

Article

Not peer-reviewed version

Fuzzy Fractional Brownian Motion: Review and Extension

[Georgy Urumov](#) , Panagiotis Chountas , [Thierry Chaussalet](#) *

Posted Date: 31 May 2024

doi: 10.20944/preprints202405.2078.v1

Keywords: fuzzy; fractional; Brownian motion



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Fuzzy Fractional Brownian Motion: Review and Extension

Georgy Urumov, Panagiotis Chountas and Thierry Chausalet *

School of Computer Science and Engineering, University of Westminster; 115 New Cavendish Street, London, W1W 6UW, United Kingdom; g.urumov@westminster.ac.uk (G.U.)
p.i.chountas@westminster.ac.uk (P.C.)

* Correspondence: chausst@westminster.ac.uk

Abstract: In traditional finance, option prices are typically calculated using crisp sets of variables. However, in a proposed new approach, the assumption is made that these parameters have a degree of fuzziness or uncertainty associated with them. This allows participants to estimate option prices based on their risk preferences and beliefs, considering a range of possible values for the parameters. To determine the belief degree associated with a particular option price, an interpolation search algorithm was proposed recently[1]. Interpolation is a mathematical technique used to estimate values between known data points. In this case, the algorithm is applied to determine the belief degree associated with a specific option price within the range of possible prices. By incorporating fuzzy numbers and the belief degree, this approach provides a more flexible framework for practitioners to make their investment decisions. It allows them to consider their risk preferences and beliefs about the uncertain parameters when selecting an option price. In this paper, we will review a unified framework for combining fractional Brownian motion with fuzzy processes and obtain a joint measure space. We start by constructing a product measure space that captures the randomness and fuzziness inherent in the respective processes. We begin with the usual Kolmogorov filtered probability space¹ $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ represents the probability measure, Ω is the set of all possible outcomes or states that a random process can take and $\{\mathcal{F}_t, t \in [0, T]\}$ is an information filtration. We will review the work of Merton (1976) on option pricing when the underlying stock returns have jumps [2] and expand it to fuzzy fractional processes.

Keywords: fuzzy; fractional; brownian motion

1. Introduction

The aim of this paper is to review the state-of-the-art methods in fuzzy fractional Brownian motion, complete and extend many novel proofs and theoretical considerations presented in the recent works of the other authors. In future work we will present examples, applications and introduce a novel proof dealing with the constraint of $h \ll \frac{1}{2}$. Traditional option pricing models rely on precise, fixed values for variables like expected return and volatility. However, a new approach recognizes the inherent uncertainty surrounding these parameters, incorporating fuzziness into the calculations. This allows investors to estimate option prices based on their individual risk tolerance and beliefs about the potential range of values for these variables. This paper delves into a unified framework that merges fractional Brownian motion with fuzzy processes, creating a combined measure space. We begin by constructing a product measure space that encompasses both the randomness inherent in fractional Brownian motion and fuzziness of the associated parameters. The underlying asset price S_t is assumed to follow the following dynamics:

$$dS/S = \mu dt + \sigma dB(t) + \sigma dBH(t) + d(\Sigma(V_i - 1)).$$

¹ Kolmogorov space is a mathematical framework used in probability theory to define and study random processes. It consists of three main components: a sample space Ω , a set of events (subsets of Ω), and a probability measure defined on the events.

In this equation, the terms represent various components: (i) the drift parameter μ represents the expected return rate on the value of the underlying asset. (ii) the volatility parameter σ represents the standard deviation of the return rate on the value of the underlying asset. (iii) the term $dB(t)$ represents a standard Brownian motion. (iv) The term $dB^H(t)$ represents an independent standard fractional Brownian motion with respect to a fractional Brownian motion parameterized by the Hurst exponent H . (v) the term $d(\Sigma(V_i-1))$ represents a sum of jumps, where $\{V_i, i \geq 1\}$ is a sequence of independent and identically distributed nonnegative random variables. The logarithm of V_i , denoted as $Y = \log(V)$, follows a normal distribution with a probability density function [2] –

$$f_Y(x) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x-\mu_j)^2}{2\sigma_j^2}}.$$

The expected return rate μ , volatility σ , jump intensity λ , and jump magnitudes sequence V_i in the underlying asset price dynamics cannot always be treated as constant real numbers. Therefore, a new model is proposed to incorporate both randomness and fuzziness in describing the dynamics of the process. When dealing with processes that involve non-standard or non-numeric variables, such as fuzzy numbers or other types of uncertain quantities, the traditional tools of real analysis may not be directly applicable. In such cases, alternative mathematical frameworks or extensions of real analysis are often employed to handle the specific characteristics and uncertainties present in the financial domain. In the next sections, we prove the no arbitrage condition in the presence of uncertain or non-standard variables. This involves extending the mathematical framework to accommodate the specific characteristics of the variables under consideration. Muzzioli et al. [3] conducted a literature review and categorised more than fifty papers on option pricing in a fuzzy setting, providing insights into different approaches and best practices for utilising fuzziness in option pricing. Nowak and Pawłowski [4], [5] studied the pricing problems of European options with the application of Lévy processes under fuzzy environments. Li et al. [6] developed a nonlinear fuzzy-parameter partial differential equation model for obtaining fuzzy option prices and derived dominating optimal hedging strategies under fuzzy environments. Wang et al. [7] studied the pricing problem of n -fold compound options using the pricing model of Nowak and Romaniuk [8], martingale method, and fuzzy sets theory. Zhang et al. [9] discussed Asian option pricing when the stock price, risk-free interest rate, and volatility were trapezoidal fuzzy numbers. Muzzioli et al. [10] used fuzzy quadratic regression methods to estimate the implied volatility smile function. Nowak and Romaniuk [11] (2010) assumed the underlying asset followed a Lévy process and fuzzified the input parameters. They derived option pricing formulas in both crisp and fuzzy cases using minimal entropy martingale measure and fuzzy arithmetic. They also proposed and verified a generalized hybrid binomial American real option pricing model. Lee et al. [12] applied fuzzy decision theory and Bayes' rule to derive a fuzzy Black-Scholes pricing model. They aimed to measure fuzziness in option pricing using fuzzy arithmetic. Wu [13] and Wu [14] developed a fuzzy version of the Black-Scholes pricing formula by considering triangular fuzzy numbers for the underlying asset price, volatility, and risk-free interest rate. Wu [15] also introduced a bisection search algorithm to calculate the membership degree associated with a given option price. Zhang et al. [1] employed stochastics and fuzzy set theory to develop a new framework for option pricing in which the underlying process follows a mixed fuzzy fractional Brownian motion with jump when $h > 1/2$. Using our derivations, we will complete and extend the work of Zhang, Wu, Lee, Nowak and Romaniuk here.

In summary, this paper proposes the utilisation and extension of the mixed fuzzy fBm as a replacement for Brownian motion in stochastic process modelling, specifically focusing on its advantageous properties such as long-range dependence, self-similarity, and stationarity of increments. Furthermore, we complete and extend proofs and theoretical framework for some novel equations, namely the mixed fuzzy stochastic Brownian motion with jumps to address the modelling needs of hybrid systems.

2. Methods

In this section we review the works of [1] to extend Merton's jump diffusion model to include fBm and fuzzy processes and present an alternative variation of the proof. In the fuzzy fractional jump-diffusion model (FFJDM), the underlying process is:

$$d\hat{S}/\hat{S} = \hat{\mu}dt + \hat{\sigma}dB_t + \hat{\sigma}dB_t^H + d\left(\sum_1^{\hat{N}}(\tilde{V}_i - 1)\right).$$

The inclusion of the fuzzy Poisson process $\hat{N}$ with fuzzy intensity λ and $\hat{V}_i$ (sequence of independent and identically distributed fuzzy random variables) makes this a jump-diffusion process, as it allows for random jumps to occur at random times, in addition to the continuous stochastic components represented by the Brownian motion terms. In this model, all the fuzzy parameters are denoted by a hat symbol, such as $\hat{\mu}$ and $\hat{\sigma}$, representing the fuzzy expected return rate and fuzzy volatility, respectively. The model includes additional components such as B_t (standard Brownian motion), B_t^H (independent standard fractional Brownian motion). The subtraction of 1 in the jump process is a convention used to represent the relative change or jump size with respect to the value of the process before the jump occurs. In the context of jump-diffusion models, the jump term is typically expressed as the sum of individual jump sizes, i.e., the magnitude of the jump experienced by the process. The jump size, V_i , represents the absolute change in the process value at the i -th jump. By subtracting 1 from V_i , you are effectively expressing the relative change or percentage change in the process due to the jump. By expressing the jump size as a relative change, the impact of jumps is effectively scaled based on the current value of the process. It allows for a more interpretable representation of jumps, as they are related to the percentage change rather than absolute changes. For example, a jump of $V_i = 0.01$ when the process is at $S = 100$ would represent a 1% increase, whereas the same jump with $V_i = 1$ when the process is at $S = 10$ would represent a 10% increase. By using the relative change, we take into account the current level of the process, which is often more meaningful in modelling and risk assessment. The fuzzy normal distribution is used to describe the fuzziness in the expected return rate and jump magnitudes. The model assumes independence among the sources of randomness and fuzziness, including $B_t, B_t^H, \tilde{N}_t$, and $\tilde{V}_i$. We will now apply Itô's lemma to FFJDM. If we make a Taylor expansion and ignore $(dt)^2$ and $dt dx$ because they are negligible compared to the dt -term.

$$df = (\partial f/\partial t) dt + (\partial f/\partial S) dS + (1/2)(\partial^2 f/\partial S^2)(dS)^2.$$

The first term $(\partial f/\partial t) dt$ is zero as f does not have an explicit time dependence. The second term $(\partial f/\partial \hat{S}) d\hat{S}$ can be computed by taking the derivative of $f(\hat{S}) = \ln(\hat{S})$ with respect to $\hat{S}$: $\partial f/\partial \hat{S} = 1/\hat{S}$. Multiplying this by $d\hat{S}$, we obtain: $(\partial f/\partial \hat{S}) d\hat{S} = (1/\hat{S}) d\hat{S}$. Now, let's consider the third term. Since $(d\hat{S})^2 = 0$ (as $d\hat{S}$ is a differential form), this term can be ignored. Now, substitute the given expression for $d\hat{S}$: $d\hat{S}/\hat{S} = (\hat{\mu}dt + \hat{\sigma}dB_t + \hat{\sigma}dB_t^H + d(\sum_1^{\hat{N}}(\tilde{V}_i - 1)))$. Expanding this expression, we have:

$$d\hat{S}/\hat{S} = \hat{\mu}dt + \hat{\sigma}dB_t + \hat{\sigma}dB_t^H + d\left(\sum_1^{\hat{N}}(\tilde{V}_i - 1)\right)$$

Integrate both sides of the expression we get:

$$\ln(\hat{S}_t) = \ln(\hat{S}(0)) + \int_0^t \hat{\mu}dt + \int_0^t \hat{\sigma}dB_t + \int_0^t \hat{\sigma}dB_t^H + \int_0^t \frac{1}{2} \left\{ -\frac{1}{\hat{S}^2} \right\} \hat{\sigma}^2 \hat{S}^2 dt + \int_0^t \frac{1}{2} \left\{ -\frac{1}{\hat{S}^2} \right\} \hat{\sigma}^2 \hat{S}^2 dt^H + \int_0^t d\left(\sum_1^{\hat{N}}(\tilde{V}_i - 1)\right) dV.$$

We used Itô isometry to show that $(dX)^2 = (\mu dt)^2 + 2\mu\sigma dt dW + \sigma(dW)^2 = \sigma dt$

Now, recall that $f(S) = \ln(S)$, so $\hat{S} = \exp(f(\hat{S}))$. Replacing $\hat{S}$ with $\exp(f(S))$ in the above expression:

$$\ln(\hat{S}_t) = \ln\left(\hat{S}(0)\right) + \hat{\mu}t + \hat{\sigma}B_t + \hat{\sigma}B_t^H - \frac{1}{2}\hat{\sigma}^2 t + \frac{1}{2}\hat{\sigma}^2 t^H + \int_0^t d\left(\sum_1^{\hat{N}}(\tilde{V}_i - 1)\right) dV.$$

Recall that $\tilde{V}_i$ are lognormally distributed such that $E(\tilde{V}_i) = E(e^{\tilde{V}_i}) = e^{\mu+1/2\sigma^2}$ then exponentiating both sides and simplifying:

$$\widehat{S}_t = \widehat{S(0)} e^{\widehat{B}t + \sigma B_t + \widehat{\sigma} dB_t^H - \frac{1}{2}\widehat{\sigma}^2 t + \frac{1}{2}\widehat{\sigma}^2 t^{2H}} \prod_{i=1}^{\widehat{N}} \widehat{V}_i.$$

If we now take an expectation of $\widehat{S}_t$, we have:

$$\mathbb{E}[\widehat{S}_t] = \mathbb{E}[\widehat{S(0)} e^{\widehat{B}t + \sigma B_t + \widehat{\sigma} dB_t^H - \frac{1}{2}\widehat{\sigma}^2 t + \frac{1}{2}\widehat{\sigma}^2 t^{2H}} \prod_{i=1}^{\widehat{N}} \widehat{V}_i].$$

In stochastic calculus, B_t represents a standard Brownian motion and the key property of Brownian motion is that its variance increases linearly with time. The variance of B_t at time t is given by t , which is a result of its definition and properties. The variance of a random variable X is defined as $\text{Var}(X) = E[(X - E[X])^2]$. For a standard Brownian motion B_t , its expected value $E[B_t]$ is 0 at any time t . Therefore, we have: $\text{Var}(B_t) = E[B_t^2] = \int_{-\infty}^{\infty} x^2 P_{B(t)}(x) dx$. Since B_t follows a Gaussian distribution with mean 0 and variance t , its probability density function is the Gaussian density function: $P_{B(t)} = \frac{1}{\sqrt{2\pi t}} \exp(-\frac{x^2}{2t})$. Plug this into the integral for the variance: $\text{Var}(B_t) = E[B_t^2] = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi t}} \exp(-\frac{x^2}{2t}) dx$. The integrand $\frac{1}{\sqrt{2\pi t}} \exp(-\frac{x^2}{2t})$ is the Gaussian p.d.f. with mean 0 and variance t . To integrate the Gaussian p.d.f over the entire real line: $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp(-\frac{x^2}{2t}) dx$, we will use the Gamma function. The integrand is an even function so $\int_{-\infty}^{\infty} e^{-x^2} dx = 2 \int_0^{\infty} e^{-x^2} dx$. Thus after change of variable $x = \sqrt{t}$, we turn this into Gamma function:

$$2 \int_0^{\infty} e^{-x^2} dx = 2 \int_0^{\infty} \frac{1}{2} e^{-t} t^{-1/2} dt = \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \quad (1)$$

Similarly, $\int_{-\infty}^{\infty} \frac{e^{-x^2}}{2} dx = \sqrt{2\pi}$. For standard normal distribution $\mu=0$ and $\sigma^2=1$, and changing variable $x = \sqrt{u}$, and $x^2 = u$ so $du = 2x dx$ then we have: $\int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) dx = \frac{1}{\sqrt{2\pi}} 2 \int_0^{\infty} u \frac{1}{\sqrt{2\pi}} \exp(-\frac{u}{2}) \frac{1}{2\sqrt{u}} du = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \sqrt{u} \frac{1}{\sqrt{2\pi}} \exp(-\frac{u}{2}) du$.

Next, we use integration by parts. Choose $u = \sqrt{u}$ and $dv = \exp(-\frac{u}{2}) du$. Then $du = \frac{1}{2} u^{-1/2}$ and $v = -2 e^{-u/2}$. Applying integration by parts:

$$\frac{1}{\sqrt{2\pi}} \int_0^{\infty} \sqrt{u} \frac{1}{\sqrt{2\pi}} \exp(-\frac{u}{2}) du = \frac{1}{\sqrt{2\pi}} \left[0 \left[-2u^{\frac{1}{2}} e^{-\frac{u}{2}} \right] + \left[\int_0^{\infty} u^{-1/2} e^{-1/2} \right] \right].$$

When x is 0, u is also 0, and as x approaches $+\infty$, u approaches $+\infty$ so the first integral uv evaluates to zero. We are left with $\int_0^{\infty} \sqrt{u} \frac{1}{\sqrt{2\pi}} \exp(-\frac{u}{2}) du = \frac{1}{\sqrt{2\pi}} \left[\int_0^{\infty} u^{-1/2} e^{-1/2} \right]$. To evaluate the remaining integral, we can once again use the integration by parts. Notice that the integral representation of $\int_0^{\infty} u^{-1/2} e^{-1/2}$ is the result we showed earlier in (1) which is the gamma function $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. Therefore, $\frac{1}{\sqrt{2\pi}} \left[\int_0^{\infty} u^{-\frac{1}{2}} e^{-\frac{1}{2}} \right] = \frac{1}{\sqrt{2}}$. Going back to $\mathbb{E}[\widehat{S}_t] = \mathbb{E}[\widehat{S(0)} e^{\widehat{B}t + \sigma B_t + \widehat{\sigma} dB_t^H - \frac{1}{2}\widehat{\sigma}^2 t - \frac{1}{2}\widehat{\sigma}^2 t^{2H}} \prod_{i=1}^{\widehat{N}} \widehat{V}_i]$, we need to evaluate $\mathbb{E}[e^{\sigma B_t + \widehat{\sigma} dB_t^H}]$. The variance of the sum of independent random variables is given by: $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. For $X = \sigma B_t$ and $Y = \widehat{\sigma} dB_t^H$, from Itô's Isometry we have $\mathbb{E}[e^{\sigma B_t}]$ and let $Z(t) = e^{\sigma B_t}$, then using Itô:

$$\begin{aligned} dZ(t) &= \frac{1}{2} \sigma^2 e^{\sigma B(t)} dt + \sigma e^{\sigma B(t)} dB(t) \\ &= \frac{1}{2} \sigma^2 Z(t) dt + \sigma e^{\sigma B(t)} dB(t). \end{aligned}$$

$Z(0) = 1$. Then in integral form: $Z(t) = 1 + \frac{1}{2} \sigma^2 \int_0^t Z(s) ds + \sigma \int_0^t dB(s)$. Taking the expectation will make stochastic integral vanish so the last term is zero. Then $m(t) = E[Z(t)] = 1 + \frac{1}{2} \sigma^2 \int_0^t m(s) ds$. Taking derivatives of both sides we get $\dot{m}(t) = \frac{\sigma^2}{2} m(t)$. This is an ordinary differential equation, so we can separate variables and then integrate both sides. Let's start by rearranging the equation: $m'(t)/m(t) = \sigma^2/2$. Now, we can integrate both sides with respect to t : $\int m'(t)/m(t) dt = \int (\sigma^2/2) dt$. Integrating the left side gives us the natural logarithm of the absolute value of $m(t)$:

$\ln|m(t)| = \left(\frac{\sigma^2}{2}\right) * t + C_1$, where C_1 is the constant of integration. To eliminate the absolute value, we can rewrite the equation as: $m(t) = \pm \exp\left(\left(\frac{\sigma^2}{2}\right)t + C_1\right)$. Now, let's combine the constant terms into a new constant, C_2 : $m(t) = C_2 * \exp\left(\left(\frac{\sigma^2}{2}\right)t\right)$. We know that $m(0) = 1$ so $C_2 = 1$ and we are only interested in positive processes so we can get rid off negative logarithm process. The result is: $E[e^{\sigma B(t)}] = E[Z(t)] = m(t) = e^{\frac{\sigma^2 t}{2}}$. Similarly, $E[e^{\widehat{\sigma} dB_t^H}] = e^{\frac{1}{2}\widehat{\sigma}^2 t^{2H}}$. Therefore, we get equation:

$$\begin{aligned} E[\widehat{S}_t] &= \widehat{S}(0) e^{\widehat{\mu}t + \sigma B_t + \widehat{\sigma} B_t^H - \frac{1}{2}\widehat{\sigma}^2 t - \frac{1}{2}\widehat{\sigma}^2 t^{2H}} \prod_{i=1}^{\widehat{N}} \widehat{V}_i \\ &\widehat{S}(0) e^{\widehat{\mu}t - \frac{1}{2}\widehat{\sigma}^2 t - \frac{1}{2}\widehat{\sigma}^2 t^{2H}} e^{\frac{1}{2}\widehat{\sigma}^2 t + \frac{1}{2}\widehat{\sigma}^2 t^{2H}} \prod_{i=1}^{\widehat{N}} \widehat{V}_i \end{aligned} \quad (2)$$

To evaluate the last term $\prod_{i=1}^{\widehat{N}} \widehat{V}_i$ which is the product of the i.i.d. distributed random jumps with mean μ and variance σ^2 , we will use the property that the expectation of the product of independent random variables is equal to the product of their expectations. Re-writing this expression and using the fact that $\sum \ln(V_i) = \ln(V_1) + \ln(V_2) + \dots + \ln(V_n) = \ln(V_1 * V_2 * \dots * V_n)$, we get $\prod_{i=1}^{\widehat{N}} \widehat{V}_i = e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)}$.

$$\begin{aligned} E[\widehat{S}_t] &= \widehat{S}(0) e^{\widehat{\mu}t - \frac{1}{2}\widehat{\sigma}^2 t - \frac{1}{2}\widehat{\sigma}^2 t^{2H}} e^{\frac{1}{2}\widehat{\sigma}^2 t + \frac{1}{2}\widehat{\sigma}^2 t^{2H}} e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)} = \\ &\widehat{S}_0 e^{\widehat{\mu}t} E[e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)}]. \end{aligned}$$

To evaluate the process $E[e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)}]$, we will use the fact that jump size is i.i.d then by the law of iterated expectation we can re-write the expectation as conditional expectation: $E[e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)}] = E[E[e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)} | \widehat{N}_t]]$. Recall that $\widehat{N}_t$ is Poisson distribution with parameter λ [2]. Because the distributions of the size of the jumps and the Poisson process of the number of jumps are independent, we can simplify the above expression as:

$$E[e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)}] = E[E[e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)} | \widehat{N}_t]] = \sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n}{n!} e^{-\widehat{\lambda}t} E[e^{\sum_{i=1}^n \ln(\widehat{V}_i)}].$$

If you recall, the jump size has lognormal distribution with mean value μ and variance σ^2 [2]. From the property of the normal distribution we know that:

$$\tilde{J} = E(\widehat{V}_i) = E(e^{\widehat{V}_i}) = e^{\mu + 1/2\sigma^2}.$$

So,

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n}{n!} e^{-\widehat{\lambda}t} E[e^{\sum_{i=1}^n \ln(\widehat{V}_i)}] &= \sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n}{n!} e^{-\widehat{\lambda}t} [\prod_{i=1}^n E(\widehat{V}_i)] = \sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n}{n!} e^{-\widehat{\lambda}t} [\prod_{i=1}^n E(\tilde{J})] = \\ \sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n}{n!} e^{-\widehat{\lambda}t} \tilde{J}^n &= \sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n}{n!} e^{-\widehat{\lambda}t} \tilde{J}^n = \sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n \tilde{J}^n}{n!} e^{-\widehat{\lambda}t}. \end{aligned}$$

The first part of this series is just re-written Taylor series $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, so we can simplify $\sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n \tilde{J}^n}{n!} = e^{\widehat{\lambda}t \tilde{J}}$. The whole expression becomes:

$$\sum_{n=0}^{\infty} \frac{(\widehat{\lambda}t)^n}{n!} e^{\widehat{\lambda}t \tilde{J}} E[e^{\sum_{i=1}^n \ln(\widehat{V}_i)}] = e^{\widehat{\lambda}t \tilde{J}} e^{-\widehat{\lambda}t} = e^{-\widehat{\lambda}t(\tilde{J}-1)}.$$

Then:

$$\begin{aligned} E[\widehat{S}_t] &= \widehat{S}(0) e^{\widehat{\mu}t - \frac{1}{2}\widehat{\sigma}^2 t - \frac{1}{2}\widehat{\sigma}^2 t^{2H}} e^{\frac{1}{2}\widehat{\sigma}^2 t + \frac{1}{2}\widehat{\sigma}^2 t^{2H}} e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)} = \widehat{S}_0 e^{\widehat{\mu}t} E[e^{\sum_{i=1}^{\widehat{N}} \ln(\widehat{V}_i)}] = \widehat{S}_0 e^{\widehat{\mu}t} e^{-\widehat{\lambda}t(\tilde{J}-1)}. \\ E[\widehat{S}_t] &= \widehat{S}_0 e^{\widehat{\mu}t} e^{-\widehat{\lambda}t(\tilde{J}-1)} = \widehat{S}_0 e^{\widehat{\mu}t - \widehat{\lambda}t(\tilde{J}-1)} = \widehat{S}_0 e^{t(\widehat{\mu} - \widehat{\lambda}(\tilde{J}-1))} \end{aligned} \quad (3)$$

In the next section we will review the works of [16] and show how we can combine fuzzy processes with fBm.

3. Proofs and Discussion

In the field of stochastic process modelling, the fractional Brownian motion (fBm) is often recommended as a replacement for Brownian motion as the driving process, particularly when studying phenomena characterized by long-range dependence. The fBm, with a Hurst exponent (H) ranging between 0 and 1, is a type of Gaussian process that possesses several advantageous properties, including long-range dependence, self-similarity, and stationarity of increments. Notably, when the Hurst exponent is different from 1/2, the fBm ceases to be a semi-martingale. This paper is somewhat reliant on an earlier work of Doob-Meyer decomposition theorem which provides a representation of a submartingale as the sum of a martingale and an increasing predictable process. This theorem lays foundation for decomposition of submartingales into their components. If we let $X(t)$ be a submartingale with respect to a filtration (a sequence of sigma algebras), then there exists a unique, increasing, predictable process $A(t)$ such that the martingale process $M(t)$ can be defined as $M(t) = X(t) - A(t)$. In other words, according to Doob-Meyer the submartingale can be decomposed into the sum of a martingale $M(t)$ and an increasing predictable process $A(t)$. The proof for a continuous case is very complicated and can be found here [16]. However, there is a contradiction here as it was said that fBM is not even a semi-martingale.

In this paper, we propose to review and extend proofs in mixed fuzzy stochastic Brownian motion as novel equations suitable for modelling hybrid systems. These equations provide a flexible framework that can effectively capture the dynamics of complex systems characterized by both stochastic and fuzzy components. By combining the concepts of fuzzy logic and stochastic processes, we aim to address the modelling challenges associated with hybrid systems. Applications of these processes can be made in telecommunication industry where measurements of variables suffer from multiple disturbances and financial industry where participants are operating under uncertainty.

To illustrate the applicability of our approach, we utilise the Liouville form of the fractional Brownian motion. The Liouville form represents a particular representation of the fBm that allows for enhanced analysis and interpretation of its properties. By leveraging this form, we can gain further insights into the behaviour of the mixed fuzzy stochastic Brownian motion and intuitional mixed fuzzy stochastic Brownian motion, thereby facilitating their use in a wide range of modelling scenarios.

3.1. Preliminaries

3.1.1. Fractional Brownian Motion

The fBm BH is a zero mean Gaussian process with the following covariance function:

$$C_{bh}(t, s) = E(B^H(t)B^H(s)) = \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H}) \quad (4)$$

This was derived in Kolmogorov [17] and further studied in Mandelbrot and van Ness [18], where a stochastic integral representation was derived in terms of the Brownian motion². $\int_0^t [(t-s)^{H-\frac{1}{2}}] dB(s)$ is called the Liouville form of a fBm and it holds many properties of the fBm except that it has non-stationary increments. The classical Itô theory cannot handle a stochastic integral in terms of fBm because B^H is not a semi-martingale if $H \neq \frac{1}{2}$. Two approaches have been used to evaluate stochastic integrals with respect to fBm. If $H > 1/2$, the Riemann-Stieltjes stochastic integral can be defined using Young's integral [19]. Suppose we have a financial time series representing the price of a stock over a certain period of time, denoted by $S(t)$. We want to calculate the Riemann-Stieltjes stochastic integral of the stock price with respect to a cumulative trading volume, denoted by $V(t)$, which represents the total number of shares traded up to time t . Let's assume $S(t)$ represents the closing price of a stock at time t . We use the following definitions. Integrator function - consider $V(t)$ to be the cumulative trading volume up to time t , which we obtain from historical trading data. Partition - choose a partition of the time interval into subintervals. For simplicity, let's consider a uniform partition with n subintervals. Piecewise linear approximation - approximate $V(t)$ on each subinterval using a piecewise linear function. This can be done by interpolating the cumulative trading volume between the available data points. Calculate the Riemann-Stieltjes integral using Young's integral approach. For each subinterval, compute the integral of $S(t)$ with respect to the piecewise linear approximation of $V(t)$. This can be done by multiplying the stock price at each subinterval by the change in the approximated cumulative trading volume. Refinement- repeat the above steps with finer and finer partitions, refining the piecewise linear approximation of $V(t)$ and calculating the Riemann-Stieltjes stochastic integral using Young's integral. By repeating the process with increasingly finer partitions, the Riemann-Stieltjes stochastic integral will converge to a well-defined value, representing the integral of the stock price with respect to the cumulative trading volume over the given time period. This can provide insights into the relationship between trading activity and stock price movements. The main difference between the Riemann-Stieltjes and the Young integral is that the former one uses a fixed partition of the time interval, while the latter one allows for an arbitrary partition of the time interval. This makes the Young's integral more flexible and easier to use in stochastics. The Riemann-Stieltjes integral cannot

² $B_H(t) = \frac{1}{\Gamma(H+\frac{1}{2})} (\int_{-\infty}^0 [(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}] dB(s) + \int_0^t [(t-s)^{H-\frac{1}{2}}] dB(s))$

be used to define an integral of a function with respect to fBm, because the Stieltjes measure is *sometimes* not defined for stochastic processes with non-stationary increments like fBm. A Stieltjes measure is a generalisation of the Lebesgue measure to allow for integration over sets with irregular boundaries. A Stieltjes measure is defined on a measurable space $(X, \mathcal{A})$, where X is a set and $\mathcal{A}$ is a σ -algebra of subsets of X . A Stieltjes measure is a function $\mu: \mathcal{A} \rightarrow [0, \infty]$ that satisfies the following properties:

- i. $\mu(\emptyset) = 0$.
- ii. $\mu(A \cup B) = \mu(A) + \mu(B)$ for all disjoint sets $A, B \in \mathcal{A}$.
- iii. If $A_1 \subset A_2 \subset \dots$ is a sequence of sets in $\mathcal{A}$ such that $A_n \uparrow A$, then $\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$.

The third property is called the continuity property. It ensures that the Stieltjes measure is not sensitive to small changes in the sets that are being measured. Stieltjes measures can be used to define integrals of functions with respect to sets with irregular boundaries. For example, the Stieltjes measure can be used to define the integral of a function with respect to a Cantor set. These are fractional sets that are constructed by repeatedly removing a middle third of a line segment. Cantor sets have no Lebesgue measure, but they do have a Stieltjes measure. This means that we can define the integral of a function with respect to a Cantor set. Here is an example of how to use a Stieltjes measure to define the integral of a function with respect to a set with an irregular boundary:

Let $X = [0, 1]$ and let $\mathcal{A}$ be the σ -algebra of all Borel subsets of X . Let $g(t)$ be the Cantor function. The Cantor function is a continuous function that is defined on the interval $[0, 1]$ and has the following properties:

- i. $g(0) = g(1) = 0$.
- ii. $g(t)$ is constant on each interval of the form $[a, b]$, where a and b are points of the Cantor set.
- iii. The value of $g(t)$ on each interval of the form $[a, b]$ is equal to the average of the values of $g(a)$ and $g(b)$.

The Cantor function is a pathological function, which means it has some unusual features. The Cantor function has no Lebesgue measure, but it does have a Stieltjes measure. It is a set of points that are sparse but not empty. We can define the Stieltjes measure for the Cantor function as follows:

$\mu(A) = g(b) - g(a)$, where A is a Borel subset of X and $[a, b]$ is the smallest interval of the form $[a, b]$ that contains A . We can now use the Stieltjes measure to define the integral of a function $f(t)$ with respect to the Cantor function: $\int f(t) dg(t) [0,1] = \mu(\int f(t) dt) [0,1]$. This integral is defined for all functions $f(t)$ that are measurable with respect to the Lebesgue measure. The Cantor function is not measurable with respect to the Lebesgue measure, but we can still define the Stieltjes measure for the Cantor function with respect to the Lebesgue measure. This is because the Stieltjes measure is a more general measure than the Lebesgue measure.

The Lebesgue measure is a measure that is defined on the Borel sets of the real line. The Borel sets of the real line are the sets that can be formed from the open sets of the real line using the operations of countable union, countable intersection, and complementation. The Stieltjes measure is a measure that is defined on the Borel sets of the real line, but it is not restricted to the Borel sets. The Stieltjes measure can be defined on any set of real numbers, even sets that are not Borel sets.

When we define the Stieltjes measure for the Cantor function with respect to the Lebesgue measure, we are using the Lebesgue measure to define the sets that the Stieltjes measure is defined on. However, we are not using the Lebesgue measure to define the values of the Stieltjes measure. The values of the Stieltjes measure are defined by the Cantor function. The Stieltjes measure for the Cantor function can be used to define the integral of a function with respect to the Cantor function. The integral of a function with respect to the Cantor function is defined as follows: $\int f(t) dg(t) [0,1] = \mu(\int f(t) dt) [0,1]$,

where $f(t)$ is the function that we want to integrate; $g(t)$ is the Cantor function; μ is the Stieltjes measure for the Cantor function.

The integral of a function with respect to the Cantor function can be used to calculate the area under the graph of the Cantor function. The area under the graph of the Cantor function is equal to the Stieltjes measure of the Cantor set.

In summary, both Young's integral and Stieltjes measures are valuable mathematical tools, but their applicability depends on the characteristics of the stochastic process being studied and the mathematical framework in which they are used. Young's integral is tailored to the needs of fBm and related processes, while Stieltjes measures are more general and can be applied to a wider range of scenarios.

If $H < 1/2$, Malliavin calculus techniques are used to approximate $B_H(t)$ by a semi-martingale process [20]. The scope for this paper is to derive results for MFFBM processes when $h > 1/2$. The theoretical framework we are relying on when $h < 1/2$ is from the Molchan-Golosov integral and adopted by [21]. It allows us to change the Hurst exponent using a specified transformation. This will be the main contribution in the future work. Returning to the proposition that if $H > 1/2$, then Riemann-Stieltjes stochastic integral can be defined using Young's integral [19].

$$B_{H,\varepsilon}(t) = \int_0^t [(t-s+\varepsilon)^{H-\frac{1}{2}}] dB(s), \quad \varepsilon > 0, a = H - 1/2 \quad (5)$$

$$B_{H,\varepsilon}(t) = a \int_0^t \varphi^\varepsilon(s) ds + \varepsilon^a B(t), \quad \text{where } \varphi^\varepsilon(s) = \int_0^t (t-s+\varepsilon)^{a-1} dB(s)$$

Proof using integration by parts technique: $\int uv = u \int v - \int u' \int v$. For an alternative proof of this theorem please see Thao (2006) Theorem 2.1 [22].

So,

$$\begin{aligned} B_{H,\varepsilon}(t) &= \int_0^t (t-s+\varepsilon)^a dB(s) \\ &= (t-s+\varepsilon)^a \int_0^t dB(s) - \int_0^t a(t-s+\varepsilon)^{a-1} \int_0^t dB(s) \\ &= [(t-s+\varepsilon)^a B(t)]_0^t - a \int_0^t (t-s+\varepsilon)^{a-1} \int_0^t dB(s) dt = (t-t+\varepsilon)^a B(t) - (0-s+\varepsilon)^a B(0) + a \int_0^t \int_0^t (t-s+\varepsilon)^{a-1} dB(s) ds \\ &= \varepsilon^a B(t) + a \int_0^t \int_0^t (t-s+\varepsilon)^{a-1} dB(s) ds = a \int_0^t \varphi^\varepsilon(s) ds + \varepsilon^a B(t), \quad \text{where } \varphi^\varepsilon(s) = \int_0^t (t-s+\varepsilon)^{a-1} dB(s) \end{aligned}$$

The process $B_{H,\varepsilon}(t)$ converges to $B_H(t)$ in $L^2(\Omega)$ when ε tends to zero because it is square integrable [22].

3.1.2. Fuzzy Random Variable, Fuzzy Stochastic Process, and Integral Overview

This section provides some necessary background for fuzzy processes. The material mainly comes from the fuzzy set theory of Zadeh [23]. Here we will define a number of measures that are used in this paper. The Lebesgue measure is widely used to measure lengths, areas, volumes, and higher-dimensional extents of sets. It is defined on the Borel σ -algebra of subsets of Euclidean space. For example, the Lebesgue measure of an interval $[a, b]$ in one dimension is simply the length of the interval, given by $[b - a]$. Counting measure - in combinatorics and discrete mathematics, the counting measure is a fundamental measure that assigns the number of elements in a set. It is defined on any set and is particularly useful for counting finite sets. For instance, the counting measure of a finite set A is equal to the cardinality (number of elements) of A . Probability measure - in probability theory, probability measures assign probabilities to events in a sample space. A probability measure is a measure that is defined on a σ -algebra of subsets of the sample space and satisfies the properties of non-negativity, assigning probability 0 to the empty set, and countable additivity. For example, the uniform distribution on the interval $[0, 1]$ is a probability measure that assigns equal probability to all sub-intervals of $[0, 1]$. Hausdorff measure - in geometric measure theory, the Hausdorff measure is used to quantify the "size" or "extent" of fractional sets. It is defined on metric spaces and captures the concept of dimension. For instance, the Hausdorff measure of a fractional set like the Cantor set or the Sierpinski triangle measures the "length" or "area" of the set in a fractional sense.

Let $K(R)$ be the family of all nonempty, compact, and convex subsets of R (field). The Hausdorff metric d_H is defined by the following:

$$d_H(A, B) = \max\left\{ \sup_{a \in A, b \in B} \inf |a - b|, \sup_{b \in B, a \in A} \inf |a - b| \right\} \quad (6)$$

A metric space is said to be separable if there exists a countable dense subset. If $A, B, C \subset K(R)$ then:

$$d_H(A + C, B + C) = d_H(A, B) \quad (7)$$

Let Ω, A, P be a probability space. The mapping $F: \Omega \rightarrow K(R)$ is called A -measurable if it satisfies $\{w \in \Omega : F(w) \cap C \neq \emptyset\} \in A$, for every closed set $C \subset R$. This means that for any element w that belongs to space Ω such that $F(w)$ is disjoint with C , it is not equal to zero for every closed set $C \subset R$. A closed set means that all limiting and interior points belong to a set. A multifunction $F \subset M$ is said to be L^p – integrably bounded if p^{th} power is integrable for $p \geq 1$, iff (if and only if) there exists $h \in L^p(\Omega, A, P, R_+)$ such that the norms metric (the norm $\| \cdot \|$ induces a metric on F , known as the metric induced by the norm or the norm metric) such that $\|F\| \leq h$ P-a.e (almost everywhere - event holds for almost all elements in a set, with the exception of a negligible subset of elements according to the probability measure P), $R_+ = [0, \infty)$ and $\|F\| = d_H(F, \{0\}) = \sup |f|, f \in F$. This means the norms in terms of metric is the greatest value of F . For example, imagine we have a function F that takes some inputs and gives us sets (groups of things) as outputs. Now, this function F is special in two ways:

- i) it doesn't make sets that are too big or too wild.
- ii) we can measure how big these sets are.

The first point means that if we look at the sets produced by this function, they're not too wild. They don't stretch out forever or become enormous. The second point is about measuring. We can measure the size of these sets by giving them a number that says how big they are. Now, this function F can't be too unpredictable. It has to follow certain rules. If we square the sizes of its sets and add them all up, that should be a reasonable number (integrable). This is where the " L^p " comes in. " p " is just a number, and when it's 1 or bigger, it means things are behaving well.

To make sure these sets aren't too big, there's another function (let's call it " h ") that keeps an eye on them. It's like a supervisor. This supervisor function, " h ," is also measured in a special way, and it makes sure our function's sets don't get too big. So, in simple terms, an L^p -integrably bounded multifunction is like a machine that produces sets. These sets are controlled and supervised to make sure they're not too big, and we can measure their sizes using numbers.

The membership function u on an interval closed set: $R \rightarrow [0,1]$ is defined for a fuzzy set $u \subset R$, where $u(x)$ is the degree of membership of x in the fuzzy set u . Let us denote by $F(R)$ the fuzzy set $u: R \rightarrow [0,1]$ such that $[u]^a \subset K(R)$ for every $a \in [0,1]$, where $[u]^a = \{x \in R: u(x) \geq a\}$. Define domain $d_\infty: F(R) \times F(R) \rightarrow [0, \infty)$ codomain /space of images and $F(R)$ with d_∞ is a complete metric space by $d_\infty(u, v) = \sup d_H([u]^a, [v]^a)$, where $a \in [0,1]$. This means that we are defining a function d_∞ that takes two inputs from the set $F(R)$ which represents the set of all real valued functions and returns a value in the range $[0, \infty)$ – the non-negative real numbers. We will now show that d_∞ is a metric in $F(R)$ and $(F(R), d_\infty)$ is a complete metric space. To accomplish this, we are considering $F(R)$ equipped with the metric d_∞ to establish if it is a complete metric space. A metric space consists of a set equipped with a metric, which is a function that measures the distance between elements in the set. In this case, $F(R)$ is the set, and d_∞ is the metric we are defining. The metric $d_\infty(u, v)$ is defined as the supremum or the least upper bound of the Hausdorff distance between the equivalence classes $[u]^a$ and $[v]^a$, where a belongs to the interval $[0,1]$. Here $[u]^a$ and $[v]^a$ represent equivalence classes of functions u and v , respectively, under the relation defined by the parameter a . The Hausdorff distance measures the "closeness" between two sets by considering the furthest distance from each set to the other. It is the maximum distance of a set to the nearest point in the other set. In this case, the sets $[u]^a$ and $[v]^a$ are being compared, and their Hausdorff distance is computed. The supremum is then taken over all values of a in the interval $[0,1]$. By defining the metric d_∞ this way, we establish that $F(R)$ is a complete metric space. Completeness means that every Cauchy sequence in $F(R)$ converges to a limit within $F(R)$. In other words, $F(R)$ with the metric d_∞ satisfies the property that any sequence of functions that gets arbitrarily close to each other eventually has a limit within the same set. So, we have defined the metric d_∞ on the set of real valued functions $F(R)$ and established $F(R)$ equipped with d_∞ as a complete

metric space. The metric measures the Hausdorff distance between equivalence classes of functions under a parameter a in the interval $[0,1]$. For every $u, v, w, z \in F(R)$, $\lambda \in R$, we have the following properties:

- i. additive property $d_\infty(u+w, v+w) = d_\infty(u, v)$
- ii. distribution property $d_\infty(u+v, w+z) = d_\infty(u, w) + d_\infty(v, z)$
- iii. transitive(triangular) property $d_\infty(u, v) \leq d_\infty(u, w) + d_\infty(w, v)$
- iv. scalar multiplication property $d_\infty(\lambda u, \lambda v) = |\lambda| d_\infty(u, v)$
- v. non-negativity $d_\infty(u, v)$ is non-negative for all u, v in $F(R)$
- vi. identity of indiscernibles $d_\infty(u, v) = 0$ if and only if $u = v$
- vii. symmetry $d_\infty(u, v) = d_\infty(v, u)$ for all u, v in $F(R)$
- viii. consistency $d_\infty(u, v) = d_\infty(v, u) = 0$ implies $u = v$.

Expanding the work of [16], particularly in 6 and 7 below, we will now define a fuzzy random variable as a function that assigns fuzzy sets to elements in probability space. We will introduce the concept of measurability with respect to different metrics such as d_∞ and d_s and describe the properties and relationships between these metrics and notion of a fuzzy random variable. We will break it down step by step:

1. Let (Ω, A, P) be a probability space, where Ω represents the sample space, A is a σ -algebra of subsets of Ω , and P is a probability measure defined on A . Given a set Ω , a σ -algebra of subsets of Ω , denoted as Σ , is a collection of subsets of Ω that satisfies certain properties. It is a fundamental concept in measure theory and serves as a foundation for defining measurable sets and functions. To formally define a σ -algebra, we require it to satisfy the following properties:

- The empty set ($\emptyset$) is in Σ : $\emptyset \in \Sigma$.
- Σ is closed under complementation: If A is in Σ , then its complement, denoted as A^c , is also in Σ .
- Σ is closed under countable unions: If $A_1, A_2, A_3, \dots$ are subsets of Ω and each A_i is in Σ , then the union of all the A_i , denoted as $\bigcup A_i$, is also in Σ .

In other words, a σ -algebra is a collection of subsets of Ω that includes the empty set, is closed under taking of complements, and is closed under countable unions.

The purpose of defining a σ -algebra is to identify a set of subsets of Ω that can be considered "measurable" in a certain sense. The elements of Σ are often referred to as measurable sets, and they form a structure that allows us to define measures (such as probability measures) on these sets.

By specifying a σ -algebra on Ω , we establish a framework for defining and working with measurable functions, integrating functions, and constructing measures. The σ -algebra provides a systematic way to determine which subsets of Ω are considered "measurable" within the given context, allowing us to analyse and reason about these subsets in a rigorous mathematical manner. In summary, a σ -algebra of subsets of Ω is a collection of subsets that satisfies specific properties, including containing the empty set, being closed under complementation, and being closed under countable unions. It provides a foundation for defining measurable sets and functions, allowing us to study measures and perform various mathematical operations on these sets.

2. A fuzzy random variable is defined as a function $X: \Omega \rightarrow F(R)$, where $F(R)$ represents the set of fuzzy sets over the real numbers. This means that X assigns each element ω in Ω a fuzzy set $X(\omega)$ in $F(R)$.
3. For every $\alpha \in [0, 1]$, the mapping $[X]^\alpha: \Omega \rightarrow K(R)$ (where $K(R)$ denotes the set of fuzzy sets over R) should be A -measurable. This implies that the pre-images of measurable sets in $K(R)$ under $[X]^\alpha$ should be measurable sets in A .
4. Consider a metric ϱ on the set $F(R)$ and let B_ϱ be the σ -algebra generated by the topology induced by ϱ . A fuzzy random variable X can be seen as a measurable mapping between the measurable spaces (Ω, A) and $(F(R), B_\varrho)$, and we refer to X as being A/B_ϱ -measurable.
5. The metric $d_s(u, v)$ is defined as follows:

$$d_s(u, v) := \inf \max \{ \sup |\lambda(t) - t|, \sup d_H(X_u(t), X_v(\lambda(t))) \} \text{ where } \lambda \in \Phi, t \in [0,1] \text{ and } t \in [0,1], \Phi$$

denotes the set of strictly increasing continuous functions $\lambda: [0,1] \rightarrow [0,1]$ satisfying $\lambda(0) = 0$ and

- $\lambda(1) = 1$. X_u and X_v are càdlàg (right-continuous with left limits) representations of the fuzzy sets u and v in $F(R)$, respectively. d_H represents the Hausdorff distance between fuzzy sets.
6. The space $(F(R), d_\infty)$ is complete (meaning all Cauchy sequences converge) and non-separable (does not contain a countable dense subset), while the space $(F(R), d_s)$ is a Polish metric space (a complete and separable metric space). The reason for the difference in these properties lies in their respective definitions and metrics. The metric d_∞ , defined as the supremum of the Hausdorff distances between equivalence classes $[u]^a$ and $[v]^a$, is a robust way to measure the “closeness” of fuzzy sets. It can be shown that the space $(F(R), d_\infty)$ is complete, meaning all Cauchy sequences converge. However, d_∞ doesn’t lend itself to finding a countable dense subset within the space $(F(R), d_\infty)$. This makes $(F(R), d_\infty)$ non-separable because it lacks the property of separability, even though it is complete. On the other hand, the metric d_s is defined differently, and results in a metric space that is both complete and separable. This metric is designed in such a way that it allows for the existence of a countable subset within the space $(F(R), d_s)$. Because $(F(R), d_s)$ is complete (all Cauchy sequences converge) and separable (contains a countable dense subset), it meets the criteria to be classified as a Polish metric space.
 7. For a mapping $X: \Omega \rightarrow F(R)$ on the probability space (Ω, A, P) , X is a fuzzy random variable iff X is $A|B_{d_s}$ -measurable and if X is $A|B_{d_\infty}$ -measurable, then it is a fuzzy random variable, but the opposite is not necessarily true. For X to be considered a fuzzy random variable, it must meet the condition of being $A|B_{d_s}$ -measurable. This means that X can be measured using d_s metric. The d_s metric is a specific way to compare how close or similar two fuzzy sets are. If X is also $A|B_{d_\infty}$ -measurable, it indicates that X can be analysed using the d_∞ metric. These two metrics, d_s and d_∞ , are used to assess the closeness of fuzzy sets, but they do so differently. d_s looks at the behaviour of the fuzzy sets through continuous λ function and their values at different points. In contrast, d_∞ considers the worst case scenario by looking at the maximum difference between the fuzzy set at any point. It is important to note that while $A|B_{d_\infty}$ -measurable is more inclusive, the opposite is not necessarily true. Just because X is $A|B_{d_\infty}$ -measurable doesn’t automatically mean it is $A|B_{d_s}$ -measurable. Therefore, $A|B_{d_s}$ -measurable is a more specific and essential criterion for defining a complete fuzzy random variable. In summary, the given explanations define a fuzzy random variable as a function that assigns fuzzy sets to elements in probability space. It introduces the concept of measurability with respect to different metrics, such as d_s and d_∞ , and describes the properties and relationships between these metrics and notion of a fuzzy random variable.

Next, we expand the work of [16] in fuzzy random variable and fuzzy stochastic process. We define a fuzzy random variable $X: \Omega \rightarrow F(R)$ is said to be L^p -integrably bounded for $p \geq 1$ if the equivalence class $[X]^\alpha$ belongs to the space $L^p(\Omega, A, P; K(R))$ for every $\alpha \in [0, 1]$. Here, the random variables $X, Y \in L^p(\Omega, A, P; K(R))$ are identical if $P(d_\infty(X, Y) = 0) = 1$. This means that for X and Y to be considered identical, the probability that their distance d_∞ is equal to zero should be 1. d_∞ measures the “distance” or dissimilarity between two fuzzy variables X and Y . When the distance d_∞ between X and Y is zero, it indicates that X and Y are indistinguishable from each other in terms of their fuzzy characteristics. For a fuzzy random variable $X: \Omega \rightarrow F(R)$ and $p \geq 1$, the following conditions are equivalent: a) $X \in L^p(\Omega, A, P; F(R))$. b) The equivalence class $[X]^0$ belongs to $L^p(\Omega, A, P; F(R))$. c) The norm of $[X]^0$ denoted as $\| [X]^0 \|$ belongs to $L^p(\Omega, A, P; R^+)$, where R^+ represents the non-negative real numbers. The statement then asserts that these three conditions (a, b, and c) are equivalent. In other words, if any one of these conditions is true, the other two conditions will also be true, and vice versa. This demonstrates the relationships between the integrability of a fuzzy random variable, its equivalence class, and the norm of that equivalence class in different spaces. The notation $I := [0, T]$ signifies the interval from 0 to T , which means it is a closed interval from 0 to T inclusive and (Ω, A, P) represents a complete probability space equipped with a filtration $\{A_t\}_{t \in I}$. A filtration is a sequence or family of sub- σ -algebras $\{A_t\}_{t \in I}$ defined on a probability space. In this case, the filtration is defined for each t in the interval $I = [0, T]$. A complete probability space refers to a probability space where every subset of a null set (a set with measure zero) is also measurable and has measure zero. A complete probability space means that not only is the sample space, the σ -algebra, and the probability measure defined, but it also has an additional property that every subset

of a null set (a set with measure zero) is also measurable. In other words, if you have a set of outcomes that is so unlikely (e.g., its probability measure is zero), then any subset of that unlikely set is also considered measurable. Being measurable means that it's part of the σ -algebra, so you can assign probabilities to it. The requirement that these subsets of null sets are measurable ensures that there are no hidden or unmeasurable sets that could potentially lead to issues in probability theory. It essentially fills in any gaps or irregularities in the probability space, making it more robust.

The filtration is an increasing and right-continuous family of sub- σ -algebras of A , which contains all P-null sets. The filtration $\{A_t\}_{t \in I}$ is said to be increasing, which means that as t increases, the sub- σ -algebra A_t becomes larger or contains more sets. In other words, if $s \leq t$, then A_s is a subset of A_t . This property ensures that the information or knowledge available at time s is also available at any later time $t \geq s$. The filtration $\{A_t\}_{t \in I}$ is also right-continuous, which means that for any t in the interval I , the sub- σ -algebra A_t is equal to the intersection of all sub- σ -algebras A_s with $s > t$. In other words, the information available at time t is already available at any slightly earlier time $t - \varepsilon$, where ε is a small positive value. The filtration $\{A_t\}_{t \in I}$ is required to contain all P-null sets. A P-null set is a subset of the sample space Ω with measure zero. By including all P-null sets in each sub- σ -algebra A_t , the filtration captures all negligible information about the underlying probability space.

A fuzzy stochastic process X is said to be d_∞ -continuous if almost all its trajectories, denoted as $X(\cdot, \omega)$, are d_∞ -continuous functions. Here, the trajectories refer to the mappings $X(\cdot, \omega)$ that associate each time point t in the interval I with a fuzzy set $X(t, \omega)$ in $F(R)$. The d_∞ -continuity of these trajectories means that they exhibit continuity with respect to the d_∞ metric, which measures the dissimilarity between fuzzy sets. A fuzzy stochastic process X is considered measurable if the equivalence class $[X]^\alpha: I \times \Omega \rightarrow K(R)$ is a measurable multifunction for all α in the interval $[0, 1]$. Here, $[X]^\alpha$ represents the equivalence class of X with respect to the parameter α , and $K(R)$ denotes the set of fuzzy sets over the real numbers. A measurable multifunction $[X]^\alpha: I \times \Omega \rightarrow K(R)$ maps a measurable set in the time domain I and the sample space Ω to a set of measurable fuzzy sets in the value domain R . For example, Let $X(t)$ be a fuzzy stochastic process that represents the price of a stock at time t . The parameter α can represent the level of confidence that we have in the price prediction. For example, $\alpha = 0.95$ could represent a 95% confidence level. We can define the equivalence class $[X(t)]^\alpha$ as follows: $[X(t)]^\alpha = \{Y \in K(R) \mid Y(x) \geq \alpha \text{ for all } x \in R\}$. The parameter α represents the level of confidence in the price prediction. It quantifies the minimum membership value required for a fuzzy set to be considered as part of the equivalence class $[X(t)]^\alpha$. This equivalence class contains all fuzzy sets that are at least as precise as $X(t)$ with respect to the confidence level α . The equivalence class $[X(t)]^\alpha$ is defined as the set of fuzzy sets Y in $K(R)$ that have a membership value greater than or equal to α for all x in the real numbers (R) . In other words, $[X(t)]^\alpha$ contains all fuzzy sets that are at least as precise as $X(t)$ with respect to the confidence level α . We can then show that the multifunction $[X(t)]^\alpha: I \times \Omega \rightarrow K(R)$ is measurable. This means that the fuzzy stochastic process $X(t)$ is measurable. By stating that the multifunction $[X(t)]^\alpha: I \times \Omega \rightarrow K(R)$ is measurable, we are asserting that the mapping from the Cartesian product of the index set I and the sample space Ω to the set of fuzzy sets $K(R)$ is a measurable multifunction. This means that the pre-images of measurable sets in $K(R)$ under $[X(t)]^\alpha$ are measurable sets in $I \times \Omega$. The pre-images, in this context, refer to the subsets of the domain $I \times \Omega$ that get mapped to a specific set in the target space $K(R)$ under the mapping $[X(t)]^\alpha: I \times \Omega \rightarrow K(R)$. To put it simply, if we have a set A in $K(R)$, the pre-image of A under $[X(t)]^\alpha$ is the set of all points (i, ω) in the domain $I \times \Omega$ such that $[X(t)]^\alpha(i, \omega)$ belongs to A .

In other words, the pre-image of a set A in $K(R)$ consists of all the elements (i, ω) in $I \times \Omega$ that get mapped to A under the mapping $[X(t)]^\alpha$. By saying that the pre-images of measurable sets in $K(R)$ under $[X(t)]^\alpha$ are measurable sets in $I \times \Omega$, it means that if we consider a measurable set A in $K(R)$, the pre-image of A under $[X(t)]^\alpha$ is a measurable set in $I \times \Omega$. This implies that the behaviour of the fuzzy stochastic process $X(t)$ with respect to such pre-images is well-behaved and can be characterised by measurable sets in the domain $I \times \Omega$. For example, we have a fuzzy stochastic process $X(t)$ that represents the temperature at a particular location over time. The parameter α represents the confidence level in our temperature predictions. The domain of the process is the Cartesian product of the time index set I and the sample space Ω . Let's assume $I = \{1, 2, 3\}$ represents three time points

and $\Omega = \{A, B\}$ represents two possible outcomes (e.g., sunny, or cloudy days). The target space $K(\mathbb{R})$ represents the set of fuzzy sets over the real numbers. Equivalence class- for a specific time point t , let's say $X(t)$ generates a fuzzy set that represents the temperature with respect to the confidence level α . Measurable sets in $K(\mathbb{R})$, let's consider a measurable set A that represents a specific range of temperatures. Now, let's define the pre-image of A under $[X(t)]^\alpha$: the pre-image of A under $[X(t)]^\alpha$ is the set of all (i, ω) in $I \times \Omega$ such that $[X(t)]^\alpha(i, \omega)$ belongs to A . In other words, it consists of all the time points and sample outcomes for which the fuzzy set generated by $X(t)$ at that particular time and outcome falls within the range represented by A . For example, let's assume A represents temperatures between 20 and 30 degrees Celsius. If $[X(t)]^\alpha(2, B)$ represents a fuzzy set with membership values indicating temperatures around 25 degrees Celsius, then $(2, B)$ would be in the pre-image of A because the temperature falls within the range represented by A . In this case, the pre-image of the measurable set A under $[X(t)]^\alpha$ would be a subset of $I \times \Omega$ containing all the time points and sample outcomes where the fuzzy sets generated by $X(t)$ fall within the temperature range represented by A .

We propose that the condition that $[X]^\alpha$ be measurable for all α in the interval $[0, 1]$ ensures that the fuzzy stochastic process X can be integrated with respect to time and the sample space.

$B(I)$ is the Borel σ -algebra, which consists of all subsets of the interval I . To determine measurability, $[X]^\alpha: I \times \Omega \rightarrow K(\mathbb{R})$ should be $B(I) \otimes A$ -measurable. Here, $B(I) \otimes A$ represents the product σ -algebra, which is generated by the Borel σ -algebra $B(I)$ and the σ -algebra A . This requirement ensures that the equivalence class $[X]^\alpha$ is measurable with respect to the joint σ -algebra generated by $B(I)$ and A . In this context, $K(\mathbb{R})$ denotes the set of fuzzy sets over the real numbers. Fuzzy sets are sets that assign membership degrees or degrees of truth to elements, allowing for a more flexible representation of uncertainty or partial truth. $B(I)$ refers to the Borel σ -algebra, which is the σ -algebra generated by the open sets in the interval I . The Borel σ -algebra consists of all subsets of I that can be formed by taking unions, intersections, and complements of open sets. $B(I) \otimes A$ represents the product σ -algebra, which is the smallest σ -algebra generated by the sets of the form $B \times A$, where B is a Borel set from $B(I)$ and A is a set from the σ -algebra A . The requirement of $[X]^\alpha$ being $B(I) \otimes A$ -measurable ensures that the equivalence class $[X]^\alpha$ is measurable with respect to the joint σ -algebra generated by $B(I)$ and A . This joint σ -algebra allows us to analyse and reason about the measurability properties of the fuzzy stochastic process X across both the time domain (Borel sets in I) and the probability space (sets in A). Suppose we have two sets: $B(I)$, which represents the Borel σ -algebra generated by the open sets in the interval I , and A , which represents a σ -algebra defined on a sample space Ω . We want to generate a σ -algebra that captures the joint measurability of sets from $B(I)$ and A . The product σ -algebra $B(I) \otimes A$ is defined as the smallest σ -algebra generated by sets of the form $B \times A$, where B is a Borel set from $B(I)$ and A is a set from the σ -algebra A . Let's consider a specific example to illustrate this concept. Assume that I is the interval $[0, 1]$ and A is the power set of $\Omega = \{a, b\}$, which contains all possible subsets of Ω . $B(I)$: $B(I)$ is the Borel σ -algebra on I , generated by the open sets in I . It includes sets such as open intervals, closed intervals, half-open intervals, and countable unions or intersections of such intervals. A : A is the power set of Ω , which includes all subsets of Ω . In this case, $A = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$. To construct the product σ -algebra $B(I) \otimes A$, we consider all possible combinations of Borel sets from $B(I)$ and sets from A . For example, let's take $B = (0.2, 0.8)$ (an open interval in I) and $A = \{a\}$ (a singleton set in A). The set $B \times A$ is the Cartesian product of B and A , which results in the set $\{(x, a) \mid 0.2 < x < 0.8\}$. Similarly, we can consider other combinations of Borel sets from $B(I)$ and sets from A . The product σ -algebra $B(I) \otimes A$ is then defined as the smallest σ -algebra that includes all such sets $B \times A$ for all possible choices of B from $B(I)$ and A from A . In this example, the product σ -algebra $B(I) \otimes A$ would consist of sets such as $B \times A$, where B represents Borel sets from $B(I)$ and A represents sets from A . It captures the joint measurability of sets from $B(I)$ and A , allowing us to reason about sets that involve both the real number domain and the sample space. Imagine we want to analyse rainfall data for different cities over time. We have a set of cities, say $\{\text{London, Los Angeles, Paris}\}$, and we want to consider different time intervals for our analysis. Our set of cities is like the interval I in our previous discussion. The time intervals could represent events or outcomes (such as days with rainfall data) in our probability space, denoted by the σ -

algebra A . Now, we want to understand the measurability of rainfall data across both the cities and time intervals. This is where the product σ -algebra $B(I) \otimes A$ comes into play. For each city, we can create Borel sets, which are sets formed from open intervals. In this context, these could represent different subsets of time periods for each city. For example, we might have the Borel set for London's rainy days, the Borel set for Los Angeles' rainy days, and so on. σ -Algebra A - this represents our probability space, which includes various outcomes or events in our analysis. In this case, each outcome or event is a different time interval (e.g., a specific date or range of dates). Now, the product σ -algebra $B(I) \otimes A$ is the smallest σ -algebra that contains all possible combinations of sets formed by pairing a Borel set from $B(I)$ with a set from σ -algebra A .

In our example, it means we're looking at measurable sets that are related to both specific cities (Borel sets) and specific time intervals (σ -algebra A). It allows us to answer questions like: "How does rainfall in London relate to specific time intervals, and how does it compare to rainfall in Los Angeles during those times?" So, $B(I) \otimes A$ helps us explore the measurability of events that occur at the intersection of these two domains: cities and time intervals.

A process X is said to be nonanticipating if, for every $\alpha \in [0, 1]$, the multifunction $[X]^\alpha$ is measurable with respect to the σ -algebra N . The σ -algebra N is defined as follows: $N := \{A \in B(I) \otimes A : A^t \in A \text{ for every } t \in I\}$, where A^t represents the set of outcomes ω such that the pair (t, ω) belongs to the set A . In other words, N consists of subsets of $I \times \Omega$ for which the corresponding time t belongs to the set A .

Before proceeding further, we need to state a few theorems.

Definition 1

The Fubini's theorem [24], also known as the Fubini-Tonelli theorem, is a fundamental result in measure theory and integration theory. It provides a condition under which the iterated integrals of a measurable function over a product space can be computed by performing separate integrations. Formally, let (Ω_1, A_1, μ_1) and (Ω_2, A_2, μ_2) be two measure spaces, where Ω_1 and Ω_2 are sets, A_1 and A_2 are σ -algebras of subsets of Ω_1 and Ω_2 , respectively, and μ_1 and μ_2 are measures defined on A_1 and A_2 , respectively. Consider the product measure space (Ω, A, μ) defined as the product of the two measure spaces, where $\Omega = \Omega_1 \times \Omega_2$, $A = A_1 \otimes A_2$ (the product σ -algebra), and μ is the product measure. Fubini's theorem states that if $f: \Omega \rightarrow [0, +\infty]$ is a non-negative measurable function, then the iterated integrals $\int_{\Omega_1} \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_1(\omega_1) d\mu_2(\omega_2)$ and $\int_{\Omega_2} \int_{\Omega_1} f(\omega_1, \omega_2) d\mu_2(\omega_2) d\mu_1(\omega_1)$ are both well-defined and equal, provided at least one of them is finite. In other words, under appropriate conditions, the order of integration can be interchanged without affecting the result.

Definition 2

The Aumann integral, also known as the integral with respect to a fuzzy measure, is a generalization of traditional integration theory that applies to fuzzy sets. It was introduced by Robert J. Aumann, a mathematician and economist, to extend the concept of integration to accommodate uncertainty and partial truth [25]. In traditional integration theory, the integral of a function measures the "area under the curve" or the accumulation of a quantity over a given range. However, when dealing with fuzzy sets, where membership degrees represent degrees of truth or uncertainty, the concept of area or accumulation needs to be generalised. The Aumann integral addresses this by considering the integration of a fuzzy set with respect to a fuzzy measure. A fuzzy measure assigns a degree of "importance" or "size" to subsets of a set, similar to how a traditional measure assigns a value to subsets. The Aumann integral captures the concept of "accumulation" of fuzzy sets over a range of values. Formally, let X be a fuzzy set defined on a set Ω , and let μ be a fuzzy measure defined on the power set of Ω . The Aumann integral of X with respect to μ , denoted as $\int X d\mu$, is a generalized notion of integration that takes into account the degrees of truth or membership of X . The Aumann integral is typically defined in terms of a level-wise integral. For each α in the interval $[0, 1]$, the Aumann integral $\int X d\mu$ is equal to the integral of the α -cut of X (i.e., the subset of elements with membership degree at least α) with respect to the α -cut of the fuzzy measure μ . The Aumann integral allows for the incorporation of fuzzy sets and fuzzy measures into integration theory, enabling the quantification of accumulation and aggregation of fuzzy information. For example, let's consider a product space where X is the set of real numbers $\mathbb{R}$, and L is the lattice of closed intervals $[a, b]$ on the

real line, ordered by inclusion. Suppose we have a function $f: \mathbb{R} \times [0, 1] \rightarrow \mathbb{R}$ defined as $f(x, [a, b]) = x^2$. Here, x represents a point in the real line, and $[a, b]$ represents a closed interval. To compute the Aumann integral of f , we divide the lattice of closed intervals into levels based on the length of the intervals, such as intervals of length 1, intervals of length 1/2, intervals of length 1/4, and so on. We then integrate the function $f(x, [a, b]) = x^2$ over each level, considering the measure on the real line. Finally, we combine these integrals according to lattice operations to obtain the Aumann integral. The Aumann integral of the function $f(x, [a, b]) = x^2$ over the product space $X \times L$, where X is the set of real numbers and L is the lattice of the closed intervals on the real line is given by:

$$\int_{X \times L} f(x, [a, b]) d\mu(x) d\lambda([a, b]) = \int_{-\infty}^{\infty} \int_0^1 x^2 d\lambda([a, b]) d\mu(x)$$

where μ is the Lebesgue measure on the real line and λ is the counting measure on the lattice L . Since the counting measure gives a weight of 1 to each interval, the integral over L is simply the sum of the values of $f(x, [a, b])$ over all intervals $[a, b] \in L$. Therefore, the Aumann integral becomes:

$$\int_{X \times L} f(x, [a, b]) d\mu(x) d\lambda([a, b]) = \int_{-\infty}^{\infty} \sum_{a \leq b} x^2 d\mu(x)$$

The sum over all intervals $[a, b]$ can be replaced by an integral over the real line, since the intervals are contiguous and non-overlapping.

$$\int_{X \times L} f(x, [a, b]) d\mu(x) d\lambda([a, b]) = \int_{-\infty}^{\infty} \int_{-\infty}^x x^2 dy dx$$

Evaluating the inner integral, we get:

$$\int_{X \times L} f(x, [a, b]) d\mu(x) d\lambda([a, b]) = \int_{-\infty}^{\infty} \frac{x^2}{3} dx$$

Evaluating the outer integral, we get:

$$\int_{X \times L} f(x, [a, b]) d\mu(x) d\lambda([a, b]) = \infty$$

The reason the Aumann integral is infinite lies in the nature of the counting measure and the unbounded function. For any given x , there are infinitely many intervals $[a, b]$ in L that contain x . Since the function $f(x, [a, b]) = x^2$ depends only on x and is unbounded for large values of x , summing its value over infinitely many intervals will inevitably lead to an infinite result. Therefore, the Aumann integral of the function $f(x, [a, b]) = x^2$ over the product space $X \times L$ is infinite. To obtain a finite Aumann integral, we could restrict the lattice L to a subset of intervals with finite measure or specific properties.

Following works of [16], we will now use Fubini's theorem and Aumann integral to build the fuzzy integral and discuss its properties. A fuzzy stochastic process X is called L^p -integrably bounded ($p \geq 1$) if there exists a real-valued stochastic process $h \in L^p(I \times \Omega, \mathcal{N}; \mathbb{R}^+)$, where I is the interval and Ω is the probability space, such that the fuzzy set $\| |X(t, \omega)|^0 \|$ is bounded by $h(t, \omega)$ for almost all (t, ω) in $I \times \Omega$. The notation $\| |X(t, \omega)|^0 \|$ represents the magnitude or absolute value of the equivalence class $[X(t, \omega)]^0$. $L^p(I \times \Omega, \mathcal{N}; F(\mathbb{R}))$ denotes the set of nonanticipating and L^p -integrably bounded fuzzy stochastic processes. Nonanticipating means that the process X does not rely on future information, but only on the current and past information. By employing the Fubini theorem, the fuzzy integral of a fuzzy stochastic process X , denoted as $\int_0^T X(s, \omega) ds$ is defined. Here, T is a fixed time point, and the integral is taken over the interval from 0 to T . The integral is defined for ω in the set Ω minus N_x , where N_x is a subset of Ω belonging to the σ -algebra $\mathcal{A}$ with $P(N_x) = 0$, meaning it has measure zero. The fuzzy integral $\int_0^T X(s, \omega) ds$ can be defined level-wise. For each α in the interval $[0, 1]$ and for each ω in the set Ω minus N_x , the Aumann integral $\int_0^T [X(s, \omega)]^\alpha ds$ belongs to the set $K(\mathbb{R})$, which represents fuzzy sets over the real numbers. Therefore, the fuzzy random variable $\int_0^T X(s, \omega) ds$ belongs to the set $F(\mathbb{R})$ for every ω in the set $\Omega \setminus N_x$.

Now we define the fuzzy stochastic Lebesgue-Aumann integral of $X \in L^1(I \times \Omega, \mathcal{N}; F(\mathbb{R}))$.

$$L_x(t, w) = \begin{cases} \int_0^T 1_{[0, t]}(s) X(s, w) ds & \text{for every } w \in \Omega \setminus N_x \\ \langle 0 \rangle & \text{for every } w \in N_x \end{cases} \quad (8)$$

We will now discuss properties of this integral L_x . First, we introduce some more terminology. The material comes from Kolmogorov [26] and Zadeh [23]. The d^∞ metric, also known as the supremum or uniform metric, is a mathematical measure of dissimilarity or distance between fuzzy sets. It quantifies the maximum discrepancy between the membership degrees of corresponding elements in two fuzzy sets. Formally, let A and B be two fuzzy sets defined on a common universe of discourse. The d^∞ metric between A and B , denoted as $d^\infty(A, B)$, is defined as the supremum or maximum of the absolute differences in membership degrees: $d^\infty(A, B) = \sup |\mu_A(x) - \mu_B(x)|$, where $\mu_A(x)$ and $\mu_B(x)$ represent the membership degrees of the element x in the fuzzy sets A and B , respectively. The supremum is taken over all elements x in the universe of discourse. The d^∞ metric provides a measure of the largest difference in membership degrees between corresponding elements of two fuzzy sets. It captures the maximum discrepancy or dissimilarity between the two sets, indicating how much they differ in terms of their membership degrees. Now we will define L^1 norm. L^1 refers to a mathematical space or norm in the context of functions and integrals. Specifically, L^1 represents a function space of integrable functions with a finite integral norm. Formally, L^1 is a function space defined over a given measure space (Ω, A, μ) . It consists of all measurable functions $f: \Omega \rightarrow \mathbb{R}$ (or $f: \Omega \rightarrow \mathbb{C}$) such that the integral of the absolute value of f , denoted as $\int |f| d\mu$, is finite: $L^1(\Omega, A, \mu) = \{f: \Omega \rightarrow \mathbb{R} \text{ (or } \mathbb{C}) \mid \int |f| d\mu < \infty\}$.

In simpler terms, L^1 consists of functions whose absolute value is integrable over the measure space, meaning that the area under the curve of the absolute value function is finite. The norm associated with the L^1 space, denoted as $\|f\|_1$, is defined as the integral of the absolute value of f : $\|f\|_1 = \int |f| d\mu$. The L^1 norm measures the size or magnitude of the function f in terms of its integral, capturing how spread out or concentrated the function is. L^1 represents a function space of integrable functions, where the absolute value of the function is integrable. The associated norm, $\|f\|_1$, measures the size of the function based on its integral. In the context of stochastic processes, the L^1 space refers to a function space that consists of processes that are integrable with respect to a given measure. Formally, let (Ω, A, P) be a probability space and let X be a stochastic process defined on that space. The L^1 space, denoted as $L^1(\Omega, A, P)$, consists of all stochastic processes X that satisfy the condition of integrability. This means that the expected value or integral of the absolute value of X is finite: $L^1(\Omega, A, P) = \{X: \Omega \rightarrow \mathbb{R} \text{ (or } \mathbb{C}) \mid E[|X|] < \infty\}$. Here, $E[\cdot]$ denotes the expected value or the integral with respect to the probability measure P . The absolute value $|X|$ ensures that both positive and negative values contribute to the integrability condition. In simpler terms, a stochastic process X belongs to the L^1 space if the expected value of its absolute value is finite. This indicates that the process is well-behaved in terms of its average behaviour and does not exhibit extreme fluctuations or infinite values. Now we define the L^p space. The L^p space, also known as a p -norm space, is a function space that consists of functions with finite p -norms. Formally, the L^p space, denoted as $L^p(\Omega, A, \mu)$, is defined over a given measure space (Ω, A, μ) . It consists of all measurable functions $f: \Omega \rightarrow \mathbb{R}$ (or $\mathbb{C}$) such that the p -th power of the absolute value of f , denoted as $|f|^p$, is integrable with respect to the measure μ : $L^p(\Omega, A, \mu) = \{f: \Omega \rightarrow \mathbb{R} \text{ (or } \mathbb{C}) \mid \int |f|^p d\mu < \infty\}$. Here, p is a positive real number greater than or equal to 1. The p -norm of a function f , denoted as $\|f\|_p$, is defined as the p -th root of the integral of $|f|^p$: $\|f\|_p = (\int |f|^p d\mu)^{1/p}$. The L^p space is a function space that captures functions whose p -th power is integrable, indicating that the function is well-behaved in terms of its magnitude and spread. The norm associated with the L^p space, the p -norm, measures the size or magnitude of the function f with respect to its integral, emphasizing the contribution of higher values due to the exponent p . The L^p spaces have various properties, including completeness, which means that every Cauchy sequence of functions converges to a limit within the space. Different values of p lead to different L^p spaces. For instance, when $p = 1$, we have the L^1 space, and when $p = 2$, we have the L^2 space. In summary, the L^p space consists of functions with finite p -norms, where the p -th power of the absolute value is integrable. It provides a framework for studying well-behaved functions in terms of their magnitudes and spread, with the p -norm quantifying their sizes. In the context of stochastic processes, the L^p space refers to a function space that consists of processes that are p -integrable with respect to a given measure. Formally, let (Ω, A, P) be a probability space and let X be a stochastic process defined on that space. The L^p space, denoted as $L^p(I \times \Omega, A \times N; F(R))$, consists of all stochastic processes X that

satisfy the condition of p -integrability. This means that the p -th power of the absolute value of X , denoted as $|X|^p$, is integrable with respect to the product measure $A \times N: L^p(I \times \Omega, A \times N; F(R)) = \{X: I \times \Omega \rightarrow F(R) \mid \iint |X(t, \omega)|^p dP(t, \omega) < \infty\}$. Here, p is a positive real number greater than or equal to 1. The integral is taken over the product measure P , which is the joint measure defined on the σ -algebra generated by the Cartesian product of the intervals I and the probability space (Ω, A, P) . The p -integrability condition ensures that the p -th power of the absolute value of the process X is well-behaved and has a finite integral.

Having defined integrability conditions, we can now expand the work of [16] with exploring the Lebesgue-Aumann integral. If $X \in L^p(I \times \Omega, N; F(R))$ for $p \geq 1$, then $L_X(\cdot, \cdot) \in L^p(I \times \Omega, N; F(R))$: This property states that if the fuzzy stochastic process X belongs to the L^p space, which consists of processes that are p -integrable, then the fuzzy stochastic Lebesgue-Aumann integral L_X , defined over the same probability space, also belongs to the L^p space. In other words, the integrability of X carries over to the integrability of L_X . Let $X \in L^1(I \times \Omega, N; F(R))$, then $\{L_X(t)\}_{t \in I}$ is d_∞ -continuous. This property states that if the fuzzy stochastic process X belongs to the L^1 space, which consists of processes that are integrable, then the trajectories of the fuzzy stochastic Lebesgue-Aumann integral $\{L_X(t)\}_{t \in I}$ are d_∞ -continuous. In other words, the integral $L_X(t)$ is a d_∞ -continuous function of time t for each fixed outcome ω . This implies that the fuzzy integral exhibits continuity with respect to the d_∞ metric, capturing the dependence and smoothness of the integral over time. We will now define an inequality that provides a bound on the maximum difference between fuzzy sets $L_{t,x}(u)$ and $L_{t,y}(u)$ over the interval $[0, t]$ in terms of the cumulative difference between fuzzy sets $X(s)$ and $Y(s)$ over the sample interval. The inequality captures the relationship between the supremum and integral of the d_∞ metric almost everywhere (a.e.). This inequality states that the supremum of the d_∞ metric between $L_{t,x}(u)$ and $L_{t,y}(u)$ for u in the interval $[0, t]$ is bounded by the scaled integral of the d_∞ metric between $X(s)$ and $Y(s)$ for s in the interval $[0, t]$ with the scaling factor depending on the value of p .

$$\sup_{u \in [0, t]} d_\infty^p(L_{t,x}(u), L_{t,y}(u)) \leq t^{p-1} \int_0^t d_\infty^p(X(s), Y(s)) ds, \text{ a.e.} \quad (9)$$

The L.H.S. represents the supremum (maximum) of the d_∞ metric between $L_{t,x}(u)$ and $L_{t,y}(u)$, where t is a fixed value and u ranges from 0 to t . Here, $d_\infty^p(a, b)$ denotes the d_∞ metric between two fuzzy sets or elements a and b . $d_\infty^p(L_{t,x}(u), L_{t,y}(u))$ is the d_∞ metric between the fuzzy sets $L_{t,x}(u)$ and $L_{t,y}(u)$ at each u . It measures the maximum difference in membership degrees between corresponding elements of the fuzzy sets at a given u . The R.H.S. represents an integral involving the d_∞ metric between the fuzzy sets $X(s)$ and $Y(s)$, where s ranges from 0 to t . The integral is scaled by t^{p-1} , where p is a positive real number. It is the integral of the d_∞ metric between $X(s)$ and $Y(s)$ at each s . It measures the cumulative difference in membership degrees between corresponding elements of the fuzzy sets over the interval $[0, t]$. t^{p-1} scales the integral by a factor that depends on the value of p . The value of p determines how the integral contributes to the overall inequality. The inequality holds almost everywhere (a.e), meaning it holds for all values of t except possibly for a set of measure zero.

We will now define the fuzzy stochastic Itô integral by using fuzzy random variable $\int_0^T X(s) dW(s)$, where W is a Wiener process. Let $X \in L^2(I \times \Omega, N; R)$, then $\{\langle \int_0^T X(s) dW(s) \rangle\}$, where $t \in I$. This represents a fuzzy stochastic process defined over the interval I . For each fixed t , $\langle \int_0^T X(s) dW(s) \rangle$ is a fuzzy set. The process captures the evolution of these fuzzy sets as t varies. Here, $dW(s)$ represents the differential of a Wiener process or Brownian motion. The next part indicates that the fuzzy stochastic process $\langle \int_0^T X(s) dW(s) \rangle$ belongs to the L^2 space over the product space $I \times \Omega$. The fuzzy sets generated by the integral have finite second moments or variances and take values in the space of fuzzy sets over the real numbers, denoted as $F(R)$. The last proposition states that if X belongs to a family of square integrables then the next integration will become fuzzy stochastic differential equation and then that integration also belongs to the family of square integrables. Formally, let $X \in L^2(I \times \Omega, N; R)$, then X belongs to the L^2 space, which consists of processes that have finite second moments or variances. The process X takes real-valued values and $\langle \int_0^T X(s) dW(s) \rangle, t \in I$ is d_∞ -continuous. In other words, as t varies within the interval I , the fuzzy sets $\langle \int_0^T X(s) dW(s) \rangle$ exhibit continuity with respect to the d_∞ metric. As discussed earlier, the d_∞ metric, also known as the

supremum or uniform metric, measures the maximum difference in membership degrees between corresponding elements of fuzzy sets. Thus, the statement asserts that the fuzzy sets $\langle \int_0^T X(s) dW(s) \rangle$, obtained by integrating X with respect to the Wiener process, show continuity with respect to the d_∞ metric as t varies. This continuity property is significant because it indicates that the fuzzy stochastic process $\langle \int_0^T X(s) dW(s) \rangle, t \in I$ has a smooth and well-behaved evolution with respect to the d_∞ metric. It suggests that the variations in the fuzzy sets, induced by the integral, are gradual and predictable as t changes within the interval I . In summary, the given statement states that if the stochastic process X belongs to the L^2 space, then the fuzzy stochastic process $\langle \int_0^T X(s) dW(s) \rangle, t \in I$, obtained by integrating X with respect to a Wiener process, exhibits continuity with respect to the d_∞ metric. This property implies a smooth and predictable evolution of the fuzzy sets as t varies.

3.1.3. Fuzzy Stochastic Differential Equation when $H > \frac{1}{2}$

We will now look at the special case of $H > \frac{1}{2}$ and use Liouville form fBm [22] [27]. This is a type of fractional Gaussian process that is defined by a covariance function that exhibits a power-law scaling. Specifically, the covariance function is given by: $C(t, s) = |t - s|^{(2H - 2)}$, where $0 < H < 1$ is the Hurst exponent. Now, in this case, $H > \frac{1}{2}$ for the process to have finite moments of order greater than or equal to two. If $H < \frac{1}{2}$, the process has infinitely many moments and the covariance function becomes singular. This means that the covariance function doesn't satisfy the conditions of positive definiteness that are typically expected for covariance functions. When $H > \frac{1}{2}$, Liouville fBm does not have mean-reversion. When $H < \frac{1}{2}$, the fBm process does not possess the self-similarity and stationarity properties required for the Liouville form. A fractional Brownian motion with $H < \frac{1}{2}$ is often referred to as "rough" or "pathologically irregular." It exhibits strong long-range dependence, which means that distant points in the process are highly correlated. In the Liouville form, fBm is expressed as a stochastic integral involving a deterministic function and a standard Brownian motion. It has the form [22] :

$$X(t) = X_0 + \int_0^t f(s, X(s)) ds + \langle \int_0^t g(s, X(s)) dB_H(s) \rangle, \quad (10)$$

$$X_0 = X(0)$$

where the FSDE is given by equation (10), $X(t)$ is the solution of the equation at time t , X_0 is the initial condition of X at time 0, $f(s, X(s))$ is a function that describes the drift of X , and $g(s, X(s))$ is a function that describes the diffusion of X , and B_H is the Liouville fBm that drives the equation. The integral term in equation represents the stochastic integral of $g(s, X(s))$ with respect to $B_H(s)$, which is a type of integral that is defined in terms of the Itô integral and the covariance structure of the fBm. The angle brackets around the integral indicate that it is taken in the sense of the Stratonovich integral. The corresponding approximation equation to (10) is given by $X^\varepsilon(t)$, where $X^\varepsilon(t)$ is an approximation to $X(t)$ that is obtained by discretising the integral term using a numerical scheme. The integral term in the approximation equation is approximated using a discrete version of the stochastic integral, which is defined in terms of the Riemann sum and the covariance structure of the discretised fBm, denoted by $B_H^\varepsilon(s)$ [27].

$$X^\varepsilon(t) = X_0 + \int_0^t f(s, X^\varepsilon(s)) ds + \langle \int_0^t g(s, X^\varepsilon(s)) dB_H^\varepsilon(s) \rangle \quad (11)$$

This section is discussing a class of stochastic differential equations that are driven by a Liouville fBm, which is a type of fractional Gaussian process with a power-law covariance structure. Building on [16], we will now make a couple of assumptions to prove that (11) has a strong unique solution. First assumption(A1) is that for every $u, v \in F(R)$ and every $t \in I$, there exists $L > 0$ such that $\max\{d_\infty(f(t, w, u), f(t, w, v)), |g(t, w, u) - g(t, w, v)|\} \leq L d_\infty(u, v)$. This assumption helps estimate upper bound using constant L . The second assumption (A2) is the Cauchy-Schwarz inequality. Suppose there is $C > 0$ such that $\max\{d_\infty(f(t, w, u(0)), |g(t, w, u)|\} \leq C(1 + d_\infty(u, \langle 0 \rangle))$. Last assumption (A3) is that the mapping $f: I \times \Omega \times F(R) \rightarrow F(R)$ and $g: I \times \Omega \times F(R) \rightarrow F(R)$ are measurable with respect (w.r.t) to Standard Brownian motion. Formally, $N \otimes B_{ds} | B_{ds}$ - measurable and $N \otimes B_{ds} | B(R)$ - measurable, respectively. The product sigma-algebra is a way to combine two sigma-algebras (sets of events) from

different sources. In this work, $\mathcal{N}$ represents the sigma-algebra associated with the underlying probability space, which captures the randomness or uncertainty inherent in the system. $\mathcal{B}_{ds}|\mathcal{B}_{ds}$ denotes the sigma-algebra generated by the Brownian motion process B_{ds} . It contains all the events or information that can be described solely in terms of the Brownian motion process. The symbol $\otimes$ represents the product operation between sigma-algebras, which creates a larger sigma-algebra that combines the information from both sources. The product sigma-algebra $\mathcal{N} \otimes \mathcal{B}_{ds}|\mathcal{B}_{ds}$ captures the combined information from the underlying probability space and the Brownian motion process. It allows us to consider events or random variables that depend on both the inherent randomness of the system (represented by $\mathcal{N}$) and the randomness introduced by the Brownian motion process (represented by $\mathcal{B}_{ds}|\mathcal{B}_{ds}$).

This construction is important in the theory of stochastic processes and stochastic differential equations because it provides a way to define and work with stochastic integrals, which are integrals with respect to Brownian motion or other stochastic processes. The measurability condition (A3) in the context provided ensures that the drift and diffusion functions are well-defined and can be integrated with respect to the Brownian motion process, taking into account both the underlying probability space and the Brownian motion itself.

Next, we will consider the bound on the distance between two stochastic integrals involving Wiener process $W(s)$. Let $X(s)$ and $Y(s)$ be two stochastic processes and let $d^2_{\infty}(\cdot, \cdot)$ denote L_2 distance between two processes. Then by Burkholder and Gundy inequality [28]:

$$E \sup_{u \in [0, t]} d^2_{\infty}(\langle \int_0^u X(s) dW(s) \rangle, \langle \int_0^u Y(s) dW(s) \rangle) \leq 4E \int_0^t d^2_{\infty}(\langle X(s) \rangle, \langle Y(s) \rangle) ds \quad (12)$$

Here, $\langle \cdot \rangle$ denotes the quadratic variation of a stochastic process, which is a measure of the total variation of the process. The integral $\int_0^u X(s) dW(s)$ represents the stochastic integral of $X(s)$ with respect to the Wiener process up to time u . The inequality states that the distance between the stochastic integrals of $X(s)$ and $Y(s)$ with respect to $W(s)$ is bounded by a multiple of the integral of their L_2 distance over time. The constant 4 in the inequality is a universal constant that arises from the properties of the Wiener process. (12) is a special case of the Burkholder-Davis-Gundy (BDG) inequality which we will explore further down the line and on which we will rely in other proofs. We will be using (12) and the assumptions above to show that (11) has a strong unique solution. We need to establish two points: 1. Existence of a strong solution. There exists a process $X(t)$ that solves the FSDE a.s. for all $t \geq 0$. 2. Uniqueness of the strong solution. The solution $X(t)$ is unique a.s. Using equation (5) and integration by parts technique, we can re-write (11):

$$\begin{aligned} X^\varepsilon(t) = & X_0 + \int_0^t f(s, X^\varepsilon(s)) ds + \langle \int_0^t (a [\int_0^t (t-s+\varepsilon)^{H-\frac{1}{2}} dB(s)] g(s, X^\varepsilon(s)) ds + \\ & \int_0^t \varepsilon^a g(s, X^\varepsilon(s)) dW(s) \rangle = X_0 + \int_0^t f(s, X^\varepsilon(s)) ds + \langle \int_0^t (a \varphi^\varepsilon(s) g(s, X^\varepsilon(s)) ds + \\ & \int_0^t \varepsilon^a g(s, X^\varepsilon(s)) dW(s) \rangle \end{aligned} \quad (13)$$

where $\varphi^\varepsilon(s) = \int_0^t (t-s+\varepsilon)^{a-1} dB(s)$, $\varepsilon > 0$, $a = H - 1/2$.

The proof begins by introducing the Picard iterations [24], denoted as $X_n^\varepsilon(t)$, which are obtained by iteratively applying the equation $X_n^\varepsilon(t) = X_0 + \int_0^t f(s, X_{n-1}^\varepsilon(s)) ds + \langle \int_0^t (a \varphi^\varepsilon(s) g(s, X_{n-1}^\varepsilon(s)) ds + \int_0^t \varepsilon^a g(s, X_{n-1}^\varepsilon(s)) dW(s) \rangle$, almost everywhere (a.e.). These iterations aim to find an approximation for the solution of the equation. To approximate $X(t)$, the Picard iteration scheme defines a sequence of iterates $X_n^\varepsilon(t)$ that should converge to $X(t)$ as n goes to infinity. Starting with $X_0^\varepsilon(t) = X_0$, we calculate $X_1^\varepsilon(t)$ by plugging X_0 into the right-hand side. Then $X_2^\varepsilon(t)$ is calculate using $X_1^\varepsilon(t)$ and so on. The key idea is that as n increases, the sequence $X_n^\varepsilon(t)$ should get closer and closer to the true solution $X(t)$, provided some condition on f , g and the initial data are satisfied to ensure convergence. The parameters ε and a are introduced to control the convergence behaviour, while φ is some approximating kernel. The Picard iterations provide a recursive way to get successively better approximations to the original equation's solution by iteratively updating an integral equation approximation. Define the sequence $j_n(t) = E \sup_{u \in [0, t]} d^2_{\infty}(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))$. The sequence measures the distance between two consecutive iterations. When we take $n=1$, and by the assumption that there is

$C > 0$ such that $\max\{d^\infty(f(t, w, u(0)), |g(t, w, u)|)\} \leq C(1 + d^\infty(u, \langle 0 \rangle))$ and from Theorem 4.1 in Llendeto (2007), we know that:

$$\begin{aligned} E|X^\varepsilon(t)|^2 &\leq E|X_0|^2 + \int_0^t f(s, X^\varepsilon(s))ds + \int_0^t g(s, X^\varepsilon(s))dB(s) \big|^2 \\ &\leq 3E|X_0|^2 + 3E \left| \int_0^t f(s, X^\varepsilon(s))ds \right|^2 \\ &\quad + 3E \left| \int_0^t g(s, X^\varepsilon(s))dB(s) \right|^2 \end{aligned}$$

Therefore, we have a constant 3 in:

$$\begin{aligned} J_1(t) &= E \sup_{u \in [0, t]} d_\infty^2 \left(\int_0^u f(s, X_0^\varepsilon(s))ds \right. \\ &\quad + \left\langle \int_0^u (a\varphi^\varepsilon(s)g(s, X_0^\varepsilon(s)))ds \right. \\ &\quad + \left. \int_0^u \varepsilon^a g(s, X_0^\varepsilon(s))dW(s) \right\rangle, \langle 0 \rangle \big) \\ &\leq 3E \sup_{u \in [0, t]} d_\infty^2 \left(\int_0^u f(s, X_0^\varepsilon(s))ds, \langle 0 \rangle \right) \\ &\quad + 3E \sup_{u \in [0, t]} d_\infty^2 \left(\left\langle \int_0^u (a\varphi^\varepsilon(s)g(s, X_0^\varepsilon(s)))ds, \langle 0 \rangle \right\rangle \right) \\ &\quad + 3E \sup_{u \in [0, t]} d_\infty^2 \left(\int_0^u \varepsilon^a g(s, X_0^\varepsilon(s))dW(s), \langle 0 \rangle \right) \end{aligned} \quad (14)$$

Using proposition (9) $\sup_{u \in [0, t]} d_\infty^p(L_{t,x}(u), L_{t,y}(u)) \leq t^{p-1} \int_0^t d_\infty^p(X(s), Y(s))ds$, a. e., proposition (12), linearity of the expectations operator, and since our metric is d_∞^p , so we can pull out constants as square, we can restate:

$$\begin{aligned} J_1(t) &\leq 3tE \left(\int_0^t d_\infty^2 f(s, X_0^\varepsilon(s)), \langle 0 \rangle \right) ds \\ &\quad + 3a^2 E \sup_{u \in [0, t]} d_\infty^2 \left(\left\langle \int_0^u (\varphi^\varepsilon(s)g(s, X_0^\varepsilon(s)))ds, \langle 0 \rangle \right\rangle + 3 \right. \\ &\quad \left. * 4 \varepsilon^{2a} E \left(\int_0^t d_\infty^2 g(s, X_0^\varepsilon(s)), \langle 0 \rangle \right) ds \right) \end{aligned}$$

Using assumption $\max\{d^\infty(f(t, w, u(0)), |g(t, w, u)|)\} \leq C(1 + d^\infty(u, \langle 0 \rangle))$, for $C > 0$, and properties of the expectation operator, we substitute this into the following integral expression to obtain the following. Also remember that $d_\infty^2(X_0^\varepsilon(s), \langle 0 \rangle)$ measures the squared maximum distance between some fuzzy set X_0^ε at a specific time s and zero fuzzy set, where all membership values are zero and utilise metric inner product. For a different version of this proof see [16].

$$\begin{aligned}
J_1(t) &\leq 3(T + 4\epsilon^{2a})[E(\int_0^t d_\infty^2 f(s, X_0^\epsilon(s)), \langle 0 \rangle) ds + E(\int_0^t d_\infty^2 g(s, X_0^\epsilon(s)), \langle 0 \rangle) ds \\
&\quad + 3a^2 E \sup_{u \in [0, t]} d_\infty^2 \left(\left\langle \int_0^u (\varphi^\epsilon(s) g(s, X_0^\epsilon(s))) ds, \langle 0 \rangle \right\rangle \right) \\
&\leq 3(T + 4\epsilon^{2a}) \left(\int_0^t E(C(1 + d_\infty^2(X_0^\epsilon), \langle 0 \rangle)) ds \right. \\
&\quad \left. + \int_0^t E(C(1 + d_\infty^2(X_0^\epsilon), \langle 0 \rangle)) ds \right. \\
&\quad \left. + 3a^2 E \sup_{u \in [0, t]} d_\infty^2 \left(\left\langle \int_0^u (\varphi^\epsilon(s) g(s, X_0^\epsilon(s))) ds, \langle 0 \rangle \right\rangle \right) \right) \quad (15) \\
&= 3 * 2C^2(T + 4\epsilon^{2a})[(1 + E\|X^\epsilon\|^0)^2]t \\
&\quad + 3a^2 E \sup_{u \in [0, t]} d_\infty^2 \left(\left\langle \int_0^u (\varphi^\epsilon(s) g(s, X_0^\epsilon(s))) ds, \langle 0 \rangle \right\rangle \right) \\
&= 6C^2(T + 4\epsilon^{2a})[(1 + E\|X^\epsilon\|^0)^2]t \\
&\quad + 3a^2 E \sup_{u \in [0, t]} d_\infty^2 \left(\left\langle \int_0^u (\varphi^\epsilon(s) g(s, X_0^\epsilon(s))) ds, \langle 0 \rangle \right\rangle \right)
\end{aligned}$$

This works only for $a = H - \frac{1}{2} > 0$ which was our assumption earlier and which is why it only works for $H > \frac{1}{2}$. For $H < \frac{1}{2}$ we will be using different integration methods. Next, we will utilise the following results. The d_H^2 distance, or Hausdorff distance, measures the maximum discrepancy between two sets by considering the minimum distance from each point in one set to the other set. This means that the d_H^2 distance accounts for the smallest possible distance between the two sets. On the other hand, the d_∞^2 distance, or supremum distance, measures the maximum difference between the values of two functions or sets. It calculates the maximum absolute difference between corresponding points of the two functions or sets. Since the d_H^2 distance considers the minimum distances, it takes into account the smallest possible difference between the two sets. Therefore, the d_H^2 distance will always be greater than or equal to the d_∞^2 distance, which measures the maximum difference. Using the above and by metric inner product:

$$\begin{aligned}
E \sup_{u \in [0, t]} d_\infty^2 \left(\left\langle \int_0^u (\varphi^\epsilon(s) g(s, X_0^\epsilon(s))) ds, \langle 0 \rangle \right\rangle \right) \\
\leq E \sup_{u \in [0, t]} d_H^2 \left(\left\{ \int_0^u (\varphi^\epsilon(s) g(s, X_0^\epsilon(s))) ds, \{0\} \right\} \right) \\
\leq E \sup_{u \in [0, t]} \left(\left(\int_0^u (\varphi^\epsilon(s) g(s, X_0^\epsilon(s))) ds \right) \right)^2 \quad (16)
\end{aligned}$$

We will now use the famous Hölder inequality [19]:

$$\int_a^b |f(x)g(x)|dx \leq (\int_a^b |f(x)|^p dx)^{1/p} (\int_a^b |g(x)|^q dx)^{1/q} \text{ and re-arrange terms r with s.}$$

$$\begin{aligned}
E \sup_{u \in [0, t]} \left(\left(\int_0^u (\varphi^\epsilon(s) g(s, X_0^\epsilon(s))) ds \right) \right)^2 \\
= E \sup_{u \in [0, t]} \left(\int_0^u \int_0^t (s - r + \epsilon)^{a-1} dW(r) g(s, X_0^\epsilon(s)) ds \right)^2
\end{aligned}$$

Since the integration variables are independent, we can use Fubini's theorem [24] to rearrange the integration variables. This is the rule for swapping independent integration variables. If the function being integrated is continuous over a rectangular region $R = [a, b] \times [c, d]$, and the integral is of the form $\int \int_R f(x, y) dx dy$, then the order of integration can be swapped, resulting in $\int \int_R f(x, y) dy dx$ or $\int \int_R f(x, y) dx dy$. For example, consider the following double integral: $\int \int_R \exp(x^2 + y^2) dx dy$, where R is the region defined by $0 \leq x \leq 1$ and $0 \leq y \leq 2$. If we integrate with respect to x first, we have: $\int \int_R \exp(x^2 + y^2) dx dy = \int_0^2 \int_0^1 \exp(x^2 + y^2) dx dy$. If we integrate

with respect to y first, we have: $\int \int_R \exp(x^2 + y^2) dx dy = \int_0^1 \int_0^2 \exp(x^2 + y^2) dy dx$. In this example, the integrand $\exp(x^2 + y^2)$ is continuous over the rectangular region R , and therefore, we can swap the order of integration without affecting the result. It is important to note that not all integrals allow for swapping the order of integration. The conditions for applying Fubini's theorem include continuity of the integrand, boundedness of the region of integration, and absolute integrability. If these conditions are not met, swapping the order of integration may lead to incorrect results.

We will now expand the work of [16]. So first we will use Fubini to re-arrange r with s in this integral. Then we will apply Itô isometry or the Itô integral rule. It is a key tool in stochastic calculus and is used to calculate the expected value of stochastic integrals involving Brownian motion or other forms of stochastic processes. The identity essentially says that the square of the difference between two points in a Brownian motion path, raised to the power of $(\alpha-1)$, can be expressed as the double stochastic integral of $(\alpha-2)$ powers of the differences between the same points and the endpoints of the integration range. Intuitively, this identity can be understood as follows: the Brownian motion process is continuous and has infinite variation, meaning that the difference between two points in a path can take on any value in a certain range with positive probability. The Itô isometry [29] allows us to calculate the expected value of the square of these differences, which is a key step in deriving stochastic integrals and differential equations. In the rearrangement of terms given in the original question, the Itô isometry is used to express the term $(s - r + \varepsilon)^{\alpha-1}$ as a double stochastic integral, which allows for the interchange of the order of integration and ultimately results in the desired form of the expression. $(W(t) - W(s))^{\alpha-1} = \int_0^s (r - s + \varepsilon)^{\alpha-2} dW(r) \int_0^t (u - t + \varepsilon)^{\alpha-2} dW(u)$. Use this identity with $t=r$ and $s=s$, we have, $(s - r + \varepsilon)^{\alpha-1} = \int_0^r (u - r + \varepsilon)^{\alpha-2} dW(u) \int_0^s (v - s + \varepsilon)^{\alpha-2} dW(v)$ substitute this into the integral, we can now interchange the order of integration between s and r . There are three integrals in the intermediate step because the stochastic integral involving the Brownian motion process requires two integrals: one with respect to time and one with respect to the Brownian motion itself. There are two stochastic integrals involving Brownian motion: one over the range $[u, s]$ and one over the range $[r, t]$. This requires a total of three integrals: one over the range $[u, t]$ with respect to time, and two with respect to Brownian motion over the ranges $[u, s]$ and $[r, t]$, respectively. When the order of integration is interchanged between s and r , the resulting expression involves a double stochastic integral over the ranges $[u, s]$ and $[r, t]$, respectively. This requires two integrals with respect to Brownian motion and one integral with respect to time, which is consistent with the three integrals in the original expression.

The identity introduced with $t = s$ and $s = v$ allows for the simplification of the expression by expressing the term $(s - r + \varepsilon)^{\alpha-1}$ as a double stochastic integral over the ranges $[s, u]$ and $[r, v]$, respectively. This ultimately leads to the desired form of the expression with only two integrals involving Brownian motion and one integral with respect to time.

$$\begin{aligned} E \sup_{u \in [0, t]} \left(\int_0^u \int_0^s (s - r + \varepsilon)^{\alpha-1} dW(r) g(s, X_0^\varepsilon(s)) ds \right)^2 \\ = E \sup_{u \in [0, t]} \left(\int_0^u \int_s^u g(r, X_0^\varepsilon(r)) dr (r - s + \varepsilon)^{\alpha-1} dW(s) \right)^2 \\ = \end{aligned}$$

Using Itô isometry [29] i.e., that $(dW)^2 = dt$, and for upper bound we simply multiply the expectation by 4 [24], we can simplify further the stochastic integral:

$$\begin{aligned}
& E \sup_{u \in [0, t]} \left(\int_0^u \int_s^u g(r, X_0^\varepsilon(r)) dr (r-s+\varepsilon)^{a-1}] dW(s) \right)^2 \\
& \leq 4E \left(\int_0^t \int_s^t g(r, X_0^\varepsilon(r)) (r-s+\varepsilon)^{a-1}] dr \right)^2 ds \\
& \leq 4E \int_0^t \left(\int_s^t g^2(r, X_0^\varepsilon(r)) (r-s+\varepsilon)^{a-1}] dr \right) \left(\int_s^t (r-s+\varepsilon)^{a-1}] dr \right) ds \\
& \leq 4E \int_0^t g^2(r, X_0^\varepsilon(r)) (r-s+\varepsilon)^{a-1}] dr \left[\frac{(r-s+\varepsilon)^a}{a} \right]_s^t ds \\
& \leq 4E \int_0^t g^2(r, X_0^\varepsilon(r)) (r-s+\varepsilon)^{a-1}] dr \left[\frac{(t-s+\varepsilon)^a}{a} - \frac{\varepsilon^a}{a} \right] ds, \text{ since } t, s, \varepsilon \geq 0 \\
& \therefore \frac{(t-s+\varepsilon)^a}{a} - \frac{\varepsilon^a}{a} \leq \frac{(t+\varepsilon)^a}{a} \\
& \therefore \text{following similar argument:}
\end{aligned} \tag{17}$$

$$\begin{aligned}
& 4E \int_0^t g^2(r, X_0^\varepsilon(r)) (r-s+\varepsilon)^{a-1}] dr \left(\frac{(t+\varepsilon)^a}{a} \right) ds \leq \\
& 4 \frac{(t+\varepsilon)^a}{a} E \int_0^t g^2(r, X_0^\varepsilon(r)) (r-s+\varepsilon)^{a-1}] ds dr \leq \\
& 4 \frac{(t+\varepsilon)^a}{a} E \int_0^t g^2(r, X_0^\varepsilon(r)) \left[\frac{(t+\varepsilon)^a}{a} \right] dr \leq \frac{4(t+\varepsilon)^{2a}}{a^2} E \int_0^t g^2(r, X_0^\varepsilon(r)) dr.
\end{aligned}$$

Because this is a linear integral, we can take the term $\frac{(t+\varepsilon)^a}{a}$ outside the integral. Now using Cauchy-Schwarz inequality [24]:

$\max\{d^\infty(f(t, w, u(0)), |g(t, w, u)|)\} \leq C(1 + d^\infty(u, \langle 0 \rangle))$, we can evaluate the last expression:

$$\begin{aligned}
& \frac{4(t+\varepsilon)^{2a}}{a^2} E \int_0^t g^2(r, X_0^\varepsilon(r)) dr \\
& \leq \frac{4(T+\varepsilon)^{2a}}{a^2} \left(\int_0^t E(C(1 + d_\infty^2(X_0^\varepsilon), \langle 0 \rangle)) ds \right. \\
& \quad \left. + \int_0^t E(C(1 + d_\infty^2(X_0^\varepsilon), \langle 0 \rangle)) ds \right) \\
& \leq 2C^2 \frac{4(T+\varepsilon)^{2a}}{a^2} [(1 + E\|X_0^\varepsilon\|^2)t] \\
& \leq C^2 \frac{8(T+\varepsilon)^{2a}}{a^2} [(1 + E\|X_0^\varepsilon\|^2)t]
\end{aligned} \tag{18}$$

We will now substitute this result in $J_1(t)$, simplify and take exponential out of the bracket.

$$\begin{aligned}
J_1(t) & \leq 6C^2(T + 4\varepsilon^{2a})[(1 + E\|X_0^\varepsilon\|^2)t] \\
& \quad + 3a^2C^2 \frac{8(T+\varepsilon)^{2a}}{a^2} [(1 + E\|X_0^\varepsilon\|^2)t] \\
& \leq (6C^2(T + 4\varepsilon^{2a}) + 3C^28(T+\varepsilon)^{2a})[(1 + E\|X_0^\varepsilon\|^2)t] \\
& \leq 6C^2(T + 4\varepsilon^{2a} + 4(T+\varepsilon)^{2a})[(1 + E\|X_0^\varepsilon\|^2)t]
\end{aligned}$$

We will explain what this all means and why we need it. The first term $J_1(t)$ in the estimate is an upper bound on the error between the true solution $X(t)$ and the numerical solution $X^\varepsilon(t)$ obtained at the n th time step. It represents the contribution to the error due to the approximation made at the current time step. The inequality for $J_1(t)$ follows from the error estimate for the numerical method used to compute $X^\varepsilon(t)$ at the current time step. This estimate indicates that the error in $X^\varepsilon(t)$ at the current time step is proportional to the time step size t and to a constant factor that depends on the parameters of the numerical method and on the regularity of the solution $X(t)$. Similarly, we can obtain the same result for $J_{n+1}(t)$ which in the estimate is an upper bound on the error between the numerical solution $X^\varepsilon(t)$ obtained at the $(n+1)$ th time step and the numerical solution $X^\varepsilon(t)$ obtained

at the n th time step. It represents the contribution to the error due to the propagation of the error from the current time step to the next time step. The inequality for the next time step which we will prove now is:

$$J_{n+1}(t) \leq 3(t + 4\varepsilon^{2a} + 4(T + \varepsilon)^{2a})L^2E \int_0^t d_\infty^2(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))ds$$

It follows from a stability estimate for the numerical method used to compute $X^\varepsilon(t)$ at the $(n+1)$ th time step. This estimate indicates that the error in $X^\varepsilon(t)$ at the $(n+1)$ th time step is proportional to the time step size t , to a constant factor that depends on the parameters of the numerical method and on the regularity of the solution $X(t)$, and to the norm of the difference between the numerical solutions for $X_n^\varepsilon(t)$ and $X_{n-1}^\varepsilon(t)$ obtained at the n th and $(n-1)$ th time steps respectively. Heuristically, this makes sense because as the time steps are reduced, the numerical method is able to capture more accurately the behaviour of the true solution $X(t)$ at each time step. Second, let's consider the constant factor that depends on the parameters of the numerical method and on the regularity of the solution $X(t)$. This constant factor represents the accuracy of the numerical method itself and is determined by the choice of the numerical scheme, the order of accuracy of the numerical method, and other parameters specific to the method. The regularity of the solution $X(t)$ also plays a role in determining the accuracy of the numerical method, since more regular solutions can be approximated more accurately than less regular ones. Finally, let's consider the norm of the difference between the numerical solutions $X_n^\varepsilon(t)$ and $X_{n-1}^\varepsilon(t)$ obtained at the n th and $(n-1)$ th time steps, respectively. This term accounts for the propagation of the error from the previous time step to the current time step. If there is a large difference between the numerical solutions at the two-time steps, this indicates that the error has been propagated and has accumulated over time, leading to a larger error in the numerical solution at the current time step. The norm of this difference is used to quantify the magnitude of the error that has been propagated from the previous time step. Therefore, the estimate for the error in $X_n^\varepsilon(t)$ at the $(n+1)$ th time step takes into account the time step size, the accuracy of the numerical method, and the propagation of the error from the previous time step. In summary, the term $J_{n+1}(t)$ accounts for the propagation of the error from the current time step to the next time step and provides an upper bound on the total error in the numerical solution $X^\varepsilon(t)$ at the $(n+1)$ th time step. By summing up the contributions from all time steps, we obtain an estimate for the total error in the numerical solution $X^\varepsilon(t)$ over the entire time interval $[0, T]$.

This is our proof of the above. Let $X_n^\varepsilon(t)$ be the numerical solution obtained at n -th time step and $X_{n-1}^\varepsilon(t)$ be the numerical solution obtained at the $(n-1)$ th time step. Then the error between these two solutions can be written as:

$\varepsilon_n(t) = X_n^\varepsilon(t) - X_{n-1}^\varepsilon(t)$, using the difference equation for numerical method, we can write: $\varepsilon_{n+1}(t) = X^\varepsilon(t + \Delta t) - X_n^\varepsilon(t)$, where $X^\varepsilon(t + \Delta t)$ is the true solution at $(n+1)$ th step and Δt is the time step size. Then the error in the numerical solution $X^\varepsilon(t)$ at $(n+1)$ th time step can be written as: $\varepsilon_{n+1}(t) - \varepsilon_n(t) = X^\varepsilon(t + \Delta t) - X_n^\varepsilon(t) - (X_n^\varepsilon(t) - X_{n-1}^\varepsilon(t)) = X^\varepsilon(t + \Delta t) - 2X_n^\varepsilon(t) + X_{n-1}^\varepsilon(t)$.

Taking the norm $||\cdot||$ of both sides and using triangle inequality property of norms in vector spaces, $||u + v|| \leq ||u|| + ||v||$, we obtain:

$$||\varepsilon_{n+1}(t)|| \leq ||X^\varepsilon(t + \Delta t) - 2X_n^\varepsilon(t)|| + ||(X_n^\varepsilon(t) - X_{n-1}^\varepsilon(t))||$$

A function $f(x)$ is said to be Lipschitz continuous if there exists a constant K such that:

$|f(x_1) - f(x_2)| \leq K|x_1 - x_2|$ for all x_1 and x_2 in the function's domain [24]. Intuitively, this means that the function's rate of change is bounded by a constant factor K , which is independent of the size of the interval over which the function is measured. Using Lipschitz continuity of the function f in the differential equation, we can write:

$$X^\varepsilon(t + \Delta t) - 2X_n^\varepsilon(t) \leq L^2\Delta t ||f(t, X_n^\varepsilon(t), u) - f(t, X_{n-1}^\varepsilon(t), u)||$$

Where L^2 is the Lipschitz constant. It is the smallest possible value of K that satisfies the Lipschitz condition $|f(x_1) - f(x_2)| \leq K|x_1 - x_2|$. In other words, it is the smallest possible upper bound on the rate of change of the function over its entire domain and it provides a measure of how rapidly the function changes over its domain. We will use it in the following context. If the right-hand side of a differential equation is Lipschitz continuous with a constant K , then numerical methods that a time step t , will converge to the true solution at a rate proportional to t/K . Using the assumption on the distance between the initial condition u_0 and the true solution $X(t)$, we can write:

$$\|f(t, X_n^\varepsilon(t), u) - f(t, X_{n-1}^\varepsilon(t), u)\| \leq d_\infty^2(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))$$

Therefore, we can write the following and integrate both sides over time interval $[t(n), t(n+1)]$:

$$\begin{aligned} \|\varepsilon_{n+1}(t)\| &\leq L^2 d_\infty^2(X_n^\varepsilon(t) - X_{n-1}^\varepsilon(t)) + \|X^\varepsilon(t) - X_{n-1}^\varepsilon(t)\| \\ \int_0^t \|\varepsilon_{n+1}(t)\| dt &\leq L^2 \Delta t \int_0^t d_\infty^2[(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))]^{1/2} \left[\int_0^t \|X_n^\varepsilon(t) - X_{n-1}^\varepsilon(t)\|^2 dt \right]^{1/2} \end{aligned}$$

Taking the expectation on both sides and using the fact that the expectation operator is linear, we obtain:

$$\begin{aligned} E \int_0^t \|\varepsilon_{n+1}(t)\| dt &\leq L^2 \Delta t E \int_0^t d_\infty^2[(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))]^{1/2} E \left[\int_0^t \|X_n^\varepsilon(t) - X_{n-1}^\varepsilon(t)\|^2 dt \right]^{1/2} \end{aligned}$$

Use the fact that the square of the norm of the difference between two solutions is bounded by 3 times the sum of the squares of their individual norms. This is a special case of Cauchy-Schwarz inequality $\|u - v\|^2 \leq 3(\|u\|^2 + \|v\|^2)$ [24]. This is true because of the dot product and the norm. The dot product is $u \cdot v = \|u\| \|v\| \cos(b)$, where b is the angle between u and v . Using this expression, we can write the square of the norm of the difference between u and v as: $\|u - v\|^2 = (u - v) \cdot (u - v) = \|u\|^2 - 2(u \cdot v) + \|v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\| \|v\| \cos(b)$, where we have used the distributive property of dot product. Note that the $\cos(b)$ is bounded by -1 and 1 so: $-2\|u\| \|v\| \leq -2\|u\| \|v\| \cos(b) \leq 2\|u\| \|v\|$. Adding this inequality to the expression $\|u - v\|^2$, we get: $\|u - v\|^2 + 2\|u\| \|v\| = \|u\|^2 + \|v\|^2 - 2(u \cdot v) + 2\|u\| \|v\| = \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| - 2(u \cdot v)$ we obtain: $\|u - v\|^2 + 2\|u\| \|v\| \leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\|$. Then take the square root we get: $\|u - v\| \leq \|u\| + \|v\|$ and use the fact that $\|u - v\|$ is non-negative, and that $(\|u\| + \|v\|)^2 \leq 3(\|u\|^2 + \|v\|^2)$ for all real numbers $\|u\|$ and $\|v\|$, to get $\|u - v\|^2 \leq (\|u\| + \|v\|)^2 \leq 3(\|u\|^2 + \|v\|^2)$. Now we continue with the original proof.

$$\begin{aligned} E \int_0^t \|\varepsilon_{n+1}(t)\| dt &\leq 3\Delta t [(T_n + 4^{2a} + 4(T_n + \varepsilon)^{2a})^{\frac{1}{2}} L^2 E \int_0^t d_\infty^2[(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))]^{1/2} dt]^{1/2} \end{aligned} \quad (19)$$

Therefore, we have derived the expression for $j_{n+1}(t)$ term in the error estimate:

$$\begin{aligned} j_{n+1}(t) &\leq 3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2 E \int_0^t d_\infty^2[(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))]^{1/2} dt]^{1/2} \\ &\leq 3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2 E \int_0^t d_\infty^2[(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))]^{1/2} dt]^{1/2} \\ &\leq 3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2 E \int_0^t d_\infty^2[(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))]^{1/2} dt]^{1/2} \\ &\leq 3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2 E \int_0^t d_\infty^2[(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))]^{1/2} dt]^{1/2} \\ &\leq 3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2 E \int_0^t d_\infty^2[(X_n^\varepsilon(t), X_{n-1}^\varepsilon(t))]^{1/2} dt]^{1/2} \end{aligned} \quad (20)$$

This inequality holds because the function $j_n(s)$ is defined as the supremum distance between values of the process X over the interval from n to $n+1$, as measured by the metric function d_∞^2 . Therefore, integral of $j_n(s)$ over the interval t_0 to s gives an upper bound on the supremum distance between values of the process X over the same interval. The left-hand side of the inequality gives an upper bound on the function $j_{n+1}(t)$ while the right hand side gives a lower bound.

We will now apply Chebyshev inequality to the metric function [24]. The Chebyshev inequality is a probabilistic inequality that relates the probability of a random variable deviating from its mean to its variance. Specifically, the Chebyshev inequality states that for any random variable X with finite mean μ and finite variance σ^2 , the probability that X deviates from its mean by more than k standard deviations is bounded by $1/k^2$. $P(|X - \mu| > k\sigma) \leq 1/k^2$. We will apply this inequality to the metric function

$d_\infty^2 [(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))]$ to obtain a probabilistic bound on the supremum distance between values of the process X over the interval n and $n+1$.

Proof. Define Y to be the random variable such that: $Y = \sup_{u \in [0, s]} d_\infty^2 [(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))]$. Then we have:

$$\begin{aligned}\mu &= E(Y) = E\left(\sup_{u \in [0, s]} d_\infty^2 [(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))]\right) \text{ and} \\ \sigma^2 &= \text{Var}(Y) = \text{Var}\left(\sup_{u \in [0, s]} d_\infty^2 [(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))]\right).\end{aligned}$$

Since the metric function is a distance measure, it is non-negative, and therefore Y is also non-negative. We can choose $k=2^n$ to obtain $P(Y > 2^n \mu) \leq \text{Var}(Y)/(2^n \mu)^2$, and since Y is non-negative, we have $E(Y) = \mu \geq 0$, therefore $(2^n \mu)^2 \geq 0$. This means that the denominator of the fraction is non-negative. We can multiply and invert the inequality to obtain:

$$\begin{aligned}P\left(\sup_{u \in [0, s]} d_\infty^2 [(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))] \geq \frac{1}{2^n}\right) &= P(Y > 2^n \mu) \leq \frac{1}{4^n} \\ &\leq (2^n)^2 \frac{\text{Var}\left(\sup_{u \in [0, s]} d_\infty^2 [(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))]\right)}{2^n E\left(\sup_{u \in [0, s]} d_\infty^2 [(X_n^\varepsilon(u), X_{n-1}^\varepsilon(u))]\right)^2} \\ &= 2^n \frac{\text{Var}(J_n(T))}{E(J_n(T))^2} \leq 2^n J_n(t)\end{aligned}\quad (21)$$

We will now use Borel-Cantelli lemma [30]. Let $\{E_n\}$ be a sequence of events in the sample space Ω then if the sum of the probabilities of the events in the sequence is finite then the probability of the intersection of all the events is zero i.e., $\sum_{n=1}^{\infty} \text{Pr}(E_n) < \infty$ then $\text{Pr}(E^s) = 0$ where $E^s = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m$. If the events $\{E_n\}$ are independent, then:

$$\sum_{n=1}^{\infty} \text{Pr} = \infty \text{ then } \text{Pr}(E^s) = 1 \text{ where } E^s = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m.$$

Conversely, if the events are independent and the sum of their probabilities diverges, then the probability of the intersection of all the events is 1. To understand this, consider the first case where the sum probabilities is finite. We can define a new sequence of events $\{F_n\}$ as follows: $F_n = \bigcup_{m \geq n} E_m$. That is the union of all the events in the sequence E_n starting from index n . Intuitively, F_n represents the event that at least one of the events E_n occurs starting from index n . Since the sum of the probabilities of the events in the sequence is finite, we have: $\sum_{n=1}^{\infty} P(E_n) < \infty$. This implies that the sum of the probabilities of the events in the sequence is also finite since F_n is the union of a finite number of events E_n . Therefore, we have:

$$\begin{aligned}\sum_{n=1}^{\infty} \text{Pr}(E_n) &= \sum_{n=1}^{\infty} \left(\bigcup_{m \geq n} E_n\right) \leq \sum_{n=1}^{\infty} \sum_{m \geq n} P(E_n) \\ &= \sum_{n=1}^{\infty} P(E_n) + P(E_{n+1}) + \dots\end{aligned}$$

The last step follows from the fact that the events E_n are non-negative and therefore the sum can be rearranged. Now, let's consider the event $E^s = \bigcap_{n=1}^{\infty} F_n$. Intuitively, E^s represents the event that infinitely many of the events E_n occur. We can write E^s as a countable intersection of the events $\{F_n\}$, i.e.: $E^s = \bigcap_{n=1}^{\infty} F_n = \limsup F_n$ where $\limsup$ is the limit supremum of the sequence $\{F_n\}$. In other words, E^s is the set of all outcomes that belong to infinitely many of the events F_n . Since the sum of the probabilities of the events in the sequence $\{F_n\}$ is finite, we can apply the first Borel-Cantelli lemma to conclude that: $\text{Pr}(\limsup F_n) = 0$. Therefore, we have shown that $\text{Pr}(E^s) = \text{Pr}(\limsup F_n) = 0$, which completes the proof of the first case. The second case we don't need here because $J_n(T)$ is convergent. Having shown the convergence, we will now apply the Borel-Cantelli lemma. Let $\{f_n(x)\}$ be a sequence of measurable functions on $[0,1]$ with $|f_n(x)| < \infty$ for a.e. $x \in [0,1]$. Then there exists a sequence $\{c_n\}$ of positive real numbers such that $f_n(x)/c_n \rightarrow 0$ a.e. $x \in [0,1]$. Suppose there is a sequence such that $m(E_n) \leq 2^{-n}$. Then from lemma it follows that $P(|f(x)| > N/n) > 2^{-n}$. Going back to our proof, and using Borel-Cantelli lemma:

$P \sup_{u \in [0, s]} d_{\infty}^2 [(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u)) \geq \frac{1}{\sqrt{2}^n} \text{ infinitely often}] = 0$. This means that the probability of the event that the supremum of the d^2 distance between $(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u))$ is greater than $\frac{1}{\sqrt{2}^n}$ for infinitely many values of n is zero. Following from this, we can say that:

$$\sup_{u \in [0, s]} d_{\infty}^2 [(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u)) \leq \frac{1}{\sqrt{2}^n} \quad (22)$$

We will now formally prove this important result. The value $1/\sqrt{2}$ comes from the fact that we are interested in the probability of the event that the supremum of the d^2 distance between $(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u))$ is greater than $(1/\sqrt{2})^n$ for infinitely many values of n . The value $1/\sqrt{2}$ is chosen because it is a convenient threshold that ensures the convergence of the series $\sum_{n=1}^{\infty} J_n(t)$ in the statement of the problem. In particular, the convergence of the series implies that the sequence $X_n^{\varepsilon}(u)$ converges almost surely in the d^2 metric, which is a stronger condition than the convergence almost surely in the Euclidean metric since the latter is a special case of d^2 metric. The choice of the threshold $(1/\sqrt{2})^n$ is motivated by the fact that it ensures that the rate of growth of $E(\sup_{u \in [0, s]} d_{\infty}^2 [(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u))$ is

bounded by $\sqrt{2}$ times the logarithm of n . The logarithm comes from the asymptotic behaviour of the harmonic series, which is related to the rate of growth of the sum of the variances of the random variables in the sequence $X_n^{\varepsilon}(u)$. The sum of the variances of the random variables is equal to n , since they are independent standard normal random variables. Therefore, the sum of the variances of the random variables in the sequence $\{X_n^{\varepsilon}(u)\}$ up to time n is equal to: $\text{Var} \sum_{j=1}^n \varepsilon_{jn}(T) u_j = \sum_{j=1}^n u_j^2 \text{Var}[\varepsilon_j(T)] = n \sum_{j=1}^n u_j^2$.

We can use the fact that $\sum_{j=1}^n u_j^2$ is bounded by 1 for all n , since u belongs to the unit ball in $\mathbb{R}^d$. Therefore, we have $\text{Var} \sum_{j=1}^n \varepsilon_j^n(T) u_j \leq n$. Using this result, we can re-write:

$$E(\sup_{u \in [0, s]} d_{\infty}^2 [(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u)) \leq \sqrt{2} \sum_{n=1}^{\infty} [\text{Var}[\sum_{j=1}^n \varepsilon_j^n(T) u_j]]^{1/2} \leq \sqrt{2} \sum_{n=1}^{\infty} \sqrt{n}, \text{ where the last}$$

inequality follows from the fact that the sum of the variances of the random variables in the sequence $\{X_n^{\varepsilon}(u)\}$ up to time n is bounded by n . We also use a well-known result in mathematics originally proved by Euler that the sum of the harmonic series i.e., the sum of the reciprocals of the positive integers is asymptotically bound by the natural logarithm of the number of terms in the series. $1 + 1/2 + 1/3 + \dots + 1/n \approx \ln(n) + \gamma$ as $n \rightarrow \infty$, where γ is the Euler-Mascheroni constant. So, in this context above the sum of $\sqrt{n}$ grows slowly compared to the harmonic series and is bounded by $\ln(n)$. So $\sum_{n=1}^{\infty} \frac{\sqrt{n}}{\ln(n)} = \infty$ which follows from the fact that the integral $1/\sqrt{x}$ from 1 to ∞ is infinite, and the comparison test for series. Heuristically, the numerator grows faster than the denominator, so the sum of the series is infinite. This is why we choose the threshold $\frac{1}{\sqrt{2}^n}$ in the statement of the Borel-Cantelli lemma to ensure that the rate of growth of $E(\sup_{u \in [0, s]} d_{\infty}^2 [(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u)) \leq$

$\sqrt{2} \sum_{n=1}^{\infty} [\text{Var}[\sum_{j=1}^n \varepsilon_{jn}(T) u_j]]^{1/2} \leq \sqrt{2} \sum_{n=1}^{\infty} \sqrt{n} \leq \sqrt{2} \ln(n)$ is bounded. Now, recall that the sequence $\{X_n^{\varepsilon}(u)\} = X_{n-1}^{\varepsilon}(u) + \sum_{j=1}^n \varepsilon_j^n(t)$ represents independent standard normal variables. Therefore, the distance between $d^2(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u)) = \sum_{j=1}^n (\varepsilon_j^n(t) - \varepsilon_j^{n-1}(t))^2 u_j^2$. Now we will use the fact that these are independent random variables with mean 0 and variance 1. Therefore $\text{Var}(\sum_{j=1}^n (\varepsilon_j^n(t) - \varepsilon_j^{n-1}(t))^2)$ is equal to the sum of their variances which is 2 and the covariance term is zero. Using this we can write that: $E[d^2(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u))] = \sum_{j=1}^n (\varepsilon_j^n(t) - \varepsilon_j^{n-1}(t))^2 u_j^2 = 2 \sum_{j=1}^n u_j^2$ taking square root, we have:

$$E(\sup_{u \in [0, s]} d_{\infty}^2 [(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u)) \leq \sqrt{2} E[\sum_{n=1}^{\infty} u_j^2]^{1/2} \leq \sqrt{2} \sum_{j=1}^n E(|u_j|) \leq \sqrt{2} \ln(n)$$

Therefore, by the Borel-Cantelli lemma, we can conclude that the probability of the event that the supremum of the d^2 distance between $X_n^{\varepsilon}(u)$ and $X_{n-1}^{\varepsilon}(u)$ is greater than $\frac{1}{\sqrt{2}^n}$ for infinitely many values of n is zero. So we have proven that:

$$\sup_{u \in [0, s]} d_{\infty}^2 [(X_n^{\varepsilon}(u), X_{n-1}^{\varepsilon}(u)) \leq \frac{1}{\sqrt{2}^n}$$

We will now complete the proof of the existence of strong and unique solution. We know that as n goes to infinity then $X_n^\varepsilon(t) \rightarrow \text{approaches to } X^\varepsilon(t)$ so $d_\infty^2(X_n^\varepsilon(t), X^\varepsilon(t)) = 0$ as $d(a,a) = a - a = 0$. In other words, $X_n^\varepsilon(t)$ and $X^\varepsilon(t)$ become the same. Let's take $X^\varepsilon(t)$ as some continuous fuzzy stochastic process. Then, by the above $d_\infty^2[(X_n^\varepsilon(t))^a, X_{n-1}^\varepsilon(u)^a] \rightarrow 0$ as $n \rightarrow \infty$. Hence, as n goes to infinity:

$$E \sup_{t \in I} [d_\infty^2(X_n^\varepsilon(t), X^\varepsilon(t)) + d_\infty^2(X_n^\varepsilon(t), X_0^\varepsilon) + \int_0^t f(s, X^\varepsilon(s)) ds + \langle \int_0^t g(s, X^\varepsilon(s)) dB_H^\varepsilon(s) \rangle]^2 \Rightarrow 0$$

This shows the existence of strong solution.

$$\sup_{t \in I} [d_\infty^2[(X^\varepsilon(t), X_0^\varepsilon) + \int_0^t f(s, X^\varepsilon(s)) ds + \langle \int_0^t g(s, X^\varepsilon(s)) dB_H^\varepsilon(s) \rangle]] = 0 \quad (23)$$

Now, we will show the uniqueness of the solution. For this we need Gronwall's lemma and inequality [31]. Let $y(t)$, $f(t)$, $g(t)$ be non-negative functions on a closed interval $[a, b]$, where $a < b$. If $f(t)$ satisfies inequality: $f(t) \leq g(t) + \int_a^b h(s)f(s)ds$ for all $t \in [a, b]$ then it follows that:

$f(t) \leq g(t) + \int_a^b h(s)f(s)e^{\int_s^t h(u)du}ds$ for all $t \in [a, b]$. A special case of Gronwall's lemma arises when $g(t)$ is identically zero on the interval $[a, b]$ which then simplifies to: $f(t) \leq \int_a^b h(s)f(s)ds$ which provides an upper bound on $f(t)$. Gronwall's inequality provides an upper bound on the function $f(t)$ in terms of the constant A and the exponential of the integral of $g(s)$ over the interval $[a, t]$. The inequality guarantees that the growth of $f(t)$ is controlled by the exponential function. The general form of the inequality is as follows: Let $f(t)$ be a non-negative, continuous function on a closed interval $[a, b]$, where $a < b$. If $f(t)$ satisfies the inequality: $f(t) \leq A + \int_a^b f(s)g(s)ds$, where A is a constant and g is a non-negative, continuous function, then it follows that: $f(t) \leq A \int_a^b e^{g(s)ds}$. Going back to our prove of uniqueness. Consider that:

$$J(t) = E \left[\sup_{u \in [0, t]} d_\infty^2(X^\varepsilon(u), Y^\varepsilon(u)) \right], \text{ where we already have an estimation or upper bound}$$

which is:

$$j(t) \leq 3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2 E \int_0^t d_\infty^2[(X^\varepsilon(s), Y^\varepsilon(s))] ds] \\ \leq 3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2 \int_0^t j_n(s) ds]$$

In the context of the $j(t)$ function, we apply Gronwall's inequality with $A = 0$ and $B = 3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2]$. Using $f(t) \leq A \int_a^b e^{g(s)ds}$, we can state that,

$j(t) \leq 0 * e^{3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2] t} = 0$. Therefore, $j(t) = 0$ for all t within the given interval. This bound shows that the function $j(t)$ grows exponentially in time, with rate of growth determined by the constant $3[(T + 4^{2a} + 4(T + \varepsilon)^{2a}) L^2]$. However, the growth of $j(t)$ is limited by its initial value $j(0)$. If $J(t)=0$ for all t in I , this means that the distance between the two solutions of the stochastic differential equation is zero for all times in the interval I . But since $j(t)$ is non-negative, this implies that $j(t)=0$ for all $t \geq t_0$ where t_0 is the endpoint of the interval I . Therefore, the two solutions of the equation coincide for all times $t \geq t_0$ which means they become arbitrarily close to each other as time goes to infinity. Prove of uniqueness is completed and

$$J(t) = E \left[\sup_{u \in [0, t]} d_\infty^2(X^\varepsilon(u), Y^\varepsilon(u)) \right] = 0 \text{ a. e.,}$$

We will next prove that $X^\varepsilon(t)$ converges to the solution $X(t)$ in L^2 space as $\varepsilon \rightarrow 0$ and $t \in [0, T]$ which is fundamental if we want to apply Itô's lemma later on. However, first we need to prove that for every $\varepsilon > 0$ and $0 < a < 1/2$,

$$\int_s^t ((r - s + \varepsilon)^{a-1} - (r - s)^{a-1}) dr \leq \frac{a+1}{a} \varepsilon^a$$

We will need this inequality to use in the former proof. We will employ the finite-increments formula in the proof. Let $f(x)$ be a differentiable function defined on an interval $[a, b]$ and let δ be a small increment. The formula states that: $f(x + \delta) - f(x) = f'(x)\delta + \theta\delta$, where $f'(x)$ represents the first derivative of f evaluated between some point x and $x + \delta$, and θ is a value between 0 and 1. In

other words, the change in the function's value over the interval $[x, x+\delta]$ is approximately equal to the derivative of the function times the increment δ , plus an additional term $\theta\delta$ accounting for the difference between the true change and the linear approximation. In the context of this proof, we will use the finite-increments formula to analyse the function $f(x) = x^{a-1}$ to obtain $(x+\varepsilon)^{a-1} - x^{a-1} = [(a-1)(x+\theta\varepsilon)^{(a-2)}\varepsilon]$. By applying formula to this function, we can derive relationships between the increments in the function's values and the derivative of the function evaluated at intermediate points. These relationships are then used to establish the inequality for the integral in the proof. In summary, the finite-increments formula allows us to approximate the change in a function over small increments and provides a useful tool for bounding the behaviour of the functions. Here are steps of the proof. (i) Apply the finite-increments formula to obtain an upper bound for the absolute value of the following: $|((r-s+\varepsilon)^{a-1} - (r-s)^{a-1})| \leq |a-1||r-s|^{a-2}\varepsilon$. Using this inequality, we will split the integral over the time interval $[s,t]$ into two parts, one from s to $s+\varepsilon$ and the second from $s+\varepsilon$ to t .

$$\begin{aligned} & \int_s^t ((r-s+\varepsilon)^{a-1} - (r-s)^{a-1}) dr \\ & \leq \int_s^{s+\varepsilon} |((r-s+\varepsilon)^{a-1} - (r-s)^{a-1})| dr \\ & \quad + \int_{s+\varepsilon}^t |(r-s+\varepsilon)^{a-1} - (r-s)^{a-1})| dr \end{aligned}$$

Notice that $(r-s+\varepsilon)^{a-1} - (r-s)^{a-1}$ can be written as $[(r-s)+\varepsilon]^{a-1} - (r-s)^{a-1} = [(r-s)^{a-1} + (a-1)(r-s)^{a-2}\varepsilon + \dots + \varepsilon^{a-1}] - (r-s)^{a-1} \leq |(a-1)(r-s)^{a-2}\varepsilon + \varepsilon^{a-1}|$ because $(r-s) \leq \varepsilon$; $(a-1)(r-s)^{a-2}\varepsilon \leq (a-1)|2(r-s)|^{a-1} - (r-s)^{a-1} |dr$.

Just to clarify $(r-s) \leq \varepsilon$ comes from the fact that we are integrating over the interval $[s,t]$ and splitting it into two parts: $[s, s+\varepsilon]$ and $[s+\varepsilon, t]$. Since r is in the interval $[s, s+\varepsilon]$, $r-s+\varepsilon \leq s+\varepsilon - s = \varepsilon$, so $r-s \leq \varepsilon$ and since r is in the interval $[s+\varepsilon, t]$, we have $r-s+\varepsilon \leq t-s$. Since we are interested in bounding the integrand over the entire interval $[s,t]$, we can take the maximum of the upper bounds for $(r-s)$ over both parts, which gives $(r-s) \leq \max\{\varepsilon, t-s\}$.

We can now use these bounds to re-write the integrals:

$$\begin{aligned} & \int_s^{s+\varepsilon} |((r-s+\varepsilon)^{a-1} - (r-s)^{a-1})| dr + \int_{s+\varepsilon}^t |(r-s+\varepsilon)^{a-1} - (r-s)^{a-1})| dr \leq \\ & \int_s^{s+\varepsilon} |(2(r-s))^{a-1} - (r-s)^{a-1}| dr + |a-1|\varepsilon \int_{s+\varepsilon}^t |r-s|^{a-2} dr \leq \int_s^{s+\varepsilon} (r-s)^{a-1} dr + \\ & |a-1|\varepsilon \int_{s+\varepsilon}^t |r-s|^{a-2} dr \leq \left[\frac{(r-s)^a}{a}\right]_s^{s+\varepsilon} + (a-1)\varepsilon \left[\frac{(r-s)^{a-1}}{a-1}\right]_{s+\varepsilon}^t \leq \frac{\varepsilon^a}{a} + \varepsilon^a \end{aligned} \quad (24)$$

We will now use this result to prove the convergence of the solution and show how to apply it for the purposes of this paper. We want to prove that $X^\varepsilon(t) = X_0 + \int_0^t f(s, X^\varepsilon(s))ds + \langle \int_0^t g(s, X^\varepsilon(s))dB_H^\varepsilon(s) \rangle$ converges to $X(t) = X_0 + \int_0^t f(s, X(s))ds + \langle \int_0^t g(s, X(s))dB_H(s) \rangle$. We will utilise equation (12) and Euler-Maruyama numerical method for that. This is a reminder of the equation (12):

$$\begin{aligned} & \sup_{u \in [0,t]} d_\infty^2(\langle \int_0^u X(s)dW(s) \rangle, \langle \int_0^u Y(s)dW(s) \rangle) \\ & \leq 4E \int_0^t d_\infty^2(\langle X(s) \rangle, \langle Y(s) \rangle) ds \end{aligned}$$

Euler-Maruyama is a modification of the standard Euler method for solving ordinary differential equations (ODEs) to solve stochastic differential equations (SDEs). The solution of an SDE is itself a random process, and therefore cannot be computed exactly in closed form. The Euler-Maruyama approximates a solution of an SDE as a piecewise linear function. At each step, the approximation is updated by taking a small step in time and adding a random perturbation. Specifically, it takes the form: $dX(t) = f(t, X(t))dt + g(t, X(t))dB_H(t)$. Choose a time step size Δt and define the time points $t_k = k\Delta t$ for $k = 0, 1, 2, \dots, n$. Initialise the approximation by setting $X_0 = X(t_0)$. For $k = 1, 2, \dots, n$, compute the approximation at time t_k as: $X(t_k) = X(t_{k-1}) + f(t_{k-1}, X(t_{k-1}))\Delta t + g(t_{k-1}, X(t_{k-1}))(B^H(t_k) - B^H(t_{k-1}))$, where $B^H(t_k) - B^H(t_{k-1})$ is the increment of the Brownian motion over the time interval $[t_{k-1}, t_k]$. Now, to obtain a bound on the distance between $X(t)$ and $X^\varepsilon(t)$, we use the triangle inequality and the fact that the distance between the two drift functions $f(X)$ and $f(X^\varepsilon)$ can be bounded by the integral of the distance between them over time. Specifically, we have:

$$\begin{aligned}
& E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] \\
& \leq 2E \sup_{u \in [0, t]} \int_0^u d_{\infty}^2 (f(s, X(s)), f(s, X^{\varepsilon}(s))) ds \\
& + 2E \sup_{u \in [0, t]} d_{\infty}^2 \left(\int_0^u g(s, X(s)) dB_H(s), \int_0^u g(s, X^{\varepsilon}(s)) dB_H^{\varepsilon}(s) \right) \quad (25)
\end{aligned}$$

To get this result, we applied triangle inequality to the distance between $X(u)$ and $X^{\varepsilon}(u)$ to bound the first term.

$$\begin{aligned}
E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] \\
\leq E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] \\
+ E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X^{\varepsilon}(u), X(u)) \right]
\end{aligned}$$

Similarly, we can bound the second term and substitute these bounds we obtain (25). Then, we will apply the Burkholder-Davis-Gundy (BDG) inequality in our proof. BDG is a tool used to bound the moments of stochastic integrals involving martingales. It provides for a way to control the growth of moments of stochastic integrals. In general form, the inequality states that: $E \left[\sup_{0 \leq s \leq t} |M_s|^p \right] \leq C_p E[(M_t^*)^p]$, where M_t is a continuous square-integrable martingale, $p > 1$ is a real positive number and $M_t^* = E \left[\sup_{0 \leq s \leq t} |M_s| \right]$ is the maximum value of the absolute value of the martingale up to time t . In other words, BDG inequality relates the supremum of the p th power of a continuous square-integrable martingale M_t to the p th moment of its maximum value M_t^* . To apply BDG, we first need to express the stochastic integrals in terms of the increments of the Brownian motion (so they are martingale). In particular, we use equation (5) derivation of the fractional Brownian motion $B_{H, \varepsilon}(t)$:

$$B_{H, \varepsilon}(t) = a \int_0^t \varphi^{\varepsilon}(s) ds + \varepsilon^a B(t)$$

where $\varphi^{\varepsilon}(s) = \int_0^t (t-s+\varepsilon)^{a-1} dB(s)$ and $B(t)$ is a standard Brownian motion.

Using this representation, we can express the stochastic integral involving the difference between X and X^{ε} as:

$$\begin{aligned}
\int_0^u (g(s, X(s)) (dB_H^{\varepsilon}(s) - dB_H(s))) &= \int_0^u (g(s, X^{\varepsilon}(s)) (dB_H^{\varepsilon}(s) - \\
& dB_H(s))) + \int_0^u (g(s, X(s)) - g(s, X^{\varepsilon}(s))) dB_H^{\varepsilon}(s).
\end{aligned}$$

We can similarly express the other integrals involving X and X^{ε} in terms of the increments of the Brownian motion. Now that we have our representations, we can apply BDG inequality to bound the supremum of the stochastic integrals involving Brownian motion increments.

Recall that a special case of BDG that we used earlier:

$$\begin{aligned}
E \sup_{u \in [0, t]} d_{\infty}^2 \left(\int_0^u X(s) dW(s), \int_0^u Y(s) dW(s) \right) \\
\leq 4E \int_0^t d_{\infty}^2 (\langle X(s), Y(s) \rangle) ds
\end{aligned}$$

Then,

$$\begin{aligned}
& E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] \\
& \leq 2E \sup_{u \in [0, t]} \int_0^u d_{\infty}^2 (f(s, X(s)), f(s, X^{\varepsilon}(s))) ds \\
& + 4E \sup_{u \in [0, t]} d_{\infty}^2 \left\langle \int_0^u g(s, X(s)) dB_H(s), \int_0^u g(s, X(s)) dB_H^{\varepsilon}(s) \right\rangle \\
& + 4E \sup_{u \in [0, t]} d_{\infty}^2 \left\langle \int_0^u g(s, X(s)) dB_H^{\varepsilon}(s), \int_0^u g(s, X^{\varepsilon}(s)) dB_H^{\varepsilon}(s) \right\rangle \\
& = 2E \sup_{u \in [0, t]} \int_0^u d_{\infty}^2 (f(s, X(s)), f(s, X^{\varepsilon}(s))) ds \\
& + 4E \sup_{u \in [0, t]} d_{\infty}^2 \left\langle \int_0^u g(s, X(s)) dB_H(s), \int_0^u g(s, X(s)) dB_H^{\varepsilon}(s) \right\rangle \\
& + 4E \sup_{u \in [0, t]} d_{\infty}^2 \left\langle \int_0^u g(s, X(s)) dB_H^{\varepsilon}(s), \int_0^u g(s, X^{\varepsilon}(s)) dB_H^{\varepsilon}(s) \right\rangle \\
& = 2E \sup_{u \in [0, t]} \int_0^u d_{\infty}^2 (f(s, X(s)), f(s, X^{\varepsilon}(s))) ds \\
& + 4E \sup_{u \in [0, t]} \left| \int_0^u g(s, X(s)) (dB_H(s) - dB_H^{\varepsilon}(s)) \right|^2 \\
& + 4E \sup_{u \in [0, t]} \left| \int_0^u (g(s, X(s)) - g(s, X^{\varepsilon}(s))) dB_H^{\varepsilon}(s) \right|^2
\end{aligned}$$

We will now substitute (5) $B_{H,\varepsilon}(t) = a \int_0^t \varphi^{\varepsilon}(s) ds + \varepsilon^a B(t)$ where $\varphi^{\varepsilon}(s) = \int_0^s (s-r+\varepsilon)^{a-1} - (s-r)^{a-1} dB(r)$ so $B_{H,\varepsilon}(t) = a \int_0^t [(s-r+\varepsilon)^{a-1} - (s-r)^{a-1}] dB(s)$ and $B(t) = a(s-r)^{a-1} dB(s)$, and $\varepsilon > 0, a = H - 1/2$ into above and use the (BDG) inequality in our proof. We need to replace the fBm with normal Brownian motion. We will use the following properties of the fBm. The increments of the fractional Brownian motion $B_H(s)$ are stationary with mean zero and variance $|t-s|^{2H}$. The increments of the fractional Brownian motion $B_H^{\varepsilon}(s)$ are also stationary and have mean zero and variance $|t|^{2H}$. Using these properties we can write:

$$\begin{aligned}
E(dB_H^{\varepsilon}(s) - dB_H(s))^2 &= E(B_H(s+\varepsilon) - B_H^{\varepsilon}(s) - B_H(s))^2 \\
&= 2E(B^2(s+\varepsilon)) + 2E(B^{2\varepsilon}) - 4E(B^{s+\varepsilon}B^{\varepsilon}) \\
&\quad - 4E(B^{s+\varepsilon}B^{\varepsilon}) + 4E(B^{\varepsilon})(B(s))
\end{aligned}$$

So,

$$\begin{aligned}
& E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] \\
& \leq 2E \sup_{u \in [0, t]} \int_0^u d_{\infty}^2 (f(s, X(s)), f(s, X^{\varepsilon}(s))) ds \\
& + 8e^{2a} E \sup_{u \in [0, t]} \left| \int_0^u g(s, X^{\varepsilon}(s)) (dB(s)) \right|^2 \\
& + 8a^2 E \sup_{u \in [0, t]} \left| \int_0^u (g(s, X^{\varepsilon}(s)) \left(\int_0^s [(s-r+\varepsilon)^{a-1} - (s-r)^{a-1}] dB(s) \right) ds \right|^2 \right. \\
& \quad \left. + 8a^2 E \sup_{u \in [0, t]} \left| \int_0^u (g(s, X^{\varepsilon}(s)) \left(\int_0^s [(s-r+\varepsilon)^{a-1} dB(s)) \right) ds \right|^2 \right. \\
& \quad \left. + 8e^{2a} E \sup_{u \in [0, t]} \left| \int_0^u (g(s, X(s)) - g(s, X^{\varepsilon}(s))) (dB(s)) \right|^2 \right.
\end{aligned}$$

We now have:

$$\begin{aligned}
& \mathbb{E} \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] \\
& \leq 2 \mathbb{E} \sup_{u \in [0, t]} \int_0^u d_{\infty}^2 (f(s, X(s)), f(s, X^{\varepsilon}(s))) ds \\
& \quad + 8 \varepsilon^{2a} \mathbb{E} \sup_{u \in [0, t]} \left| \int_0^u g(s, X^{\varepsilon}(s)) (dB(s)) \right|^2 \\
& \quad + 8 a^2 \mathbb{E} \sup_{u \in [0, t]} \left| \int_0^u \int_s^u (g(r, X^{\varepsilon}(r)) [(r-s+\varepsilon)^{a-1} \right. \\
& \quad \left. - (r-s)^{a-1}] dr dB(s)) \right|^2 \\
& \quad + 8 a^2 \mathbb{E} \sup_{u \in [0, t]} \left| \int_0^u \int_s^u (g(r, X^{\varepsilon}(r)) \right. \\
& \quad \left. - g(s, X^{\varepsilon}(s)) [(r-s+\varepsilon)^{a-1} dr dB(s)) \right|^2 \\
& \quad + 8 \varepsilon^{2a} \mathbb{E} \sup_{u \in [0, t]} \left| \int_0^u (g(s, X(s)) - g(s, X^{\varepsilon}(s))) (dB(s)) \right|^2
\end{aligned}$$

We will now apply the Doob's inequality for Brownian motion to the 2nd, 3rd, 4th and 5th term which states that $\mathbb{E} \left[\sup_{u \in [0, t]} d_{\infty}^2 |B_t| \right]^2 \leq 4 \mathbb{E} |B_t|^2$, Holder's inequality to the 3rd and 4th term $f \in L^p$, $g \in L^q$,

$\int f g \leq \|f\|_p \|g\|_q$ and Itô isometry to the 2nd, 3rd, 4th and last terms $\mathbb{E} (\int_0^t \int_s^t f_t dW_t)^2 = \mathbb{E} (\int_0^t (f_t)^2 dt)$.

$$\begin{aligned}
& \mathbb{E} \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] \\
& \leq 2 \mathbb{E} \sup_{u \in [0, t]} \int_0^t d_{\infty}^2 (f(s, X(s)), f(s, X^{\varepsilon}(s))) ds \\
& \quad + 32 \varepsilon^{2a} \mathbb{E} \int_0^u g^2(s, X^{\varepsilon}(s)) ds \\
& \quad + 32 a^2 \mathbb{E} \int_0^t \left[\int_s^t g^2(r, X^{\varepsilon}(r)) \left[(r-s+\varepsilon)^{a-1} \right. \right. \\
& \quad \left. \left. - (r-s)^{a-1} \right] dr \int_s^t [(r-s+\varepsilon)^{a-1} - (r-s)^{a-1}] dr \right] ds \\
& \quad + 32 a^2 \mathbb{E} \int_0^t \int_s^t (g(r, X^{\varepsilon}(r)) - g(s, X^{\varepsilon}(s)))^2 (r-s \\
& \quad + \varepsilon)^{a-1} dr \int_s^t (r-s+\varepsilon)^{a-1} dr ds \\
& \quad + 32 \varepsilon^{2a} \mathbb{E} \int_0^u (g(s, X(s)) - g(s, X^{\varepsilon}(s)))^2 ds
\end{aligned}$$

Using (17), (18), (23) and assumptions A1-A3:

$$\begin{aligned}
& E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] \\
& \leq 2L^2 E \int_0^t d_{\infty}^2 (X(s), X^{\varepsilon}(s)) ds \\
& + 64\varepsilon^{2a} C^2 \int_0^t (1 + E ||[X^{\varepsilon}(s)]^0||^2) ds \\
& + 64C^2 (T + \varepsilon)^{2a} \frac{a+1}{a} \varepsilon^a \int_0^t (1 \\
& + E ||[X^{\varepsilon}(s)]^0||^2) ds \\
& + 32L^2 (T + \varepsilon)^{2a} E \int_0^t d_{\infty}^2 (X(s), X^{\varepsilon}(s)) ds \\
& + 32L^2 \varepsilon^{2a} E \int_0^t d_{\infty}^2 (X(s), X^{\varepsilon}(s)) ds \\
& \leq (2L^2 + 32L^2 (T + \varepsilon)^{2a} \\
& + 32L^2 \varepsilon^{2a}) \int_0^t E \sup_{u \in [0, s]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) ds \\
& + (64\varepsilon^{2a} C^2 \\
& + 64C^2 (T + \varepsilon)^{2a} \frac{a+1}{a} \varepsilon^a) \int_0^t (1 \\
& + E ||[X^{\varepsilon}(s)]^0||^2) ds
\end{aligned} \tag{26}$$

We will now show the convergence of the solution of a fuzzy stochastic differential equation driven by fBm. The first part of (26) provides a bound on the supremum of the distance between the solution $X(u)$ and a perturbed version of the solution $X^{\varepsilon}(u)$ over the interval $[0, t]$. Here $\sup_{u \in [0, s]} d_{\infty}^2 (X(u), X^{\varepsilon}(u))$ denotes the supremum distance between the two processes, L is a Lipschitz constant, T is the final time, ε is small perturbation parameter, and a is the Hurst parameter of the fBm. The second part of (26) provides a bound on the second moment of the perturbed solution $X^{\varepsilon}(s)$. Here C is a constant and $[X^{\varepsilon}(s)]^0$ is the initial value of the perturbed process. We will apply Gronwall's lemma to the inequality to show that the supremum of the distance between

$X(u)$ and $X^{\varepsilon}(u)$ converges to zero as ε approaches zero, which implies the convergence of the solution $X(u)$ to a limit process as the perturbation parameter ε goes to zero. The equation provides an estimate for the convergence rate of the solution of the fuzzy stochastic differential equation driven by a fBm. To apply Gronwall's lemma to the inequality, we first define $Y(s) = E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right]$ and rewrite the inequality as: $Y(s) \leq (2L^2 + 32L^2 (T + \varepsilon)^{2a} + 32L^2 \varepsilon^{2a}) \int_0^t Y(s) ds + (64\varepsilon^{2a} C^2 + 64C^2 (T + \varepsilon)^{2a} \frac{a+1}{a} \varepsilon^a) \int_0^t (1 + E ||[X^{\varepsilon}(s)]^0||^2) ds$.

We can now apply Gronwall's lemma, which states that if $f(t)$ and $g(t)$ are nonnegative continuous functions on $[0, T]$ and if there exists a nonnegative constant K such that $f(t) \leq K + \int_0^t g(s) f(s) \forall t \in [0, T]$ then $f(t) \leq K e^{\int_0^t g(s) f(s)} \forall t \in [0, T]$. Applying this lemma we have:

$$\begin{aligned}
Y(s) & \leq (2L^2 + 32L^2 (T + \varepsilon)^{2a} + 32L^2 \varepsilon^{2a}) \int_0^t Y(s) ds + (64\varepsilon^{2a} C^2 \\
& + 64C^2 (T + \varepsilon)^{2a} \frac{a+1}{a} \varepsilon^a) \int_0^t (1 + E ||[X^{\varepsilon}(s)]^0||^2) ds \\
K & = (64\varepsilon^{2a} C^2 + 64C^2 (T + \varepsilon)^{2a} \frac{a+1}{a} \varepsilon^a) \int_0^t (1 + E ||[X^{\varepsilon}(s)]^0||^2) ds \\
g(s) & = 2L^2 + 32L^2 (T + \varepsilon)^{2a} + 32L^2 \varepsilon^{2a}
\end{aligned}$$

Then, by lemma $f(t) \leq K e^{\int_0^t g(s) f(s)} \forall t \in [0, T]$, we obtain:

$$\begin{aligned}
E \left[\sup_{u \in [0, t]} d_{\infty}^2 (X(u), X^{\varepsilon}(u)) \right] &\leq (64\varepsilon^{2a} C^2 \\
&+ 64C^2 (T + \varepsilon)^{2a} \frac{a+1}{a} \varepsilon^a) \int_0^t (1 \\
&+ E[|X^{\varepsilon}(s)|^0]^2) ds e^{2L^2 + 32L^2(T+\varepsilon)^{2a} + 32L^2\varepsilon^{2a}}
\end{aligned}$$

We can use the fact that the space of continuous functions equipped with the supremum norm is complete i.e., all Cauchy sequences converge, to conclude that the solution $X(u)$ converges to a limit process as the perturbation parameter ε goes to zero. As ε approaches zero, the second and thirds terms in the exponent tend to zero faster than the first term which dominates for large t . Therefore, the exponent approaches zero as ε approaches zero, which implies that the supremum of the distance between $X(u)$ and $X^{\varepsilon}(u)$ also approaches zero. This completes the proof.

4. Conclusions

This paper reviews a new approach to option pricing that incorporates fuzziness or uncertainty in the parameters like expected return rate, volatility, jump intensity, and jump magnitudes. Traditional option pricing models like Black-Scholes assume these parameters are crisp values. The proposed model treats them as fuzzy numbers to allow participants to estimate option prices based on their risk preferences and beliefs about the uncertain parameters. In Section 2, we build on [1] to show an alternative proof to extend Merton's jump-diffusion model for underlying asset price dynamics to include fractional Brownian motion and fuzzy processes. The asset price follows a stochastic differential equation driven by Brownian motion, fractional Brownian motion, and a fuzzy jump process. Detailed mathematical derivations are provided to incorporate the fuzziness into the asset price dynamics using fuzzy arithmetic and fuzzy stochastic calculus. An expectation formula is derived for the fuzzy asset price process, which depends on the fuzzy drift, volatility, jump parameters, and the expectation of the log jump sizes. The article builds on and extends prior work on fuzzy option pricing models. We start by considering the usual probability space $(\Omega, \mathcal{F}, P)$ that captures the randomness via the probability measure P on the sample space Ω and a filtration $\{\mathcal{F}_t\}$. This represents the standard stochastic/Brownian motion components. To incorporate fuzziness, we propose to construct a product measure space by taking the product of the above probability space with another measure space that captures the fuzziness inherent in the model parameters like expected return rate μ , volatility σ , jump intensity λ etc. Essentially, we represent the fuzzy parameters using fuzzy numbers or fuzzy sets defined on appropriate measure spaces. By taking products of these measure spaces with the original probability space, we obtain a new joint "product measure space" that simultaneously captures both the randomness from the stochastic processes and the fuzziness from the fuzzy parameters. On this product space, we can then define stochastic processes that have fractional, jump-diffusion dynamics given by SDEs driven by Brownian motions, but where the coefficients (drift, volatility etc.) are now fuzzy numbers instead of constants.

The use of product measure spaces allows us to apply fuzzy arithmetic, fuzzy calculus and fuzzy stochastic analysis on these mixed "fractional fuzzy" processes in a theoretically consistent manner within the standard measure-theoretic probability framework.

In summary, the key idea is to start with classical probability spaces for the stochastic components, construct separate measure spaces for the fuzzy parameters, and then form product measure spaces that merge the randomness and fuzziness into a unified model for asset price dynamics. This enables extending stochastic models like fractional Brownian motion to incorporate fuzzy coefficients and parameters.

Future work will present real-world examples, potential applications and a novel way of dealing with the proofs when $h \ll \frac{1}{2}$.

Author Contributions: Conceptualization, G.U.; methodology, G.U.; validation, G.U.; formal analysis, G.U.; writing- original draft preparation, G.U.; writing – review and editing, G.U., P.C. and T.C; supervision, P.C. and T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created.

Acknowledgments: The authors wish to thank all reviewers for their useful comments and discussions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Z. L. Y.-J. L. Y. Z. Wei-Guo Zhang, "Pricing European Option Under Fuzzy Mixed Fractional Brownian Motion Model with Jumps," *Computational Economics*, vol. 58, pp. 483-515, 2021.
2. R. Merton, "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics*, vol. 3, no. 1-2, pp. 125-144, 1976.
3. S. & D. B. B. Muzzioli, "Fuzzy approaches to option price modelling," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 2, pp. 392-401, 2017.
4. P. & P. M. Nowak, "Option pricing with application of Levy processes and the minimal variance equivalent martingale measure under uncertainty," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 2, pp. 402-416, 2017.
5. P. & P. M. Nowak, "Pricing European options under uncertainty with application of Levy processes and the minimal Lq equivalent martingale measure," *Journal of Computational and Applied Mathematics*, vol. 345, pp. 416-433, 2019.
6. H. W. A. L. D. Y. G. & S. A. Li, "The application of nonlinear fuzzy parameters PDE method in pricing and hedging European options," *Fuzzy Sets and Systems*, vol. 331, pp. 14-25, 2018.
7. X. D. & H. J. M. Wang, "A geometric Levy model for n-fold compound option pricing in a fuzzy framework," *Journal of Computational and Applied Mathematics*, vol. 306, pp. 246-264, 2016.
8. P. & R. M. Nowak, "Application of Levy processes and Esscher transformed martingale measures for option pricing in fuzzy framework," *Journal of Computational and Applied Mathematics*, vol. 263, pp. 129-151, 2014.
9. W. G. X. W. L. K. W. T. & Z. Y. Zhang, "Fuzzy pricing of geometric Asian options and its algorithm," *Applied Soft Computing*, vol. 28, pp. 360-367, 2015.
10. S. R. A. & D. B. B. Muzzioli, "A comparison of fuzzy regression methods for the estimation of the implied volatility smile function," *Fuzzy Sets and Systems*, vol. 266, pp. 131-143, 2015.
11. P. & R. M. Nowak, "Computing option price for Levy process with fuzzy parameters," *European Journal of Operational Research*, vol. 201, no. 1, pp. 206-210, 2010.
12. C. F. T. G. H. & W. S. Y. Lee, "A new application of fuzzy set theory to the Black-Scholes option pricing model," *Expert Systems with Applications*, vol. 29, no. 2, pp. 330-342, 2005.
13. H. C. Wu, "Pricing European options based on the fuzzy pattern of Black-Scholes formula," *Computers & Operations Research*, vol. 31, no. 7, pp. 1069-1081, 2004.
14. H. C. Wu, "European option pricing under fuzzy environments," *International Journal of Intelligent Systems*, vol. 20, no. 1, pp. 89-102, 2005.
15. H. C. Wu, "Using fuzzy sets theory and Black-Scholes formula to generate pricing boundaries of European options," *Applied Mathematics and Computation*, vol. 185, no. 1, pp. 136-146, 2007.
16. P. Meyer, "A decomposition theorem for supermartingales," *Illinois Journal of Mathematics*, vol. 6, pp. 193-205, 1962.
17. A. Kolmogorov, "Wiener's spirals and some other interesting curves in Hilbert space," *C. R. (Doklady) Acad. USSR(N.S.)*, vol. 26, pp. 115-118, 1940.
18. B. B. a. J. W. V. Ness, "Fractional Brownian Motions, Fractional Noises and Applications," *Society for Industrial and Applied Mathematics*, vol. 10, no. 4, pp. 422-447, 1968.
19. L. Young, "An inequality of the Hölder type, connected with Stieltjes integration," *Acta Mathematica*, vol. 67, pp. 251-282, 1936.
20. E. M. O. N. D. Alòs, "Stochastic calculus with respect to fractional Brownian motion with Hurst parameter less than 1/2," *Stochastic Processes and their Applications*, vol. 86, pp. 121-139, 2000.
21. C. Jost, "Transformation formulas for fractional brownian motion," *Stochastic Processes and their Applications*, vol. 116, no. 10, pp. 1341-1357, 2006.
22. T. Thao, "An approximate approach to fractional analysis for finance," *Nonlinear Analysis: Real World Applications*, vol. 7, no. 1, pp. 124-132, 2006.
23. L. Zadeh, "Fuzzy Sets," *Journal of Information and Controls*, vol. 8, pp. 338 - 353, 1965.
24. E. A. Llendeto, *Modeling with Itô Stochastic Differential Equations*, Berlin: Springer, 2007.

25. R. J. Aumann, "Integrals of set-valued functions," *Journal of Mathematical Analysis and Applications*, vol. 12, no. 1, pp. 1-12, 1965.
26. A. a. F. S. Kolmogorov, *Elements of the Theory of Functions and Functional Analysis*, Moscow: Mir, 1957-1961.
27. H. M. M. & E. M. Jafari, "Fuzzy stochastic differential equations driven by fractional Brownian motion," *Advances in Difference Equations*, vol. 16, pp. 1-17, 2021.
28. G. R. Burkholder. D.L., "Extrapolation and interpolation for convex functions of operators on martingales," *Acta Math*, vol. 124, pp. 249-304, 1979.
29. S. P. Lalley, "University of Chicago," 14 November 2016. [Online]. Available: <http://galton.uchicago.edu/~lalley/Courses/385/ItoIntegral.pdf>. [Accessed September 2022].
30. E. a. S. R. Stein, *Real Analysis*, Oxford: Princeton University Press, 2005.
31. E. & L. N. Coddington, *Theory of ordinary differential equations*, New York: McGraw-Hill, 1955.
32. T. Björk, *Arbitrage Theory in Continuous Time*, vol. 3, Oxford: Oxford University press, 2009.
33. S. Y. a. S. Watanabe, "On the uniqueness of solutions of stochastic differential equations," *Journal of Mathematics of Kyoto University*, vol. 11, no. 1, pp. 155-167, 1971.
34. L. Arnold., *Stochastic Differential Equations: Theory and Applications*, Wiley, 1974.
35. J. J. T. & R. M. Gatheral, "Volatility is rough," *Quantitative Finance*, vol. 18, pp. 933-949, 2014.
36. P. Cheridito, "Mixed fractional Brownian Motion," *Bernoulli*, vol. 7, pp. 913-934, 2001.
37. B. Mandelbrot, A. Fisher and L. and Calvet, "A Multifractal Model of Asset Returns," *Cowles Foundation For Research in Economics*, vol. 1412, pp. 1-39, September 1997.
38. R. D. G. W. Ruipeng Liu, "Volatility forecasting with bivariate multifractal models," *Journal of Forecasting*, vol. 39, no. 2, pp. 155-167, 2019.
39. R. L. Thomas Lux, "Generalized Method of Moment estimation of multivariate multifractal models," *Economic Modelling*, vol. 67, pp. 136-148, 2017.
40. R. G. Ruipeng Liu, "Investors' Uncertainty and Forecasting Stock Market Volatility," *Journal of Behavioral Finance*, vol. 23, pp. 327-337, 2022.
41. S. A. A. P. A. Thavaneswarana, "Weighted possibilistic moments of fuzzy numbers with applications to GARCH modeling and option pricing," *Mathematical and Computer Modelling*, vol. 49, pp. 352-368, 2009.
42. B. Y. Fang Jia, "Forecasting Volatility of Stock Index: Deep Learning Model with Likelihood-Based Loss Function," *Hindawi Complexity*, vol. 2021, p. 13, 2021.
43. R. K. T. J. F. Z. Z. a. M. S. A. Thavaneswaran, "Fuzzy Option Pricing Using a Novel Data-Driven Feed Forward Neural Network Volatility Model," 2019.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.