Nonlinear Parabolic Partial Differential Equations with Branching Diffusion Jump Processes

MATH4400 Capstone Project

WONG Po Chai

1155127163

Department of Mathematics
the Chinese University of Hong Kong
Hong Kong

22 December 2021

Contents

0.1 Introduction

It is well known that the Feynman-Kac Formula can be used to solve a certain class of nonlinear parabolic partial differential equations. However, the class applicable for that formula is restrictive, and therefore, with real-world financial application in mind, Labordére's worked on a larger class of non-linearity and provided a numerical algorithm using branching diffusion processes which outperforms the classical Monte Carlo Simulation method. As an extension, Belak embedded branching diffusion processes with jumps to solve a related PDE with an even wider class of uncertainties. This paper is to conclude their results and try to extend them in an experimental attitude.

The article is organised as follows:

Chapter 1 presents the preliminaries of the probability theory, the Brownian motion and stochastic differential equations. Chapter 2 lays down the discoveries related to the *Feynman-Kac Formula.* Chapter 3 gives a probabilistic account of the solution to the Feynman-Kac formula by combining diffusion with the branching processes. There, we will evolve from the simplest such counting process, *Galton-Watson Processes* and present some of their related theories. Our goal is to modify the process formally to suit a wider class of partial differential equations. The rigorous construction will be illustrated in a more general setting, *Marked Branching Diffusion with Jumps* in Chapter 4. We present the construction and the intuitive proofs related and showed the relationships between PDE and the branching representation of an appropriate class of stochastic processes. We also will present numerical simulations to verify some of the results.

Chapter 5 deserves a whole chapter of treatment since it enlists my slight contribution to the theories in the two ways: (1) looking for sufficient conditions to look for the PDE related to the moments of the estimator in the original main theorem; and (2) we limit our attention to exponential split time and verify the sufficient condition for the applicability with exponential distributions using our own MATLAB codes.

Chapter 1

Preliminaries

To capture the whole sense of the theories, with slight bothersomeness, we start from the very beginning of the probability theory in a succinct manner, then move to the Brownian motion and their related stochastic differential equations. Here we assume some basic notions from measure theory. The approach used here is in reference to Flesher's work.

1.1 Probability Theory

In probability, we study randomness in a probability space, a triplet $(\Omega, \mathcal{F}, \mathbb{P})$, a σ-finite measure space with $\mathbb{P}(\Omega) = 1$, (i.e. the space can be decompose into countably many subsets, each with finite measure under $\mathbb{P}$). Here Ω is called the *sample space*, $\mathcal{F}$ the *event space*, $\mathbb{P}$ the *probability measure*.

A *random variable* is a measurable function $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ-algebra of $\mathbb{R}^d$. On the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the *expected value* of X is defined as

$$\mathbb{E}[X] = \int_\Omega X(\omega)\mathbb{P}(d\omega).$$

Definition 1.1.1 (Stochastic Process)**.** A *stochastic process* is a collection of random variables $\{X_t : t \in I\}$, where I here is an index set. A *sample path* is a map $t \mapsto X_t(\omega)$ where $\omega \in \Omega$ is fixed. A *filtration* is a non-decreasing family $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ with $0 < s < t < \infty$, of sub-σ-field of $\mathcal{F}$. A stochastic process X_t os $\mathcal{F}-$measurable if the function

$$\begin{aligned}([0,\infty) \times \Omega, \mathcal{B}([0,\infty) \otimes \mathcal{F}) &\to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))\\ (t,\omega) &\longmapsto X_t(w)\end{aligned}$$

is measurable. A process is *adapted* to a filtration $\{\mathcal{F}_t\}$ if for each t, X_t is an $\mathcal{F}_t$-measurable random variable.

For the rest of the chapter, we assume the filtration $\mathcal{F}_t$ is right-continuous and all $\mathbb{P}$-negligible events in $\mathcal{F}$.

Definition 1.1.2 (Martingale)**.** A *martingale* is an adapted process X that satisfies

$$\mathbb{E}[\,X_t |\, X_s] = X_s \text{ for any } s < t.$$

Remark 1. Intuitively, this refers to stochastic processes that the best we can guess based on the present information is the value at the last available moment.

We say that a martingale is *square-integrable* if the expectation of its second moment is finite at any time. The *norm* of a square-integrable martingale is

$$\|X\|_t = \sqrt{\mathbb{E}[X_t^2]}$$

With this norm, we can construct a metric space $\mathcal{M}_2, \|\cdot\|$ of all square-integrable martingales X with $X_0 = 0$. It can be proved that this space is a complete metric space, and the subset $\mathcal{M}_2^c$ of continuous square-integrable martingales is closed.

We now have a space that has a lot of nice properties, and with closed-ness, we can talk about the asymptotics of some sequence of processes later.

Definition 1.1.3 (Stopping Time)**.** On $(\Omega, \mathcal{F})$ with filtration $\mathcal{F}_t$, a *stopping time* T is an $\mathcal{F}$-measurable random variable taking values in $[0, \infty]$ such that $\{T \leq t\} \in \mathcal{F}_t$.

Basically, it means that "stopping at time t" is allowed in $\mathcal{F}$. We therefore denote $X_T(\omega) = X_{T(\omega)}(\omega)$ for $\omega \in \Omega$.

Definition 1.1.4 (Quadratic Variation)**.** If X is a square integrable martingale, the *quadratic variation* of X is the unique martingale $\langle X \rangle$ such that $X^2 - \langle X \rangle$ is a martingale. Alternatively, suppose $0 = t_0 \leq t_1 \leq \cdots \leq t_m = t$ be a partition of $[0, t]$. The quadratic variation is the limit of $\sum_{k=1}^{m} (X_{t_k} - X_{t_{k-1}})^2$ as $\lim_{m\to\infty} \|\{t_i\}\| = 0$.

Definition 1.1.5 (Covariation)**.** For square-integrable processes X, Y, the *covariation* of X and Y is

$$\langle X, Y \rangle_t = \frac{1}{2}(\langle X + Y \rangle - \langle X \rangle - \langle Y \rangle).$$

Now we are ready to talk about one important notion, *Brownian Motion.*

Definition 1.1.6 (Brownian Motion)**.** A one-dimensional *standard Brownian Motion* is a continuous adapted process $\{W_t, \mathcal{F}_t, t \in \mathbb{R}^+\}$ such that

1. $W_0 = 0$ a.s.
2. $\forall 0 \leq s < t$, $W_t - W_s$ is independent of $\mathcal{F}_s$
3. $\forall 0 \leq s < t$, $W_t - W_s \sim \mathcal{N}(0, t - s)$, where $\mathcal{N}$ is the Gaussian distribution.

Remark 2. Intuitively, Brownian motion can be thought of as the limit of a random walk, as the step size tends to 0 and taking infinitely many steps. There are a lot of nice properties of Brownian motions. We do not explicate them here but one important property is that it is a continuous square integrable martingale. We will use this fact a lot.

Another related name for the differential of this process is *white noise.* Its name really does it justice since it captures the sense that this is *stationary in the wide sense.* One can simply take the Fourier transform and notice that the *spectral density* of it is flat. Just like white light, which is even in the color spectrum.

A natural question is about the existence of Brownian motion. The construction is rather non-trivial yet brute-force with some clever setting using the Schrauder's functions.

The approach is to look for some cleverly selected orthonormal basis in $\mathcal{M}_2$, construct a series, prove its convergence and verify that it's indeed the Brownian motion. The details are omitted here.

With Brownian motion, we are at the position to construct a *stochastic integral.*

1.2 Stochastic Integral

The main motivation is of stochastic integrals is that, when we want to express a dynamical system, it is often natural to write a differential equation. But when uncertainties are involved, one needs to develop calculus for stochastic processes, so that the two theories merge. There are a lot of variations in the construction, and one rule is *Itô's Calculus*.

In classical calculus, we some define integrals using Riemann approximation of series or the Lebesgue-Stieltjes integral. One can attempt doing so for stochastic integrals, by summing infinitely terms of stochastic processes. But the attempt would be in vain because martingales may not have bounded variations. From this, we see the necessity of a new means to construct stochastic integrals. But luckily, one can prove that continuous square-integrable martingales admit quadratic variation. We can define the integral $\int_0^T X_t(\omega) d\left\langle M\right\rangle_t(\omega)$.

To do this, suppose $M \in \mathcal{M}_2^c$, we define a measure

$$\mu_M(A) := \int_0^\infty \chi_A((t,\omega)) d\left\langle M\right\rangle_t(\omega)$$

on $([0,\infty) \times \Omega, \mathcal{B}([0,\infty) \otimes \mathcal{F}))$.

Let $\mathcal{L}$ be the set of measurable, $\mathcal{F}_t$-adapted processes that

$$[X]_T^2 := \mathbb{E}\left[\int_0^\infty X_t^2 d\left\langle M\right\rangle_t\right] < \infty \text{ for } T > 0$$

This is then an L^2-norm on $[0,T] \times \Omega$ under μ_M. Notice the space $\mathcal{L}$ depends on M.

Also, we define another metric

$$[X] := \sum_{n=1}^\infty \frac{\min(1, [X]_n)}{2^n}.$$

We further suppose that the map $t \mapsto \left\langle M\right\rangle_t(\omega)$ is *absolutely continuous* with respect to the Lebesgue measure for almost all ω. We now can follow the approach in Lebesgue-Stieltjes integral.

Definition 1.2.1 (Simple Process)**.** A process H is *simple* if it admits the form

$$H_t(\omega) = \xi_0(\omega)\chi_{\{0\}}(t) + \sum_{i=0}^\infty \xi_i(\omega)\chi_{(t_i,t_{i+1}]}(t)$$

where $\{t_i\}_{i=0}^\infty$ is a strictly increasing sequence with $t_0 = 0$, $t_n \to \infty$, $\{\xi_n\}$ is a sequence of random variables that are uniformly bounded. This class of simple process is denote by $\mathcal{L}_0$.

One can prove that $\mathcal{L}_0$ is dense in $\mathcal{L}$ with respect to the metric $[\cdot]$. The approach here is to construct a similar series using simple processes

$$I_t(X) = \sum_{i=0}^\infty \xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge t_i}).$$

Then, prove that every element in $\mathcal{L}$ can be approximated by a sequence in $\mathcal{L}_0$. The sequence can be proved to be Cauchy, and thus under I, the resultant sequence also converges to $I(X)$ since $\mathcal{M}_2^c$ is complete and closed. Uniqueness also follows from a similar approach.

Without much of the technical details, we now can define stochastic integrals.

Definition 1.2.2 (Stochastic Integral). For $X \in \mathcal{L}$, the *stochastic integral* of X with respect to M is the unique process $I(X)$ such that $\|I(X^{(n)}) - I(X)\| \to 0$ as $n \to \infty$ for all sequences of simple functions $X^{(n)}$ with $\left[X^{(n)} - X\right] \to 0$. Denote

$$I_t(X) = \int_0^t X_s dM_s.$$

1.3 Itô's Lemma

We now move to the central topic of *Itô's lemma*.

Definition 1.3.1 (Continuous Local Martingale). A *continuous local martingale* is a continuous adapted process M_t such that there is a non-decreasing sequence of stopping times $\{T_n\}$ with $\mathbb{P}(T_n \to \infty) = 1$ and that $\{M_{t \wedge T_n}; \mathcal{F}_t\}$ is a martingale for each n.

Definition 1.3.2 (Continuous Semimartingale). A *continuous semimartingale* is an adapted process of the form

$$X_t = X_0 + M_t + A_t,$$

where M is a local continuous martingale, and A_t is a finite variation process.

Theorem 1 (Itô's Lemma). *Suppose $X^1, \cdots, X^d$ are continuous semimartingales, and that $f : \mathbb{R}^+ \times \mathbb{R}^d$ be a $\mathcal{C}^{1,2}$ function. Then, with $\mathbb{P}-$a.s.,*

$$\begin{aligned} f(t, X_t) =& f(0, X_0) + \int_0^t \frac{\partial}{\partial t} f(s, X_s) ds + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(s, X_s^1, \cdots, X_s^d) dX_s^i \\ &+ \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j} f(s, X_s^1, \cdots, X_s^d) d\left\langle X^i, X^j \right\rangle_s. \end{aligned}$$

Essentially, this is the chain rule in stochastic calculus.

1.4 Stochastic Differential Equations

The most important result here is the existence and uniqueness of the solution to a stochastic differential equations.

Theorem 2 (Existence and Uniqueness of Solution to a stochastic differential equation). *Suppose $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $B : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{m \times n}$ are continuous and satisfy the following: there exist some constant $L > 0$*

$$\|b(t, x) - b(t, y)\| + \|B(t, x) - B(t, y)\| \leq L\|x - y\| \text{ for all } t \in [0, T], x, y \in \mathbb{R}^n$$

$$\|b(t, x)\| + \|B(t, x)\| \leq L(1 + \|x\|) \text{ for all } t \in [0, T], x \in \mathbb{R}^n$$

Suppose X_0 is a $\mathbb{R}^n$-valued random variable such that

$$\mathbb{E}[|X_0|^2] < \infty$$

and

$$X_0 \text{ is independent of the increments}$$

where $W_0(\cdot)$ is a given m-dimensional Brownian motion.

Then, there exists a unique solution $X \in \mathbb{L}_n^2(0,T)$ of the stochastic differential equation:

$$\begin{cases} dX & = b(t,X)dt + B(t,X)dW \ \text{for } t \in [0,T] \\ X(0) & = X_0 \end{cases}$$

Remark 3. The first condition states the uniform Lipschitiz continuity of b and B in terms of x. The "uniqueness" here refers to the almost sure equality of random variables.

Proof. The proof of the uniqueness makes use of the Gronwall's Lemma mainly; while the existence requires the approximation scheme similar to the Picard's iteration, but in the context of random variables. The proof is omitted here. □

1.5 Diffusion Processes

Diffusion refers to the stochastic processes that solve a stochastic differential equation. Only a brief little would be said here.

First is that, for an Itô's diffusion process X, we represent in the following differential:

$$dX_t = \mu dt + \sigma dW_t$$

where μ and σ are measurable functions called *drift* and *volatility* respectively. Drift refers to the general overall trend of the process whereas volatility refers to the local abrupt changes possible.

From this equation, along with some initial condition, we can simulate the whole process. In particular, if the process is a martingale (in fact a local martingale will do if we consider a finite interval), the drift term would have to be zero. In fact, the *Girsanov theorem* tells us that, under certain conditions, we can find a measure under which the drift term vanishes. We will see an example in the next chapter.

Notice that the μ takes on the deterministic part while σ the stochastic one along with white noise (differential of Brownian motion). And recall that Brownian motion has independent and stationary increments. We can heuristically conclude that diffusion processes are Markov process. This allows us to relate to the *infinitesimal generator*.

Definition 1.5.1 (Infinitesimal Generator)**.** For a diffusion process X and a function $f : \mathbb{R}^d \to \mathbb{R}$, the *infinitesimal generator* of X is a second order partial differential operator $\mathcal{L}$ defined as

$$\mathcal{L}[f](x) := \lim_{t \to 0^+} \frac{\mathbb{E}_x[f(X_t)] - f(x)}{t}.$$

For f such that is limit exists for all $x \in \mathbb{R}^d$, and suppose $dX_t = \mu dt + \sigma dW_t$ we have directly

$$\mathcal{L}[f](x) = \sum_{i=1}^{d} \mu_i \frac{\partial}{\partial x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^{d} (\sigma\sigma^T) \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Intuitively, this represents how the function would change infinitesimal with respect to the stochastic process. This is why we see a quantum of resemblance between this expression and the heat equation.

Chapter 2

Feynman-Kac Formula

2.1 First Glance at Black-Scholes Formula

One common use of the Feynman-Kac formula is in Financial Mathematics, namely the *Black-Scholes Formula*, in option pricing theory.

There are lots of derivations of the formula. Here we present a nice one using the martingale theory. A direct application of what we introduced just now.

Theorem 3 (Martingale Approach to the Black-Scholes Formula)**.** *Suppose the stock price S_t follows the geometric Brownian motion with drift μ and volatility σ.*

$$dS_t = S_t \left(\mu(t, S_t)dt + \sigma(t, S_t)dB_t\right)$$

And the interest rates R_t follows

$$dR_t = r(t, S_t)R_t dt.$$

Suppose there is a probability measure $\mathbb{Q}$ that is mutually absolutely continuous with respect to $\mathbb{P}$, the realistic probability, such that under $\mathbb{Q}$,

$$\tilde{S}_t := \frac{S_t}{R_t}$$

the discounted stock price is a martingale.

Further suppose that there is an option at time T with payoff value $F(S_T)$ such that

$$\mathbb{E}\left[R_T^{-1}\left|F(S_T)\right|\right] < \infty.$$

Then, the arbitrage-free price of the option at time $t < T$ is

$$V_t = R_t \mathbb{E}\left[R_T^{-1} F(S_T)\middle|\mathcal{F}_t\right].$$

Remark 4. We should note that the second assumption about a mutually absolutely continuous measure is stated from the first *Fundamental Theorem of Asset Pricing (FTAP)*, while its uniqueness comes from the second FTAP about the completeness of the market. Another assumption implicitly made here is the *self-financing principle*, which simply means no other money flow is allowed during the process.

Remark 5. Since the condition on R_t is deterministic, we can solve

$$dR_t = r(t, S_t)R_t dt$$

like an ODE, giving us

$$R_t = \exp\left\{\int_0^t r(v, S_v)dv\right\}$$

conditional on $R_0 = 1$. In this case, the arbitrage-free price is actually

$$V_t = \mathbb{E}\left[\exp\left\{-\int_t^T r(v, S_v)dv\right\} F(S_t)\middle| \mathcal{F}_t\right]$$

The essence of the formula is that, under the two FTAPs, the discounted value of the asset is essentially a martingale under a certain probability measure, namely risk-neutral measure. The details can be covered in a standard textbook of Financial Mathematics.

The main point here is about its use in the Black-Scholes formula.

By assuming $\phi(t, S_t) = V_t = \tilde{V}_t R_t$ where $\tilde{V}_t$ is the discounted claimed value of the asset at time t, and by applying the Itô's formula, we have

$$\begin{aligned}
d\phi(t, S_t) &= \partial_t \phi(t, S_t)dt + \phi'(t, S_t)dS_t + \frac{1}{2}\phi"(t, S_t)d\left\langle S\right\rangle_t \\
d\tilde{V}_t &= d\left[R_t^{-1}\phi(t, S_t)\right] \\
&= R_t^{-1}\left[\partial_t\phi(t, S_t) + \frac{\sigma(t, S_t)^2 S_t^2}{2}\phi"(t, S_t) + r(t, S_t)S_t\phi'(t, S_t) - r(t, S_t)\phi(t, S_t)\right]dt \\
&+ \left[R_t^{-1}S_t\sigma(t, S_t)\phi'(t, S_t) - \tilde{S}_t\sigma(t, S_t)\phi'(t, S_t)\right]dW_t^{\mathbb{Q}}
\end{aligned}$$

where $\left\langle S\right\rangle_t$ is the quadratic variation of the random variable S_t. The fact that V_t is a martingale entails that the dt term vanishes.Hence, we have our *Feynman-Kac formula*

$$\partial_t\phi(t, S_t) + \frac{\sigma(t, S_t)^2 S_t^2}{2}\phi"(t, S_t) + r(t, S_t)S_t\phi'(t, S_t) - r(t, S_t)\phi(t, S_t) = 0$$

Obviously, this approach assumes enough smoothness of ϕ which may not be the case here.

However, this can be overcome by proving that ϕ corresponds to a local martingale and therefore we can obtain the same result by the Optional Sampling Theorem. The verification is omitted here.

2.2 Second Glance at the Heat Equation

But the link to the branching diffusion process is quite vague from the above approach. The original discovery was actually due to McKean, regarding a question arisen from physics about wave equation. However, we here states a more accessible approach.

Definition 2.2.1 (Heat Equation with Dissipation)**.** Let $u : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ be a $\mathcal{C}^{1,2}$ function,

$c : [0, \infty)\times\mathbb{R}^d \to \mathbb{R}$ be the *dissipation rate*. Suppose u is continuous at each point $x \in \{0\}\times\mathbb{R}^d$ and $u(0, x) = f(x)$ for any $x \in \mathbb{R}^d$. Suppose

$$\partial_t u(t, x) = \frac{1}{2}\Delta u(t, x) - c(t, x)u(t, x) \text{ on } (0, \infty) \times \mathbb{R}^d$$

.

Theorem 4 (Feynman-Kac Formula I.). *Under the above setting, suppose that c is bounded and u be a solution to the Heat Equation with Dissipation c on every $[0,t] \times \mathbb{R}^d$. Then, uniquely determined,*

$$u(t,x) = \mathbb{E}_x \left[\exp \left\{ - \int_0^t c(t-r, W_r) dr \right\} f(W_t) \right]$$

We can state the formula even formally with the existence theorem.

Theorem 5 (Feynman-Kac Formula II). *Let $f : \mathbb{R}^d \to \mathbb{R}$, $c : \mathbb{R}^d \to \mathbb{R}^+$ and $g : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be continuous functions. Let $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be a function which is $\mathcal{C}^{1,2}$ on $[0,T) \times \mathbb{R}^d$ such that*

$$\begin{gathered} -\frac{\partial v}{\partial t} + cv = \frac{1}{2}\Delta v + g \text{ on } [0,T) \times \mathbb{R}^d \\ v(T,x) = f(x) \text{ for } x \in \mathbb{R}^d. \end{gathered}$$

Furthermore, suppose

$$\max_{0 \le t \le T} |v(t,x)| + \max_{0 \le t \le T} |g(t,x)| \le K e^{a|x|^2}$$

for any $x \in \mathbb{R}^d$ and some $K > 0$, $0 < a < \frac{1}{2dT}$. Then, we can write

$$v(t,x) = \mathbb{E}_x \left[f(W_{T-t}) e^{-\int_0^{T-t} c(W_s) ds} + \int_0^{T-t} g(t+r, W_r) e^{-\int_0^r c(W_s) ds} dr \right].$$

Remark 6. The key lies with the connection between Heat Equation and Brownian motions. The heat equation states the distribution of heat energy over time. To relate this to particles, we can imagine in an open bounded space, a lot of negligibly-weighted particles with a fixed total mass are released at some point in the space. As time goes, their move according to Brownian motions. Dissipation also plays an important role here. It means the particle may vanish along their journey, which is why it is also called the killing rate.

The conditional expectation here refers to the expected distribution of the particles on the boundary of this space, conditional on the even that they all start at the point x a.s.

We see that there is a probabilistic view to solving the PDE, namely *Branching Diffusion Processes.* In fact, this can be used to solve a larger class of PDE with non-linearity.

The rest of this report surrounds this topic and their related theories.

Chapter 3

Branching Diffusion Processes

We start with the simple counting process.

3.1 Galton-Watson Processes

Definition 3.1.1 (Galton-Watson Process)**.** A *Galton-Watson process* is a Markov chain $\{Z_n\}_{n\in\mathbb{N}}$ on non-negative integers. Along with its transition function $P(i,j)$ defined using probabilities $\{p_k\}$ as

$$P(i,j) = \mathbb{P}\{\, Z_{n+1} = j |\, Z_n = i\} = \begin{cases} p_j^{*i} & \text{if } i \geq 1,\ j \geq 0 \\ \delta_{0,j} & \text{if } i = 0,\ j \geq 0 \end{cases}$$

where $\{p_k^{*i}\}_{k\in\mathbb{N}}$ is the i-fold convolution of $\{p_k\}_{k\in\mathbb{N}}$.

Intuitively, starting with i particles, each particle may split into k particles, $k = 0, 1, \cdots, M$, with the distribution $\{p_k\}_{k=0,\cdots,M}$.

The study of this process can be done via the generating function:

$$f(s) = \sum_{k=0}^{\infty} p_k s^k \text{ for } |s| \leq 1.$$

One important remark worth noting here is that suppose $Z(\omega)$ is on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, this setting does not tell us enough after what happens after the splits. In this case, we need to extend Ω to contain all "family trees".

One important property of G-W processes is the *Additive Properties*: $\{Z_n^{(i)}; n = 0, 1, \cdots\}$ is the sum of i independent copies of the branching process. In other words, ...

We can classify the process into three main cases:

$$\begin{aligned} \text{Subcritical case } &: \mathbb{E}[Z_n] < 1 \\ \text{Critical case } &: \mathbb{E}[Z_n] = 1 \\ \text{Supercritical case } &: \mathbb{E}[Z_n] > 1 \end{aligned}$$

The importance of this classification is the existence of solutions and the asymptotic behaviors. In the subcritical cases, the process would arrive at 0 within a finite time while it would be infinite asymptotically in the super-critical cases. We can go further. In fact, in the sub-critical cases, at any time t, conditioned on that the process is not at 0, we can find the distribution of the particles. In the super-critical cases, it can be proven that the number of particles has an exponential growth. In particular, the process will converge to a certain stochastic process. By some manipulation, we can create a related martingale and expand our theories from that. But the theories are rather complicated and thus are omitted from this article with only reference to Arthreya's work. The main point is that there exists such a well studied process whose intuition coincides with our prototype here.

3.2 Branching Diffusion Processes

Going back to our original goal, we combine branching processes with diffusion processes. In the last section, G-W processes only concern about the number of particles and the time stamps are in discrete units. One extension is to look at the distribution of the particles in a continuous-time setting.

We will officially construct such a process rigorously in the next chapter in a more useful variation. But the main mechanism of such a process is that, in an $\mathbb{R}^d$ space, we start with a particle at a point x_0, it moves according to a diffusion process with drift μ and volatility σ.

Up to a certain time point τ_1, called a *split time*, that follows the distribution γ, it dies and produces $k \in [0 : m]$ descendants according to the law $\{p_k\}$, each follows the same procedure of movements and splits.
Continue this fashion until the process reaches the *mature time* T.

Here we moved from Galton-Watson processes to branching diffusion process. The converse also works. This is the embeddability problem. A natural idea is to enforce discrete time stamps on the time interval and treat the number of particle at each time stamp as a state. This can be stated as a theorem:

Theorem 6. *For every $\delta > 0$, $\{Z_n(\omega)\} := \{Z(n\delta, \omega)\}$ is a Galton Watson process with offspring probability generating function*

$$g(s) = F(s, \delta) = \sum_{k=0}^{\infty} \mathbb{P}\{\, Z(t) = k |\, Z(0) = 1\} s^k$$

Under this setting, we are empowered to study continuous-time processes using the limit results in G-W processes. However, this is not going to be our main focus, since it only captures the number of particles. What we are also interested in is the position of each particle.

3.3 Marked Branching Diffusion

Now with the intuition of Branching Diffusion, we are at the position to introduce a useful variation.

3.3.1 Notations

From now on, we will use the following notations.

1. We fix $T > 0$ as the *mature time* and $m \in \mathbb{Z}_{>0}$ as the *particle type*.

2. $\psi : \mathbb{R}^d \to \mathbb{R}$ is a smooth function.

3. We consider the following PDE

$$\text{(PDE): } \partial_t u + \mathcal{L}[u] + f = 0$$

   with the boundary condition

$$u(T, x) = \psi(x)$$

   where $u : [0, T] \times \mathbb{R}^d \to \mathbb{R}$.

4. $$\mathcal{L}[u](t, x) := \mu(t, x)^T D_x u(t, x) + \frac{1}{2} tr\left[\sigma(t, x)\sigma(t, x)^T D_x^2 u(t, x)\right]$$

   is the infinitesimal generator of the corresponding family $\{W^{(k)}\}$ of independent diffusion processes with noise in $\mathbb{R}^d$-valued Brownian motions.

5. $$f : [0, T] \times \mathbb{R}^d,\ f(t, x) := \sum_{k=0}^{m} a_k u(t, x)^k$$

   is the non-linearity term with constants a_k.

Remark 7. This means, at every split time, the particle is given a *type* $k \in [0 : m]$.

This modification is particularly important. If we apply the concept of particle types to the original G-W processes, we obtain the study of multi-type G-W processes. One interesting result is that, conditioning on non-extinction, we can grasp the asymptotic behaviours of the process. In fact, this result is more apparent in the continuous-time Markov process; while similarly but weaker results can be found for arbitrary split time distribution (the process is then called *age-dependent processes*).

3.3.2 Branching Diffusion Representation

Our estimator in the end becomes

$$u(t, x) = \mathbb{E}_{t,x}[\Psi] = \mathbb{E}_{t,x}\left[\prod_{i=1}^{N_T} \psi(z_T^i) \prod_{k=0}^{m} \left(\frac{a_k}{p_k}\right)^{\omega_k}\right]$$

where ω_k is the number of k-type particles, N_T is the number of alive particles at time T, and z_T^i is the position of the i-th particle alive at time T.

Justification was given in Labordére's, which mainly relies on re-arrangement of the terms in the Feynman-Kac Formula we mentioned above.

In particular, the version of Feynman-Kac formula used there concerned only when the split times follow exponential distribution. But in fact, that is not necessary under some other conditions. We will cover that in the next chapter.

Also, in this case, the non-linearity is expressed in the form of a polynomial with constant coefficients a_k. This is also not necessary. As we see later, the coefficients can be dependent on (t, x).

Indeed, the motivation originally came from McKean who studied the case when $m = \infty$. But in practice, this version provides a more realistic alternative in which $m < \infty$.

3.4 Optimal Probabilities

Labordére also proposed the optimal probabilities that minimize the variance upper bound. Here we simply state the results:

Proposition 1 (Optimal Probabilities). *Suppose the estimator exists, then the optimal probability of the $k-$type particle is given by*

$$p_k = \frac{|a_k| \|\psi\|_\infty^k}{\sum_{k=0}^M |a_k| \|\psi\|_\infty^k}.$$

Proof. Under the existence condition,

$$\begin{aligned}\tilde{u}(0, x) &\leq \mathbb{E}_{0,x}\left[\prod_{k=0}^M \left(\frac{|a_k|}{p_k}\right)^{\omega_k} \|\psi\|_\infty^{N_T}\right] \\ &= \|\psi\|_\infty \tilde{\mathbb{P}}\left(T, -\log\frac{|a_k|}{p_k} - \log\|\psi\|_\infty^{k-1}\right),\end{aligned}$$

where $\tilde{\mathbb{P}}$ is the Laplace transform of $\mathbb{P}$,

$$\tilde{\mathbb{P}}(T, c) = \mathbb{E}\left[\prod_{k=0}^M e^{-c_k \omega_k}\right].$$

The proposed $\{p_k\}$ are obtained by minimizing the last term. $\square$

Chapter 4

Marked Branching Diffusion with Jumps

In this chapter, we are concerned about the following PDE:

$$\text{(PDE): } \partial_t u(t,x) + \mathcal{L}\left[u\right](t,x) + \int_{\Xi} f(t,x,\xi,\mathcal{J}\left[u\right](t,x,\xi))\gamma(d\xi) = 0$$

$$\text{(BC): } u(T,x) = \psi(x)$$

To rigorously describe the probabilistic account in addition to the last chapter, we need to work on the *Branching Mechanism* (the distribution of the offspring) and the *Diffusion Dynamics* (how the particles move in between splits).

4.1 Branching Mechanism

Some new notations:

1. $\mathcal{N}_n := \bigcup_{i=1}^n \mathbb{N}^i$ is the set of all $\mathbb{N}$-words with length at most n and we let $\mathcal{N} := \bigcup \mathcal{N}_n$.
2. γ is the jump distribution on an abstract measurable space $(\Xi, \mathcal{B})$.
3. $\mathcal{I} \subseteq \mathbb{N}_0^m$ is the set of multi-indices. For each $I \in \mathcal{I}$, $I := (I_1, I_2, \cdots, I_m)$ where each I_l is the number of the l-type offspring, which appears at split with probability p_l.
   From this $\mathcal{I}$, we have $\{I^{(k)} \in \mathcal{I}\}_{k\in\mathcal{N}}$ representing the distribution of offspring in each generation k.
4. $$\mathcal{J}\left[u\right] := (\mathcal{J}_1\left[u\right], \mathcal{J}_2\left[u\right], \cdots, \mathcal{J}_m\left[u\right])$$
   with each $\mathcal{J}_l\left[u\right](t,x,\xi) := u(t, \Gamma_l(t,x,\xi)))$ where $\Gamma_l : [0,T] \times \mathbb{R}^d \times \Xi \to \mathbb{R}^d$ are measurable jump maps for each l.
5. We specify the family $\{W^{(k)}\}_{k\in\mathcal{N}}$ to be i.i.d. random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
6. $\{\Delta^{(k)}\}_{k\in\mathcal{N}}$ is a family of i.i.d. $\mathcal{R}^d$-valued random variables with distribution γ.
7. $\{\tau^{(k)}\}_{k\in\mathcal{N}}$ is a family of i.i.d. $\mathbb{R}_+$-valued random variables with distribution F, which admits a continuous density $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying
   $$F(t) := \int_t^\infty \rho(s)ds > 0, \text{ for all } t > 0.$$

8. Like what we said previously, we consider a more general non-linearity,

$$f : [0,T] \times, \mathbb{R}^d \times \Xi \times \mathbb{R}^m, (t,x,\xi,y) \longmapsto \sum_{I\in\mathcal{I}} c_I(t,x,\xi) y^I.$$

   where c_I are bounded measurable.

Remark 8. As mentioned in the theory of Galton-Watson processes, Ω here is the set of all family trees.

We first assume that the number of particles will not explode.

Assumption 1. For all $I \in \mathcal{I}$,

$$\mathbb{E}[|I|] = \sum_{I\in\mathcal{I}} |I|\, p_I < \infty$$

We now formalise the branching mechanism.
We fix $t \in [0,T]$ to be the starting point and consider $\omega \in \Omega$.

1. For a particle of generation $n \in \mathbb{N}$ labeled $k = (k_1, \cdots, k_n) \in \mathbb{N}^n$, denote its parent by $k- := (k_1, \cdots, k_{n-1}) \in \mathbb{N}^{n-1}$.
2. The split time of particle k is $T^{(k)} := \left(T^{(k-)+\tau^{(k)}} \wedge T\right)$.
3. If $\{T^{(k)} < T\}$, at $T^{(k)}$, the particle k dies and gives $I_l^{(k)}$ l-type particles for $l \in [0:m]$, respectively labeled by $(k_1, \cdots, k_n, k_{n+1})$ with $k_{n+1} \in \left[1 : \left|I^{(k)}\right|\right]$.
4. The jump types of the k-th generation $\mathcal{J}^{(k)} = \left(\mathcal{J}_1^{(k)}, \cdots, \mathcal{J}_{|I^{(k)}|}^{(k)}\right)$ is defined by $\mathcal{J}_i^{(k)} = l$ if the i-th particle is l-type.
5. The process starts with

$$(1)- := () = \emptyset \text{ and } T^{\emptyset} = t$$

6. We denote the set of particles of generation $n \in \mathbb{N}$ alive at time $s \in [t,T]$ by

$$\mathcal{K}_t^n(s) := \begin{cases} \{k \in \mathbb{N}^n : k \text{ is a particle and } T^{(k-)} \le s < T^{(k)}\} & \text{if} s \in [t,T) \\ \{k \in \mathbb{N}^n : k \text{ is a particle and } T^{(k)} = T\} & \text{if} s = T \end{cases}$$

7. We also define

$$\bar{\mathcal{K}}_t^n(s) := \bigcup_{r\in[t,s]} \mathcal{K}_t^n(r)$$

   to be the set of all particles alive before or at s regardless of their generations.

   For simplicity, When we do not specify n in the expression, we consider the union over all $\{n \in \mathbb{N}\}$. When we do not specify the s, we are considering the mature time $\mathcal{K}_t^n(T)$.

The notations above describe the whole picture regarding the splits. We now mention the dynamics in between split times.

4.2 Diffusion Dynamics

Assumption 2. The functions

$$\mu : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \text{ and } \sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$$

are measurable and satisfy the following Lipschitz and linear growth conditions: There exists a constant $L > 0$ such that

$$\begin{aligned} |\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| &\leq L\,|x-y|,\ t \in [0,T]\,, x,y \in \mathbb{R}^d \\ |\mu(t,x)|^2 + |\sigma(t,x)|^2 &\leq L^2\left(1+|x|^2\right),\ t \in [0,T]\,, x \in \mathbb{R}^d \end{aligned}$$

Remark 9. As we have seen before, this assumption ensures the existence and uniqueness of the solution to the stochastic differential equation for each $(t,x) \in [0,T] \times \mathbb{R}^d$,

$$\begin{aligned} X_s^{t,x} &= x && s \in [0,t] \\ dX_s^{t,x} &= \mu(s, X_s^{t,x})ds + \sigma(s, X_s^{t,x})dW_s && s \in [t,T] \end{aligned}$$

Considering the whole $[0,T] \times \mathbb{R}^d$ at the same time, we can obtain a random field $\bar{X}_s^{t,x}$ and can directly replace the $X_s^{t,x}$ in the equations.

We define $\bar{\mathcal{F}}^{t,x} := \{\bar{\mathcal{F}}_s^{t,x}\}_{s\in[0,T]}$ to be the natural filtration of $\bar{X}^{t,x}$.

We now take for granted the following properties of diffusion:

Lemma 7. *We can choose a certain $\{\bar{X}_s^{t,x}\}_{s,t\in[0,T],x\in\mathbb{R}^d}$ that satisfies*

1. *It solves the stochastic differential equation above.*

2. $$\bar{X} : [0,T] \times \mathbb{R}^d \times [0,T] \to \mathbb{R}^d,\ (t,x,s) \mapsto \bar{X}_s^{t,x}$$

   is almost surely continuous.

3. *For all $(t,x) \in [0,T] \times \mathbb{R}^d$ and any $[t,T]$-valued $\bar{\mathcal{F}}^{t,x}$-stopping time τ,*

   $$\bar{X}_{\tau+s}^{\tau,\bar{X}_\tau^{t,x}} 1_{\{\tau+s\leq T\}} = \bar{X}_{\tau+s}^{t,x} 1_{\{\tau+s\leq T\}},\ s \in [0,\infty)\,.$$

4. *For all $(t,x) \in [0,T] \times \mathbb{R}^d$ and $s \in [0,\infty)$, any $[t,T]$-valued $\bar{\mathcal{F}}^{t,x}$-stopping time τ, and any bounded measurable function $h : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$,*

   $$\mathbb{E}\left[h(\tau+s, \bar{X}_{\tau+s}^{t,x}) 1_{\{\tau+s\leq T\}} \middle| \bar{\mathcal{F}}_\tau^{t,x}\right] = \mathbb{E}\left[h\left(\tau+s, \bar{X}_{\tau+s}^{t,x}\right) \middle| \left(\tau, \bar{X}_\tau^{t,x}\right)\right]$$

Remark 10. Essentially, the first point is due to the a.s. continuity from Brownian motions. The second point states that every stopping time can be considered a starting point of the mechanism. The third point is to confirm that this process is a strong Markov process, with which we can deduce recurrence and transience of states if we want.

These are very nice properties and exactly describes the diffusion we need for branching diffusion. However, this is not what we need for branching diffusion with jumps because this description does not consider the jumps.

More precisely, this is what we need to consider:

$$\begin{aligned} X^{(k)}_{T^{(k-)}} &= \Gamma_{\mathcal{J}^{(k)}}\left(T^{(k-)}, X^{(k-)}_{T^{(k-)}}, \Delta^{(k-)}\right) \\ dX^{(k)}_s &= \mu\left(s, X^{(k)}_s\right) ds + \sigma\left(s, X^{(k)}_s\right) dW^{(k)}_s,\ s \in \left(T^{(k-)}, T^{(k)}\right] \end{aligned}$$

where we can see the starting point of each diffusion is the point after the jump, which is selected according to the jump type $\mathcal{J}^{(k)}$.

4.3 Branching Diffusion Representation

Now we present the main result. Notations:

1. Let
$$\mathcal{F}^n := \sigma\left(W^{(k)}, \tau^{(k)}, \Delta^{(k)}, I^{(k)} : k \in \mathcal{N}_n\right)$$
for $n \in \mathbb{N}$. This includes all the information up to the n-th generation.
2. We also write
$$\mathcal{F}^0 := \{\emptyset, \Omega\}$$
for the trivial σ-algebra.
3. Define
$$\mathcal{B}^n := \mathcal{F}^n \vee \sigma\left(\tau^{(k)} : k \in \mathcal{N}_{n+1}\right).$$
to denote the information up to n-th generation with an additional future split time.

Remark 11. A key observation here is that for $n \in \mathbb{N}$ and $k \in \bar{\mathcal{K}}^n_t$,

$$\text{Conditional on } \mathcal{B}^{n-1}\text{, the laws of } X^{(k)} \text{ and } \bar{X}^{T^{(k-)}, X^{(k)}_{T^{(k-)}}} \text{ on } \left[T^{(k-)}, T^{(k)}\right] \text{ are identical,}$$

but that they are driven by independent noises.

Assumption 3. The function $\psi : \mathbb{R}^d \to \mathbb{R}$ and $c_I : [0,T] \times \mathbb{R}^d \times \Xi \times \to \mathbb{R}$ are bounded and measurable.

Definition 4.3.1 (Classical Solution)**.** A continuous function

$$u : [0,T] \times \mathbb{R}^d \to \mathbb{R},\ (t,x) \mapsto u(t,x)$$

is a classical solution of (PDE) with (BC) if

1. $u \in \mathcal{C}^{1,2}\left([0,T) \times \mathbb{R}^d\right)$
2. for each $(t,x) \in [0,T] \times \mathbb{R}^d$,
$$\sum_{I \in \mathcal{I}} \int_{\Xi} \left| c_I(t,x,\xi) \mathcal{J}[u](t,x,xi)^I \right| \gamma(d\xi) < \infty$$
3. u satisfies (PDE) for each $(t,x) \in [0,T) \times \mathbb{R}^d$ and (BC) for each $x \in \mathbb{R}^d$.

We claim the result to be

$$\mathbb{E}\left[\Psi^{t,x}\right], \text{ where } \Psi^{t,x} := \lim_{n\to\infty} \Psi^{t,x}_n := \lim_{n\to\infty} \prod_{k \in \mathcal{K}^n} \frac{\psi(X^{(k)}_T)}{F(T - T^{(k-)})} \prod_{k \in \bar{\mathcal{K}}^n \setminus \mathcal{K}^n} \frac{c_I(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)})}{\rho(T^{(k)} - T^{(k-)}) p_I}$$

Proof. We only give the outline of the proof here, since the method is similar to what was presented in Pierre's.

It can be proven that, under the condition of our following theorem, the Feynman-Kac Formula can be expressed in this way:

$$\begin{aligned} u(T^{(k-)}, X^{(k)}_{T^{(k-)}}) &= \mathbb{E}\left[\psi(\bar{X}_T) + \int_{T^{(k-)}}^{T} f\left(r, \bar{X}_r, \Delta, \mathcal{J}[u]\left(r, \bar{X}_r, \Delta\right)\right) dr \middle| \mathcal{F}^{n-1} \right] \\ &= \mathbb{E}\left[1_{\{T^{(k)}=T\}} \frac{\psi\left(\bar{X}_T\right)}{F(T-T^{(k-)})} + 1_{\{T^{(k)}<T\}} \frac{c_I(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta)}{\rho(T^{(k)}-T^{(k-)})p_I} \mathcal{J}[u](T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta)^I \middle| \mathcal{F}^{n-1} \right] \end{aligned}$$

Then by iterating the second term, we obtain the results. □

Remark 12. Intuitively, this form coheres with the attempt to iterate all cases in accordance to the distribution of the split time $T^{(\cdot)}$. The first term refers to the case where the particle survives til the end whilst the latter undergoes splits and jumps.

For the technicality, we need one more definition.

Definition 4.3.2 (Uniform Integrable)**.** A collection of random variable $\mathcal{F}$ is *uniformly integrable* if

$$\lim_{K\to\infty} \sup\{\mathbb{E}[1_{|X|>K}|X|] : X \in \mathcal{F}\} = 0$$

We now state the result in theorem.

Theorem 8 (Branching Representation of Classical Solutions)**.** *With the assumptions made before, let* $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ *be a classical solution of (PDE) with boundary condition (BC) and fix* $(t,x) \in [0,T] \times \mathbb{R}^d$*. For each* $n \in \mathbb{N}_0$*, iteratively define*

$$\begin{aligned} \mathcal{G}^{t,x}_0 &= \mathcal{C}^{t,x}_0 = 1 \\ \mathcal{G}^{t,x}_n &= \mathcal{G}^{t,x}_{n-1} \prod_{k\in\mathcal{K}^n} \frac{g(X^{(k)}_T)}{F(T-T^{(k-)})} \\ \mathcal{C}^{t,x}_n &= \mathcal{C}^{t,x}_{n-1} \prod_{k\in\bar{\mathcal{K}}^n\setminus\mathcal{K}^n} \frac{c_{I^{(k)}}(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)})}{\rho(T^{(k)}-T^{T^{(k-)}})p_{I^{(k)}}} \\ \mathcal{R}^{t,x}_n &= \prod_{k\in\mathcal{K}^{n+1}} \frac{\psi(X^{(k)}_T)}{F(T-T^{(k-)})} \times \prod_{k\in\bar{\mathcal{K}}^{n+1}\setminus\mathcal{K}^{n+1}} \frac{c_{I^{(k)}}(T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)})\mathcal{J}[u](T^{(k)}, X^{(k)}_{T^{(k)}}, \Delta^{(k)})^{I^{(k)}}}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}} \\ \Psi^{t,x}_n &= \mathcal{G}^{t,x}_n \mathcal{C}^{t,x}_n \mathcal{R}^{t,x}_n \text{ and } \Psi^{t,x} = \lim_{n\to\infty} \Psi^{t,x}_n \end{aligned}$$

Suppose that

1. $\{\Psi^{t,x}_n\}$ *is uniformly integrable*
2. *for every* $(s,y) \in [t,T] \times \mathbb{R}^d$,

$$\sum_{I\in\mathcal{I}} \mathbb{E}\left[\int_s^T \int_\Xi \left|c_I(r, \bar{X}^{s,y}_r, \xi)\mathcal{J}[u](r, \bar{X}^{s,y}_r, \xi)^i\right| \gamma(d\xi)dr\right] < \infty$$

3. *for every* $(s,y) \in [t,T] \times \mathbb{R}^d$*, the local martingale*

$$M^{s,y}_\cdot := \int_s^\cdot D_x u(r, \bar{X}^{s,y}_r)\sigma(r, \bar{X}^{s,y}_r)d\bar{W}_r$$

is a martingale.

Then, $\Psi^{t,x}$ is integrable and u admits the branching diffusion representation

$$u(t,x) = \mathbb{E}\left[\Psi^{t,x}\right]$$

Remark 13. The formulation of $\Psi^{t,x}$ looks daunting but actually intuitive. This continues our previous remark on the intuitive understanding of the result. The crux then lies with how to iterate the second term, which has a lot to do with the term $\mathcal{R}_n^{t,x}$.

4.3.1 Sufficient Condition for Uniform Integrability

For the first condition in the theorem, we provide some sufficient condition.

Theorem 9 (Integrability Condition)**.** *Let $\mathcal{K} \in (1,\infty)$, define*

$$C_1 = \frac{\|\psi\|_\infty^{\mathcal{K}}}{F(T)^{\mathcal{K}-1}} \text{ and } C_2 = \sup_{I\in\mathcal{I}, t\in[0,T]} \left(\frac{\|c_I\|_\infty}{\rho(t)p_I}\right)^{\mathcal{K}}.$$

Then, the family $\{\Psi^{t,x}\}_{(t,x)\in[0,T]\times\mathbb{R}^d}$ is bounded in $L^{\mathbb{K}}(\mathbb{P})$ and in particular uniformly integrable in either of the following cases:

1. $\frac{C_1}{F(T)} \vee C_2^{\frac{\mathcal{K}}{\mathcal{K}-1}} \leq 1$.
2. $\sum_{I\in\mathcal{I}} \|c_I\|_\infty x^{|I|}$ *is nonzero and has an infinite radius of convergence, and the mature time $T>0$ satisfies*

$$T < \int_{C_1}^{\infty} \left(C_2 \sum_{I\in\mathcal{I}} \|c_I\|_\infty x^{|I|}\right)^{-1} dx.$$

Proof. For the first point, the proof is straightforward:

$$\mathbb{E}\left[\left|\Psi^{t,x}\right|^{\mathcal{K}}\right] \leq \mathbb{E}\left[\prod_{k\in\mathcal{K}} \frac{C_1}{F(T)} \prod_{k\in\bar{\mathcal{K}}\setminus\mathcal{K}} C_2^{\frac{\mathcal{K}}{\mathcal{K}-1}}\right] \leq \mathbb{E}\left[\prod_{k\in\bar{\mathcal{K}}} \frac{C_1}{F(T)} \vee C_2^{\frac{\mathcal{K}}{\mathcal{K}-1}}\right] \leq 1.$$

For the second condition, the main idea is to treat an upper bound of the estimator as a solution of another branching diffusion. Namely,

$$\mathbb{E}\left[\left|\Psi^{t,x}\right|^{\mathcal{K}}\right] \leq \mathbb{E}\left[\prod_{k\in\mathcal{K}} \frac{C_1}{F(T-T^{(k-)})} \prod_{k\in\bar{\mathcal{K}}\setminus\mathcal{K}} \frac{C_2\|c_{I^{(k)}}\|_\infty}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}}\right]$$

This coheres with the ODE:

$$\dot{\eta}(t) + C_2 \sum_{I\in\mathcal{I}} \|c_I\|_\infty \eta(t)^{|I|} = 0 \text{ for } t\in[0,T]$$

$$\eta(T) = C_1$$

The second condition ensures the existence of such a solution η on $[0,T]$. With this, we can define

$$\mathcal{X}_n := \prod_{k\in\bigcup_{m=1}^n \mathcal{K}^m} \frac{C_1}{F(T-T^{(k-)})} \prod_{k\in\bigcup_{m=1}^n \bar{\mathcal{K}}^m\setminus\mathcal{K}^m} \frac{C_2\|c_{I^{(k)}}\|_\infty}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}} \prod_{k\in\bar{\mathcal{K}}^{n+1}} \eta(T^{(k-)})$$

and

$$\mathcal{X}_\infty := \lim_{n\to\infty} \mathcal{X}_n := \prod_{k\in\mathcal{K}} \frac{C_1}{F(T-T^{(k-)})} \prod_{k\in\bar{\mathcal{K}}\setminus\mathcal{K}} \frac{C_2\|c_{I^{(k)}}\|_\infty}{\rho(T^{(k)}-T^{(k-)})p_{I^{(k)}}}$$

Notice that their existences come from our main theorem in branching diffusion. Then, by the similar argument in the proof of theorem 8, and by Fatou's lemma,

$$\mathbb{E}\left[\left|\Psi^{t,x}\right|^{\mathcal{K}}\right] \le \mathbb{E}\left[\mathcal{X}_\infty\right] \le \liminf_{n\to\infty} \mathbb{E}\left[\mathcal{X}_n\right] \le \eta(t) \le \sup_{t\in[0,T]} \eta(t) < \infty$$

□

4.4 Simulation Results

Here, as a verification of our MATLAB code, we replicated the example presented in Belak.

$$\partial_t u(t,x) + \frac{1}{2d^2}\Delta u(t,x) + \int_\Xi \sum_{I\in\mathcal{I}} c_I(t,x)u(t,x)^{I_1}u(t,x+\xi)^{I_2}\gamma(d\xi) = 0$$
$$u(T,x) = \sin(1_d^T x)$$

γ is the uniform distribution on $\{-(\pi/2)e_i \in \mathbb{R}^d : i \in [1:d]\}$. The final time $T=1$. $\mathcal{I} := \{I \in \mathbb{N}_0^2 : |I| \le 2\}$. The coefficient functions are:

$$\begin{aligned}
c_{0,0}(t,x) &= (\alpha + 1/(2d))cos(1_d^T x)exp(\alpha(T-t)) + \cos(1_d^T x)^2/d - 1/(2d)\\
c_{1,0}(t,x) &= (-1)\cdot\cos(1_d^T x)exp(-\alpha(T-t))/d\\
c_{0,1}(t,x) &= (-1)\cdot c_{1,0}(t,x)\\
c_{2,0}(t,x) &= c_{0,2}(t,x) = exp(-2\alpha(T-t))/(2d)\\
c_{1,1}(t,x) &= (-2)\cdot c_{2,0}(t,x)
\end{aligned}$$

where $\alpha = 0.2$. The analytic solution is

$$u(t,x) = \cos(1_d^T x)e^{\alpha(T-t)} = \cos(\sum_{i=1}^d e^{\alpha(T-t)}) \text{ for } (t,x)\in[0,T]\times\mathbb{R}^d$$

Here we used $\Gamma(0.5, 2.5)$ as the split distribution.

p_I	value
$p_{0,0}$	1/3
$p_{1,0}, p_{0,1}, p_{1,1}$	1/6
$p_{2,0}, p_{0,2}$	1/12

Table 4.1: The distribution of offspring

The results are as follows:

4.5 Viscosity Solution

We have been covering the branching diffusion representation of the classical solution to a PDE. The converse can also be studied, with the obvious short conclusion that not every branching diffusion with jumps can be a solution

Number of trials	$d=3$	$d=5$	$d=10$	$d=50$
10^3	-1.14 (0.0031)	0.26934 (0.00213)	-0.94085 (0.0016952)	1.14 (0.0016324)
10^5	-1.1568 (0.015041)	0.26995 (8.54E-05)	-0.94231 (0.00014353)	1.1415 (0.00014896)
10^7	-1.1453 (5.99E-06)	0.26926 (1.11E-06)	-0.9425 (1.69E-06)	1.1415 (1.52E-06)

Table 4.2: Simulation results with their corresponding variance (in brackets).

to a PDE – the case where the regularity condition on uniform integrability is not fulfilled.

To further study this, we introduce *Viscosity Solution*.

Definition 4.5.1 (Viscosity Solution)**.** Suppose that ψ an c_I are bounded and measurable. Let $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be a continuous function such that

$$\sum_{I \in \mathcal{I}} \int_{\Xi} \left| c_I(t,x,\xi) \mathcal{J}\left[u\right](t,x,\xi)^I \right| \gamma(d\xi) < \infty \text{ for all } (t,x) \in [0,T) \times \mathbb{R}^d$$

1. u is called a *viscosity subsolution* of the PDE if for all $(t,x) \in [0,T) \times \mathbb{R}^d$ and all test functions $\phi \in \mathcal{C}^{1,2}([0,T) \times \mathbb{R}^d)$ with $\phi(t,x) = u(t,x)$ and $\phi \geq u$, we have

$$-\partial_t \phi(t,x) - \mathcal{J}[\phi](t,x) - \int_{\Xi} f(t,x,\xi,\mathcal{J}[u](t,x,\xi))\gamma(d\xi) \leq 0$$

2. u is called a *viscosity supersolution* of the PDE if for all $(t,x) \in [0,T) \times \mathbb{R}^d$ and all test functions $\phi \in \mathcal{C}^{1,2}([0,T) \times \mathbb{R}^d)$ with $\phi(t,x) = u(t,x)$ and $\phi \geq u$, we have

$$-\partial_t \phi(t,x) - \mathcal{J}[\phi](t,x) - \int_{\Xi} f(t,x,\xi,\mathcal{J}[u](t,x,\xi))\gamma(d\xi) \geq 0$$

3. u is called a *viscosity solution* of PDE if it is both a viscosity sub- and super-solution.

Theorem 10 (Viscosity Property of the Branching Representation)**.** *For all $(t,x) \in [0,T] \times \mathbb{R}^d$, let $\Psi^{t,x}$ be defined as in theorem 8, and define*

$$u : [0,T] \times \mathbb{R}^d \to \mathbb{R},\ (t,x) \mapsto u(t,x) := \mathbb{E}\left[\Psi^{t,x}\right].$$

Along with assumptions 1 and 2, assume that $\mu, \sigma, c_I, I \in \mathcal{I}, \psi$ are all continuous. Furthermore, for $(t,x) \in [0,T] \times \mathbb{R}^d$

1. *there exists $\epsilon > 0$ such that $\{\Psi^{s,y}\}_{(s,y) \in \mathcal{B}_\epsilon(t,x)}$ is uniformly integrable.*

2. *there exists $\delta > 0$ and a measurable function $\zeta : \mathcal{I} \times \mathbb{R}^d \to \mathbb{R}$ such that*

$$\left| c_I(t,x,\xi) \mathcal{J}\left[u\right](s,y,\xi)^{|I|} \right| \leq \zeta(I,\xi) \text{ for all } (s,y) \in \bar{\mathcal{B}}_\delta(t,x),$$

with

$$\sum_{I \in \mathcal{I}} \int_{\Xi} \left|\zeta(I,\xi)\right| \gamma(d\xi) < \infty$$

3. *It holds that*

$$\sum_{I \in \mathcal{I}} \mathbb{E}\left[\int_t^T \int_{\Xi} \left| c_I(s, \bar{X}_s^{t,x}, \xi) \mathcal{J}\left[u\right](x, \bar{X}_s^{t,x}, \xi)^I \right| \gamma(d\xi)\right] < \infty$$

Then u is a viscosity solution to the PDE.

Proof. The proof is referred to Belak. ☐

Chapter 5

Extension on the Moments

5.1 Moments

In this subsection, we give a slight modification. We explore on the variance of the estimator $\Psi^{t,x}$, which for simplicity we denote as

$$\Psi^{t,x} = \prod_{k\in\mathcal{K}} \frac{\psi(X_T^{(k)})}{F(T-T^{(k-)})} \prod_{k\in\bar{\mathcal{K}}\setminus\mathcal{K}} \frac{c_I(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)})}{\rho(T^{(k)}-T^{(k-)})p_I}$$

.

Then,

$$\Psi^2 = \prod_{k\in\mathcal{K}} \frac{\psi(X_T^{(k)})^2}{F(T-T^{(k)})^2} \prod_{k\in\bar{\mathcal{K}}\setminus\mathcal{K}} \frac{c_I(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)})^2}{\rho(T^{(k)}-T^{(k-)})^2 p_I^2}$$

. We want to find the corresponding PDE for this variance.

The solution is difficult for a general distribution γ, but surprisingly easy in the Markovian case, which we present here.

Imagine we are re-defining a distribution $\tilde{F}$ of the split time F which has the following properties:

1. There exists a continuous density $\tilde{\rho} : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ of $\tilde{F}$ such that

$$\tilde{F}(t) = \int_t^\infty \tilde{\rho}(s)ds, \, \forall t > 0$$

2. $\forall t > 0, \tilde{F}(t) = F(t)^2$.

3. $\tilde{F}(0) = 1$.

Now we claim that

$$\tilde{\rho}(t) = 2F(t)\rho(t), \, t > 0$$

would satisfy these conditions.

Proof. For (1), since F and ρ are continuous on $\mathbb{R}_+$, so is $\tilde{\rho}$.
For (2), by integration by part and the Fundamental Theorem of Calculus,

$$d\tilde{F} = -\tilde{\rho}(t)dt$$

. Then, immediately we obtained,

$$\begin{aligned}\tilde{F}(t) &= \int_t^\infty \tilde{\rho}(s)ds \\ &= \int_t^\infty 2F(s)\rho(s)ds \\ &= -2\ F(s)^2\big|_t^\infty - \int_t^\infty 2F(s)\rho(s)ds \\ &= F(t)^2\end{aligned}$$

(3) then follows from (2) and the properties of F. □

Also, we need another set of probabilities $\{\tilde{p}_I\}$ and boundary condition $\tilde{\psi}$. Define

$$\tilde{p}_I := \frac{p_I^2}{\sum p_I^2}$$

, for any I and

$$\tilde{\psi}(x) = \psi(x)^2$$

. Obviously $\{\tilde{p}_I\}$ is a probability measure since I is finite.

Now, substituting all these into the variance expression, assuming that they are "well-behaving" (leave it for later), we obtained

$$\mathbb{E}\left[\Psi^2\right] = \mathbb{E}\left[\prod_{k\in\mathcal{K}} \frac{\tilde{\psi}(X_T^{(k)})}{\tilde{F}(T-T^{(k)})} \prod_{k\in\bar{\mathcal{K}}\setminus\mathcal{K}} \left(\frac{c_I(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)})}{\tilde{\rho}(T^{(k)}-T^{(k-)})\tilde{p}_I} \cdot \frac{2F(T^{(k)}-T^{(k-)})}{\rho(T^{(k)}-T^{(k-)})(\sum p_I^2)}\right)\right]$$

As we want to construct a new $\tilde{c_I}$ term that is independent of the number of branches but the jump types. So everything looks nice, except for the term $\frac{2F}{\rho}$.

The ideal case is for it to be a constant. Suppose that is the case, for $t > 0$, let

$$a := \frac{2F(t)}{\rho(t)} > 0$$

. If we further assume ρ to be differentiable on $\mathbb{R}_+$,

$$\begin{aligned}F(t) &= \frac{a}{2}\rho(t) \\ \rho(t) &= -\frac{a}{2}\rho'(t) \\ \rho(t) &= A\exp\{-\frac{2}{a}t\}\end{aligned}$$

where A is a constant here. From $\tilde{F}(0) = 1$, we have $A = 2/a$.
This means that in this ideal case, F has to be an exponential distribution for an arbitrary parameter $\frac{2}{a} > 0$.

If $F \sim \text{Exp}(\alpha)$, then we are done by letting a "well-behaving"

$$\tilde{c_I} := \frac{2c_I^2}{\alpha\sum p_I^2}$$

In fact, as long as $F \sim \text{Exp}(\alpha)$ and is "well-behaving", it works on Ψ^n for an arbitrary n, with

$$\begin{aligned}
\tilde{p}_I &= \frac{p_I^n}{\sum p_I^n} \\
\tilde{\rho}(t) &= nF(t)^{n-1}\rho(t) \text{ for } t \in [0,\infty) \\
\tilde{c_I} &= \frac{nc_I^n}{a^{n-1}\sum p_I^n}
\end{aligned}$$

Proof. The arguments are the direct treatment for the case $n = 2$. □

In fact, n does not need to be an integer. The whole argument works for any positive n. For negative n, all the integrability conditions need to be taken care, which we should consider a topic different from here.

We are going to mainly focus on this case: $F \sim \text{Exp}(\alpha)$.

5.2 Convergence Issues

Here two issues arise.

1. What "well-behaving" properties are required?
2. Does it work for arbitrary $\alpha > 0$, the parameter of the exponential distribution?

For the first question, Belak has already answered in Theorem 11. But as we can see, the uniform integrability condition becomes stronger:

$$\left\{ \Psi^n := \prod_{k\in\mathcal{K}} \frac{\psi(X_T^{(k)})^n}{F(T)^n} \prod_{k\in\bar{\mathcal{C}}\setminus\mathcal{K}} \frac{c_I(T^{(k)}, X_{T^{(k)}}^{(k)}, \Delta^{(k)})^n}{\rho(T^{(k)} - T^{(k-)})^n p_I^n} \right\} \text{ is uniformly integrable.}$$

For the second question, it can be proven that, under a certain condition, uniform integrability can be fulfilled by exponential distribution.

For this, we need the following proposition:

Proposition 2. *If* $\sup\{\mathbb{E}[|X|^p] : X \in \mathcal{F}\} < \infty$ *for some* $p > 1$*, then* $\mathcal{F}$ *is uniformly integrable.*

Now we can look at the question for exponential distribution.

Suppose $F(t) := \int_t^\infty \alpha e^{-\alpha s} ds = e^{-\alpha t}$. Then,

$$\begin{aligned}
C_1 &= \frac{\|\psi\|^{\mathcal{K}}}{e^{-\alpha(\mathcal{K}-1)T}} = \|\psi\|^{\mathcal{K}} e^{\alpha(\mathcal{K}-1)T} \\
C_2 &= \sup_{I\in\mathcal{I}, t\in[0,T]} \left(\frac{\|c_I\|}{\alpha e^{-\alpha t} p_I}\right)^{\mathcal{K}-1} = \sup_{I\in\mathcal{I}} \left(\frac{\|c_I\|}{p_I}\right)^{\mathcal{K}-1} \frac{e^{\alpha(\mathcal{K}-1)T}}{\alpha^{\mathcal{K}-1}}
\end{aligned}$$

With this, we have the following proposition especially for exponential split times.

Proposition 3. *If* $\|\psi\|_\infty \leq \frac{1}{e}$*, further suppose that*

$$T \leq \frac{1}{e \sup_{I\in\mathcal{I}} \left(\frac{\|c_I\|_\infty}{p_I}\right)},$$

then, there exists some $\alpha > 0$ such that with $\rho(t) = \exp(-\alpha t)$ on $[0, T]$.
In particular, $\alpha = 1/T$ is one such parameter of the distribution.

Proof. We based our proposition on the first sufficient condition for uniform integrability provided by Belak.

$$\begin{aligned}
C_1 &= \frac{\|\psi\|_\infty}{e^{-\alpha T}} = \|\psi\|_\infty e^{\alpha T} \leq 1 \text{ if } \alpha = 1/T \\
C_2 &= \sup_{I\in\mathcal{I}} \frac{\|c_I\|_\infty}{\alpha e^{-\alpha T} p_I} \\
&= \sup_{I\in\mathcal{I}} \left(\frac{\|c_I\|_\infty}{p_I}\right) \frac{e^{\alpha T}}{\alpha} \\
&\leq \sup_{I\in\mathcal{I}} \left(\frac{\|c_I\|_\infty}{p_I}\right) \frac{e}{1/T} \\
&\leq 1
\end{aligned}$$

If the assumptions here hold, then the first condition in the sufficient conditions holds automatically for any $\mathcal{K} > 1$. In fact, this is logically equivalent to the first condition, since it also holds for other $\mathcal{K} \in (1, \infty)$. □

In the case that $\|\psi\|_\infty > 1/e$, we can re-scale the whole PDE as the following:

$$\begin{cases} \partial_t u + \mathcal{L}[u] + \int_\Xi \sum_{I\in\mathcal{I}} c_I(t,x,\xi)\mathcal{J}[u]^{|I|}\gamma(d\xi) = 0 \\ u(T,x) = \frac{\psi(x)}{\|\psi\|_\infty e} \end{cases}$$

Then we undergo the same procedures to attain the required results.

5.3 Simulation Results

As an example, we consider the following:

$$\partial_t u(t,x) + \partial_x u(t,x) + \frac{1}{d}\partial_{xx}u(t,x) + \int_\Xi \sum_{I\in\mathcal{I}} c_I(t,x)u(t,x)^{I_1}u(t,x+\xi)^{I_2}\gamma(d\xi) = 0$$
$$u(T,x) = \frac{1}{e}\sin(1_d^T x)$$

γ is the uniform distribution on $\{-(\pi/2)e_i \in \mathbb{R}^d : i \in [1:d]\}$. The final time $T = 1$. $\mathcal{I} := \{I \in \mathbb{N}_0^2 : |I| \leq 3\}$.

$$c_I(t,x) = \frac{1}{e\sqrt{2}}\left(\sin(1_d^T x)^{i_1} + \cos(1_d^T x)^{i_2}\right) \text{ where } I = (i_1, i_2).$$

It is rather straight-forward to verify that the above setting fulfills the hypotheses in proposition 3.

Similar results are obtained by running 100 and 1000 times, as shown in the appendices. The heat map shows the light color (values < 1) and dark color (values > 1). For those with light color, the estimator converges. We see a general trend of the inverse proportion between the parameter α and the mature time T. In particular, the claim that at least $\alpha = 1/T$ is an appropriate parameter is verified here.

p_I	value
$p_{0,0}$	$1/15$
$p_{1,0}, p_{0,1}$	$1/10$
$p_{2,0}, p_{0,2}, p_{1,1}$	$2/15$
$p_{3,0}, p_{0,3}, p_{1,2}, p_{2,1}$	$1/6$

Table 5.1: The distribution of offspring

T \ α	5.00E-01	1.00E+00	1.50E+00	2.00E+00	2.50E+00	3.00E+00	3.50E+00	4.00E+00	4.50E+00	5.00E+00
1.00E-01	0.00021282	0.00058647	0.00044561	0.00053339	0.00041079	0.001414	0.0012617	0.00043554	0.0015448	0.0015259
3.00E-01	0.015119	0.0015558	0.0042359	0.0099504	0.0052902	0.0045203	0.010243	0.023746	0.0121	0.023013
5.00E-01	0.010719	0.0022155	0.00662	0.0054956	0.013039	0.026088	0.026719	0.052311	0.092697	0.20813
7.00E-01	0.0010826	0.042524	0.014484	0.016787	0.04551	0.24066	0.48361	0.64811	1.0834	1.7821
9.00E-01	0.018204	0.018684	0.041274	0.13342	0.19105	0.35655	0.61052	37.8833	27.1071	29.0105
1.10E+00	0.022457	0.014448	0.027732	0.11315	0.46625	2.2168	13.7775	58.3867	30109.4757	33650.209
1.30E+00	0.09522	0.10352	0.080871	0.21738	3.1019	4564.1811	4292.2111	245033.186	24980.773	2.1274E+11
1.50E+00	0.55243	0.070045	0.28931	1.6467	14.7107	8502.5831	725279.689	4002737.34	1079457.99	1.4395E+11
1.70E+00	0.053617	0.045456	0.1321	92.0496	101.9443	167300.324	5.6046E+11	6.3998E+15	7667408263	1.7588E+23
1.90E+00	0.024667	0.0098902	0.63975	52.9298	9377.7222	1363747724	2.543E+10	2154038206	1.2505E+13	7.4012E+10
2.10E+00	0.50005	0.098402	1.2932	42.5151	10428872.1	1.9889E+10	3.1503E+12	6.6174E+12	1.2361E+16	1.699E+19
2.30E+00	0.020367	0.24372	11.014	3886.4889	630305492	2.3017E+22	7.8998E+25	8.0913E+21	1.3201E+25	2.3652E+31
2.50E+00	0.051613	0.26294	29.3743	246817584	7.3651E+11	6.2806E+18	2.1001E+26	5.4812E+25	9.4074E+25	2.5243E+27
2.70E+00	0.017136	0.40495	10.7692	291477.974	9.0416E+13	4.9689E+25	5.052E+30	1.8281E+33	1.9894E+34	8.2478E+83
2.90E+00	0.11915	0.1026	2052308.54	387420310	7.2007E+14	1.9704E+26	3.5596E+55	9.0775E+50	2.4488E+39	2.2552E+35

Figure 5.1: The variance of the simulation results with the corresponding final time (T) and the parameter (α) of the exponential distribution, running 10 trials

As the number of trials progresses, the differentiation becomes bigger, since the power of the estimator (namely the number of particles alive at maturity) is proportional to the time as we have seen from the theory of the Galton-Watson Processes.

5.4 Conclusion

Most of this article has been devoted to dissecting the current progress on building and applying stochastic processes to numerical methods to solving a certain type of partial differential equations. The origin of the geniuses underneath comes from the physical interpretation of the heat equation.

Through evolving and modifying the model accordingly, we have witnessed the extension of the type of PDEs applicable for the approach. Our slight contribution mainly surrounds the generalisation of the branching diffusion processes with jumps. We have also discovered the power of the Markov property in this context.

Our results, though a simple extension, gives a rather strict constraint to the class of PDEs. One main difficulty is the existence of the solution, which has a lot to do with uniform integrability. When studying the upper bound of the estimator, we can transform the problem to one in the Multi-type Continuous time Branching Processes. We attempted discretizing the process using the method in chapter 3 and studied the moments of the estimator. During the journey, we have discovered the power of Markov property: under a certain condition, we can grasp the asymptotic behaviors of the processes and obtain an implicit formula for the Laplace transform of the number of particles of each type. However, such an existence still requires the estimator to be bounded by 1. The main difficulty is that, the distribution of the number of particles at time t is complicated, with which we cannot analytically prove the existence of its generating function on $\mathbb{R}$. Though its asymptotic behaviours are highly studied. One can refer to Arthreya's work. The existence of the estimator beyond $[-1, 1]$ is still an open problem to us.

One natural extension is the variation of the non-linearity. It turns out that the non-linearity dependent of the first derivative of the solution has also been studied. This variation therefore implies that, if the non-linearity is smooth enough, using Taylor's expansion, we can reduce the problem applicable to that version.

Another extension is the variation using Lévy processes. As well-known, uncertainties in a dynamical system can be modelled using Brownian motions or Lévy processes. We also know that the Laplace differential operator in the PDE in this paper is different if we implement the Lévy version. This was studied by Glau.

It is indeed a pity that the existence the generating functions cannot be fully understood over the course of this project. It is indeed worth the time to further study it. From Arthreya's work, we can see that the asymptotic behavior of the multi-type continuous time Markov processes (or even age dependent processes) are understood to some extent. I suspect that this is the right direction to research upon.

Appendix A

Simulation Results

T \ α	5.00E-01	1.00E+00	1.50E+00	2.00E+00	2.50E+00	3.00E+00	3.50E+00	4.00E+00	4.50E+00	5.00E+00
1.00E-01	0.00075191	9.4412E-05	3.8743E-05	6.1228E-05	9.1953E-05	2.8866E-05	3.0549E-05	9.7459E-05	4.9167E-05	8.8711E-05
3.00E-01	0.00051945	0.00044162	0.00042004	0.0005172	0.0005873	0.00054961	0.00084871	0.0016023	0.0017123	0.0017407
5.00E-01	0.003726	0.00061865	0.00049524	0.0010323	0.00255	0.0029844	0.0040208	0.017657	0.024112	0.017926
7.00E-01	0.0041359	0.0015493	0.0023651	0.0048456	0.0036727	0.0074611	0.0079567	0.041887	0.12147	0.29829
9.00E-01	0.012538	0.0014925	0.0039004	0.012815	0.017947	0.062545	0.22595	0.62543	31.4123	7.0532
1.10E+00	0.021923	0.0027717	0.013523	0.014367	0.14448	0.88769	1.8987	213.8373	1668.2515	45967.166
1.30E+00	0.010198	0.0069638	0.0076066	0.039629	0.54199	4.2327	5333.5574	4281.1237	385272.727	74557429.8
1.50E+00	0.009059	0.0068295	0.007478	0.29355	1942.2107	16331.9066	1071738.32	13129757.1	7.142E+10	2.5276E+12
1.70E+00	0.039575	0.0063988	0.064241	1.4614	2149.1321	289046.217	12985333.7	1.2358E+14	2.0335E+13	2.2186E+18
1.90E+00	0.0078467	0.0098945	0.12152	1452.0001	45739.5018	91319615.8	6.2942E+14	2.7305E+15	3.5284E+29	1.3042E+24
2.10E+00	0.027944	0.009294	0.43149	324.6248	24655703.9	2.8326E+17	1.396E+22	1.2684E+33	5.5232E+41	8.0666E+48
2.30E+00	0.018281	0.023608	2.9719	378054011	8.6555E+28	4.4654E+26	1.0968E+33	3.9336E+33	8.0251E+37	1.9689E+41
2.50E+00	0.012061	0.1244	42.0827	5.6166E+11	8.2121E+18	1.6421E+23	3.9468E+29	4.4962E+40	9.3947E+42	1.4352E+73
2.70E+00	0.043796	0.16631	512.7592	166200722	6.4327E+22	2.6562E+28	4.5288E+46	6.7283E+54	7.8692E+83	4.0084E+54
2.90E+00	NaN	0.048918	7528.2625	6.4703E+17	4.9351E+18	9.7728E+51	1.3997E+66	7.6615E+50	2.6293E+70	3.7296E+68

Figure A.1: The variance of the simulation results with the corresponding final time (T) and the parameter (α) of the exponential distribution, running 100 trials

T \ α	5.00E-01	1.00E+00	1.50E+00	2.00E+00	2.50E+00	3.00E+00	3.50E+00	4.00E+00	4.50E+00	5.00E+00
1.00E-01	4.2111E-05	2.936E-06	5.8293E-06	3.0197E-06	4.9516E-06	1.0206E-05	8.1014E-06	9.7012E-06	4.9423E-06	8.4615E-06
3.00E-01	0.00018606	4.5117E-05	2.2302E-05	5.1046E-05	0.00003055	3.2497E-05	4.9917E-05	0.00012046	0.00012495	0.00038891
5.00E-01	0.0001227	5.7763E-05	7.4047E-05	9.8882E-05	0.00018898	0.00042027	0.00024415	0.00040548	0.00081977	0.0020221
7.00E-01	0.00057974	6.7074E-05	0.00022244	0.00093718	0.00053101	0.0010727	0.0051833	0.0049644	0.0066923	0.01947
9.00E-01	0.00078095	0.00023194	0.00060831	0.0010265	0.0011309	0.0094017	0.010333	0.096315	1.1708	158.0136
1.10E+00	0.00049351	0.00022697	0.00049358	0.0019193	0.0064722	7.1477	0.38465	3292.6825	8131.2895	123545.958
1.30E+00	0.00084309	0.00030059	0.0011538	0.0095826	0.055449	62.2171	11938.1812	1234710.49	6907039.37	4034896046
1.50E+00	0.0026884	0.00036426	0.003262	2.6856	227.9056	338645.29	91687760.7	4.8367E+12	3.238E+20	1.3535E+19
1.70E+00	0.0018204	0.0017085	0.0031191	2.195	5525145.74	999951271	3.3173E+15	6.917E+26	8.15E+22	3.3951E+23
1.90E+00	0.0021556	0.00078381	0.030894	526.3768	1830786.79	4.897E+12	3.9126E+22	1.8927E+23	1.0747E+37	2.904E+35
2.10E+00	0.0020331	0.0025668	0.15608	1160.8622	5.7781E+13	5.4948E+23	6.075E+19	6.7956E+31	1.6168E+46	6.6872E+37
2.30E+00	0.0019188	0.001288	0.82436	27216969	4.5508E+15	6.1793E+22	8.4585E+38	5.1365E+43	2.3203E+41	2.4497E+61
2.50E+00	0.0014882	0.0014697	44.7349	1.5216E+10	9.9976E+21	2.7994E+29	1.6106E+43	4.5792E+52	6.9894E+53	5.3993E+59
2.70E+00	NaN	0.008852	430715.515	6.2902E+14	2.0206E+27	2.1229E+40	1.7603E+51	3.424E+59	9.3568E+75	1.7235E+70
2.90E+00	NaN	0.0090804	5565.0884	2.1908E+15	1.823E+35	2.2436E+69	9.9917E+69	1.3948E+65	5.4371E+77	8.1866E+93

Figure A.2: The variance of the simulation results with the corresponding final time (T) and the parameter (α) of the exponential distribution, running 1000 trials

Bibliography

[1] Abraham R., Delmas J.F., "An introduction to Galton-Watson trees and their local limits", 2015.

[2] Arthreya K.B., Ney P.E., "Branching Processes", 1970.

[3] Belak C., Hoffman D., Seifried F.T., "Branching Diffusions with Jumps and Valuation with Systemic Counterparties", 2020.

[4] Evans L.C., "An Introduction to Stochastic Differential Equations", 2013.

[5] Glau K.,"Feynman-Kac formula for Lévy processes with discontinuous killing rate". 2015.

[6] Labordére P.H., "Counterparty Risk Valuation: a Marked Branching Diffusion Approach", 2012.

[7] Labordére P.H., Oudjane N., Tan X., Touzi N., Warwin X, "Branching diffusion representation of semilinear PDEs and Monte Carlo approximation", 2016.

[8] Lawler G.F., "Stochastic Calculus: An Introduction with Applications", 2014.