|y_0|^2+1\right) < \infty$, ν-a.s.

Finally, to get rid of the assumption that $K_0\left(|y_0|^2\right) < \infty$ note that Corollary 3.4 of [13] ensures that if $K_0(1) < \infty$, then at any time $t > 0$, K_t (and hence X_t, X_t') is compactly supported. Therefore, letting $S_r = \{K_r \neq 0\}$ we see that

$$\begin{aligned}\mathbb{P}\left(\lim_{t\uparrow\eta} \bar{Y}_t \text{ does not exist}, S^c\right) &= \mathbb{P}\left(\bigcup_{r\in\mathbb{N}} \left\{\lim_{t\uparrow\eta} \bar{Y}_t \text{ does not exist}, S_{1/r}, S^c\right\}\right) \\ &\leq \sum_{r\in\mathbb{N}} \mathbb{P}\left(\mathbb{P}_{K_{1/r}}\left(\lim_{t\uparrow\eta} \bar{Y}_t \text{ does not exist}, S^c\right) \mathbf{1}\left(S_{1/r}\right)\right) \\ &= 0\end{aligned}$$

by 3.10 since $K_{1/r}$ a.s. compact implies that $K_{1/r}\left(|y_{1/r}|^2\right) < \infty$ holds. This completes the proof for the convergence of $\bar{Y}$ on S^c in its full generality.

The convergence of $\bar{Z}_t$ now follows from the convergence of $\bar{Y}_t$ and Equation 3.5.

□

Theorem 3.4.2. *On S the following hold:*

(a) If $\gamma > 0$

$$\bar{Z}_t \xrightarrow{a.s.} 0 \text{ and } \bar{Y}_t \xrightarrow{a.s.} \gamma \int_0^\infty \bar{Z}_s ds,$$

and this integral is finite almost surely.

(b) If $\gamma = 0$, then $\bar{Z}_t = \bar{Y}_t$ converges almost surely.

Proof. Let $\gamma \geq 0$ and as in the previous theorem assume $\mathbb{P}\left(\tilde{K}_0\left(|y_0|^2\right)\right) < \infty$. Also, without loss of generality, assume that $d = 1$ for this proof. By Theorem 3.3.5, $\bar{Y}_t$ is a continuous local martingale with decomposition given by $\bar{Y}_t = \bar{Y}_0 + M_t(Y)$ where

$$[M(Y)]_t = \int_0^t \frac{\overline{Y_s^2} - \bar{Y}_s^2}{K_s(1)} ds \equiv \int_0^t \frac{V(Y_s)}{K_s(1)} ds.$$

Theorem IV.34.12 of [18] shows that on the set $\{[M(Y)]_\infty < \infty\} \cap S$, $M_t(Y)$ a.s. converges.

Note that by Lemma 3.3.7, for a.s. $\omega \in S$, $W(\omega) > 0$, recalling that $W = \lim_{t\to\infty} e^{-\beta t} K_t(1)$. Hence it follows that $[M(Y)]_\infty < \infty$ on S if

$$\int_0^\infty e^{-\beta s} V(Y_s) ds < \infty.$$

Then

$$\begin{aligned}\mathbb{P}\left(\int_0^\infty \frac{V(Y_s)}{e^{\beta s}} ds; S\right) &\leq \mathbb{P}\left(\int_0^\infty \frac{\overline{Y_s^2}}{e^{\beta s}} ds; S\right) \\ &\leq \int_0^\infty \frac{A(\gamma, s)}{e^{\beta s}} ds \\ &< \infty\end{aligned}$$

by Cauchy-Schwartz and Lemma 3.3.9 since $\gamma \geq 0$. Therefore on S, $\bar{Y}_t$ converges a.s. to some limit $\bar{Y}_\infty$. Note that if $\gamma = 0$, Remark 3.3.4 (b) gives $\bar{Y}_t = \bar{Z}_t$ and so (b) holds.

That $\bar{Z}_t$ converges on S for $\gamma > 0$ follows from the fact that $\bar{Y}_t$ converges and

Equation 3.5 by setting

$$\begin{aligned}\bar{Z}_t &= \bar{Y}_t - \gamma \int_0^t e^{-\gamma(t-s)} \bar{Y}_s ds \\ &= \bar{Y}_t - \bar{Y}_\infty + \gamma \int_0^t e^{-\gamma(t-s)} \left(\bar{Y}_\infty - \bar{Y}_s\right) ds + e^{-\gamma t} \bar{Y}_\infty \\ &\to 0 \text{ as } t \to \infty.\end{aligned}$$

By Remark 3.3.4 (b) we see that for $\gamma > 0$ since $\bar{Y}_t = \bar{Z}_t + \gamma \int_0^t \bar{Z}_s ds$, $\bar{Z}_t \xrightarrow{\text{a.s.}} 0$ and $\bar{Y}_t \xrightarrow{\text{a.s.}} \gamma \int_0^\infty \bar{Z}_s ds$.

Now argue by conditioning as in the end of Theorem 3.4.1 to get the full result.

□

Theorem 3.4.3. *On the extinction set, S^c,*

$$\tilde{X}_t \to \delta_F \text{ and } \tilde{X}'_t \to \delta_{F'}$$

as $t \uparrow \eta < \infty$ a.s., where F and F' are $\mathbb{R}^d$-valued random variables such that

$$F' = F + \gamma \int_0^\eta \bar{Z}_s ds.$$

Proof. As in the previous theorems, note that we need only consider the case that $\mathbb{P}\left(\tilde{K}_0\left(|y_0|^2\right)\right) < \infty$.

We will follow the proof of Theorem 1 of Tribe [20] here. Define

$$\zeta(t) = \int_0^t \frac{1}{K_s(1)} ds, t < \eta.$$

It is known by the work of Konno and Shiga [10] in the case where $\beta = 0$, that $\zeta : [0, \eta) \to [0, \infty)$ homeomorphically (recall that $\eta < \infty$ a.s. in that case). This latter result also holds when $\beta > 0$ on the extinction set, S^c, by a Girsanov argument.

To see this, suppose that on some probability space,

$$z_t = z_0 + \int_0^t \sqrt{z_s} dB_s, \mathbb{Q}\text{-a.s.}$$

Then using Girsanov's Theorem, (Theorem V.27.1 of [18]) we see that there is a probability measure $\mathbb{Q}'$ such that, $\frac{d\mathbb{Q}}{d\mathbb{Q}'}|_{\mathcal{F}_t} = N_t$ where

$$N_t = \exp\left(\int_0^t \beta dz_s - \frac{\beta^2}{2}\int_0^t z_s ds\right)$$

is an exponential martingale. Furthermore, under $\mathbb{Q}'$, z_t solves

$$z_t = z_0 + \int_0^t \sqrt{z_s} dB_s + \int_0^t \beta z_s ds, \mathbb{Q}'\text{-a.s.}$$

Let $S_t^c = \{\omega : \exists s \leq t, z_s = 0\}$. Then

$$\int_{S_t^c} 1\left(\int_0^\eta \frac{1}{z_s} ds < \infty\right) d\mathbb{Q}' = \int_{S_t^c} 1\left(\int_0^\eta \frac{1}{z_s} ds < \infty\right) N_t d\mathbb{Q} = 0,$$

by the result of Konno and Shiga. Note then that on $\cup S_t^c$, $\int_0^\eta z_s^{-1} ds = \infty$, $\mathbb{Q}'$-a.s. Since $K_t(1)$ is a diffusion that satisfies the previous SDE with respect to $\mathbb{P}$, ζ satisfies the required properties on S^c, $\mathbb{P}$-a.s. as well as on S (where it is trivial).

Define $D : [0, \zeta(\eta-)) \to [0, \eta)$ as the unique inverse of ζ (on S^c, this defines the inverse on $[0, \infty)$) and for $t \geq \zeta(\eta-)$, let $D_t = \infty$. Let

$$X_t^D = X'_{D_t}, \tilde{X}_t^D = \frac{X_t^D}{X_t^D(1)} \text{ and } \mathcal{G}_t = \mathcal{F}_{D_t}$$

and define

$$L_t\phi(x) = \gamma(\bar{Y}_t - x) \cdot \nabla\phi(x) + \frac{1}{2}\Delta\phi(x).$$

Let

$$T_N = \int_0^{\eta_N} \frac{1}{K_s(1)} ds,$$

where $\eta_N = \inf\{s : K_s(1) \leq 1/N\}$. Then note that $T_N \uparrow \zeta(\eta-)$ and each T_N is a

$\mathcal{G}_t$-stopping time. On S^c for $\phi \in C_b^2$, Theorem 3.3.5 implies,

$$\begin{aligned}
\tilde{X}^D_{t\wedge T_N}(\phi) &= \tilde{X}_0(\phi) + \int_0^{D_{t\wedge T_N}} \tilde{X}'_s(L_s\phi)ds + M_{D_{t\wedge T_N}}(\phi) \\
&= \tilde{X}_0(\phi) + \int_0^{t\wedge T_N} \tilde{X}^D_s(L_{D_s}\phi)X^D_s(1)ds + N_{t\wedge T_N}(\phi) \\
&= \tilde{X}_0(\phi) + \int_0^{t\wedge T_N} X^D_s(L_{D_s}\phi)ds + N_{t\wedge T_N}(\phi) \qquad (3.11)
\end{aligned}$$

since $dD_t = X^D_t(1)dt$ and where $N_t = M_{D_t}$. It follows that $N_{t\wedge T_N}$ is a $\mathcal{G}_t$-local martingale. Then, by Theorem 3.3.5,

$$\begin{aligned}
\left[\tilde{X}^D(\phi)\right]_{t\wedge T_N} &= \int_0^{D_{t\wedge T_N}} \frac{\tilde{X}'_s(\phi^2) - \tilde{X}'_s(\phi)^2}{X_s(1)} ds \\
&= \int_0^{t\wedge T_N} \tilde{X}^D_s(\phi^2) - \tilde{X}^D_s(\phi)^2 ds,
\end{aligned}$$

which is uniformly bounded in N. Hence, sending $N \to \infty$, one sees that $N_{t\wedge\zeta(\eta-)}$ is a $\mathcal{G}_t$-martingale.

Note that on S^c, $\zeta(\eta-) = \infty$ and hence on that event,

$$\begin{aligned}
\int_0^\infty X^D_s(|L_{D_s}\phi|)ds &\le \int_0^\infty \int |\gamma(\bar{Y}_{D_s} - x)\cdot\nabla\phi(x)| + \frac{1}{2}|\Delta\phi(x)| X^D_s(dx)ds \\
&\le \int_0^\infty |\gamma|\,\|\nabla\phi\|\, K_{D_s}(|\bar{Y}_{D_s} - Y_{D_s}|) + \frac{1}{2}\|\Delta\phi\|\, X^D_s(1)ds \\
&= \int_0^\infty X^D_s(1)\left(|\gamma|\|\nabla\phi\|\tilde{K}_{D_s}(|\bar{Y}_{D_s} - Y_{D_s}|) + \frac{1}{2}\|\Delta\phi\|\right)ds \\
&\le \int_0^\infty X^D_s(1)\left(|\gamma|\|\nabla\phi\|\left(\overline{|Y_{D_s}|^2}\right)^{\frac{1}{2}} + \frac{1}{2}\|\Delta\phi\|\right)ds
\end{aligned}$$

where in the second line we have used the definition of X' and the Cauchy-Schwartz inequality in the fourth. Using the definition of D_s yields

$$\begin{aligned}\int_0^\infty X_s^D(|L_{D_s}\phi|)ds &\le \left(|\gamma|\|\nabla\phi\|\sup_{s<\eta}\left(\overline{|Y_s|^2}\right)^{\frac{1}{2}}+\frac{1}{2}\|\Delta\phi\|\right)\int_0^\infty X_s^D(1)ds \\ &= \left(|\gamma|\|\nabla\phi\|\sup_{s<\eta}\left(\overline{|Y_s|^2}\right)^{\frac{1}{2}}+\frac{1}{2}\|\Delta\phi\|\right)\eta \\ &< \infty \qquad (3.12)\end{aligned}$$

as $\phi\in C_b^2$ and $\overline{|Y_s|^2}$ is continuous on $[0,\eta)$ (which follows from Theorem 3.3.5). Hence, this implies that for ϕ positive, on S^c

$$N_t(\phi) > -\tilde{X}_0(\phi) - \int_0^\infty X_s^D(|L_{D_s}\phi|)ds$$

for all t and hence by Corollary IV.34.13 of [18], N_t converges as $t\to\infty$. Therefore by (3.11) and (3.12), $\tilde{X}_t^D(\phi)$ converges a.s. as well.

Denote by $\tilde{X}_\infty^D(\phi)$ the limit of $\tilde{X}_t^D(\phi)$. It is immediately evident that $\tilde{X}_\infty^D(\cdot)$ is a probability measure on $\mathbb{R}^d$. To show that $\tilde{X}_\infty^D(\cdot)=\delta_{F'}$ where F' is a random point in $\mathbb{R}^d$, we now defer to the proof of Theorem 1 in Tribe [20], as it is identical from this point forwards.

Similar (but simpler) reasoning holds to show $\tilde{X}_t\to\delta_F$ a.s. on S^c where F is a random point in $\mathbb{R}^d$. Let $f(t)=\gamma\int_0^t\bar{Z}_s ds$. Note that f is independent of y and that $f(t)\to f(\eta)$ a.s. when $t\uparrow\eta$ because $\bar{Y}_t=\bar{Z}_t+f(t)$ and both $\bar{Y}_t$ and $\bar{Z}_t$ converge a.s. by Theorem 3.4.1. Then for ϕ bounded and Lipschitz,

$$\begin{aligned}\left|\int\phi(x-f(t))\tilde{X}_t'(dx)-\int\phi(x-f(\eta))\tilde{X}_t'(dx)\right| &\le C|f(\eta)-f(t)| \\ &\xrightarrow{\text{a.s.}} 0.\end{aligned}$$

as $t\uparrow\eta$. Therefore it is enough to note that since $f(\eta)$ depends only on ω, the convergence of $\tilde{X}_t'$ gives

$$\int\phi(x-f(\eta))\tilde{X}_t'(dx)\xrightarrow{\text{a.s.}}\phi\left(F'-f(\eta)\right)$$

and hence

$$\int \phi(x-f(t))\tilde{X}_t'(dx) \xrightarrow{\text{a.s.}} \phi\left(F'-f(\eta)\right).$$

By Remark 3.3.4 (a),

$$\begin{aligned}
\int \phi(x-f(t))\tilde{X}_t'(dx) &= \int \phi\left(Y_t(y)-f(t)\right)\tilde{K}_t(dy)\\
&= \int \phi\left(Z_t(y)\right)\tilde{K}_t(dy)\\
&= \int \phi(x)\tilde{X}_t(dx)\\
&\xrightarrow{\text{a.s.}} \phi\left(F\right)
\end{aligned}$$

as $t \uparrow \eta$. Since there exists a countable separating set of bounded Lipschitz functions $\{\phi_n\}$, and the above holds for each ϕ_n,

$$F' = F + \gamma \int_0^\eta \bar{Z}_s ds,$$

a.s. □

Remark 3.4.4. (a) Theorem 3.4.3 holds in the critical branching case. That is, if $\beta = 0$,

$$X_t \xrightarrow{\text{a.s.}} \delta_F \text{ and } X_t' \xrightarrow{\text{a.s.}} \delta_{F'}$$

where

$$F' = F + \int_0^\eta \bar{Z}_s ds.$$

The convergence of the critical ordinary SOU process to a random point follows directly from Tribe's result. That this holds for the SOU process with attraction to the COM follows from the calculations above.

(b) The distribution of the random point F has been identified in Tribe [20] by approximating with branching particle systems. In fact, the law of F can be identified as x_η, where x_t is an Ornstein-Uhlenbeck process with initial distribution given by $\tilde{X}_0$ and η is the extinction time. Finding the distribution of F' remains an open problem however.

The next few results are necessary to establish the almost sure convergence of

$\tilde{X}_t$ on the survival set. This will in turn be used to show the almost sure convergence of $\tilde{X}_t'$ using the correspondence of Remark 3.3.4 (a).

Define $\mathrm{Lip}_1 = \left\{\psi \in C(\mathbb{R}^d) : \forall x, y, |\psi(x) - \psi(y)| \leq |x - y|, \|\psi\| \leq 1\right\}$. Define P_t as the standard Ornstein-Uhlenbeck semigroup (with attraction to the origin). Note that $P_t \to P_\infty$ in norm where

$$P_\infty \phi(x) = \int \phi(z) \left(\frac{\gamma}{\pi}\right)^{\frac{d}{2}} e^{-\gamma|z|^2} dz,$$

which is independent of x. Recall that $W = \lim_{t\to\infty} e^{-\beta t} X_t(1)$ and $S = \{W > 0\}$ a.s. from Lemma 3.3.7.

For the following pair of lemmas we will, with a slight abuse of notation, allow M to denote the orthogonal martingale measure generated by the martingale problem for X. Let A be the infinitesimal generator for an Ornstein-Uhlenbeck process, and hence recall that for $\phi \in C^2(\mathbb{R}^d)$,

$$A\phi(x) = -\gamma x \cdot \nabla \phi(x) + \frac{\Delta}{2} \phi(x).$$

Lemma 3.4.5. *If $\gamma > 0$, $\mathbb{P}\left(X_0\left(|x|^4\right)\right) < \infty$ and $\mathbb{P}\left(X_0(1)^4\right) < \infty$, then on S, for any $\phi \in Lip_1$, $e^{-\beta t} X_t(\phi) \xrightarrow{L^2} W P_\infty \phi$ and*

$$\mathbb{P}\left(\left|e^{-\beta t} X_t(\phi) - W P_\infty \phi\right|^2\right) \leq C e^{-\zeta t}$$

where C depends only on d and X_0, and ζ is a positive constant dependent only on β and γ.

Proof. Let $\phi \in \mathrm{Lip}_1$. By the extension of the martingale problem for X given in Proposition II.5.7 of [14], for functions $\psi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that ψ satisfies the definition before that Proposition,

$$X_t(\psi_t) = X_0(\psi_0) + \int_0^t \int \psi_s(x) dM(x,s) + \int_0^t X_s\left(A\psi_s + \beta\psi_s + \dot{\psi}_s\right) ds$$

where M is the orthogonal martingale measure derived from the martingale problem for the SOU process. It is not difficult to show that $\psi_s = P_{t-s}\phi$ where ϕ as above satisfies requirements for Proposition II.5.7 of [14]. Plugging this in gives

$$X_t(\phi) = X_0(P_t\phi) + \int_0^t \int P_{t-s}\phi(x)dM(s,x) + \int_0^t \beta X_s(P_{t-s}\phi)ds$$

since $\frac{\partial}{\partial s}P_s\phi = AP_s\phi$. Multiplying by $e^{-\beta t}$ and integrating by parts gives

$$\begin{aligned} e^{-\beta t}X_t(\phi) &= e^{-\beta t}X_t(\psi_t) \\ &= X_0(\psi_0) - \int_0^t \beta e^{-\beta s}X_s(\psi_s)ds + \int_0^t e^{-\beta s}dX_s(\psi_s) \\ &= X_0(P_t\phi) + \int_0^t \int e^{-\beta s}P_{t-s}\phi(x)dM(s,x) \qquad (3.13) \end{aligned}$$

Note that as the OU-process has a stationary distribution P_∞ where $P_t \to P_\infty$ in norm. When s is large in (3.13), $P_{t-s}\phi(x)$ does not contribute much to the stochastic integral and hence we expect the limit of $e^{-\beta t}X_t(\phi)$ to be

$$X_0(P_\infty\phi) + \int_0^\infty \int e^{-\beta s}P_\infty\phi(x)dM(s,x), \qquad (3.14)$$

which is a well defined, finite random variable as

$$\left[\int_0^\cdot \int e^{-\beta s}P_\infty\phi(x)dM(s,x)\right]_\infty < \|\phi\|^2 \int_0^\infty e^{-2\beta s}X_s(1)ds,$$

which is finite in expectation. As $P_\infty\phi(x)$ does not depend on x, it follows that

$$\begin{aligned} (3.14) &= (P_\infty\phi)X_0(1) + (P_\infty\phi)\int_0^\infty \int e^{-\beta s}dM(s,x) \\ &= WP_\infty\phi. \end{aligned}$$

Given this decomposition for $WP_\infty\phi$, we write

$$\begin{aligned} \mathbb{P}\left(\left(e^{-\beta t}X_t(\phi) - WP_\infty\phi\right)^2\right) &\le 3\mathbb{P}\left(\left(\int_t^\infty e^{-\beta s}P_\infty\phi(x)dM(s,x)\right)^2\right) \\ &+ 3\mathbb{P}\left(\left(\int_0^t \int e^{-\beta s}\left(P_{t-s}\phi(x) - P_\infty\phi(x)\right)dM(s,x)\right)^2 + X_0(P_\infty\phi - P_t\phi)^2\right) \end{aligned}$$

If z_t is a d-dimensional OU process satisfying $dz_t = -\gamma z_t dt + dB_t$, where B_t is a d-dimensional Brownian motion, then

$$z_t = e^{-\gamma t} z_0 + \int_0^t e^{-(t-s)\gamma} dB_t$$

and hence z_t is Gaussian, with mean $e^{-\gamma t} z_0$ and covariance matrix $\frac{1}{2\gamma}(1-e^{-2\gamma t})I$. Evidently, z_∞ is also Gaussian, mean 0 and variance $\frac{1}{2\gamma}I$. We use a simple coupling: suppose that w_t is a random variable independent of z_t such that $z_\infty = z_t + w_t$ (i.e. w_t is Gaussian with mean $-e^{-\gamma t} z_0$ and covariance $\frac{1}{2\gamma} e^{-2\gamma t} I$). Then using the fact that $\phi \in \mathrm{Lip}_1$ and the Cauchy-Schwartz inequality, followed by our coupling with $z_0 = x$ gives

$$\begin{aligned}
X_0(P_\infty \phi - P_t \phi)^2 &= \left(\int \mathbb{E}^x \left(\phi(z_\infty) - \phi(z_t) \right) X_0(dx) \right)^2 \\
&\leq \int \mathbb{E}^x \left(|z_\infty - z_t| \right)^2 X_0(dx) X_0(1) \\
&= \int \mathbb{E}^x \left(|w_t|^2 \right) X_0(dx) X_0(1) \\
&= \int e^{-2\gamma t} \left(|x|^2 + \frac{d}{2\gamma} \right) X_0(dx) X_0(1) \\
&\leq c e^{-2\gamma t} \left(\int |x|^2 X_0(dx) X_0(1) + X_0(1)^2 \right).
\end{aligned}$$

Taking expectations and using Cauchy-Schwartz and the assumptions on X_0 gives exponential rate of convergence for the above term.

Since we can think of $\int_0^r \int e^{-\beta s} P_{t-s}\phi(x) dM(s,x)$ as a martingale in r up until time t, various martingale inequalities can be applied to get bounds for the terminal element, $\int_0^t \int e^{-\beta s} P_{t-s}\phi(x) dM(s,x)$. Note that this process is not in general a martingale in t. Therefore, we have

$$\begin{aligned}
&\mathbb{P}\left[\left(\int_0^t \int e^{-\beta s} P_\infty \phi(x) dM(s,x) - \int_0^t \int e^{-\beta s} P_{t-s}\phi(x) dM(s,x) \right)^2 \right] \qquad (3.15) \\
&\quad = \mathbb{P}\left[\left(\int_0^t \int e^{-\beta s} (P_\infty \phi(x) - P_{t-s}\phi(x)) dM(s,x) \right)^2 \right] \\
&\quad \leq \mathbb{P}\left[\int_0^t e^{-2\beta s} \int (P_\infty \phi(x) - P_{t-s}\phi(x))^2 X_s(dx) ds \right].
\end{aligned}$$

Then as ϕ Lipschitz, by the coupling above,

$$\begin{aligned}(3.15) &\le \mathbb{P}\left[\int_0^t e^{-2\beta s}\int e^{-2\gamma(t-s)}\left(|x|^2+\frac{d}{2\gamma}\right)X_s(dx)ds\right]\\ &= \int_0^t e^{-2\beta s-2\gamma(t-s)}\mathbb{P}\left[\int |x|^2+\frac{d}{2\gamma}X_s(dx)\right]ds\\ &= \int_0^t e^{-2\beta s-2\gamma(t-s)}\mathbb{P}\left[K_s\left(|Z_s|^2\right)+\frac{d}{2\gamma}X_s(1)\right]ds.\end{aligned}$$

Applying the Cauchy-Schwartz inequality followed by Remark 3.3.10(b) gives

$$\begin{aligned}\mathbb{P}\left(K_s\left(|Z_s|^2\right)\right) &= \mathbb{P}\left(\overline{|Z_s|^2}K_s(1); s<\eta\right)\\ &\le \mathbb{P}\left(\overline{|Z_s|^2}^2; s<\eta\right)^{\frac{1}{2}}\mathbb{P}\left(X_s(1)^2\right)^{\frac{1}{2}}\\ &\le B(s,\gamma,4)^{\frac{1}{2}}\mathbb{P}\left(X_s(1)^2\right)^{\frac{1}{2}}\\ &\le cB(s,\gamma,4)^{\frac{1}{2}}e^{\beta s}\mathbb{P}\left(X_0(1)^2+\frac{1}{\beta}X_0(1)\right)^{\frac{1}{2}}\end{aligned}$$

where the last line follows by first noting that

$$e^{-\beta t}X_t(1) = X_0(1)+\int_0^t\int e^{-\beta s}dM(s,x)$$

is a martingale. That is,

$$\begin{aligned}e^{-2\beta s}\mathbb{P}\left(X_s(1)^2\right) &\le 2\mathbb{P}\left(X_0(1)^2+\left(\int_0^s e^{-\beta r}dM(r,x)\right)^2\right) \qquad (3.16)\\ &= 2\mathbb{P}\left(X_0(1)^2+\left[\int_0^\cdot e^{-\beta r}dM(r,x)\right]_s\right)\\ &= 2\mathbb{P}\left(X_0(1)^2+\int_0^s e^{-2\beta r}X_r(1)dr\right)\\ &= 2\mathbb{P}\left(X_0(1)^2\right)+2\int_0^s e^{-\beta r}\mathbb{P}\left(e^{-\beta r}X_r(1)\right)dr\\ &= 2\mathbb{P}\left(X_0(1)^2\right)+2\int_0^s e^{-\beta r}\mathbb{P}\left(X_0(1)\right)dr\\ &\le 2\mathbb{P}\left(X_0(1)^2+\frac{1}{\beta}X_0(1)\right).\end{aligned}$$

Therefore,

$$
\begin{aligned}
(3.15) &\leq \int_0^t e^{-2\beta s-2\gamma(t-s)}\left[e^{\beta s}B(s,\gamma,4)^{\frac{1}{2}}\mathbb{P}\left(X_0(1)^2+\frac{1}{\beta}X_0(1)\right)^{\frac{1}{2}}\right]ds \\
&\quad + \int_0^t e^{-2\beta s-2\gamma(t-s)}e^{\beta s}\frac{d}{2\gamma}\mathbb{P}(X_0(1))ds \\
&\leq \int_0^t e^{-\beta s-2\gamma(t-s)}\left[B(s,\gamma,4)^{\frac{1}{2}}\mathbb{P}\left(X_0(1)^2+\frac{1}{\beta}X_0(1)\right)^{\frac{1}{2}}+\frac{d}{2\gamma}\mathbb{P}(X_0(1))\right]ds \\
&< Ce^{-\zeta_1 t},
\end{aligned}
$$

where $\zeta_1 = \min(\beta, 2\gamma) - \varepsilon$ where ε is arbitrary small and comes from the polynomial term in the integral.

Finally,

$$
\begin{aligned}
\mathbb{P}\left(\left(\int_t^\infty e^{-\beta s}P_\infty\phi(x)dM(s,x)\right)^2\right) &= (P_\infty\phi)^2\,\mathbb{P}\left(\int_t^\infty e^{-2\beta s}X_s(1)ds\right) \\
&\leq (P_\infty\phi)^2\int_t^\infty e^{-\beta s}\mathbb{P}\left(e^{-\beta s}X_s(1)\right)ds \\
&\leq \frac{(P_\infty\phi)^2}{\beta}e^{-\beta t}\mathbb{P}\left(X_0(1)\right),
\end{aligned}
$$

since $e^{-\beta s}X_s(1)$ is a martingale. Therefore, since $\zeta_1 < \beta$, we see that $\zeta = \zeta_1$ gives the correct exponent. $\square$

Remark 3.4.6. As the L^2 convergence in Lemma 3.4.5 is exponentially fast, it follows from the Borel-Cantelli Lemma and Chebyshev inequality that for a strictly increasing sequence $\{t_n\}_{n=0}^\infty$ where $\left|\{t_n\}\cap[k,k+1)\right| = \lfloor e^{\zeta k/2}\rfloor$, for $\phi \in \mathrm{Lip}_1$,

$$e^{-\beta t_n}X_{t_n}(\phi) \to WP_\infty\phi \text{ a.s. as } n\to\infty.$$

The idea is to use the above remark to bootstrap up to almost sure convergence in Lemma 3.4.5 with some estimates on the modulus of continuity of the process $e^{-\beta t}X_t(\phi)$.

Lemma 3.4.7. *Suppose* $\gamma > 0$, $\mathbb{P}\left(X_0\left(|x|^8\right)\right) < \infty$ *and* $\mathbb{P}\left(X_0(1)^8\right) < \infty$. *If* $\phi \in Lip_1$

and $h > 0$, then

$$\mathbb{P}\left(\left[e^{-\beta(t+h)}X_{t+h}(\phi) - e^{-\beta t}X_t(\phi)\right]^4\right) \leq C(t)h^2e^{-\zeta^* t}$$

where ζ^* is a positive constant depending only on β and γ and C is polynomial in t, and depends on γ, β and d.

Proof. The proof will follow in a manner very similar to the proof of the previous lemma. From the calculations above we see that

$$\begin{aligned}
e^{-\beta(t+h)}X_{t+h}(\phi) - e^{-\beta t}X_t(\phi) &= X_0(P_{t+h}\phi - P_t\phi) \\
&+ \int_0^{t+h}\int e^{-\beta s}P_{t+h-s}\phi(x)dM(s,x) - \int_0^t\int e^{-\beta s}P_{t-s}\phi(x)dM(s,x) \\
&= X_0(P_{t+h}\phi - P_t\phi) + \int_0^t\int e^{-\beta s}\left((P_{t+h-s} - P_{t-s})\phi(x)\right)dM(s,x) \\
&+ \int_t^{t+h}\int e^{-\beta s}P_{t+h-s}\phi(x)dM(s,x) \\
&\equiv I_1 + I_2 + I_3
\end{aligned}$$

where P_t is the OU semigroup. Using the Cauchy-Schwartz inequality, we can find bounds for $\mathbb{P}(|I_k|^4), k = 1,2,3$ separately.

$$\begin{aligned}
|I_1|^4 &= X_0(P_{t+h}\phi - P_t\phi)^4 \\
&\leq \left[X_0\left((P_{t+h}\phi - P_t\phi)^2\right)X_0(1)\right]^2
\end{aligned}$$

Recalling the simple coupling in the previous lemma to see that

$$\begin{aligned}
(P_{t+h}\phi(x) - P_t\phi(x))^2 &\leq \mathbb{E}^x\left(|\phi(z_{t+h}) - \phi(z_t)|^2\right) \\
&\leq \mathbb{E}^x\left(|z_{t+h} - z_t|^2\right) \\
&\leq \mathbb{E}^x\left(|w_{t,t+h}|^2\right)
\end{aligned}$$

where z is as above, an OU process started at x, and $w_{s,t}$ is independent of z_s but such that $z_t = z_s + w_{s,t}$ Hence $w_{s,t}$ is Gaussian with mean $x(e^{-\gamma t} - e^{-\gamma s})$ and covariance matrix $\frac{I}{2\gamma}(e^{-2s\gamma}-e^{-2t\gamma})$. Therefore,

$$\begin{aligned}(P_{t+h}\phi(x)-P_t\phi(x))^2 &\le |x|^2\left(e^{-\gamma(t+h)}-e^{-\gamma t}\right)^2+\frac{d}{2\gamma}\left(e^{-2t\gamma}-e^{-2(t+h)\gamma}\right)\\ &= e^{-2\gamma t}\left(|x|^2\left(1-e^{-\gamma h}\right)^2+\frac{d}{2\gamma}\left(1-e^{-2h\gamma}\right)\right).\end{aligned}$$

Hence

$$\begin{aligned}&\mathbb{P}(|I_1|^4)\\ &=\mathbb{P}\left[\left(e^{-2\gamma t}\left(1-e^{-\gamma h}\right)^2\int|x|^2X_0(dx)X_0(1)+\frac{d}{2\gamma}e^{-2\gamma t}\left(1-e^{2\gamma h}\right)X_0(1)^2\right)^2\right]\\ &\le C_1(d,\gamma)h^2e^{-4\gamma t}\end{aligned}$$

where C_1 is a constant dependent on $X_0(\cdot)$, and is finite by assumptions on the initial measure.

To get bounds on the expectation of I_2 we use martingale inequalities. Note that $\int_0^{\cdot}\int e^{-\beta s}\left((P_{t+h-s}-P_{t-s})\phi(x)\right)dM(s,x)=N(\cdot)$ is a martingale until time t. Therefore, using the Burkholder-Davis-Gundy inequality and the coupling above gives

$$\begin{aligned}\mathbb{P}\left(|I_2|^4\right) &\le c\mathbb{P}([N]_t^2)\\ &= c\mathbb{P}\left[\left(\int_0^t\int e^{-2\beta s}\left((P_{t+h-s}-P_{t-s})\phi(x)\right)^2X_s(dx)ds\right)^2\right]\\ &\le c\mathbb{P}\left[\left(\int_0^t e^{-2\beta s}\left(e^{-\gamma(t+h-s)}-e^{-\gamma(t-s)}\right)^2\int|x|^2X_s(dx)ds\right)^2\right.\\ &\quad\left.+\left(\int_0^t\frac{de^{-2\beta s}}{2\gamma}\left(e^{-2(t-s)\gamma}-e^{-2(t+h-s)\gamma}\right)X_s(1)ds\right)^2\right]\\ &\le c\mathbb{P}\left[\left(t\int_0^t e^{-4\beta s}\left(e^{-\gamma(t+h-s)}-e^{-\gamma(t-s)}\right)^4\left(\int|x|^2X_s(dx)\right)^2ds\right)\right.\\ &\quad\left.+t\left(\frac{d}{2\gamma}\right)^2\int_0^t e^{-4\beta s}\left(e^{-2(t-s)\gamma}-e^{-2(t+h-s)\gamma}\right)^2X_s(1)^2ds\right]\end{aligned}$$

$$= ct\int_0^t e^{-2\beta s}\left(e^{-\gamma(t+h-s)} - e^{-\gamma(t-s)}\right)^4 \mathbb{P}\left[\left(\int |x|^2 e^{-\beta s} X_s(dx)\right)^2\right] ds$$
$$+ t\left(\frac{d}{2\gamma}\right)^2 \int_0^t e^{-2\beta s}\left(e^{-2(t-s)\gamma} - e^{-2(t+h-s)\gamma}\right)^2 \mathbb{P}\left[\left(e^{-\beta s} X_s(1)\right)^2\right] ds.$$

Since $X_s(|x|^2) = K_s(|Z_s|^2)$, by Remark 3.3.10 (b),

$$\begin{aligned}
\mathbb{P}\left(e^{-2\beta s} X_s(|x|^2)^2\right) &\le \mathbb{P}\left[\overline{Z_s^2}^2 \left(e^{-\beta s} X_s(1)\right)^2; s < \eta\right] \\
&\le \mathbb{P}\left(\overline{Z_s^2}^4; s < \eta\right)^{1/2} \mathbb{P}\left(e^{-4\beta s} X_s(1)^4\right)^{1/2} \\
&\le \mathbb{P}\left(\overline{Z_s^8}; s < \eta\right)^{1/2} \mathbb{P}\left(e^{-4\beta s} X_s(1)^4\right)^{1/2} \\
&\le cB(\gamma, s, 8)^{1/2} \left(\mathbb{P}\left(X_0(1)^4 + sX_0(1)^2 + sX_0(1)\right)\right)^{1/2}.
\end{aligned}$$

The bound on the expectation of $e^{-4\beta s} X_s(1)^4$ follows by an application of the Burkholder-Davis-Gundy inequality to $e^{-\beta s} X_s(1) = X_0(1) + \int_0^s e^{-\beta r} dM(r,x)$:

$$\begin{aligned}
\mathbb{P}\left(\left(e^{-\beta s} X_s(1)\right)^4\right) &\le c\mathbb{P}\left(X_0(1)^4 + \left(\int_0^s e^{-\beta r} dM(r,x)\right)^4\right) \\
&\le c\mathbb{P}\left(X_0(1)^4 + \left[\int_0^\cdot e^{-\beta r} dM(r,x)\right]_s^2\right) \\
&\le c\mathbb{P}\left(X_0(1)^4 + \left(\int_0^s e^{-2\beta r} X_r(1) dr\right)^2\right),
\end{aligned}$$

where the last line follows as $[M(\phi)]_t = \int_0^t X_s\left(\phi^2\right) ds$. Then, using the Cauchy-

Schwartz Inequality followed by Fubini's Theorem,

$$\begin{aligned}
\mathbb{P}\left(\left(e^{-\beta s}X_s(1)\right)^4\right) &\le c\mathbb{P}\left(X_0(1)^4 + s\int_0^s e^{-4\beta r}X_r(1)^2 dr\right)\\
&\le c\mathbb{P}\left(X_0(1)^4\right) + cs\int_0^s e^{-2\beta r}\mathbb{P}\left(e^{-2\beta r}X_r(1)^2\right)dr\\
&= c\mathbb{P}\left(X_0(1)^4\right) + cs\int_0^s e^{-2\beta r}\mathbb{P}\left(X_0(1)^2 + \frac{1}{\beta}X_0(1)\right)dr\\
&\le c\mathbb{P}\left(X_0(1)^4 + \frac{s}{2\beta}X_0(1)^2 + \frac{s}{2\beta^2}X_0(1)\right).
\end{aligned}$$

where in the third line we have substituted the calculation from (3.16). Therefore,

$$\begin{aligned}
\mathbb{P}\left(|I_2|^4\right) \le\ & ct\int_0^t e^{-2\beta s}B(\gamma,s,8)^{1/2}\left(e^{-\gamma(t+h-s)} - e^{-\gamma(t-s)}\right)^4\\
&\qquad\cdot\left(\mathbb{P}\left(X_0(1)^4 + sX_0(1)^2 + sX_0(1)\right)\right)^{1/2} ds\\
&+ \frac{ct}{4\gamma^2}\int_0^t e^{-2\beta s}\left(e^{-2(t-s)\gamma} - e^{-2(t+h-s)\gamma}\right)^2\mathbb{P}\left(X_0(1)^2 + \frac{1}{\beta}X_0(1)\right)ds\\
&\le ctB(\gamma,t,8)^{1/2}\left(\mathbb{P}\left(X_0(1)^4 + tX_0(1)^2 + tX_0(1)\right)\right)^{1/2}\\
&\qquad\cdot\left(e^{-\gamma h} - 1\right)^4\int_0^t e^{-2\beta s}e^{-4\gamma(t-s)}ds\\
&+ \frac{ct}{4\gamma^2}\left(e^{-2\gamma h} - 1\right)^2\mathbb{P}\left(X_0(1)^2 + \frac{1}{\beta}X_0(1)\right)\int_0^t e^{-2\beta s}e^{-4(t-s)\gamma}ds\\
&\le C_2(t,\gamma,\beta)h^2 e^{-\zeta_1 t}
\end{aligned}$$

where C_2 is polynomial in t. By another application of the BDG Inequality, and noting that $\|\phi\| = 1$,

$$
\begin{aligned}
\mathbb{P}\left(|I_3|^4\right) &= \mathbb{P}\left[\left(\int_t^{t+h}\int e^{-\beta s}P_{t+h-s}\phi(x)dM(s,x)\right)^4\right] \\
&\leq c\mathbb{P}\left[\left(\int_t^{t+h}\int\left(e^{-\beta s}P_{t+h-s}\phi(x)\right)^2 X_s(dx)ds\right)^2\right] \\
&\leq c\mathbb{P}\left[\left(\int_t^{t+h}\int e^{-2\beta s}X_s(dx)ds\right)^2\right] \\
&\leq ch\mathbb{P}\left[\int_t^{t+h} e^{-4\beta s}X_s(1)^2 ds\right] \\
&= che^{-2\beta t}\int_t^{t+h}\mathbb{P}\left[e^{-2\beta s}X_s(1)^2\right]ds \\
&\leq che^{-2\beta t}\int_t^{t+h}\mathbb{P}\left[X_0(1)^2 + X_0(1)/\beta\right]ds \\
&\leq C_3(\beta)h^2 e^{-2\beta t}
\end{aligned}
$$

where the second last line follows from the same calculations performed in estimating moments of I_2. Note that the constant C_3 does not depend on t here.

Putting the pieces together shows that there exists a function C polynomial in t and a positive constant ζ^* such that

$$
\mathbb{P}\left[\left(e^{-\beta(t+h)}X_{t+h}(\phi) - e^{-\beta t}X_t(\phi)\right)^4\right] \leq C(t,\gamma,\beta,d)h^2 e^{-\zeta^* t}
$$

□

The following is a very useful result of Garsia, Rodemich, and Rumsey [7].

Let $\Psi : \mathbb{R}^d \to \mathbb{R}$, $p : \mathbb{R}_+ \to \mathbb{R}$ be positive, continuous functions. Further, suppose Ψ is symmetric about 0 and convex with $\lim_{|x|\to\infty}\Psi(x) = \infty$ and $p(x)$ is increasing with $p(0) = 0$.

Proposition 3.4.8. *If f is a measurable function on $[0,1]$ such that*

$$
\int\int_{[0,1]^2}\Psi\left(\frac{f(t)-f(s)}{p(|t-s|)}\right)dsdt = B < \infty \tag{3.17}
$$

then there is a set K of measure 0 such that if $s,t \in [0,1] \setminus K$ *then*

$$|f(t) - f(s)| \leq 8 \int_0^{|t-s|} \Psi^{-1}\left(\frac{B}{u^2}\right) dp(u). \tag{3.18}$$

If f is also continuous, then K is in fact the empty set.

With this result in hand, we can now bring everything together to prove convergence of $\tilde{X}_t$. Let d denote the Vasserstein metric on the space of finite measures on $\mathbb{R}^d$. I.e., for μ, ν finite measures,

$$d(\mu, \nu) = \sup_{\phi \in \mathrm{Lip}_1} \int \phi(x) d(\mu - \nu)(x),$$

recalling that $\mathrm{Lip}_1 = \left\{\psi \in C(\mathbb{R}^d) : \forall x, y, |\psi(x) - \psi(y)| \leq |x-y|, \|\psi\| \leq 1\right\}$.

Theorem 3.4.9. *Suppose* $X_0(1) < \infty$ *a.s. and* $\gamma > 0$*. Then on S,*

$$d(\tilde{X}_t, P_\infty) \xrightarrow{a.s.} 0,$$

where P_t *is the semigroup of an Ornstein-Uhlenbeck process with attraction to* 0 *at rate* γ*.*

Proof. The strategy for this proof is simple: We use Remark 3.4.6 to see that we can lay down an increasingly (exponentially) denser sequence $e^{-\beta t_n} X_{t_n}(\phi)$ which converges almost surely, and that we can use Lemma 3.4.7 to get a modulus of continuity on the process $e^{-\beta t_n} X_{t_n}(\phi)$, which then implies that if the sequence is converging, then the entire process must be converging.

Assume that $\mathbb{P}\left(K_0\left(|y_0|^8\right)\right) = \mathbb{P}\left(X_0\left(|x|^8\right)\right) < \infty$ and $\mathbb{P}\left(X_0(1)^8\right) < \infty$ and argue as in Theorem 3.4.1 in the general case. Let $\phi \in \mathrm{Lip}_1$. Denote $e^{-\beta t} X_t$ by $\mathring{X}_t$ for the remainder of the proof. Let $T > 0$, and let $\Psi(x) = |x|^4$ and $p(t) = |t|^{3/4}\left(\log\left(\frac{\lambda}{t}\right)\right)^{1/2}$ where $\lambda = e^4$. Let $B_T(\omega)$ be the constant B that appears in Proposition 3.4.8, with aforementioned functions Ψ and p, for the path $\mathring{X}_{Tt}(\omega), t \in [0,1]$.

Then note that

$$
\begin{aligned}
\mathbb{P}(B_T) &\equiv \mathbb{P}\left[\int\int_{[0,1]^2} \Psi\left(\frac{\mathring{X}_{Tt}-\mathring{X}_{Ts}}{p(|t-s|)}\right) dsdt\right] \\
&= \int\int_{[0,1]^2} \frac{\mathbb{P}\left[\left|\mathring{X}_{Tt}-\mathring{X}_{Ts}\right|^4\right]}{|t-s|^3 \log^2\left(\frac{\lambda}{|t-s|}\right)} dsdt \\
&\leq \int\int_{[0,1]^2} \frac{C(T(s\wedge t))e^{-\zeta^*(s\wedge t)}T^2|t-s|^2}{|t-s|^3 \log^2\left(\frac{\lambda}{|t-s|}\right)} dsdt \\
&= 2T^2 \int_0^1 \int_0^t \frac{C(Ts)e^{-\zeta^*(Ts)}|t-s|^2}{|t-s|^3 \log^2\left(\frac{\lambda}{|t-s|}\right)} dsdt \\
&\leq 2C(T)T^2 \int_0^1 \int_0^t \frac{1}{|t-s| \log^2\left(\frac{\lambda}{|t-s|}\right)} dsdt \\
&\leq \frac{C(T)T^2}{2e^4}
\end{aligned}
$$

where C is the polynomial term that appears in Lemma 3.4.7. Since $\mathring{X}_t$ is continuous, by Garsia-Rodemech-Rumsey, for all $s,t \leq 1$,

$$
\begin{aligned}
\left|\mathring{X}_{Tt}-\mathring{X}_{Ts}\right| &\leq 8\int_0^{|t-s|} \left(\frac{B_T}{u^2}\right)^{\frac{1}{4}} dp(u) \\
&\leq AB_T^{\frac{1}{4}}|t-s|^{\frac{1}{4}} \left(\log \frac{\lambda}{|t-s|}\right)^{\frac{1}{2}}
\end{aligned}
$$

where A is a constant independent of T (see Corollary 1.2 of Walsh [22] for this calculation). Rewriting the above,

$$
\left|\mathring{X}_t-\mathring{X}_s\right| \leq D_T|t-s|^{\frac{1}{4}} \left(\log \frac{\lambda T}{|t-s|}\right)^{\frac{1}{2}} \forall s<t\leq T, \tag{3.19}
$$

where $D_T \equiv A\left(\frac{B_T}{T}\right)^{\frac{1}{4}}$. Note that $\mathbb{P}(D_T^4) = \frac{A^4TC(T)}{2e^4}$, which is polynomial in T of fixed degree $d_0 > 1$. Let Ω_0 be the set of probability 1 such that for all positive integers T Equation (3.19) holds and $D_T \leq T^{d_0}$ for T large enough. To see $P(\Omega_0) =$

1, use Borel Cantelli:

$$\mathbb{P}\left(D_T \geq T^{d_0}\right) = \mathbb{P}\left(D_T^4 \geq T^{4d_0}\right)$$
$$\leq \frac{\mathbb{P}(D_T^4)}{T^{4d_0}}$$
$$\leq \frac{c}{T^{3d_0}}$$

which is summable over all positive integers T.

Suppose $\omega \in \Omega_0$. Let $\delta^T(\omega)$ be such that $\delta^{\frac{-1}{8}}\left(\log\frac{\lambda}{\delta}\right)^{\frac{-1}{2}} = T^{d_0}$. Then for all integral $T > T_0(\omega)$, and $s,t \leq T$ with $|t-s| \leq \delta$, $\left|\mathring{X}_t - \mathring{X}_s\right| \leq |t-s|^{\frac{1}{8}}$.

Now let $\left\{\mathring{X}_{t_n}\right\}$ be a sequence of the form in Remark 3.4.6, with the additional condition that $\{t_n\} \cap [k,k+1)$ are evenly spaced within $[k,k+1)$ for each $k \in \mathbb{Z}_+$ (i.e. $t_{n+1} - t_n = ce^{-\zeta k/2}$ for $t_n \in \{t_n\} \cap [k,k+1)$). Evidently $\mathring{X}_{t_n}$ converges a.s. to a limit $\mathring{X}_\infty$. Without loss of generality, assume convergence of the sequence on the set Ω_0.

There exists $T_1(\omega)$ such that for all $T > T_1$, $ce^{-\zeta T/2} < \delta$. Hence, for all t such that $T_1 \vee T_0 < t \leq T$ there exists $t_n' \in \{t_n\}$ such that $|t - t_n'| < ce^{-\zeta \lfloor t \rfloor/2} < \delta$ and hence

$$\left|\mathring{X}_t - \mathring{X}_\infty\right| \leq \left|\mathring{X}_t - \mathring{X}_{t_n'}\right| + \left|\mathring{X}_{t_n'} - \mathring{X}_\infty\right|$$
$$\leq \left|t - t_n'\right|^{\frac{1}{8}} + \left|\mathring{X}_{t_n'} - \mathring{X}_\infty\right|$$
$$\leq ce^{-\frac{\zeta\lfloor t \rfloor}{16}} + \left|\mathring{X}_{t_n'} - \mathring{X}_\infty\right|.$$

Sending $t \to \infty$ gives almost sure convergence of $e^{-\beta t}X_t(\phi)$ to $\mathring{X}_\infty = WP_\infty(\phi)$ by Theorem 3.4.5, since $t_n' \to \infty$ with t. Note that this implies for $\phi \in \mathrm{Lip}_1$

$$\tilde{X}_t(\phi) \xrightarrow{\text{a.s.}} P_\infty(\phi) \tag{3.20}$$

since on S, $\frac{e^{\beta t}}{K_t(1)} \to W^{-1}$ a.s.

By Exercise 2.2 of [6] on $\mathcal{M}_1(\mathbb{R}^d)$, the space of probability measures on $\mathbb{R}^d$, the Prohorov metric of weak convergence is equivalent to the Vasserstein metric.

Note that the class of functions

$$\Theta = \left\{ \psi : \psi(x) = \sum_{i=1}^{n} p_i(x) e^{-q_i (x-x_i)^2 + b_i}, \right.$$
$$\left. p_i \text{ polynomial with rational coefficients}, q_i \in \mathbb{Q}_+, b_i \in \mathbb{Q}, x_i \in \mathbb{Q}^d, n < \infty \right\}$$

is a countable algebra of Lipschitz functions that is strongly separating (see p. 113 of [6]). Hence by Theorem 3.4.5(b) of [6], Θ is convergence determining. Since there exists a set $S_0 \subset S$ with $\mathbb{P}(S \backslash S_0) = 0$ such that on S_0, Equation (3.20) holds simultaneously for all $\phi \in \Theta$,

$$\tilde{X}_t(\cdot) \to P_\infty(\cdot)$$

in the Vasserstein metric, for $\omega \in S_0$ because Θ is convergence determining.

To drop the dependence on the eighth moment, we argue as in Theorem 3.4.1, where we make use of the Markov Property and the Compact support property for Historical Brownian Motion.

□

Remark 3.4.10. It is possible to show that Theorem 3.4.9 holds for a more general class of superprocesses. That is, if the underlying process has an exponential rate of convergence to a stationary distribution, then the above theorem goes through. One can appeal to, for example, Theorem 4.2 of Tweedie and Roberts [17] for a class of such continuous time processes.

Recall that $X'_t(\cdot) = K_t(Y_t \in \cdot)$ is the SOU process with attraction to its centre of mass.

Theorem 3.4.11. *Suppose $X'_0(1) < \infty$ a.s. and $\gamma > 0$. Then on S,*

$$d(\tilde{X}'_t, P_\infty^{\bar{Y}_\infty}) \xrightarrow{a.s.} 0 \tag{3.21}$$

where $P_\infty^{\bar{Y}_\infty}$ represents the OU-semigroup at infinity, with the origin shifted to $\bar{Y}_\infty$.

Proof. This follows almost immediately from Theorem 3.4.9 and the representation given in Remark 3.3.4 (a). Let $\phi \in \text{Lip}_1$, then

$$\begin{aligned}\tilde{X}_t'(\phi) &= \tilde{K}_t\left(\phi\left(Y_t\right)\right) \\ &= \tilde{K}_t\left(\phi\left(Z_t + \gamma\int_0^t \bar{Z}_s ds\right)\right) \\ &= \int \phi\left(x + f(t,\omega)\right) d\tilde{X}_t(dx)\end{aligned}$$

where $f(t) = \gamma\int_0^t \bar{Z}_s ds$. Remark 3.3.4 (b) gives $f(t) = \bar{Y}_t - \bar{Z}_t$, and hence $f(t) \xrightarrow{\text{a.s.}} \bar{Y}_\infty$ follows from Theorem 3.4.2. Note that

$$\begin{aligned}\left|\tilde{X}_t'(\phi) - P_\infty^{\bar{Y}_\infty}(\phi)\right| &\le \left|\int \phi\left(x + f(t)\right) d\tilde{X}_t(dx) - \int \phi\left(x + f(\infty)\right) d\tilde{X}_t(dx)\right| \\ &+ \left|\int \phi\left(x + f(\infty)\right) d\tilde{X}_t(dx) - \int \phi\left(x + f(\infty)\right) d\tilde{X}_\infty(dx)\right| \\ &\le |f(t) - f(\infty)| + d(\tilde{X}_t, \tilde{X}_\infty)\end{aligned}$$

since $\phi \in \text{Lip}_1$. Taking the supremum over ϕ and the previous theorem give

$$d(\tilde{X}_t', \tilde{X}_\infty') \le |f(t,\omega) - f(\infty,\omega)| + d(\tilde{X}_t, \tilde{X}_\infty) \xrightarrow{\text{a.s.}} 0.$$

□

3.5 The repelling case

Much less can be said for the SOU process repelling from its center of mass than in the attractive case. However we can show convergence of the center of mass, provided the rate of repulsion is not too strong. Recall that in the attractive case, this was the first step towards showing the a.s. convergence of the normalized interactive SOU process. We finish with some conjectures on the limiting measure for the repelling case.

As in the previous section, we will assume that $Z_0 = Y_0 = y_0$, unless stated otherwise.

Theorem 3.5.1. *Assuming $K_0(1) < \infty$ a.s., the following hold on S:*

(a) For $0 > \gamma > -\frac{\beta}{2}$, $\bar{Y}_t$ converges almost surely.

(b) *For* $0 > \gamma > -\frac{\beta}{2}$, $\bar{Z}_t$ *diverges exponentially fast. That is,*

$$\mathbb{P}_{X_0}\left(\lim_{t\to\infty} e^{\gamma t}\bar{Z}_t \to L \neq 0 \middle| S\right) = 1.$$

Proof. Assume that $\mathbb{P}\left(\tilde{K}\left(|y_0|^2\right)\right) < \infty$, as in the proof of Theorem 3.4.1. As in that theorem, this condition can be weakened here to just the finite initial mass condition using similar reasoning.

For part (a), note that as in Theorem 3.4.2, $\bar{Y}_t$ will converge if $\mathbb{P}([\bar{Y}]_t) < \infty$ and which holds if the following quantity is bounded:

$$\begin{aligned}\mathbb{P}\left(\int_0^\infty \frac{\overline{|Y_s|^2} - \bar{Y}_s^2}{e^{\beta s}} ds; S\right) &\leq \mathbb{P}\left(\int_0^\infty \frac{\overline{Y_s^2}}{e^{\beta s}} ds; S\right) \\ &\leq c\int_0^\infty \frac{1 + s^6 e^{-2\gamma s}}{e^{\beta s}} ds \\ &< \infty,\end{aligned}$$

by Lemma 3.3.9 and by the conditions on γ.

For (b), we require the following Lemma:

Lemma 3.5.2. *Let* $-\beta/2 < \gamma < 0$ *and* $X_0 \neq 0$. *For a measure* m *on* $\mathbb{R}^d$, *let* $\tau_a(m)$ *be* m *translated by* $a \in \mathbb{R}^d$. *That is,* $\tau_a(m)(\phi) = \int \phi(x+a)m(dx)$. *Then*

(i) *For all but at most countably many* a,

$$P_{\tau_a(X_0)}\left(e^{\gamma t}\bar{Z}_t \to L \neq 0 | S\right) = 1. \tag{3.22}$$

(ii) *For all but at most one value of* a

$$P_{\tau_a(X_0)}\left(e^{\gamma t}\bar{Z}_t \to L \neq 0 | S\right) > 0.$$

Proof of Lemma. We first note that, by the correspondence (3.5), we have that

$$\bar{Z}_t = \bar{Y}_t - \gamma e^{-\gamma t}\int_0^t e^{\gamma s}\bar{Y}_s ds.$$

Under our hypotheses $\bar{Y}_t$ converges by Theorem 3.5.1 (a), and hence

$$\lim_{t\to\infty} e^{\gamma t}\bar{Z}_t + \int_0^t \gamma e^{\gamma s}\bar{Y}_s ds = 0 \tag{3.23}$$

a.s. on S. Therefore, on S,

$$\lim_{t\to\infty} e^{\gamma t}\bar{Z}_t \text{ exists a.s.} \tag{3.24}$$

Note that one can build a solution of $(\mathrm{SE})^2$ with initial conditions given by $\tau_a(X_0)$ by seeing that if Y_t gives the solution of $(\mathrm{SE})^2_{Y_0,K}$, then $Y_t + a$ gives the solution of $(\mathrm{SE})^2_{Y_0+a,K}$, and that the projection

$$X_t'(\cdot) = \int \mathbf{1}(Y_t + a \in \cdot)K_t(dy)$$

gives the appropriate interacting SOU process.

By (3.23) and (3.24),

$$\begin{aligned}
\mathbb{P}_{X_0}\left(\left\{e^{\gamma t}\bar{Z}_t \to L \neq 0\right\}^c \Big| S\right) &= \mathbb{P}_{\tau_a(X_0)}\left(\lim_{t\to\infty}\int_0^t e^{\gamma s}\bar{Y}_s ds = 0 | S\right) \\
&= \mathbb{P}_{X_0}\left(\lim_{t\to\infty}\int_0^t e^{\gamma s}\frac{K_s\left(\tau_a(Y_s)\right)}{K_s(1)}ds = 0\Big| S\right) \\
&= \mathbb{P}_{X_0}\left(\lim_{t\to\infty}\int_0^t e^{\gamma s}\left(a + \frac{K_s(Y_s)}{K_s(1)}\right)ds = 0\Big| S\right) \\
&= \mathbb{P}_{X_0}\left(-\frac{a}{\gamma} + \lim_{t\to\infty}\int_0^t e^{\gamma s}\bar{Y}_s ds = 0\Big| S\right).
\end{aligned}$$

The random variable $\int_0^\infty e^{\gamma s}\bar{Y}_s ds$ is finite a.s. and so only a countable number of values a exist with the latter expression positive, implying the first result. The second result also follows as well since the last expression in the above display can be 1 for at most 1 value of a. □

To complete the proof of Theorem 3.5.1 (b), choose a value $a \in \mathbb{R}^d$ such that (3.22) holds. By Theorem III.2.2 of [14] and the fact that $X_0 P_s \ll \tau_a(X_0) P_t$, for all $0 < s \leq t$, for the Ornstein-Uhlenbeck semigroup P_t, we have that for all $0 < s \leq t$

$$\mathbb{P}_{X_0}\left(X_{s+\cdot} \in \cdot\right) \ll \mathbb{P}_{\tau_a(X_0)}\left(X_{t+\cdot} \in \cdot\right). \tag{3.25}$$

By our choice of a,

$$\mathbb{P}_{\tau_a(X_0)}\left(\mathbb{P}_{X_1}\left(\lim_{t\to\infty} e^{\gamma t}\bar{Z}_t = 0, S\right)\right) = 0,$$

holds, and hence by (3.25) we have

$$\begin{aligned}\mathbb{P}_{X_0}\left(\lim_{t\to\infty} e^{\gamma t}\bar{Z}_t = 0, S\right) &= \mathbb{P}_{X_0}\left(\mathbb{P}_{X_1}\left(\lim_{t\to\infty} e^{\gamma t}\bar{Z}_t = 0, S\right)\right)\\ &= 0.\end{aligned}$$

Recalling from (3.24) that $\lim_{t\to\infty} e^{\gamma t}\bar{Z}_t$ exists a.s., we are done. □

Note that for $0 > \gamma > -\frac{\beta}{2}$, this implies that even if mass is repelled at rate γ, the COM of the interacting SOU process still settles down in the long run. That is, driving Y_t away from $\bar{Y}_t$ seems to have the effect of stabilizing it. One can think of this as a situation where the mass is growing quickly enough that the law of large numbers overcomes the repelling force.

More surprising is that the COM of the ordinary SOU process diverges exponentially fast, even while the COM of the interacting one settles down. This follows from the correspondence

$$\bar{Y}_t = \bar{Z}_t + \gamma\int_0^t \bar{Z}_s ds,$$

and the cancellation that occurs in the equation due to the exponential rate of $\bar{Z}_t$.

The next lemma shows that Theorem 1 of Engländer and Winter [5] can be reformulated to yield a result for the SOU process with repulsion at rate γ (where γ is taken to be a negative parameter in our setting):

Lemma 3.5.3. *On S, for the SOU process, X, with repulsion rate $-\frac{\beta}{d} < \gamma < 0$ and compactly supported initial measure μ, and any $\psi \in C_c^+(R^d)$*

$$e^{d|\gamma|t}\tilde{X}_t(\psi) \xrightarrow{P} \xi \int_{\mathbb{R}^d} \psi(x)dx,$$

where ξ is a positive random variable on the set S.

Proof. Note that by Example 2 of Pinsky [15] it is shown that the hypotheses of

Theorem 1 of [5] hold for the SOU process with repulsion from the origin at rate $0 < -\gamma < \frac{\beta}{d}$. The theorem says that there is a function $\phi_c \in C_b^\infty\left(\mathbb{R}^d\right)$ such that

$$\frac{X_t(\psi)}{\mathbb{E}^\mu\left(X_t(\psi)\right)} \xrightarrow{P} \frac{W\xi}{\mu(\phi_c)}, \tag{3.26}$$

where W is as in Lemma 3.3.7.

Example 2 also shows that for $\psi \in C_c^+\left(\mathbb{R}^d\right)$,

$$\lim_{t\to\infty} e^{-(\beta+\gamma d)t}\mathbb{E}^\mu\left(X_t(\psi)\right) = \mu(\phi_c)m(\psi)$$

where m is Lebesgue measure on $\mathbb{R}^d$. Hence, manipulating the expression in (3.26) by using the previous equation and Lemma 3.3.7 gives

$$\begin{aligned} e^{|\gamma|dt}\tilde{X}_t(\psi) \xrightarrow{P} \frac{\xi W}{\mu(\phi_c)} \lim_{t\to\infty} \frac{e^{|\gamma|dt}\mathbb{E}^\mu\left(X_t(\psi)\right)}{X_t(1)} \\ = \frac{\xi W}{\mu(\phi_c)} \lim_{t\to\infty} \frac{e^{-(\beta+\gamma d)t}\mathbb{E}^\mu\left(X_t(\psi)\right)}{e^{-\beta t}X_t(1)} \\ = \frac{\xi W}{\mu(\phi_c)} \frac{\mu(\phi_c)m(\psi)}{W}. \end{aligned}$$

□

This lemma indicates that on the survival set, when $\gamma < 0$, one cannot naively normalize X_t by its mass since the probability measures $\left\{\tilde{X}_t\right\}$ are not tight. That is, a proportion of mass is escaping to infinity and is not seen by compact sets. Note that the lemma above implies that for X_t, the right normalizing factor is $e^{(\beta+\gamma d)t}$.

Definition. We say a measure-valued process $\mathring{X}$ undergoes local extinction if for any finite initial measure $\mathring{X}_0$ and any bounded $A \in \mathcal{B}(\mathbb{R}^d)$, there is a $\mathbb{P}_{\mathring{X}_0}$-a.s. finite stopping time τ_A so that $\mathring{X}_t(A) = 0$ for all $t \geq \tau_A$ a.s.

Remark 3.5.4. Example 2 of Pinsky [15] also shows that for $\gamma \leq -\beta/d$ the SOU undergoes local extinction (all the mass escapes). Hence for $\psi \in C_c(\mathbb{R}^d)$, there is no normalization where $X_t(\psi)$ can be expected to converge to something nontrivial.

From Remark 3.3.6, one can show that

$$\bar{Z}_t = \bar{Z}_0 + N_t - \gamma \int_0^t \bar{Z}_s ds,$$

where N is a martingale. Therefore can think of the COM of X, the SOU process repelling from origin, as being given by an exponential drift term plus fluctuations. The correspondence of Remark 3.3.4(b) implies then that

$$\bar{Y}_t = \bar{Y}_0 + N_t,$$

or in other words, the center of mass of the SOU process repelling from its COM is given by simply the fluctuations.

If one fixes a compact set $A \subset \mathbb{R}^d$, then for the ordinary SOU process, X, A is exponentially distant from the COM of the process. However with X', A will possibly lie in the vicinity of the COM of the process for all time. Therefore one might expect that A is charged by a different amount of mass by X_t' than X_t, and thus we might need to renormalize differently for the two cases. We finish with some conjectures:

Conjecture 3.5.5. On the survival set, if X_0' is fixed and compactly supported, then the following is conjectured to hold:

(a) If $0 < -\gamma < \frac{\beta}{d}$ then there exists constant $\beta + \gamma d \leq \alpha < \beta$ so that for $\phi \in C_c(\mathbb{R}^d)$,

$$e^{-\alpha t} X_t'(\phi) \xrightarrow{P} \nu(\psi)$$

where ν is a random measure depending on $\bar{Y}_\infty$.

(b) If $\beta/d \leq -\gamma$ then X_t' undergoes local extinction.

Note that we expect that $\alpha < \beta$ simply because of the repulsion from the COM built in to the model results in a proportion of mass being lost to infinity. One would expect that the limiting measure ν is a random multiple of Lebesgue measure as in the ordinary SOU process case, due to the correspondence, but it is conceivable that it is some other measure which has, for example, a dearth of mass near the limiting COM.

It is difficult to use Lemma 3.5.3 to prove this conjecture as the correspondence becomes much less useful in the repulsive case. The problem is that while the equation

$$\int \phi(x) dX_t'(x) = \int \phi\left(x + \gamma \int_0^t \bar{Z}_s ds\right) dX_t(x)$$

still holds for t finite, the time integral of $\bar{Z}_s$ now diverges.

Chapter 4

Conclusions

The significance of the above research lies in part with the use of Perkins' historical calculus to carry out our analyses. This tool has, for the most part, been underutilized by the broader mathematical community in the construction and study of interacting measure-valued processes since its introduction by Perkins in the early nineties. We have shown that it can be successfully applied in a transparent manner to many problems that lie unsolved in this area. Partially, the importance of research comes from successfully showing that one can determine the equilibrium behavior of a particular interacting measure-valued diffusion (the super Ornstein-Uhlenbeck process attracting to its center of mass).

Dawson, Li and Wang in [2] use a scaling limit method to create a superprocess in a random medium which has location dependent branching according to any given bounded Borel measurable function. While we did not implement this in Chapter 2, the work there could be extended to include variable branching in the vein of Chapter 4 of Perkins [13], although only for location dependent branching functions that are bounded and twice differentiable in space. It is not clear if this method can be extended to include arbitrary Borel measurable branching functions. However, one could use the methods of [13] to incorporate certain types of *interactive* branching, which has not been achieved in any model of this sort thus far. A location dependent branching density is interesting since it is conceivable that particles (like plankton) might give birth to more offspring in some areas of space (ocean) than others. An interactive density is even more interesting since then the rate at which particles branch could vary according to the positions of other particles. This is desirable if, for example, a particle of a particular species is more likely to give birth when there are other particles nearby than when the particle is isolated.

Chapter 2 is motivated by plankton dynamics and could find a number of applications in that area. One could go through the oceanographic literature and decide explicitly which coefficients σ_1, σ_2, and b are most appropriate to model a given population of oceanic phytoplankton. It would be especially interesting to decide what the interactions between the individual plankton should be (and these might vary for different species of plankton). Once a model is decided upon, long term behavior and equilibria could be considered. To do this one may have to overcome some obstacles in the theory of historical calculus, particularly by loosening the restrictions on the types of initial measures admitting solutions to the strong equations $(\text{SE})_{Z_0,K}$. This technical hurdle by itself would be an interesting problem, and once cleared paves the way for considering long-term behavior of different types of measure-valued processes, including those which are immersed in random media.

It would perhaps be more realistic to consider colored noise, instead of white noise, in Chapter 2 to model the dependent spatial motion. Colored noise differs from white noise in that there is a spatial correlation structure built in which complicates analysis. The first step in this direction might be to prove that given a colored noise independent of a historical process remains a colored noise (orthogonal to K_t) under the different measures $\hat{\mathbb{P}}_T$.

In Chapter 3, our most striking results come for the super Ornstein-Uhlenbeck process attracting to its center of mass. Much remains to be shown for the super Ornstein-Uhlenbeck process repelling from its center of mass: for now, a description of the limiting behavior of the process remains elusive. One would guess that the limiting measure should be a random multiple of Lebesgue measure since this is the case for the ordinary SOU process, but it is not entirely clear that this is true since the correspondence between the ordinary SOU process and the SOU repelling from its COM is no longer useful. The COM of the ordinary SOU process diverges, whereas it converges in the interacting SOU, implying that there might be a qualitative difference in the limiting behavior of the two processes.

In regards to the interactive SOU process, there are many directions for future research. First one can look at the aforementioned repelling case, for which little has been established. Secondly, one can consider the critical SOU process attracting to its COM (critical in the sense that the branching law associated with the process has mean 1), and condition it to survive indefinitely. The goal then would be to establish similar equilibrium behavior as in the supercritical case for this new process.

The conditioned interacting SOU process has properties which are highly desirable since physical systems of particles usually behave as if they are critical (and so do not have unchecked population growth) and survive for long periods of time (which is unlike critical processes). Further, this model captures an aspect of biological particle systems where growth is exponential when the population is small and growth is curtailed when the population is large.

The coupling found in Chapter 3 between an ordinary SOU and the interacting SOU can be adapted for this new setting and so the key will be to show that the COM of the interacting SOU converges in some sense. Proving this will be difficult for two reasons: We can no longer rely on it being a martingale and so the drift term will need to be controlled. Secondly, the mass generated by a superprocess conditioned to survive is only $O(t)$, making it harder to take advantage of "law of large numbers" type of behavior than in the supercritical case where there is exponential mass growth.

A different direction of research is to consider what is known as the Fleming-Viot measure-valued process with the same spatial interaction through the center of mass as considered in Chapter 3. The Fleming-Viot process arises as a scaling limit of a number of discrete models for population genetics (such as the Wright-Fisher model). The mean reversion in this process comes from the fact that in the discrete models, the genetic fitness of each individual is determined in comparison to the fitness of the entire population, which is represented by the mean (COM) of the population.

Bibliography

[1] D. Dawson. Stochastic Evolution Equations and Related Measure Processes. *J. Mult. Anal.*, 5:1–52, 1975. → pages 2

[2] D. Dawson, Z. Li, and H. Wang. Superprocesses with Dependent Spatial Motion. *Elec. J. Prob.*, 6:1–33, 2001. → pages 6, 13, 35, 40, 95

[3] J. Engländer. The Center of Mass for Spatial Branching Processes and an Application for Self-Interaction. *To Appear*, 2010. → pages 9, 45

[4] J. Engländer and D. Turaev. A Scaling Limit Theorem for a Class of Superdiffusions. *Ann. Prob.*, 30(2):683–722, 2002. → pages 48

[5] J. Engländer and A. Winter. Law of Large Numbers for a Class of Superdiffusions. *Ann. I. H. Poincaré*, 42(2):171–185, 2006. → pages 48, 91, 92

[6] S. Ethier and T. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 1986. → pages 86, 87

[7] A. Garsia, G. Rodemich, and H. J. Rumsey. A Real Variable Lemma and the Continuity of Paths of Some Gaussian Processes. *Indiana U. Math. J.*, 20: 565–578, 1970. → pages 83

[8] H. Gill. Superprocesses with Spatial Interactions in a Random Medium. *Stoch. Proc. and Appl.*, 119(12):3981–4004, 2009. → pages 5, 46

[9] H. Gill. A Super-Ornstein-Uhlenbeck Process Interacting with its Center of Mass. *To Appear*, 2011. → pages 8

[10] N. Konno and T. Shiga. Stochastic Differential Equations for Some Measure-Valued Diffusions. *Prob. Theory and Rel. Fields*, 79:201–225, 1988. → pages 69

[11] J. Le Gall. *Spatial Branching Processes, Random Snakes and Partial Differential Equations*, volume 1781 of *St. Flour Lectures on Probability Theory and Statistics*. Springer, 1999. → pages 5

[12] L. Mytnik and E. Perkins. Pathwise Uniqueness for Stochastic Heat Equations with Holder Continuous Coefficients: the White Noise Case. *Prob. Theory and Rel. Fields*, 149(1):1–96, 2009. → pages 41

[13] E. Perkins. *On the Martingale Problem for Interactive Measure-Valued Branching Diffusions*, volume 115 of *Memoirs of the American Mathematical Society*. AMS, 1995. → pages 2, 13, 35, 44, 46, 51, 57, 64, 67, 95

[14] E. Perkins. *Dawson-Watanabe Superprocesses and Measure-valued Diffusions*. Lectures in Mathematics ETH Zürich. Birkhäuser, 1999. → pages 4, 12, 13, 15, 19, 21, 24, 25, 27, 28, 30, 32, 41, 43, 46, 48, 51, 52, 53, 56, 57, 58, 59, 61, 67, 74, 75, 90

[15] R. Pinsky. Transience, Recurrence and Local Extinction Properties of the Support for Supercritical Finite Measure-Valued Diffusions. *Ann. Prob.*, 24: 237–267, 1996. → pages 91, 92

[16] A. Polyanin and A. Manzhirov. *Handbook of Integral Equations*. Chapman and Hall/CRC, 2008. → pages 54

[17] G. Roberts and R. Tweedie. Rates of Convergence of Stochastically Monotone and Continuous Time Markov Models. *J. Appl. Prob.*, 37:359–373, 2000. → pages 87

[18] L. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales*. Cambridge University Press, 1985. → pages 67, 68, 69, 72

[19] G. Skoulakis and A. Adler. Superprocesses over a Stochastic Flow. *Ann. App. Prob.*, 11:488–543, 2001. → pages 6, 7, 13

[20] R. Tribe. The Behavior of Superprocesses Near Extinction. *Ann. Prob.*, 20 (1):286–311, 1992. → pages 47, 69, 72, 73

[21] G. Uhlenbeck and L. Ornstein. On the Theory of Brownian Motion. *Phys. Rev.*, 36:823–841, 1930. → pages 8

[22] J. Walsh. *An Introduction to Stochastic Partial Differential Equations*. Lecture Notes in Mathematics. Springer, 1984. → pages 14, 18, 19, 49, 85

[23] H. Wang. State Classification for a Class of Measure-Valued Branching Diffusions in a Brownian Medium. *Probab. Th. Rel. Fields*, 109:39–55, 1997. → pages 6, 13

[24] H. Wang. A Class of Measure-Valued Branching Diffusions in a Random Medium. *Stoc. Anal. Appl.*, 16(4):753–786, 1998. → pages 6, 13

[25] S. Watanabe. A Limit Theorem of Branching Processes and Continuous State Branching Processes. *J. Math. Kyoto U.*, 8:141–167, 1968. → pages 2

[26] J. Xiong. A Stochastic Log-Laplace Equation. *Ann. Prob.*, 32(3B):2362–2388, 2004. → pages 7

[27] W. R. Young. Brownian Bugs and Superprocesses. *Proc. 12th 'Aha Huliko'a Hawaiian Winter Workshop*, 2001. → pages 6

[28] W. R. Young, A. Roberts, and G. Stuhne. Reproductive Pair Correlations and the Clustering of Organisms. *Nature*, 412:328–331, 2001. → pages 6