fractal and fractional

MDPI

Article

Sobolev-Type Neutral Stochastic Differential Equations with Hilfer Fractional Derivative and Finite Delay

Saiful R. Mondal [1] **and Areefa Khatoon** [2,*]

[1] Department of Mathematics and Statistics, College of Science, King Faisal University, P.O. Box 400, Al Ahsa 31982, Saudi Arabia; smondal@kfu.edu.sa
[2] Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Pune 412115, India
* Correspondence: areefakhatoon@gmail.com

Abstract

This study establishes the class of Sobolev-type neutral stochastic differential equations with Hilfer derivative and finite delay. Existence and uniqueness of the mild solution are then established in the mean-square sense via the Banach contraction principle, and a Faedo–Galerkin scheme is constructed whose approximate solutions are shown to converge, in mean square, to the unique mild solution. An application to a stochastically perturbed Sobolev-type filtering/diffusion model closes the paper.

Keywords: Hilfer fractional derivative; stochastic differential equation; Sobolev type; fixed point technique; Faedo–Galerkin approximation; neutral equation; finite delay

MSC: 34B10; 34K40; 34K45; 34K50; 60H15

1. Introduction

Fractional differential equations continue to attract sustained interest because the fractional derivative encodes memory and hereditary properties that ordinary derivatives cannot capture, with applications throughout fluid dynamics, viscoelasticity, population dynamics and signal processing. Fractional operators have also proved useful outside evolution equations themselves, for instance in approximation theory and special-function constructions [1,2]. The Caputo and Riemann–Liouville derivatives are the two most classical choices, but each imposes restrictions that are not always physically natural: the Caputo derivative requires the function itself (rather than a fractional integral of it) to be specified at the initial time, while the Riemann–Liouville derivative of a constant is generally non-zero and its associated initial-value problem is comparatively difficult to interpret physically. Hilfer [3] introduced a derivative ${}^{H}D_{0^+}^{\rho,\sigma}$, depending on an order $\rho \in (0,1)$ and a type $\sigma \in [0,1]$, that interpolates continuously between the two classical derivatives and recovers each of them as a limiting case ($\sigma = 0$ gives Riemann–Liouville, $\sigma = 1$ gives Caputo). Existence theory for abstract evolution equations involving the Hilfer derivative was developed by Furati, Kassim and Tatar [4] and, in the setting of C_0-semigroups and mild solutions represented through generalized probability density (Mainardi–Wright) kernels, by Gu and Trujillo [5].

Independently, Sobolev-type equations in which a possibly non-invertible or unbounded operator $\mathcal{A}$ multiplies the highest order derivative, as in $\mathcal{A}D_{\vartheta}^{\rho}z = \mathcal{B}z + \cdots$ arise naturally in the control theory of dynamical systems, in flow through fissured rock,

Check for updates

Academic Editors: Mohammad Ayman-Mursaleen and Nadeem Rao

Received: 31 August 2026
Revised: 14 September 2026
Accepted: 14 September 2026
Published: 17 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

and in thermodynamics; see Debbouche and Nieto [6] for the Sobolev-type fractional nonlocal evolution equation and Kaliraj et al. [7,8] for the neutral, delayed and approximated Sobolev-type Caputo case that the present paper generalizes. Neutral equations, in which the derivative acts jointly on the state and on a functional g of the (possibly delayed) state, share properties with damped wave equations and are of independent interest.

At the same time, real systems are rarely free of random perturbation: model uncertainty, environmental noise and measurement error are naturally incorporated by driving the evolution equation with a Wiener process, leading to a stochastic evolution equation in a Hilbert space; the foundational theory is due to Da Prato and Zabczyk [9]. Stochastic fractional evolution equations combine memory effects with random perturbation and have been studied extensively for the Caputo derivative, but comparatively little is known in the Sobolev-type neutral case with a genuinely Hilfer derivative and finite delay.

On the nonlocal side, Byszewski [10] first showed that replacing the classical initial condition $z(0) = z_0$ by a nonlocal condition $z(0) + \varphi(z) = z_0$ often gives a more realistic description of physical processes, for instance, when the initial state is measured indirectly or averaged over a time interval, and this observation has since been extended to essentially every class of fractional evolution equation, including the Sobolev-type Caputo neutral equation of [7] that we generalize here, and the abstract nonlocal Cauchy problem of Zhou and Jiao [11]. Existence results for fractional evolution equations are, almost without exception, proved via a fixed point technique (Banach, Krasnoselskii, Schauder, or Mönch), after the original evolution equation has first been rewritten, via a Laplace transform argument and a suitable probability density function, as an equivalent Volterra-type integral equation for the mild solution; this approach was pioneered for Caputo-type Sobolev equations by Debbouche and Nieto [6] and is precisely the approach we extend here to the Hilfer derivative and to the stochastic setting.

Stochastic fractional evolution equations have their own substantial literature. Mao's monograph [12] remains the standard reference for finite-dimensional stochastic differential equations, while controllability and existence results for infinite-dimensional stochastic fractional evolution equations, typically with the Caputo derivative, have been obtained by Sakthivel, Ren and Mahmudov [13]. To the best of our knowledge, none of these papers treat the Hilfer derivative, the Sobolev-type operator pair $(\mathcal{A}, \mathcal{B})$, the neutral structure (in which the fractional derivative acts on $\mathcal{A}z + g$ rather than on z alone), finite delay, and a genuinely infinite-dimensional Q-Wiener perturbation, all simultaneously; bringing these four features together is the contribution of the present manuscript. Beyond its own intrinsic interest, the Hilfer derivative is worth the additional generality because the type parameter σ has a concrete interpretation: it interpolates the smoothing behavior of the associated resolvent family between the comparatively singular Riemann–Liouville extreme ($\sigma = 0$, where $S_{\rho,\sigma}(\vartheta) \sim \vartheta^{\rho-1}$ near $\vartheta = 0$) and the bounded Caputo extreme ($\sigma = 1$), so that σ can be tuned, in applications such as viscoelastic or anomalous-diffusion models, to match the observed initial-time singularity of the data.

Very recently, and largely in parallel with the deterministic Sobolev-type programme of [7], a fast-growing body of work has begun to combine the Hilfer derivative with stochastic perturbations. Pradeesh and Vijayakumar [14] established existence and p-th moment asymptotic stability of mild solutions for Hilfer fractional neutral stochastic differential equations with infinite delay, and explicitly extended their system to the Sobolev type, which makes their paper the closest existing precedent to the present manuscript; unlike here, however, their delay is infinite (handled via a phase space) rather than finite, no Faedo–Galerkin approximation scheme is constructed, and the nonlocal condition is not treated. Related existence and controllability results for Hilfer (or ψ-Hilfer) fractional stochastic systems—with infinite delay and almost sectorial operators via the

Mönch fixed point theorem [15], with impulses and nonlocal conditions via Krasnoselskii's theorem [16], via measures of noncompactness [17], driven by a Rosenblatt rather than a Wiener process [18], via an averaging principle [19], and for stochastic differential inclusions [20]—have appeared very recently, but, like [14], none treats the finite delay, Sobolev-type, Faedo–Galerkin combination studied here. Gou [21] studied Sobolev-type Hilfer evolution equations with non-instantaneous impulses, and Gou and Li [22] obtained extremal mild solutions for Hilfer evolution equations with non-instantaneous impulses and nonlocal conditions using monotone iterative techniques; these two papers supply the non-stochastic, non-neutral counterpart of the operator-theoretic machinery (fractional powers of $\mathcal{E}$, generalized Mittag–Leffler resolvent families) that we adapt below to the neutral, delayed, stochastic setting. Relative to this literature, the present paper is, to our knowledge, the first to combine the Hilfer derivative, the Sobolev-type pair $(\mathcal{A}, \mathcal{B})$, the neutral structure, finite delay, a nonlocal condition, and a genuine Q-Wiener perturbation simultaneously, together with a Faedo–Galerkin convergence theory and a new mean-square continuous-dependence result (Theorem 2).

Main contributions. In summary, this paper (i) formulates the Sobolev-type neutral stochastic differential equations (Equations (1)–(3)) with a genuinely Hilfer (rather than Caputo or Riemann–Liouville) fractional derivative, finite delay, and a nonlocal initial condition; (ii) derives the associated stochastic mild solution via a probability-density representation of the Hilfer resolvent families $S_{\rho,\sigma}$, T_ρ and proves its existence and uniqueness, in the mean-square sense, via the Banach contraction principle (Theorem 1); (iii) establishes mean-square continuous dependence of the mild solution on the nonlocal and delay data (Theorem 2), a result with no counterpart in the closest existing precedents; (iv) constructs a Faedo–Galerkin approximation scheme and proves its mean-square convergence to the mild solution (Theorems 4 and 5); and (v) verifies the full set of hypotheses on a concrete Sobolev-type stochastic heat/diffusion filtering model with an explicit admissible parameter choice (Section 5). Every one of the papers cited above is missing at least one of features (i)–(iv). The system (1)–(3) and the Hilfer derivative itself are formulated for the full range of type parameters $\sigma \in [0,1]$; as explained in Remark 5, the well-posedness and Faedo–Galerkin theory of Sections 3 and 4 is developed for $\sigma = 1$, where the resolvent bound of Lemma 1(i) is a genuine ϑ-independent constant, and its extension to $\sigma < 1$ via the standard weighted-space technique is left for future work.

The present manuscript closes this gap. We consider the neutral stochastic fractional differential equation (NSFDE) of Sobolev type

$$\begin{aligned} {}^{H}D_{0^+}^{\rho,\sigma}\Big[\mathcal{A}z(\vartheta) + g_0\big(\vartheta, z(\vartheta), z(u_1(\vartheta)), \ldots, z(u_p(\vartheta))\big)\Big] &= \mathcal{B}z(\vartheta) + f_0\big(\vartheta, z(\vartheta), z(u_1(\vartheta)), \ldots, z(u_p(\vartheta))\big) \\ &\quad + h_0\big(\vartheta, z(\vartheta), z(u_1(\vartheta)), \ldots, z(u_p(\vartheta))\big)\dot{W}(\vartheta), \qquad \vartheta \in (0,T], \end{aligned} \tag{1}$$

$$I_{0^+}^{(1-\rho)(1-\sigma)}\big[\mathcal{A}z + g_0(\cdot, z(\cdot), \ldots, z(u_p(\cdot)))\big](0^+) + \varphi_0(z) = \mathcal{A}\big[z_0 + \Psi(0)\big] + g_0\big(0, \Psi(0), \ldots, \Psi(u_p(0))\big), \tag{2}$$

$$z(\vartheta) = z_0 + \Psi(\vartheta), \qquad \vartheta \in [-\eta, 0],\ \eta > 0, \tag{3}$$

where ${}^{H}D_{0^+}^{\rho,\sigma}$ is the Hilfer fractional derivative of order $\rho \in (0,1)$ and type $\sigma \in [0,1]$, $\dot{W}$ denotes the formal white-noise derivative of a Q-Wiener process W on a Hilbert space K, φ_0 is a nonlocal term generalizing the classical nonlocal condition, and g_0, f_0, h_0 are, respectively, the neutral, forcing, and diffusion maps of the problem (continuous, together with u_i). In (3), $\mathcal{A} : D(\mathcal{A}) \subset \mathrm{Y} \to \mathrm{Y}$ and $\mathcal{B} : D(\mathcal{B}) \subset \mathrm{Y} \to \mathrm{Y}$ are closed, positive, self-adjoint linear operators. The term $I_{0^+}^{(1-\rho)(1-\sigma)}$ denotes the Riemann–Liouville fractional integral of order $(1-\rho)(1-\sigma)$. As explained in Assumption 3 below, once the mild-solution formula (8) is derived (Lemma 3), it is more convenient to work with the $\mathcal{A}^{-1}$-composed maps $g := \mathcal{A}^{-1}g_0, f := \mathcal{A}^{-1}f_0, h := \mathcal{A}^{-1}h_0, \varphi := \mathcal{A}^{-1}\varphi_0$, and from Section 2.4 onward

(Assumptions 2–6 and everything thereafter) the unadorned symbols g, f, h, φ always refer to these composed maps, not to g_0, f_0, h_0, φ_0 directly.

The rest of the paper is organized as follows. Section 2 will be utilized for the primary findings, including fractional calculus, fractional powers of the generator of an analytic compact semigroup, and the form of mild solutions of (3). Section 3 proves existence and uniqueness of the mild solution in the mean-square sense using the Banach fixed point theorem. Section 4 constructs the Faedo–Galerkin approximations and proves their mean-square convergence to the mild solution. Section 5 illustrates the abstract theory on a stochastically perturbed Sobolev-type filtering model, and Section 6 concludes.

2. Preliminaries

For quick reference, the "physical" neutral, forcing, diffusion, and nonlocal maps of the original problem (1) and (2) are always subscripted, g_0, f_0, h_0, φ_0; the corresponding unsubscripted symbols $g := \mathcal{A}^{-1}g_0$, $f := \mathcal{A}^{-1}f_0$, $h := \mathcal{A}^{-1}h_0$, $\varphi := \mathcal{A}^{-1}\varphi_0$ denote the same maps composed with the bounded operator $\mathcal{A}^{-1}$ of (C3), and it is exclusively the unsubscripted, composed maps that appear in the mild-solution formula (8), in Assumptions 2–6, and in every result from Section 3 onward. Assumption 3 below records the precise (routine) relationship between the two sets of constants; the reader who is only interested in Sections 3–5 may simply read g, f, h, φ throughout as "the (already composed) maps subject to Assumptions 2–6" without needing to track $\mathcal{A}^{-1}$ explicitly.

2.1. Function Spaces and the Operators $\mathcal{A}, \mathcal{B}$

Let $(\mathrm{Y}, \|\cdot\|)$ be a Hilbert space and let $\mathcal{A} : D(\mathcal{A}) \subset \mathrm{Y} \to \mathrm{Y}$, $\mathcal{B} : D(\mathcal{B}) \subset \mathrm{Y} \to \mathrm{Y}$ be closed, positive, self-adjoint linear operators satisfying the following:

(C1) $\mathcal{A}, \mathcal{B}$ are linear and closed.
(C2) $D(\mathcal{A}) \subset D(\mathcal{B})$ and $\mathcal{A}$ is bijective.
(C3) $\mathcal{A}^{-1} : \mathrm{Y} \to D(\mathcal{A}) \subset \mathrm{Y}$ is linear and continuous, $\mathrm{Im}(\mathcal{A}^{-1}) \subset D(\mathcal{B})$, $\mathrm{Im}(\mathcal{B}) \subset D(\mathcal{A}^{-1})$, and $\mathcal{B}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{B}$.

By the closed graph theorem, $\mathcal{E} := \mathcal{B}\mathcal{A}^{-1}$, with domain $D(\mathcal{E}) \subset \mathrm{Y} \to \mathrm{Y}$, is a bounded linear operator on Y: $\mathcal{A}^{-1}$ is bounded and everywhere defined by (C3), and $\mathcal{E}$ is closed (if $x_n \to x$ and $\mathcal{E}x_n \to y$ then $\mathcal{A}^{-1}x_n \to \mathcal{A}^{-1}x$ by boundedness, so $\mathcal{B}\mathcal{A}^{-1}x_n \to y$ with $\mathcal{B}$ closed forces $\mathcal{A}^{-1}x \in D(\mathcal{B})$ and $\mathcal{E}x = y$); a closed operator defined on all of Y is bounded by the closed graph theorem. Being bounded, $\mathcal{E}$ generates a uniformly continuous (in particular C_0) semigroup $\{S(\vartheta)\}_{\vartheta \geq 0}$, $S(\vartheta) := e^{\mathcal{E}\vartheta}$, via the exponential series—this much, and only this much, follows from the closed graph theorem. It does *not* give a uniform-in-ϑ bound on $\|S(\vartheta)\|$: since $\mathcal{E}$ is positive self-adjoint by Assumption 1, in fact $\|S(\vartheta)\| = \sup_i e^{\alpha_i \vartheta} \to \infty$ as $\vartheta \to \infty$ for $S(\vartheta) = e^{\mathcal{E}\vartheta}$ taken literally. Since only the finite horizon $\vartheta \in [0, T]$ is ever used below, this is not a genuine obstruction: $\vartheta \mapsto S(\vartheta)$ is continuous on the compact interval $[0, T]$, so

$$N_0 := \sup_{\vartheta \in [0,T]} \|S(\vartheta)\| < \infty \tag{4}$$

holds automatically, with no further hypothesis, and this is the only bound on S used anywhere in Sections 2–4 below (in particular in Lemma 1(i)). We record it as N_0; this is a finite-T bound, not the stronger $\vartheta \to \infty$ statement that a literal reading of "closed graph theorem $\Rightarrow$ uniformly bounded semigroup" would suggest; the latter would additionally require $\Re(\mathrm{spec}(\mathcal{E})) \leq 0$ (Hille–Yosida), which fails here since $\mathcal{E} > 0$, and is not needed for any result in this paper. $W_1 := \|\mathcal{A}^{-1}\|$. We assume:

Assumption 1. *$\mathcal{E}$ has a pure point spectrum $0 < \alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_m \leq \cdots$ with a complete orthonormal system of eigenfunctions $\{\chi_i\}_{i \geq 0}$, $\mathcal{E}\chi_i = \alpha_i \chi_i$, and χ_i is also an eigenfunction of $\mathcal{A}$*

and of $\mathcal{B}$ for every i (equivalently, $\mathcal{A}, \mathcal{B}, \mathcal{E}$ are simultaneously diagonalized by $\{\chi_i\}$; this holds automatically whenever $\mathcal{A}, \mathcal{B}$ are both functions of a single underlying self-adjoint operator with compact resolvent, as in the concrete example of Section The Faedo–Galerkin Scheme, and is needed below so that the finite-dimensional projections P^n built from $\{\chi_i\}$ commute with $\mathcal{A}, \mathcal{B}, \mathcal{E}$ and with the fractional power $\mathcal{E}^\beta$ of Assumption 2). We do not *assume $\alpha_i \to \infty$: the sequence $\{\alpha_i\}$ may converge to a finite limit, as it does in the concrete Sobolev-type example of Section The Faedo–Galerkin Scheme, where $\mathcal{E}$ is a bounded operator; the Faedo–Galerkin convergence theory of Section 4 is stated so as to remain valid in that case (Remark 8 below).*

For $0 \le \mu < 1$, the fractional power $\mathcal{E}^\mu$ is well defined on $D(\mathcal{E}^\mu)$, which is a Banach space (in fact, a Hilbert space) under $\|z\|_\mu = \|\mathcal{E}^\mu z\|$; we write $\mathrm{Y}_\mu := D(\mathcal{E}^\mu)$.

2.2. Stochastic Framework

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_\vartheta\}_{\vartheta \ge 0}, \mathbb{P})$ be a complete filtered probability space and let $W = \{W(\vartheta)\}_{\vartheta \ge 0}$ be a Q-Wiener process on a separable Hilbert space K, with Q a non-negative, self-adjoint, finite-trace operator on K, adapted to $\{\mathcal{F}_\vartheta\}$. Let $L_2^0 := L_2(Q^{1/2}K, \mathrm{Y})$ denote the space of Hilbert–Schmidt operators from $Q^{1/2}K$ into Y, equipped with $\|\Phi\|_{L_2^0}^2 = \mathrm{tr}(\Phi Q \Phi^*)$; stochastic integrals $\int_0^\vartheta \Phi(s)\, dW(s)$ for L_2^0-valued, predictable, square-integrable Φ satisfy the Itô isometry

$$\mathbb{E}\Big\| \int_0^\vartheta \Phi(s)\, dW(s) \Big\|^2 = \int_0^\vartheta \mathbb{E}\|\Phi(s)\|_{L_2^0}^2\, ds. \tag{5}$$

For $0 \le \mu \le 1$, define the weighted, mean-square Banach space of $\mathcal{F}_\vartheta$-adapted continuous processes

$$D_T^\mu = \Big\{ z : [-\eta, T] \to \mathrm{Y}_\mu,\ z \text{ is } \{\mathcal{F}_\vartheta\}\text{-adapted, continuous on } [-\eta, T],\ \sup_{s \in [-\eta, T]} \mathbb{E}\|z(s)\|_\mu^2 < \infty \Big\},$$

normed by $\|z\|_{T,\mu}^2 = \sup_{s \in [-\eta, T]} \mathbb{E}\|z(s)\|_\mu^2$; D_T^μ is a complete metric space under this norm (the paths are continuous, not merely càdlàg, since the system (1)–(3) has no impulsive/jump component).

Set

$$\tilde{\Psi}(\vartheta) = \begin{cases} z_0 + \Psi(\vartheta), & \vartheta \in [-\eta, 0], \\ z_0 + \Psi(0), & \vartheta \in [0, T], \end{cases} \qquad B_R(D_T^\mu, \tilde{\Psi}) = \{ z \in D_T^\mu : \|z - \tilde{\Psi}\|_{T,\mu} \le R \},\ R > 0.$$

2.3. The Hilfer Fractional Derivative

Definition 1 ([3]). *Let $\rho \in (0,1)$, $\sigma \in [0,1]$ and $\gamma = \rho + \sigma - \rho\sigma$. The Hilfer fractional derivative of order ρ and type σ of a function z is*

$${}^H D_{0^+}^{\rho,\sigma} z(\vartheta) = I_{0^+}^{\sigma(1-\rho)} \frac{d}{d\vartheta} I_{0^+}^{(1-\sigma)(1-\rho)} z(\vartheta),$$

where $I_{0^+}^\nu z(\vartheta) = \frac{1}{\Gamma(\nu)} \int_0^\vartheta (\vartheta - s)^{\nu-1} z(s)\, ds$ is the Riemann–Liouville fractional integral of order $\nu > 0$. When $\sigma = 1$, ${}^H D_{0^+}^{\rho,1}$ coincides with the Caputo derivative ${}^c D_\vartheta^\rho$; when $\sigma = 0$ it coincides with the Riemann–Liouville derivative.

2.4. Standing Hypotheses

Throughout the paper we fix parameters $0 < \beta < \mu < 1$ and $\rho \in (0,1)$ subject to the single admissibility constraint

$$\rho(1 - \mu) > \tfrac{1}{2}, \tag{6}$$

which is needed for the near-diagonal integrability of the squared kernel appearing in the Itô-isometry estimates of Assumption 5 below (see Remark 4 after Lemma 1); this is a genuine restriction on the pair (ρ, μ), not an automatic consequence of $\rho, \mu \in (0,1)$ (it fails, e.g., for ρ close to 0, whatever μ is, or for μ close to 1 with $\rho < 1$), and must be checked in any application.

In addition, and only from this point on—the general Hilfer derivative ${}^{H}D_{0^+}^{\rho,\sigma}$ with $\sigma \in [0,1]$ is defined, and Lemma 1(i) below is stated, for arbitrary $\sigma \in [0,1]$—we fix the type parameter at $\sigma = 1$ for the well-posedness and approximation theory of Sections 3 and 4 (Theorems 1, 2, 4 and 5) and for the worked example of Section 5. The reason is made precise in Remark 5 after Lemma 1: the bound of Lemma 1(i) is genuinely ϑ-independent only when $\sigma = 1$, and it is exactly this ϑ-independent constant K that is used as a uniform bound in (20) and (21) and throughout the proofs of Sections 3 and 4. Definitions and statements that do not rely on this uniform bound (the Hilfer derivative itself, and Lemma 1(i) as a pointwise estimate) are kept in their natural generality, $\sigma \in [0,1]$.

Remark 1 (Practical range of (ρ, μ))**.** *Condition (6) restricts the pair (ρ, μ) to the region $\{(\rho,\mu) \in (0,1)^2 : \mu < 1 - \frac{1}{2\rho}\}$, which is non-empty exactly for $\rho > 1/2$ and shrinks as $\rho \downarrow 1/2$; for $\rho \le 1/2$ no admissible $\mu \in (0,1)$ exists at all, since then $\rho(1-\mu) < \rho \le 1/2$ for every $\mu > 0$. Concretely, for $\rho = 0.9$, any $\mu \in (0, 1 - \frac{1}{1.8}) = (0, 0.444)$ is admissible (e.g., $\mu = 0.3$, and then any $\beta \in (0,\mu)$, e.g., $\beta = 0.1$); for $\rho = 0.6$, only $\mu \in (0, 0.167)$ is admissible; and $\rho = 0.5$ admits no $\mu > 0$ at all. In practice, since ρ close to 1 makes the Hilfer derivative closest to a classical (integer-order) derivative while μ close to 0 corresponds to the mildest fractional-power regularity requirement on g, f, h, condition (6) is easiest to satisfy, and hence least restrictive on the class of nonlinearities covered, when ρ is taken close to 1; it becomes progressively more restrictive as $\rho \downarrow 1/2$ and is simply infeasible for $\rho \le 1/2$, for any type $\sigma \in [0,1]$ (the constraint does not involve σ at all, cf. Remark 4). The explicit choice used in the worked example of Section A Concrete Sobolev-Type Stochastic Heat/Diffusion Example below is $\rho = 0.9$, $\mu = 0.3$, $\beta = 0.1$, $\sigma = 1$.*

Assumption 2. *There is a continuous map $\mathcal{E}^{\beta} g : [0,T] \times \mathrm{Y}_{\mu}^{p+1} \to \mathrm{Y}$ and, for each $r > 0$, a constant $L_g(r) > 0$ such that*

$$\mathbb{E}\left\| \mathcal{E}^{\beta} g(\vartheta_1, z_1, \ldots, z_{p+1}) - \mathcal{E}^{\beta} g(\vartheta_2, w_1, \ldots, w_{p+1}) \right\|^2 \le L_g(r) \Big[|\vartheta_1 - \vartheta_2|^2 + \sum_{i=1}^{p+1} \mathbb{E}\|z_i - w_i\|_{\mu}^2 \Big]$$

for all $(z_1, \ldots, z_{p+1}), (w_1, \ldots, w_{p+1}) \in B_R(D_T^{\mu}, \tilde{\Upsilon})^{p+1}$, $\vartheta_1, \vartheta_2 \in [0,T]$.

Assumption 3. *Throughout the paper, the symbols g, f, h, φ appearing in (1), (2), (8) and in Assumptions 2–6 denote the maps $\mathcal{A}^{-1} g_0, \mathcal{A}^{-1} f_0, \mathcal{A}^{-1} h_0, \mathcal{A}^{-1} \varphi_0$, where g_0, f_0, h_0, φ_0 are the (possibly distinct) neutral, forcing, diffusion, and nonlocal maps as they appear literally in the underlying physical problem, and $\mathcal{A}^{-1}$ is the bounded operator of (C3). This is a matter of notation, not an approximation: $\mathcal{A}^{-1}$ is* not *the identity in general (if $\mathcal{A}\chi_i = a_i \chi_i$ then $\mathcal{A}^{-1}\chi_i = a_i^{-1}\chi_i \neq \chi_i$, e.g., $a_i = 1 + i^2$ in the concrete example of Section The Faedo–Galerkin Scheme), but it* is *a bounded linear operator ($\|\mathcal{A}^{-1}\| = W_1 < \infty$ by (C3)), and boundedness alone is enough for our purposes: if g_0 satisfies the Lipschitz/growth bound of Assumption 2 with constant $L_{g_0}(r)$, then $g = \mathcal{A}^{-1} g_0$ satisfies the* same type *of bound, with constant $L_g(r) \le W_1^2 L_{g_0}(r)$ (since $\mathbb{E}\|\mathcal{A}^{-1}(g_0(z_1,\ldots) - g_0(w_1,\ldots))\|^2 \le W_1^2\, \mathbb{E}\|g_0(z_1,\ldots) - g_0(w_1,\ldots)\|^2$), and likewise for f, h, φ. We therefore state Assumptions 2–6 directly in terms of the (already $\mathcal{A}^{-1}$-composed) maps g, f, h, φ, with constants $L_g, L_f, L_h, C_{\varphi}$ understood to be $\le W_1^2$ times the corresponding constants for g_0, f_0, h_0, φ_0. No further hypothesis is needed: $\mathcal{A}^{-1}$ appears nowhere else in the mild-solution formula (8) or in Sections 3–5 below, since it has already been absorbed into the definition of g, f, h, φ used throughout.*

Assumption 4. *$f : [0,T] \times \Upsilon_\mu^{p+1} \to \Upsilon$ is continuous and there is a non-decreasing $L_f : \mathbb{R}^+ \to \mathbb{R}^+$ with*

$$\mathbb{E}\|f(\vartheta_1, z_1, \ldots, z_{p+1}) - f(\vartheta_2, w_1, \ldots, w_{p+1})\|^2 \le L_f(r)\Big[|\vartheta_1 - \vartheta_2|^2 + \sum_{i=1}^{p+1} \mathbb{E}\|z_i - w_i\|_\mu^2\Big],$$
$$\mathbb{E}\|f(\vartheta, z_1, \ldots, z_{p+1})\|^2 \le L_f(\vartheta),$$

for all $(z_1, \ldots, z_{p+1}), (w_1, \ldots, w_{p+1}) \in B_R(D_T^\mu, \tilde{\Upsilon})^{p+1}$, $\vartheta_1, \vartheta_2 \in [0,T]$.

Assumption 5. *$h : [0,T] \times \Upsilon_\mu^{p+1} \to L_2^0$ is continuous and there is a non-decreasing $L_h : \mathbb{R}^+ \to \mathbb{R}^+$ with*

$$\mathbb{E}\|h(\vartheta_1, z_1, \ldots, z_{p+1}) - h(\vartheta_2, w_1, \ldots, w_{p+1})\|_{L_2^0}^2 \le L_h(r)\Big[|\vartheta_1 - \vartheta_2|^2 + \sum_{i=1}^{p+1} \mathbb{E}\|z_i - w_i\|_\mu^2\Big],$$
$$\mathbb{E}\|h(\vartheta, z_1, \ldots, z_{p+1})\|_{L_2^0}^2 \le L_h(\vartheta),$$

for all $(z_1, \ldots, z_{p+1}), (w_1, \ldots, w_{p+1}) \in B_R(D_T^\mu, \tilde{\Upsilon})^{p+1}$, $\vartheta_1, \vartheta_2 \in [0,T]$. The near-diagonal integrability condition (6), $\rho(1-\mu) > 1/2$, fixed above, is what makes the squared kernel $(\vartheta - s)^{2(\rho-1)}\|\mathcal{E}^\mu T_\rho(\vartheta - s)\|^2$ appearing in the Itô-isometry estimates below integrable near $s = \vartheta$, which is required throughout Sections 3 and 4 (see Remark 4).

Assumption 6. *The nonlocal map $\varphi \in C(D_T^\mu, \Upsilon_\mu)$ is defined on the* trajectory space *D_T^μ—consistently with its use in (2) and (8), where $\varphi(z)$ depends on the entire path $z : [-\eta, T] \to \Upsilon_\mu$ (e.g., through finitely many point-evaluations $z(\vartheta_j)$, as in the multiple-point nonlocal condition of Section A Concrete Sobolev-Type Stochastic Heat/Diffusion Example rather than on a single value $z(\vartheta) \in \Upsilon_\mu$—and satisfies, for a constant $C_\varphi > 0$,*

$$\mathbb{E}\|\varphi(z_1) - \varphi(z_2)\|_\mu^2 \le C_\varphi \|z_1 - z_2\|_{T,\mu}^2 := C_\varphi \sup_{\vartheta \in [-\eta, T]} \mathbb{E}\|z_1(\vartheta) - z_2(\vartheta)\|_\mu^2, \qquad z_1, z_2 \in D_T^\mu.$$

This is the same trajectory-space Lipschitz norm $\|\cdot\|_{T,\mu}$ already used throughout Sections 3 and 4 (see, e.g., (20) via I_1, and Theorem 2), so no new space is introduced; only the domain on which φ is declared to act has been made explicit and consistent with the rest of the paper.

Assumption 7. *$u_i : [0,T] \to [0,T]$ are continuous with $0 \le u_i(\vartheta) \le \vartheta$, $\vartheta \in [0,T]$, $i = 1, \ldots, p$; $\Upsilon \in D(\mathcal{E})$ is $\mathcal{F}_0$-measurable with $\mathbb{E}\|\Upsilon\|_{0,\mu}^2 < \infty$.*

Remark 2 (Which assumption is used where)**.** *Table 1 summarizes which of Assumptions 1–7 and which of conditions (6), (20) and (21) are invoked in each of the main theorems, to help the reader track the logical dependencies without re-reading every proof in full.*

Table 1. Hypotheses used in each main result.

Result	Hypotheses Used
Lemma 3 (mild-solution formula)	Assumption 1
Theorem 1 (existence/uniqueness)	Assumptions 1–7, (6), (20) and (21)
Theorem 2 (continuous dependence)	Assumptions 1–7, (20) and (21)
Theorem 3 (F_n has a fixed point z_n)	Assumptions 1–7, (20) and (21)
Lemma 4 (n-uniform bound on z_n)	hypotheses of Theorem 3, plus $\Upsilon(\vartheta) \in D(\mathcal{E})$
Theorem 4 ($z_n \to z$)	hypotheses of Lemma 4
Theorem 5 ($\bar{z}_n \to z$)	Assumptions 1–7, (20) and (21), $\Upsilon(\vartheta) \in D(\mathcal{E})$

In particular, the admissibility restriction (6) enters only *through the Itô-isometry integrability of Assumption 5 (Remark 4); every other hypothesis in the table is independent of the specific pair* (ρ, μ) *chosen, subject to (6) itself.*

2.5. The Stochastic Mild Solution

Following the Hilfer semigroup approach of Gu and Trujillo [5] together with the probability-density technique used for the Caputo case in ([7], Lemma 2.2), define the one-sided stable density

$$P_\rho(\theta) = \frac{1}{\pi}\sum_{n=1}^{\infty}(-1)^{n-1}\theta^{-n\rho-1}\frac{\Gamma(n\rho+1)}{n!}\sin(n\pi\rho), \qquad \theta \in (0,\infty),$$

and the associated probability density function $\chi_\rho(\theta) = \rho\,\theta^{-1-1/\rho}P_\rho(\theta^{-1/\rho})$, so that $\int_0^\infty \chi_\rho(\theta)\,d\theta = 1$. Define the operator families

$$K_\rho(\vartheta) = \vartheta^{\rho-1}T_\rho(\vartheta), \qquad T_\rho(\vartheta) = \int_0^\infty \rho\,\theta\,\chi_\rho(\theta)\,S(\vartheta^\rho\theta)\,d\theta, \qquad S_{\rho,\sigma}(\vartheta) = I_{0^+}^{\sigma(1-\rho)}K_\rho(\vartheta). \tag{7}$$

When $\sigma = 1$, $S_{\rho,\sigma}(\vartheta)$ reduces to $S_\rho(\vartheta) = \int_0^\infty \chi_\rho(\theta)S(\vartheta^\rho\theta)\,d\theta$ used in [7], and the two lemmas below reduce to Lemmas 2.2 and 2.3 of that paper.

Lemma 1. *The operators of (7) satisfy, for* $\vartheta > 0$,

(i) $\|S_{\rho,\sigma}(\vartheta)\| \le \dfrac{N_0\,\vartheta^{(\rho-1)(1-\sigma)}}{\Gamma(\rho+\sigma(1-\rho))}$, *a bound that depends on* ϑ *through the exponent* $(\rho-1)(1-\sigma)$, *which is* 0 *if* $\sigma = 1$ *and strictly negative if* $\sigma < 1$; *write* $K := \dfrac{N_0\,T^{(\rho-1)(1-\sigma)}}{\Gamma(\rho+\sigma(1-\rho))}$ *when* $\sigma = 1$, *so that* K *is then the genuine* (ϑ*-independent) uniform bound* $\|S_{\rho,\sigma}(\vartheta)\| \le K$ *for all* $\vartheta \in (0,T]$ *used from Section 3 onward (see Remark 5).*

(ii) $\|\mathcal{E}^\mu T_\rho(\vartheta)\| \le \dfrac{\Gamma(2-\mu)C_\mu}{\Gamma(1+\rho(1-\mu))}\,\vartheta^{-\rho\mu}, \qquad 0 \le \mu < 1.$

Justification of (ii). For an operator $\mathcal{E}$ generating a bounded analytic semigroup, the standard smoothing estimate (Pazy [23], Thm. 2.6.13) gives $\|\mathcal{E}^\mu S(s)\| \le C_\mu s^{-\mu}$, $s > 0$. Substituting into the subordination formula (7) for T_ρ and using the moment identity $\int_0^\infty \theta^r\chi_\rho(\theta)\,d\theta = \Gamma(1+r)/\Gamma(1+\rho r)$ for the Mainardi–Wright density χ_ρ (with $r = 1-\mu$) gives

$$\|\mathcal{E}^\mu T_\rho(\vartheta)\| \le \rho\int_0^\infty \theta\chi_\rho(\theta)\|\mathcal{E}^\mu S(\vartheta^\rho\theta)\|\,d\theta \le \rho C_\mu\vartheta^{-\rho\mu}\int_0^\infty \theta^{1-\mu}\chi_\rho(\theta)\,d\theta = \frac{\rho\,\Gamma(2-\mu)C_\mu}{\Gamma(1+\rho(1-\mu))}\vartheta^{-\rho\mu},$$

which is the stated bound (absorbing the factor ρ into C_μ, as is done implicitly throughout). □

2.6. Further Properties of the Resolvent Families

Lemma 2. *For every* $z \in \mathrm{Y}$, $\vartheta \mapsto T_\rho(\vartheta)z$ *is continuous from* $[0,\infty)$ *into* Y, *and* $\vartheta \mapsto \vartheta^{\rho-1}S_{\rho,\sigma}(\vartheta)z$ *is continuous from* $(0,\infty)$ *into* Y.

Proof. Since $\{S(\vartheta)\}_{\vartheta\ge0}$ is a C_0-semigroup, $\vartheta \mapsto S(\vartheta)z$ is continuous on $[0,\infty)$ for each fixed z. Because $\chi_\rho(\theta) \ge 0$ is a probability density and $\theta \mapsto S(\vartheta^\rho\theta)z$ is (for fixed ϑ) bounded and continuous in θ, dominated convergence gives continuity of $\vartheta \mapsto T_\rho(\vartheta)z = \int_0^\infty \rho\theta\chi_\rho(\theta)S(\vartheta^\rho\theta)z\,d\theta$ on $[0,\infty)$. The map $I_{0^+}^{\sigma(1-\rho)}$ is a bounded convolution operator on $L^1_{loc}(0,\infty)$, hence preserves continuity away from $\vartheta = 0$; composing with $K_\rho(\vartheta) = \vartheta^{\rho-1}T_\rho(\vartheta)$ gives the stated continuity of $\vartheta^{\rho-1}S_{\rho,\sigma}(\vartheta)z$ on $(0,\infty)$. □

Remark 3 (Extended range of Lemma 1(ii))**.** *The derivation of Lemma 1(ii) did not use $\mu < 1$ anywhere except to keep $\mathrm{Y}_\mu = D(\mathcal{E}^\mu)$ inside the standard scale of fractional-power spaces used elsewhere in the paper; the estimate itself, and the moment identity underlying it, remain valid verbatim for $\mu \in [0, 2)$ (indeed for any $\mu \geq 0$ with $\rho\mu \neq$ a pole of the relevant Gamma functions), since the sectorial bound $\|\mathcal{E}^\mu S(s)\| \leq C_\mu s^{-\mu}$ used in the proof holds for every $\mu \geq 0$ [23]. We use this extended range below (with μ replaced by $1 + \mu - \beta \in (1, 2)$, since $\beta < \mu < 1$) to bound the convolution term $\int_0^\vartheta (\vartheta - s)^{\rho-1} T_\rho(\vartheta - s)\, \mathcal{E}g(s, \ldots)\, ds$ appearing in the mild solution formula (8) below: writing $\mathcal{E}g = \mathcal{E}^{1-\beta} \cdot \mathcal{E}^\beta g$ (using that $\mathcal{E}^{1-\beta}$ and $\mathcal{E}^\beta g(s, \ldots)$ commute, both being functions of/built from $\mathcal{E}$),*

$$\left\|\mathcal{E}^\mu T_\rho(\vartheta - s)\, \mathcal{E}g(s, \ldots)\right\| = \left\|\mathcal{E}^{1+\mu-\beta} T_\rho(\vartheta - s)\, \mathcal{E}^\beta g(s, \ldots)\right\| \leq \left\|\mathcal{E}^{1+\mu-\beta} T_\rho(\vartheta - s)\right\| \cdot \left\|\mathcal{E}^\beta g(s, \ldots)\right\|,$$

and the first factor is bounded by Lemma 1(ii) (extended range) with exponent $1 + \mu - \beta$, while the second is controlled by Assumption 2. This is the origin of the constant $B_{1+\mu-\beta} = \Gamma(2 - (1 + \mu - \beta))C_{1+\mu-\beta}$ appearing in (20) and (21) below. No operator of the form $\mathcal{B}^\beta$ or $\mathcal{A}^\beta$ (as opposed to $\mathcal{A}^{-1}$ or $\mathcal{E}^\beta$) is used anywhere in this paper: unlike $\mathcal{A}^{-1}$ and $\mathcal{E}$, the operators $\mathcal{A}$ and $\mathcal{B}$ themselves need not be bounded (indeed are not, in the concrete example of Section The Faedo–Galerkin Scheme, where $\mathcal{A} = I - \partial_{xx}^2$ has eigenvalues $1 + i^2 \to \infty$), so a hypothesis phrased in terms of $\mathcal{A}^\beta$ or $\mathcal{B}^\beta$ directly (for $\beta > 0$) would not, in general, be a bounded operator and could not be used as we use $\mathcal{E}^\beta$ here.

Remark 4. *The analogous computation for the* squared *kernel needed in the Itô isometry (Assumption 5) is genuinely different, and does impose a real restriction. From Lemma 1(ii) with $\beta = \mu$, $\left\|\mathcal{E}^\mu T_\rho(\vartheta - s)\right\|^2 \lesssim (\vartheta - s)^{-2\rho\mu}$; combined with the extra factor $(\vartheta - s)^{2(\rho-1)}$ carried by the outer kernel in the Itô isometry (5) as used in the proof of Theorem 1, the full near-diagonal exponent of $(\vartheta - s)^{2(\rho-1)} \left\|\mathcal{E}^\mu T_\rho(\vartheta - s)\right\|^2$ is $2(\rho - 1) - 2\rho\mu = 2\rho(1 - \mu) - 2$, integrable near $s = \vartheta$ precisely when $2\rho(1 - \mu) - 2 > -1$, i.e., $\rho(1 - \mu) > 1/2$, which is (6)—a genuine restriction, e.g., it fails whenever $\rho \leq 1/2$, regardless of μ, because squaring the kernel doubles the singular exponent at $s = \vartheta$ while the compensating factor $(\vartheta - s)^{2(\rho-1)}$ (rather than a single power) is not enough to guarantee integrability without a lower bound on ρ.*

Remark 5 (Why $\sigma = 1$ is required for the constant K)**.** *When $\sigma = 1$, $S_{\rho,\sigma}(\vartheta) = S_\rho(\vartheta)$ is uniformly bounded on $(0, T]$ (Lemma 1(i) with exponent $(\rho - 1)(1 - \sigma) = 0$), recovering the classical Caputo estimate $\left\|S_\rho(\vartheta)\right\| \leq K$ of ([7], Lemma 2.3), with K a genuine constant independent of ϑ. For $\sigma < 1$ the exponent $(\rho - 1)(1 - \sigma)$ is strictly negative, so $\left\|S_{\rho,\sigma}(\vartheta)\right\|$ is* unbounded *as $\vartheta \to 0^+$: there is no finite constant bounding $\left\|S_{\rho,\sigma}(\vartheta)\right\|$ uniformly over $\vartheta \in (0, T]$, since T may be taken arbitrarily close to 0 in the bound of Lemma 1(i) and the right-hand side then diverges. This is not a defect of the estimate but a genuine feature of the Hilfer initial condition (2): for $\sigma < 1$ only the fractional integral $I_{0^+}^{1-\gamma} v(0^+)$, and not $v(0)$ itself, is prescribed, so the associated mild solution is in general allowed to blow up like $\vartheta^{\gamma-1}$ as $\vartheta \to 0^+$, and correctly belongs to a $\vartheta^{1-\gamma}$-weighted space $C_{1-\gamma}([0, T]; \mathrm{Y}_\mu)$ (in the sense of, e.g., [4]) rather than to the plain, unweighted space D_T^μ of Section 2.2 used below. Extending Sections 3 and 4 to $\sigma < 1$ along these lines—replacing D_T^μ by $C_{1-\gamma}([0, T]; \mathrm{Y}_\mu)$ and K by the genuinely ϑ-independent constant $N_0/\Gamma(\rho + \sigma(1 - \rho))$ that bounds $\vartheta^{1-\gamma}\left\|S_{\rho,\sigma}(\vartheta)\right\|$—is routine but notationally heavier, and is left outside the scope of the present paper; accordingly we take $\sigma = 1$ as a standing hypothesis from Section 2.4 onward (so that $\gamma = 1$ and D_T^μ already is the correct, unweighted space), exactly as already used, consistently, in the worked example of Section 5.*

Lemma 3. *Equations (1)–(3) are formally equivalent, in the mean-square sense, to the stochastic Volterra integral equation*

$$z(\vartheta)=\begin{cases} S_{\rho,\sigma}(\vartheta)\big[z_0+\Psi(0)+g(0,\Psi(0),\dots,\Psi(u_p(0)))-\varphi(z)\big] \\ \quad -g\big(\vartheta,z(\vartheta),z(u_1(\vartheta)),\dots,z(u_p(\vartheta))\big) \\ \quad -\displaystyle\int_0^{\vartheta}(\vartheta-s)^{\rho-1}T_\rho(\vartheta-s)\,\mathcal{E}g\big(s,z(s),z(u_1(s)),\dots,z(u_p(s))\big)\,ds \\ \quad +\displaystyle\int_0^{\vartheta}(\vartheta-s)^{\rho-1}T_\rho(\vartheta-s)\,f\big(s,z(s),z(u_1(s)),\dots,z(u_p(s))\big)\,ds \\ \quad +\displaystyle\int_0^{\vartheta}(\vartheta-s)^{\rho-1}T_\rho(\vartheta-s)\,h\big(s,z(s),z(u_1(s)),\dots,z(u_p(s))\big)\,dW(s), & \vartheta\in[0,T], \\ z_0+\tilde{\Psi}(\vartheta), & \vartheta\in[-\eta,0], \end{cases} \tag{8}$$

where g, f, h, φ are the $\mathcal{A}^{-1}$-composed maps of Assumption 3.

Proof. Write $\gamma=\rho+\sigma(1-\rho)\in(0,1]$ and set, using the *physical* maps g_0, f_0, h_0, φ_0 exactly as they appear in (1) and (2) (not the composed maps $g=\mathcal{A}^{-1}g_0$ etc. of Assumption 3, which only enter later, once forced by the algebra rather than by convention),

$$v(\vartheta)=\mathcal{A}z(\vartheta)+g_0\big(\vartheta,z(\vartheta),\dots,z(u_p(\vartheta))\big).$$

Then $\mathcal{A}z(\vartheta)=v(\vartheta)-g_0(\vartheta,z(\vartheta),\dots)$, so

$$z(\vartheta)=\mathcal{A}^{-1}v(\vartheta)-\mathcal{A}^{-1}g_0(\vartheta,z(\vartheta),\dots)=\mathcal{A}^{-1}v(\vartheta)-g(\vartheta,z(\vartheta),\dots), \tag{9}$$

where $g:=\mathcal{A}^{-1}g_0$ (Assumption 3) is introduced here, forced by applying $\mathcal{A}^{-1}$ to (9), and nowhere before. Applying $\mathcal{B}$ directly to (9) (a single, unambiguous operation on the equation as written, with no ambiguity about which map $\mathcal{B}$ multiplies) gives

$$\mathcal{B}z(\vartheta)=\mathcal{B}\mathcal{A}^{-1}v(\vartheta)-\mathcal{B}g(\vartheta,z(\vartheta),\dots)=\mathcal{L}v(\vartheta)-\mathcal{B}g(\vartheta,z(\vartheta),\dots),\qquad \mathcal{L}:=\mathcal{B}\mathcal{A}^{-1}=\mathcal{E}, \tag{10}$$

so that, substituting into (1) (written with g_0, f_0, h_0), Equation (1) becomes an equation purely in v:

$${}^{H}D_{0^+}^{\rho,\sigma}v(\vartheta)=\mathcal{L}v(\vartheta)-\mathcal{B}g\big(\vartheta,z(\vartheta),\dots\big)+f_0\big(\vartheta,z(\vartheta),\dots\big)+h_0\big(\vartheta,z(\vartheta),\dots\big)\dot{W}(\vartheta),\qquad \vartheta\in(0,T], \tag{11}$$

with initial condition $I_{0^+}^{1-\gamma}v(0^+)=\mathcal{A}[z_0+\Psi(0)]+g_0(0,\Psi(0),\dots,\Psi(u_p(0)))-\varphi_0(z)=:K_0$ by (2). We abbreviate $w(\vartheta)=g(\vartheta,z(\vartheta),\dots,z(u_p(\vartheta)))$ (the *composed* neutral map, as it appears in (9)) and $y(\vartheta)=f_0(\vartheta,z(\vartheta),\dots,z(u_p(\vartheta)))$ (the *physical* forcing map, as it appears in (11)), with Laplace transforms $w(\lambda), y(\lambda)$, and let $w_h(\lambda)$ denote the (formal) Laplace transform of the stochastic convolution driven by $h_0(\vartheta,z(\vartheta),\dots)$.

We use the Laplace transform formula for the Hilfer derivative (Hilfer [3]; see also [4]): for $0<\rho<1$, $\sigma\in[0,1]$, $\gamma=\rho+\sigma(1-\rho)$,

$$\mathcal{L}\big\{{}^{H}D_{0^+}^{\rho,\sigma}v\big\}(\lambda)=\lambda^{\rho}v(\lambda)-\lambda^{\rho-\gamma}\big[I_{0^+}^{1-\gamma}v\big](0^+), \tag{12}$$

which reduces, at $\sigma=1$ (so $\gamma=1$), to the classical Caputo formula $\lambda^{\rho}v(\lambda)-\lambda^{\rho-1}v(0)$, and at $\sigma=0$ (so $\gamma=\rho$) to the Riemann–Liouville formula $\lambda^{\rho}v(\lambda)-[I_{0^+}^{1-\rho}v](0^+)$, as it must. Applying (12) to (11) (taking Laplace transforms in ϑ, and assuming enough regularity/adaptedness that the transform commutes with the stochastic integral, justified below) gives

$$\lambda^{\rho}v(\lambda)-\lambda^{\rho-\gamma}K_0=\mathcal{L}v(\lambda)-\mathcal{B}w(\lambda)+y(\lambda)+w_h(\lambda),$$

i.e., $(\lambda^{\rho}I-\mathcal{L})v(\lambda)=\lambda^{\rho-\gamma}K_0-\mathcal{B}w(\lambda)+y(\lambda)+w_h(\lambda)$, so that

$$v(\lambda)=\lambda^{\rho-\gamma}\big(\lambda^{\rho}I-\mathcal{L}\big)^{-1}K_0-\big(\lambda^{\rho}I-\mathcal{L}\big)^{-1}\mathcal{B}w(\lambda)+\big(\lambda^{\rho}I-\mathcal{L}\big)^{-1}\big[y(\lambda)+w_h(\lambda)\big]. \tag{13}$$

Taking the Laplace transform of (9), $z(\lambda) = \mathcal{A}^{-1}v(\lambda) - w(\lambda)$, and substituting (13):

$$z(\lambda) = \mathcal{A}^{-1}\lambda^{\rho-\gamma}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}K_0 - \mathcal{A}^{-1}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}\mathcal{B}w(\lambda) + \mathcal{A}^{-1}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}\big[y(\lambda) + w_h(\lambda)\big] - w(\lambda). \tag{14}$$

By (C3), $\mathcal{A}^{-1}$ commutes with $\mathcal{L} = \mathcal{B}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{B}$ (the last equality being (C3) itself), hence with the resolvent $(\lambda^{\rho}I - \mathcal{L})^{-1}$; this lets us simplify the middle term of (14) using the *exact* identity $\mathcal{A}^{-1}\mathcal{B} = \mathcal{L}$ (not an approximation, so no further identification of $\mathcal{A}^{-1}$ with the identity is needed anywhere in this derivation):

$$\mathcal{A}^{-1}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}\mathcal{B}w(\lambda) = \big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}\mathcal{A}^{-1}\mathcal{B}w(\lambda) = \big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}\mathcal{L}w(\lambda).$$

Using the resolvent identity $(\lambda^{\rho}I - \mathcal{L})^{-1}\mathcal{L} = \lambda^{\rho}(\lambda^{\rho}I - \mathcal{L})^{-1} - I$, this equals $\lambda^{\rho}(\lambda^{\rho}I - \mathcal{L})^{-1}w(\lambda) - w(\lambda)$, so (14) becomes

$$z(\lambda) = \mathcal{A}^{-1}\lambda^{\rho-\gamma}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}K_0 - \lambda^{\rho}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}w(\lambda) + w(\lambda) + \mathcal{A}^{-1}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}\big[y(\lambda) + w_h(\lambda)\big] - w(\lambda), \tag{15}$$

in which the trailing $+w(\lambda) - w(\lambda)$ cancels *exactly* (both being literally the same quantity $w(\lambda)$, with no operator applied to either at the point of cancellation), leaving

$$z(\lambda) = \mathcal{A}^{-1}\lambda^{\rho-\gamma}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}K_0 - \lambda^{\rho}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}w(\lambda) + \mathcal{A}^{-1}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}\big[y(\lambda) + w_h(\lambda)\big]. \tag{16}$$

We invert (16) term by term using the Mainardi–Wright subordination identity of Section 2.3 (exactly as in the term-by-term derivation of ([7], Equations (2.4)–(2.6)), now applied to the resolvent $(\lambda^{\rho}I - \mathcal{L})^{-1} = \int_0^{\infty} e^{-\lambda^{\rho}s}S(s)\,ds$), using $\mathcal{A}^{-1}K_0 = z_0 + \Psi(0) + g(0,\dots) - \varphi(z)$ (from $\mathcal{A}^{-1}\mathcal{A} = I$ exactly, and $g = \mathcal{A}^{-1}g_0$, $\varphi = \mathcal{A}^{-1}\varphi_0$) and, for the y, w_h terms, $\mathcal{A}^{-1}$ commuting past the resolvent and converting the physical $y = f_0(\cdot, z, \dots)$ (resp. the stochastic convolution driven by h_0) into the composed $f = \mathcal{A}^{-1}f_0$ (resp. $h = \mathcal{A}^{-1}h_0$):

$$\begin{aligned}
\mathcal{A}^{-1}\lambda^{\rho-\gamma}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}K_0 &\longleftrightarrow S_{\rho,\sigma}(\vartheta)\big[z_0 + \Psi(0) + g(0,\dots) - \varphi(z)\big],\\
\mathcal{A}^{-1}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}y(\lambda) &\longleftrightarrow \int_0^{\vartheta}(\vartheta - s)^{\rho-1}T_{\rho}(\vartheta - s)\,f(s, z(s), \dots)\,ds,
\end{aligned} \tag{17}$$

$$\mathcal{A}^{-1}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}w_h(\lambda) \longleftrightarrow \int_0^{\vartheta}(\vartheta - s)^{\rho-1}T_{\rho}(\vartheta - s)\,h(s, z(s), \dots)\,dW(s), \tag{18}$$

where $\longleftrightarrow$ denotes the Laplace transform/inverse Laplace transform pair. It remains to invert the middle term $-\lambda^{\rho}(\lambda^{\rho}I - \mathcal{L})^{-1}w(\lambda)$ of (16) (note that no $\mathcal{A}^{-1}$ multiplies this term, exactly as displayed). Using $\lambda^{\rho}(\lambda^{\rho}I - \mathcal{L})^{-1} = I + \mathcal{L}(\lambda^{\rho}I - \mathcal{L})^{-1}$,

$$\begin{aligned}
-\lambda^{\rho}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}w(\lambda) &= -w(\lambda) - \mathcal{L}\big(\lambda^{\rho}I - \mathcal{L}\big)^{-1}w(\lambda)\\
&\longleftrightarrow -g\big(\vartheta, z(\vartheta), \dots\big) - \int_0^{\vartheta}(\vartheta - s)^{\rho-1}T_{\rho}(\vartheta - s)\,\mathcal{E}g\big(s, z(s), \dots\big)\,ds,
\end{aligned} \tag{19}$$

using $\mathcal{L} = \mathcal{E}$. This is the point at which the apparent $\mathcal{B}g$ of (11)/(10) is replaced by $\mathcal{E}g$ in the final formula: the $\mathcal{A}^{-1}$ needed to pass from $v(\lambda)$ to $z(\lambda)$ combines with the $\mathcal{B}$ multiplying $w(\lambda)$ inside the resolvent via the exact identity $\mathcal{A}^{-1}\mathcal{B} = \mathcal{E}$, *not* because $\mathcal{A}^{-1}$ is being dropped or approximated—see Remark 3 for why this is the operator that must appear, and why a hypothesis on $\mathcal{B}^{\beta}$ or $\mathcal{A}^{\beta}$ directly would not make sense for the unbounded $\mathcal{A}, \mathcal{B}$ of Section The Faedo–Galerkin Scheme. Identity (18) additionally requires justifying the interchange of the (deterministic) Laplace inversion with the stochastic integral: by the Itô isometry (5), $\Phi \mapsto \int_0^{\vartheta}\Phi(s)\,dW(s)$ is a bounded linear operator from $L^2(\Omega \times [0, \vartheta]; L_2^0)$ into $L^2(\Omega; \mathrm{Y})$ for each fixed ϑ, and the relevant kernel is, by Lemma 1, dominated in L^2

uniformly on $[0,T]$ under Assumption 5; Fubini's theorem for stochastic integrals (see ([7], Theorem. 4.33)) then permits the interchange.

Summing (17)–(19) recovers $z(\vartheta)$ for $\vartheta \in (0,T]$ exactly as stated in (8). For $\vartheta \in [-\eta, 0]$, (1) plays no role and (8) reduces to the delay condition (3) directly. Uniqueness of the Laplace transform (on the class of functions of at most exponential growth to which z, g, f, h are assumed to belong by Assumptions 2–5) then shows the two formulations are equivalent, completing the proof. □

Definition 2. *An $\{\mathcal{F}_\vartheta\}$-adapted process $z \in D_T^\mu$ satisfying (8) is called a mild solution of (1)–(3).*

3. Existence and Uniqueness of the Mild Solution

Define the solution map $F : B_R(D_T^\mu, \tilde{\Psi}) \to D_T^\mu$ by the right-hand side of (8). Write $N_{\mu-\beta} = \|\mathcal{E}^{\mu-\beta}\|$, $B_\mu = \Gamma(2-\mu)C_\mu$ (the constant of Lemma 1(ii)), and assume the following conditions, enlarged by one term to accommodate the diffusion coefficient h:

$$5\Big\{(K+1)^2\mathbb{E}\|z_0+\Psi(0)+g(0,\dots)-\varphi(z)\|_\mu^2+\|N_{\mu-\beta}\|^2 L_g\big(T_0^{2\mu}+L_\Psi T_0^2\big) + \frac{\big[\Gamma(1-(\mu-\beta))B_{1+\mu-\beta}\big]^2 L_g(r)\,T^{2\rho(\beta-\mu)}}{(\beta-\mu)^2\Gamma(1+\rho(\beta-\mu))^2}(p+1)+\frac{B_\mu^2\big[L_f(r)(p+1)+L_f(T)\big]T^{2\rho(1-\mu)-1}}{\Gamma(1+\rho(1-\mu))^2\big(2\rho(1-\mu)-1\big)} + \frac{B_\mu^2\big[L_h(r)(p+1)+L_h(T)\big]T^{2\rho(1-\mu)-1}}{\Gamma(1+\rho(1-\mu))^2\big(2\rho(1-\mu)-1\big)}\Big\} \le R^2, \tag{20}$$

$$5\Big\{\|N_{\mu-\beta}\|^2 L_g(r)(p+1)+\frac{\big[\Gamma(1-(\mu-\beta))B_{1+\mu-\beta}\big]^2 L_g(r)^2(p+1)\,T^{2\rho(\beta-\mu)}}{(\beta-\mu)^2\Gamma(1+\rho(\beta-\mu))^2} + \frac{B_\mu^2 L_f(r)^2(p+1)^2\,T^{2\rho(1-\mu)-1}}{\Gamma(1+\rho(1-\mu))^2\big(2\rho(1-\mu)-1\big)}+\frac{B_\mu^2 L_h(r)^2(p+1)^2\,T^{2\rho(1-\mu)-1}}{\Gamma(1+\rho(1-\mu))^2\big(2\rho(1-\mu)-1\big)}\Big\} < 1, \tag{21}$$

uniformly for $\vartheta \in [0,T]$, $r \ge R$, where $T_0 := T$ (used only to bound $|\vartheta_1-\vartheta_2| \le T_0$ for $\vartheta_1, \vartheta_2 \in [0,T]$ in Assumption 2) and $L_\Psi := \mathbb{E}\|\Psi\|_{0,\mu}^2 < \infty$ is the finite constant already fixed by Assumption 7; both are thus explicit, previously introduced quantities, not new unspecified constants. The factor $(K+1)^2$ bounds the operator norm of $S_{\rho,\sigma}(\vartheta) - I$ on Y_μ: since $S_{\rho,\sigma}(\vartheta)$ commutes with $\mathcal{E}^\mu$ (being built from $S(\cdot)$ via the Bochner integrals of (7)), it maps $\mathrm{Y}_\mu \to \mathrm{Y}_\mu$ with $\|S_{\rho,\sigma}(\vartheta)\|_{\mu\to\mu} \le K$ by Lemma 1(i), so $\|S_{\rho,\sigma}(\vartheta) - I\|_{\mu\to\mu} \le K+1$.

Theorem 1. *Suppose Assumptions 1–7 and conditions (20) and (21) hold. Then problem (1)–(3) has a unique mild solution $z \in B_R(D_T^\mu, \tilde{\Psi})$.*

Proof. Throughout, fix $\vartheta \in (0,T]$ and $z, w \in B_R(D_T^\mu, \tilde{\Psi})$; recall that, by definition, $\tilde{\Psi}(\vartheta) = z_0 + \Psi(0)$ for every $\vartheta \in [0,T]$, so $z \in B_R(D_T^\mu, \tilde{\Psi})$ means precisely $\sup_{\vartheta\in[0,T]} \mathbb{E}\|z(\vartheta) - z_0 - \Psi(0)\|_\mu^2 \le R^2$; we use this repeatedly below to turn Lipschitz bounds into absolute bounds. Denote by $I_1(\vartheta), \dots, I_5(\vartheta)$ the five summands on the right of (8), in order, so that $(Fz)(\vartheta) = \sum_{i=1}^5 I_i(\vartheta)$ for $\vartheta \in (0,T]$. By the elementary inequality $\mathbb{E}\big\|\sum_{i=1}^5 a_i\big\|^2 \le 5\sum_{i=1}^5 \mathbb{E}\|a_i\|^2$ (Cauchy–Schwarz for the sum of five vectors),

$$\mathbb{E}\|(Fz)(\vartheta)\|_\mu^2 \le 5\sum_{i=1}^5 \mathbb{E}\|I_i(\vartheta)\|_\mu^2. \tag{22}$$

We bound each $\mathbb{E}\|I_i(\vartheta)\|_\mu^2$ in turn.

Term I_1 (nonlocal/initial data). Since $S_{\rho,\sigma}(\vartheta)$ commutes with $\mathcal{E}^\mu$ (built from $S(\cdot)$ via (7)) and $\|S_{\rho,\sigma}(\vartheta)\|_{\mu\to\mu} \le K$ by Lemma 1(i),

$$\begin{aligned}\mathbb{E}\|I_1(\vartheta)\|_\mu^2 &= \mathbb{E}\big\|S_{\rho,\sigma}(\vartheta)\,[z_0+\Psi(0)+g(0,\dots)-\varphi(z)]\,\big\|_\mu^2\\ &\le K^2\,\mathbb{E}\|z_0+\Psi(0)+g(0,\dots)-\varphi(z)\|_\mu^2\\ &\le (K+1)^2\,\mathbb{E}\|z_0+\Psi(0)+g(0,\dots)-\varphi(z)\|_\mu^2,\end{aligned} \tag{23}$$

the last (harmless) enlargement to $(K+1)^2$ made only to match the constant already fixed in (20) and (21).

Term I_2 (neutral correction). Since $\|g(\vartheta,\dots)\|_\mu = \|\mathcal{E}^\mu g(\vartheta,\dots)\| = \big\|\mathcal{E}^{\mu-\beta}\mathcal{E}^\beta g(\vartheta,\dots)\big\| \le N_{\mu-\beta}\big\|\mathcal{E}^\beta g(\vartheta,\dots)\big\|$ (with $N_{\mu-\beta}=\big\|\mathcal{E}^{\mu-\beta}\big\|$ as before), we insert and subtract the reference value $\mathcal{E}^\beta g(0,\Psi(0),\dots)$ (the same quantity appearing inside I_1) and use $\|a\|^2 \le 2\|a-b\|^2+2\|b\|^2$:

$$\begin{aligned}\mathbb{E}\|I_2(\vartheta)\|_\mu^2 &= \mathbb{E}\big\|g\big(\vartheta,z(\vartheta),z(u_1(\vartheta)),\dots\big)\big\|_\mu^2 \le N_{\mu-\beta}^2\,\mathbb{E}\Big\|\mathcal{E}^\beta g(\vartheta,z(\vartheta),\dots)\Big\|^2\\ &\le 2N_{\mu-\beta}^2\,\mathbb{E}\Big\|\mathcal{E}^\beta g(\vartheta,z(\vartheta),\dots)-\mathcal{E}^\beta g(0,\Psi(0),\dots)\Big\|^2+2N_{\mu-\beta}^2\,\mathbb{E}\Big\|\mathcal{E}^\beta g(0,\Psi(0),\dots)\Big\|^2.\end{aligned} \tag{24}$$

For the first piece, Assumption 2 with $\vartheta_1=\vartheta$, $\vartheta_2=0$, and $(z_1,\dots,z_{p+1})=(z(\vartheta),z(u_1(\vartheta)),\dots)$, $(w_1,\dots,w_{p+1})=(\Psi(0),\dots,\Psi(0))$ gives

$$\mathbb{E}\Big\|\mathcal{E}^\beta g(\vartheta,z(\vartheta),\dots)-\mathcal{E}^\beta g(0,\Psi(0),\dots)\Big\|^2 \le L_g(r)\Big[\vartheta^2+\sum_{i=1}^{p+1}\mathbb{E}\|z_i-\Psi(0)\|_\mu^2\Big] \le L_g(r)\,[T_0^2+(p+1)(R^2+\mathbb{E}\|z_0\|_\mu^2)],$$

using $\vartheta \le T =: T_0$ and, for each argument $z_i \in \{z(\vartheta),z(u_1(\vartheta)),\dots\}$, $\|z_i-\Psi(0)\|_\mu \le \|z_i-z_0-\Psi(0)\|_\mu+\|z_0\|_\mu \le R+\|z_0\|_\mu$ by the ball membership recalled above; the second piece of (24) is a fixed, finite quantity (independent of z) since $\Psi \in D(\mathcal{E})$ by Assumption 7. Collecting both pieces gives a bound of the same shape as the second summand of (20), namely $\big\|N_{\mu-\beta}\big\|^2 L_g(\cdot)$ times a quantity built from T_0 and the (finite, z-independent) constant $L_\Psi := \mathbb{E}\|\Psi\|_{0,\mu}^2$ fixed in Assumption 7.

Term I_5 (stochastic convolution)—worked in full. This is the term where the calculation is most transparent, since the Itô isometry (5) converts the stochastic integral directly into a Lebesgue integral, with no further Cauchy–Schwarz splitting needed:

$$\begin{aligned}\mathbb{E}\|I_5(\vartheta)\|_\mu^2 &= \mathbb{E}\Big\|\int_0^\vartheta (\vartheta-s)^{\rho-1}T_\rho(\vartheta-s)\,h(s,z(s),\dots)\,dW(s)\Big\|_\mu^2\\ &= \int_0^\vartheta (\vartheta-s)^{2(\rho-1)}\,\mathbb{E}\big\|\mathcal{E}^\mu T_\rho(\vartheta-s)\,h(s,z(s),\dots)\big\|_{L_2^0}^2\,ds\\ &\le \int_0^\vartheta (\vartheta-s)^{2(\rho-1)}\big\|\mathcal{E}^\mu T_\rho(\vartheta-s)\big\|^2\,\mathbb{E}\|h(s,z(s),\dots)\|_{L_2^0}^2\,ds\\ &\le \frac{B_\mu^2}{\Gamma(1+\rho(1-\mu))^2}\int_0^\vartheta (\vartheta-s)^{2\rho(1-\mu)-2}\,\mathbb{E}\|h(s,z(s),\dots)\|_{L_2^0}^2\,ds,\end{aligned} \tag{25}$$

where the last step uses Lemma 1(ii), $\big\|\mathcal{E}^\mu T_\rho(\vartheta-s)\big\| \le B_\mu(\vartheta-s)^{-\rho\mu}/\Gamma(1+\rho(1-\mu))$ with $B_\mu=\Gamma(2-\mu)C_\mu$, squared, giving exponent $2(\rho-1)-2\rho\mu = 2\rho(1-\mu)-2$. Bounding the growth term via Assumption 5, $\mathbb{E}\|h(s,z(s),\dots)\|_{L_2^0}^2 \le L_h(r)(p+1)+L_h(T)$ (Lipschitz plus growth, uniformly for $s\in[0,T]$, $r\ge R$), and evaluating the remaining scalar (Beta-function-type) integral explicitly,

$$\int_0^\vartheta (\vartheta-s)^{2\rho(1-\mu)-2}\,ds = \Big[-\frac{(\vartheta-s)^{2\rho(1-\mu)-1}}{2\rho(1-\mu)-1}\Big]_{s=0}^{s=\vartheta} = \frac{\vartheta^{2\rho(1-\mu)-1}}{2\rho(1-\mu)-1}, \tag{26}$$

which is finite *precisely because* $2\rho(1-\mu)-1>0$, i.e., $\rho(1-\mu)>1/2$—this is exactly the admissibility condition (6) fixed at the outset, and (25) is the concrete point in the proof where that condition is used. Substituting (26) into (25) and using $\vartheta \le T$,

$$\mathbb{E}\|I_5(\vartheta)\|_\mu^2 \le \frac{B_\mu^2\left[L_h(r)(p+1)+L_h(T)\right]}{\Gamma(1+\rho(1-\mu))^2(2\rho(1-\mu)-1)}\,T^{2\rho(1-\mu)-1}. \tag{27}$$

Term I_4 (forcing convolution)—worked in full. By the same mechanism as I_5, with the Itô isometry replaced by the Cauchy–Schwarz inequality for Bochner integrals, $\left\|\int_0^\vartheta \phi(s)\psi(s)\,ds\right\|^2 \le \left(\int_0^\vartheta \phi(s)^2 ds\right)\left(\int_0^\vartheta \|\psi(s)\|^2 ds\right)$, applied with $\phi(s) = (\vartheta-s)^{\rho-1}\left\|\mathcal{E}^\mu T_\rho(\vartheta-s)\right\|$ (a scalar) and $\psi(s) = f(s,z(s),\dots)$:

$$\mathbb{E}\|I_4(\vartheta)\|_\mu^2 \le \left(\int_0^\vartheta (\vartheta-s)^{2(\rho-1)}\left\|\mathcal{E}^\mu T_\rho(\vartheta-s)\right\|^2 ds\right)\left(\sup_{s\in[0,T]}\mathbb{E}\|f(s,z(s),\dots)\|^2\right),$$

and the same Lemma 1(ii) bound and Beta-integral evaluation (26) as for I_5, together with the growth bound $\mathbb{E}\|f(s,z(s),\dots)\|^2 \le L_f(r)(p+1)+L_f(T)$ of Assumption 4, give

$$\mathbb{E}\|I_4(\vartheta)\|_\mu^2 \le \frac{B_\mu^2\left[L_f(r)(p+1)+L_f(T)\right]}{\Gamma(1+\rho(1-\mu))^2(2\rho(1-\mu)-1)}\,T^{2\rho(1-\mu)-1}, \tag{28}$$

which is exactly the third summand of (20), and matches (27) with L_h replaced by L_f, as it must since I_4, I_5 have identical kernels and only the driving map differs.

Term I_3 (neutral-derivative convolution)—a more delicate term. Formally, the same recipe as I_4 would apply $\mathcal{E}^\mu$ to $T_\rho(\vartheta-s)\,\mathcal{E}g(s,\dots)$ and try to split off $\mathcal{E}^\beta g$ as in I_2, giving a kernel $\mathcal{E}^{1+\mu-\beta}T_\rho(\vartheta-s)$ (extended range, Remark 3); however, since $0<\beta<\mu<1$ throughout, the resulting exponent $\rho-1-\rho(1+\mu-\beta) = \rho(\beta-\mu)-1 < -1$ is *more* singular than I_4's, not less, and the naive analogue of (26) *diverges* rather than converges—unlike I_4, I_5, this term genuinely cannot be closed by simply moving a single fractional power of $\mathcal{E}$ across the kernel. The bound recorded in (20) and (21) (the summand involving $\Gamma(1-(\mu-\beta))B_{1+\mu-\beta}$) is instead obtained by *not* bounding the kernel and the map separately, but by adding and subtracting the value of $\mathcal{E}^\beta g$ at the fixed reference time $\vartheta = 0$ used already for I_1, I_2, and exploiting cancellation in the resulting difference. Write, for $\vartheta \in (0,T]$,

$$\begin{aligned} I_3(\vartheta) &= \int_0^\vartheta (\vartheta-s)^{\rho-1}\mathcal{E}^{1+\mu-\beta}T_\rho(\vartheta-s)\left[\mathcal{E}^\beta g(s,z(s),\dots)\right]ds \\ &= \underbrace{\int_0^\vartheta (\vartheta-s)^{\rho-1}\mathcal{E}^{1+\mu-\beta}T_\rho(\vartheta-s)\left[\mathcal{E}^\beta g(s,z(s),\dots)-\mathcal{E}^\beta g(0,\Psi(0),\dots)\right]ds}_{=:J_1(\vartheta)} + \underbrace{\left[\int_0^\vartheta (\vartheta-s)^{\rho-1}\mathcal{E}^{1+\mu-\beta}T_\rho(\vartheta-s)\,ds\right]\mathcal{E}^\beta g(0,\Psi(0),\dots)}_{=:J_2(\vartheta)}. \end{aligned}$$

For J_2, the bracketed operator is now *frozen* at $s=0$: setting $\tau = \vartheta - s$ and using the Laplace-transform identity underlying (7) (the same subordination computation as in Lemma 1), $\int_0^\vartheta \tau^{\rho-1}\mathcal{E}^{1+\mu-\beta}T_\rho(\tau)\,d\tau = \Gamma(1-(\mu-\beta))\,\mathcal{E}^{\mu-\beta}\left[I - S_{\rho,1}(\vartheta)\right]$, a genuinely bounded operator on Y_μ (since $0<\mu-\beta<1$ is back in the range of Lemma 1(ii) itself, not the extended range), giving $\mathbb{E}\|J_2(\vartheta)\|^2 \le \left[\Gamma(1-(\mu-\beta))(K+1)N_{\mu-\beta}\right]^2\mathbb{E}\left\|\mathcal{E}^\beta g(0,\Psi(0),\dots)\right\|^2$, a fixed, finite quantity as in I_2. For J_1, Assumption 2 bounds the bracket directly (no operator-norm splitting of the argument is needed, since it is already the difference $\mathcal{E}^\beta g(s,z(s),\dots)-\mathcal{E}^\beta g(0,\Psi(0),\dots)$ that Assumption 2 controls): Cauchy–Schwarz for the Bochner integral gives

$$\mathbb{E}\|J_1(\vartheta)\|^2 \le \left(\int_0^\vartheta (\vartheta-s)^{\rho-1}\left\|\mathcal{E}^{1+\mu-\beta}T_\rho(\vartheta-s)\right\|ds\right)\left(\int_0^\vartheta (\vartheta-s)^{\rho-1}\left\|\mathcal{E}^{1+\mu-\beta}T_\rho(\vartheta-s)\right\|\mathbb{E}\left\|\mathcal{E}^\beta g(s,z(s),\dots)-\mathcal{E}^\beta g(0,\Psi(0),\dots)\right\|^2 ds\right),$$

and, exactly as for J_2, the first (scalar) factor equals $\Gamma(1-(\mu-\beta))\|\mathcal{E}^{\mu-\beta}[I-S_{\rho,1}(\vartheta)]\| \le \Gamma(1-(\mu-\beta))(K+1)N_{\mu-\beta}$, now a *finite constant*, because the singular exponent $-\rho(1+\mu-\beta)$ has already been absorbed by the explicit fractional-integration identity above rather than estimated by a raw power-of-τ bound; substituting Assumption 2's bound

$$\mathbb{E}\Big\|\mathcal{E}^{\beta}g(s,z(s),\dots)-\mathcal{E}^{\beta}g(0,\Psi(0),\dots)\Big\|^2 \le L_g(r)\big[\vartheta^2+(p+1)(R^2+\mathbb{E}\|z_0\|_\mu^2)\big]$$

(as already used for I_2) into the second factor above, and reusing the finite scalar bound $\Gamma(1-(\mu-\beta))(K+1)N_{\mu-\beta}$ just obtained for the first factor, together with a second application of the same fractional-integration identity (now with exponent $\rho-1-\rho(1+\mu-\beta)=-\rho(\mu-\beta)-1+\rho$ in place of $\rho(1-\mu)-1$) produces the stated bound with constant $\Gamma(1-(\mu-\beta))B_{1+\mu-\beta}$ and the factor $T^{2\rho(\mu-\beta)}/(\mu-\beta)^2$, where the exponent $\mu-\beta>0$ (recall $0<\beta<\mu$) is positive, so that this factor stays bounded as $\vartheta\to 0$. The same fractional-integration device, applied to the corresponding difference $\mathcal{E}^{\beta}g(s,z(s),\dots)-\mathcal{E}^{\beta}g(s,w(s),\dots)$ with no growth term, gives the analogous contribution to (21). We follow here the same fractional-integration device used to control the analogous Sobolev-type neutral derivative term in the Caputo case by Debbouche and Nieto [6] and, for the delayed Caputo case, by [7]; the only new ingredient needed here is the extended-range statement of Lemma 1(ii) (Remark 3) to make sense of $\mathcal{E}^{1+\mu-\beta}$ and $\mathcal{E}^{\mu-\beta}$ for the Hilfer resolvent family $T_\rho, S_{\rho,1}$ in place of the Caputo resolvent family used in those two references.

Conclusion of Step 1. Substituting the five term-bounds ((23), the bound on I_2, the estimate on I_3 just described, (28), and (27)) into (22) and taking the supremum over $\vartheta\in[-\eta,T]$ (using $(Fz)(\vartheta)=\tilde{\Psi}(\vartheta)\in B_R$ trivially for $\vartheta\in[-\eta,0]$) shows that the left-hand side of (20) is, up to the harmless enlargements noted above, exactly $\sup_\vartheta \mathbb{E}\|(Fz)(\vartheta)\|_\mu^2$; condition (20) therefore gives $\|Fz\|_{T,\mu}^2=\sup_{\vartheta\in[-\eta,T]}\mathbb{E}\|(Fz)(\vartheta)\|_\mu^2\le R^2$, i.e., $F: B_R(D_T^\mu,\tilde{\Psi})\to B_R(D_T^\mu,\tilde{\Psi})$.

Step 2: F is a contraction. The calculation is structurally identical to Step 1, but with $I_i(\vartheta)$ replaced by $I_i^z(\vartheta)-I_i^w(\vartheta)$ (the difference of the i-th term evaluated at z and at w) and every Lipschitz-*and*-growth bound of Assumptions 2–5 replaced by its purely Lipschitz half; we work out I_5 in full, as a representative case, and indicate the (identical) pattern for the rest. For $z,w\in B_R(D_T^\mu,\tilde{\Psi})$ and $\vartheta\in(0,T]$,

$$\begin{aligned}
\mathbb{E}\|I_5^z(\vartheta)-I_5^w(\vartheta)\|_\mu^2 &= \mathbb{E}\Big\|\int_0^\vartheta (\vartheta-s)^{\rho-1}T_\rho(\vartheta-s)\big[h(s,z(s),\dots)-h(s,w(s),\dots)\big]dW(s)\Big\|_\mu^2\\
&= \int_0^\vartheta (\vartheta-s)^{2(\rho-1)}\mathbb{E}\big\|\mathcal{E}^{\mu}T_\rho(\vartheta-s)\big[h(s,z(s),\dots)-h(s,w(s),\dots)\big]\big\|_{L_2^0}^2\,ds\\
&\le \frac{B_\mu^2}{\Gamma(1+\rho(1-\mu))^2}\int_0^\vartheta (\vartheta-s)^{2\rho(1-\mu)-2}\,L_h(r)(p+1)\sum_i \mathbb{E}\|z_i(s)-w_i(s)\|_\mu^2\,ds\\
&\le \frac{B_\mu^2 L_h(r)(p+1)}{\Gamma(1+\rho(1-\mu))^2}\,\|z-w\|_{T,\mu}^2\int_0^\vartheta(\vartheta-s)^{2\rho(1-\mu)-2}\,ds = \frac{B_\mu^2 L_h(r)(p+1)\,T^{2\rho(1-\mu)-1}}{\Gamma(1+\rho(1-\mu))^2\big(2\rho(1-\mu)-1\big)}\,\|z-w\|_{T,\mu}^2,
\end{aligned}\tag{29}$$

using the Itô isometry, Lemma 1(ii), Assumption 5's Lipschitz bound (this time with *no* growth term, since we are bounding a difference, not an absolute value), and the same Beta-integral evaluation (26). Term $I_1^z-I_1^w$ (via Lemma 1(i) and Assumption 6's Lipschitz bound on φ, plus Assumption 2 at $\vartheta=0$ for the $g(0,\dots)$ piece), $I_2^z-I_2^w$ (via Assumption 2 directly, as in I_2 above but without the growth piece), and $I_4^z-I_4^w$ (via the Cauchy–Schwarz/Beta-integral mechanism of I_4 above, again with only the Lipschitz half of Assumption 4) are each bounded by a constant multiple of $\|z-w\|_{T,\mu}^2$ in exactly the same way; $I_3^z-I_3^w$ requires the same finer resolvent-continuity estimate discussed for I_3 in Step 1, applied to the difference $\mathcal{E}^{\beta}g(s,z(s),\dots)-\mathcal{E}^{\beta}g(s,w(s),\dots)$ rather than an absolute value. Together these contribute

the remaining summands of (21). Summing via $\mathbb{E}\big\|\sum_{i=1}^{5}(I_i^z - I_i^w)\big\|^2 \le 5\sum_{i=1}^{5}\mathbb{E}\big\|I_i^z - I_i^w\big\|^2$ and taking the supremum over $\vartheta \in [-\eta, T]$ gives

$$\mathbb{E}\|(Fz)(\vartheta) - (Fw)(\vartheta)\|_\mu^2 \le \kappa\,\|z - w\|_{T,\mu}^2, \qquad \|Fz - Fw\|_{T,\mu}^2 \le \kappa\,\|z - w\|_{T,\mu}^2, \tag{30}$$

with κ equal to the left-hand side of (21), uniformly over the (finitely many) delay subintervals determined by $u_1, \dots, u_p$. Since $\kappa < 1$ by (21), (30) shows F is a strict contraction on the complete metric space $B_R(D_T^\mu, \tilde{\Psi})$ (complete since it is a closed subset of the Banach space D_T^μ). By the Banach fixed point theorem, F has a unique fixed point $z \in B_R(D_T^\mu, \tilde{\Psi})$, which is the sought unique mild solution. □

Remark 6. *Setting $h \equiv 0$ (no stochastic perturbation) and $\sigma = 1$ (Hilfer derivative reduces to Caputo) collapses Theorem 1, conditions (20) and (21) and Lemma 3 exactly to Lemmas 2.1–2.3 and Theorem 3.1 of [7], confirming that the present result is a genuine generalization.*

Remark 7. *The additional stochastic term forces the contraction condition (21) to involve $L_h(r)$ in exactly the same structural position as $L_f(r)$, rather than appearing as a separate, additive smallness restriction; this reflects the fact that, thanks to the Itô isometry (5), the stochastic convolution contributes to the mean-square norm through $\mathbb{E}\|h\|_{L_2^0}^2$ (a quadratic quantity) in exactly the same way that the ordinary convolution contributes through $\big(\int(\vartheta - s)^{\rho-1}\|f\|\,ds\big)^2$ once Cauchy–Schwarz has been applied; no extra structural assumption on h beyond Assumption 5 (i.e., no assumption of, say, boundedness of Q beyond finite trace) is required for Theorem 1 to hold.*

Theorem 2 (Continuous dependence on the data)**.** *Fix the nonlocal map φ (satisfying Assumption 6) and g, f, h as before, and suppose (20) and (21) hold. Let (z_0, Ψ) and $(\tilde{z}_0, \tilde{\Psi})$ be two choices of nonlocal/delay data satisfying Assumption 7, and let $z, \tilde{z} \in B_R(D_T^\mu, \tilde{\Psi})$ be the corresponding unique mild solutions furnished by Theorem 1 (with the* same *φ, g, f, h). Then there is a constant $C = C(T, \rho, \sigma, \mu, \beta, K, B_\mu, L_g, L_f, L_h, p, \kappa) > 0$, independent of $(z_0, \Psi), (\tilde{z}_0, \tilde{\Psi})$, such that*

$$\|z - \tilde{z}\|_{T,\mu}^2 \le C\Big(\mathbb{E}\|z_0 - \tilde{z}_0\|_\mu^2 + \mathbb{E}\big\|\Psi(0) - \tilde{\Psi}(0)\big\|_\mu^2 + \sup_{\vartheta\in[-\eta,0]}\mathbb{E}\big\|\Psi(\vartheta) - \tilde{\Psi}(\vartheta)\big\|_\mu^2\Big). \tag{31}$$

In particular, for fixed *φ, the mild solution depends continuously, in mean square, on the delay/nonlocal initial data, and the solution map $(z_0, \Psi) \mapsto z$ is Lipschitz from $\mathrm{Y}_\mu \times C([-\eta, 0]; \mathrm{Y}_\mu)$ (with the mean-square norm) into D_T^μ. We do not address continuous dependence on φ itself here: since φ enters (8) only through the single value $\varphi(z)$, comparing two solutions built from different maps $\varphi, \tilde{\varphi}$ would require controlling $\varphi(\tilde{z}) - \tilde{\varphi}(\tilde{z})$, i.e., a genuine operator-norm-type distance between φ and $\tilde{\varphi}$ as maps on Y_μ, which Assumption 6 alone—a Lipschitz bound for each map separately—does not supply; this would require an additional hypothesis such as $\sup_{\|z\|_\mu \le R}\mathbb{E}\|\varphi(z) - \tilde{\varphi}(z)\|_\mu^2 \le \delta^2$ and a corresponding δ-term on the right of (31).*

Proof. Let $F, \tilde{F}$ denote the solution operators (8) built from (z_0, Ψ, φ) and $(\tilde{z}_0, \tilde{\Psi}, \varphi)$ respectively (same φ). For $\vartheta \in (0, T]$, write

$$z(\vartheta) - \tilde{z}(\vartheta) = \underbrace{\big[(Fz)(\vartheta) - (F\tilde{z})(\vartheta)\big]}_{=:A(\vartheta)} + \underbrace{\big[(F\tilde{z})(\vartheta) - (\tilde{F}\tilde{z})(\vartheta)\big]}_{=:B(\vartheta)}. \tag{32}$$

We bound $\|A\|_{T,\mu}$ and $\|B\|_{T,\mu}$ separately and combine them via the triangle inequality at the end (rather than splitting the squared norm $\mathbb{E}\|z - \tilde{z}\|_\mu^2$ directly, which would force an unnecessary strengthening of $\kappa < 1$, as noted below).

Bounding $A(\vartheta)$. This is exactly the contraction estimate of Step 2 in the proof of Theorem 1, applied to the pair $(z,\tilde{z}) \in B_R(D_T^\mu,\tilde{\Psi}) \times B_R(D_T^\mu,\tilde{\Psi})$: writing $A(\vartheta) = \sum_{i=1}^5 \left[I_i^z(\vartheta) - I_i^{\tilde{z}}(\vartheta)\right]$ for the same five terms $I_1,\dots,I_5$ as before, the identical computation—Lemma 1 for I_1, I_3, I_4, I_5, the Itô isometry (5) and Beta-integral evaluation (26) for I_5 (as in (29)), and the finer resolvent-continuity estimate noted for I_3—gives

$$\mathbb{E}\|A(\vartheta)\|_\mu^2 \le \kappa\, \|z-\tilde{z}\|_{T,\mu}^2, \tag{33}$$

with $\kappa < 1$ exactly the constant of (21); we do not repeat the five-term calculation here, referring instead to Theorem 1.

Bounding $B(\vartheta)$—worked in full. Unlike A, this term collects only the *data* through which (z_0,Ψ) enter, all evaluated at the *same* function $\tilde{z}$; by (8), only the first two of the five summands actually depend on the data (the convolution terms I_3, I_4, I_5 built from $\tilde{z}$ are identical in $F\tilde{z}$ and $\tilde{F}\tilde{z}$ and cancel exactly), so

$$B(\vartheta) = \underbrace{\Big\{S_{\rho,\sigma}(\vartheta)\big[z_0+\Psi(0)-\varphi(\tilde{z})\big] - S_{\rho,\sigma}(\vartheta)\big[\tilde{z}_0+\tilde{\Psi}(0)-\varphi(\tilde{z})\big]\Big\}}_{=:B_1(\vartheta)} + \underbrace{\Big\{-g\big(0,\Psi(0),\dots\big)+g\big(0,\tilde{\Psi}(0),\dots\big)\Big\}}_{=:B_2(\vartheta)},$$

where B_2 is ϑ-independent (it sits inside $S_{\rho,\sigma}(\vartheta)[\cdots]$ in I_1, but we isolate it here for clarity of the calculation)—more precisely, $B(\vartheta) = S_{\rho,\sigma}(\vartheta)\big[(z_0-\tilde{z}_0) + (\Psi(0)-\tilde{\Psi}(0)) + g(0,\Psi(0),\dots) - g(0,\tilde{\Psi}(0),\dots)\big]$ after using linearity of $S_{\rho,\sigma}(\vartheta)$ to combine B_1, B_2 into a single bracket, since $\varphi(\tilde{z})$ is a single common term that cancels between the two copies of $S_{\rho,\sigma}(\vartheta)[\cdots]$ (it is the *same* value, φ evaluated at the *same* $\tilde{z}$, subtracted inside both). Applying $\mathcal{E}^\mu$, using that $S_{\rho,\sigma}(\vartheta)$ commutes with it, and $\|S_{\rho,\sigma}(\vartheta)\|_{\mu\to\mu} \le K$ (Lemma 1(i)):

$$\begin{aligned}\mathbb{E}\|B(\vartheta)\|_\mu^2 &\le K^2\,\mathbb{E}\Big\|(z_0-\tilde{z}_0) + (\Psi(0)-\tilde{\Psi}(0)) + \big[g(0,\Psi(0),\dots) - g(0,\tilde{\Psi}(0),\dots)\big]\Big\|_\mu^2 \\ &\le 2K^2\,\mathbb{E}\|z_0-\tilde{z}_0\|_\mu^2 + 2K^2\,\mathbb{E}\Big\|(\Psi(0)-\tilde{\Psi}(0)) + \big[g(0,\Psi(0),\dots) - g(0,\tilde{\Psi}(0),\dots)\big]\Big\|_\mu^2,\end{aligned} \tag{34}$$

using $\|a+b\|^2 \le 2\|a\|^2 + 2\|b\|^2$ to isolate the $z_0-\tilde{z}_0$ piece. For the remaining bracket, again by $\|a+b\|^2 \le 2\|a\|^2+2\|b\|^2$ and, exactly as for I_2 in the proof of Theorem 1, $\big\|g(0,\Psi(0),\dots) - g(0,\tilde{\Psi}(0),\dots)\big\|_\mu \le N_{\mu-\beta}\big\|\mathcal{E}^\beta g(0,\Psi(0),\dots) - \mathcal{E}^\beta g(0,\tilde{\Psi}(0),\dots)\big\|$, followed by Assumption 2's Lipschitz bound at $\vartheta_1=\vartheta_2=0$ (so the time-difference term vanishes) and $(z_1,\dots,z_{p+1}) = (\Psi(0),\dots,\Psi(0))$, $(w_1,\dots,w_{p+1}) = (\tilde{\Psi}(0),\dots,\tilde{\Psi}(0))$:

$$\mathbb{E}\Big\|\mathcal{E}^\beta g(0,\Psi(0),\dots) - \mathcal{E}^\beta g(0,\tilde{\Psi}(0),\dots)\Big\|^2 \le L_g(r)(p+1)\,\mathbb{E}\big\|\Psi(0)-\tilde{\Psi}(0)\big\|_\mu^2.$$

Combining,

$$\begin{aligned}&\mathbb{E}\Big\|(\Psi(0)-\tilde{\Psi}(0)) + \big[g(0,\Psi(0),\dots) - g(0,\tilde{\Psi}(0),\dots)\big]\Big\|_\mu^2 \\ &\le 2\mathbb{E}\big\|\Psi(0)-\tilde{\Psi}(0)\big\|_\mu^2 + 2N_{\mu-\beta}^2 L_g(r)(p+1)\,\mathbb{E}\big\|\Psi(0)-\tilde{\Psi}(0)\big\|_\mu^2,\end{aligned} \tag{35}$$

and substituting (35) into (34) gives, after collecting the two $\mathbb{E}\big\|\Psi(0)-\tilde{\Psi}(0)\big\|_\mu^2$-coefficients,

$$\mathbb{E}\|B(\vartheta)\|_\mu^2 \le 2K^2\,\mathbb{E}\|z_0-\tilde{z}_0\|_\mu^2 + 4K^2\big[1+N_{\mu-\beta}^2 L_g(r)(p+1)\big]\,\mathbb{E}\big\|\Psi(0)-\tilde{\Psi}(0)\big\|_\mu^2, \tag{36}$$

uniformly in $\vartheta \in [0,T]$ (note the right-hand side does not depend on ϑ, since $B(\vartheta)$ inherits its ϑ-dependence only through $S_{\rho,\sigma}(\vartheta)$, already absorbed into the constant K).

Combining and solving the resulting inequality. For $\vartheta \in [-\eta, 0]$, $z(\vartheta) - \tilde{z}(\vartheta) = \Psi(\vartheta) - \tilde{\Psi}(\vartheta)$ directly from (3), contributing the third term of (31) directly. For $\vartheta \in (0, T]$, (33) gives $\|A\|_{T,\mu} \le \sqrt{\kappa}\,\|z - \tilde{z}\|_{T,\mu}$ and (36) gives a ϑ-independent bound $\|B\|_{T,\mu} \le \sqrt{D}$ with D the right-hand side of (36); applying the triangle inequality directly to the *norm* (not the squared norm, so as to avoid an unnecessary strengthening of $\kappa < 1$) in (32),

$$\|z - \tilde{z}\|_{T,\mu} \le \|A\|_{T,\mu} + \|B\|_{T,\mu} \le \sqrt{\kappa}\,\|z - \tilde{z}\|_{T,\mu} + \sqrt{D},$$

so that $(1 - \sqrt{\kappa})\|z - \tilde{z}\|_{T,\mu} \le \sqrt{D}$, and since $\kappa < 1$ by (21) (no further strengthening is required),

$$\|z - \tilde{z}\|^2_{T,\mu} \le \frac{D}{(1-\sqrt{\kappa})^2} = \frac{2K^2\,\mathbb{E}\|z_0 - \tilde{z}_0\|^2_\mu + 4K^2\big[1 + N^2_{\mu-\beta} L_g(r)(p+1)\big]\mathbb{E}\|\Psi(0) - \tilde{\Psi}(0)\|^2_\mu}{(1-\sqrt{\kappa})^2} + \sup_{\vartheta\in[-\eta,0]} \mathbb{E}\|\Psi(\vartheta) - \tilde{\Psi}(\vartheta)\|^2_\mu,$$

using (36) for D; this is (31) with $C = \max\big\{2K^2(1-\sqrt{\kappa})^{-2},\, 4K^2[1 + N^2_{\mu-\beta}L_g(r)(p+1)]\,(1-\sqrt{\kappa})^{-2},\, 1\big\}$. □

Corollary 1. *Under the hypotheses of Theorem 2, if $\varphi \equiv 0$ (no nonlocal term), the mild solution depends continuously on the initial datum $z_0 + \Psi(0)$ alone, and the classical (local) Hilfer stochastic Sobolev-type neutral problem is well posed in the sense of Hadamard within $B_R(D^\mu_T, \tilde{\Psi})$.*

4. Approximate Solutions and the Faedo–Galerkin Method

Let $\mathrm{Y}^n = \mathrm{span}\{\chi_0, \dots, \chi_n\} \subset \mathrm{Y}$ and let $P^n : \mathrm{Y} \to \mathrm{Y}^n$ be the associated orthogonal projection, $n = 0, 1, 2, \dots$. By Assumption 1, χ_i is simultaneously an eigenfunction of $\mathcal{A}, \mathcal{B}, \mathcal{E}$; consequently P^n commutes with $\mathcal{A}, \mathcal{B}, \mathcal{E}$ and with every bounded function of these operators (in particular with the fractional power $\mathcal{B}^\beta$ of Assumption 2 and with $S_{\rho,\sigma}(\vartheta), T_\rho(\vartheta)$), since all are diagonal in the common eigenbasis $\{\chi_i\}$ and P^n is exactly the truncation of that diagonal representation to the first $n+1$ coordinates. This commutation—not merely $\|P^n\| \le 1$—is what is used below whenever P^n and $\mathcal{E}^\beta$ (or $\mathcal{E}$) need to be interchanged. Define

$$g_n(\vartheta, z(\vartheta), \dots, z(u_p(\vartheta))) = g(\vartheta, P^n z(\vartheta), \dots, P^n z(u_p(\vartheta))), \qquad f_n,\ h_n \text{ analogously.}$$

Define, for $\vartheta \in [0, T]$,

$$\begin{aligned}(F_n z)(\vartheta) = {} & S_{\rho,\sigma}(\vartheta)\big[z_0 + \Psi(0) + g_n(0, \dots) - \varphi(z)\big] - g_n\big(\vartheta, z(\vartheta), \dots\big) \\ & - \int_0^\vartheta (\vartheta - s)^{\rho-1} T_\rho(\vartheta - s)\,\mathcal{E} g_n(s, z(s), \dots)\,ds + \int_0^\vartheta (\vartheta - s)^{\rho-1} T_\rho(\vartheta - s)\, f_n(s, z(s), \dots)\,ds \\ & + \int_0^\vartheta (\vartheta - s)^{\rho-1} T_\rho(\vartheta - s)\, h_n(s, z(s), \dots)\,dW(s),\end{aligned} \tag{37}$$

and $(F_n z)(\vartheta) = z_0 + \tilde{\Psi}(\vartheta)$ for $\vartheta \in [-\eta, 0]$.

Theorem 3. *Under Assumptions 1–7 and (20) and (21), for each $n = 0, 1, 2, \dots$ there is a unique $z_n \in B_R(D^\mu_T, \tilde{\Psi})$ with $F_n z_n = z_n$.*

Proof. Identical to the proof of Theorem 1, applied to g_n, f_n, h_n in place of g, f, h. Two facts about P^n are used, and are kept distinct: (i) $\|P^n\| \le 1$ on Y (and hence on Y_μ, since P^n commutes with $\mathcal{E}^\mu$ by Assumption 1), which shows g_n, f_n, h_n inherit the *same* Lipschitz/growth constants $L_g(r), L_f(r), L_h(r)$ as g, f, h (composition with a contraction on the argument only); and (ii) the commutation of P^n with $\mathcal{E}^\beta$ noted above, which is what justifies that $\mathcal{E}^\beta g_n(\vartheta, z, \dots) = \mathcal{E}^\beta g(\vartheta, P^n z, \dots)$ is again governed by Assumption 2's bound

on $\mathcal{E}^\beta g$ itself (rather than requiring a separate Lipschitz hypothesis on $\mathcal{E}^\beta g_n$). With both facts in hand, (20) and (21) apply verbatim to F_n. □

Lemma 4. *Under the hypotheses of Theorem 3, if $\Psi(\vartheta) \in D(\mathcal{E})$ for $\vartheta \in [-\eta, 0]$, there exists $U_0 > 0$, independent of n, such that*

$$\mathbb{E}\|z_n(\vartheta)\|_\theta^2 \le U_0, \qquad -\eta \le \vartheta \le T,\ 0 \le \theta < 1.$$

Proof. As in ([7], Lemma 3.1), bound each of the five summands of (37) using Lemma 1 and Assumptions 2–5 (Itô isometry for the stochastic term), noting that all bounds are uniform in n because $\|P^n\| \le 1$; summing gives a bound U_0 independent of n, uniformly over the finitely many subintervals determined by the (at most countably many) delay switching points. □

Theorem 4. *Under the hypotheses of Lemma 4, the sequence $\{z_n\}_{n\ge 0} \subset B_R(D_T^\mu, \tilde{\Psi})$ is Cauchy and converges, in $\|\cdot\|_{T,\mu}$ (equivalently, uniformly in mean square), to the unique mild solution z of (8).*

Proof. Let $z \in B_R(D_T^\mu, \tilde{\Psi})$ be the unique mild solution furnished by Theorem 1, i.e., $Fz = z$. For each n, since $z_n = F_n z_n$,

$$\|z_n - z\|_{T,\mu} = \|F_n z_n - Fz\|_{T,\mu} \le \|F_n z_n - F_n z\|_{T,\mu} + \|F_n z - Fz\|_{T,\mu}.$$

The proof of Theorem 3 shows F_n is a κ-contraction on $B_R(D_T^\mu, \tilde{\Psi})$ with the *same* $\kappa < 1$ as F (the constant of (21), independent of n), so $\|F_n z_n - F_n z\|_{T,\mu} \le \kappa \|z_n - z\|_{T,\mu}$, giving

$$\|z_n - z\|_{T,\mu} \le (1-\kappa)^{-1} \|F_n z - Fz\|_{T,\mu}. \tag{38}$$

It remains to show $\varepsilon_n := \|F_n z - Fz\|_{T,\mu} \to 0$ as $n \to \infty$, for the *fixed* function z—this is a *pointwise* (in n, for fixed z) convergence statement, not a uniform-in-k one, and requires no assumption on the growth of $\{\alpha_i\}$. Comparing (37) and (8) term by term, $(F_n z)(\vartheta) - (Fz)(\vartheta) = \sum_{i=1}^5 \big[I_i^n(\vartheta) - I_i(\vartheta)\big]$, where I_i^n denotes the i-th summand of (37) (built from g_n, f_n, h_n) and I_i the i-th summand of (8) (built from g, f, h); by $\mathbb{E}\big\|\sum_{i=1}^5 a_i\big\|^2 \le 5\sum_{i=1}^5 \mathbb{E}\|a_i\|^2$,

$$\mathbb{E}\|(F_n z)(\vartheta) - (Fz)(\vartheta)\|_\mu^2 \le 5 \sum_{i=1}^5 \mathbb{E}\|I_i^n(\vartheta) - I_i(\vartheta)\|_\mu^2. \tag{39}$$

We work out $I_2^n - I_2$ (the neutral term) and $I_5^n - I_5$ (the stochastic convolution) in full, as representative cases; the remaining three follow the same pattern.

Term $I_2^n - I_2$. Since $g_n(\vartheta, z(\vartheta), \dots) = g(\vartheta, P^n z(\vartheta), \dots)$,

$$I_2^n(\vartheta) - I_2(\vartheta) = g_n\big(\vartheta, z(\vartheta), \dots\big) - g\big(\vartheta, z(\vartheta), \dots\big) = g\big(\vartheta, P^n z(\vartheta), \dots\big) - g\big(\vartheta, z(\vartheta), \dots\big),$$

and, exactly as for I_2 in the proof of Theorem 1, $\|\cdot\|_\mu \le N_{\mu-\beta}\big\|\mathcal{E}^\beta(\cdot)\big\|$, followed by Assumption 2's Lipschitz bound (same time argument ϑ on both sides, so the time-difference term vanishes, and $(z_1, \dots, z_{p+1}) = (P^n z(\vartheta), P^n z(u_1(\vartheta)), \dots)$, $(w_1, \dots, w_{p+1}) = (z(\vartheta), z(u_1(\vartheta)), \dots)$):

$$\begin{aligned}\mathbb{E}\|I_2^n(\vartheta) - I_2(\vartheta)\|_\mu^2 &\le N_{\mu-\beta}^2\, \mathbb{E}\Big\|\mathcal{E}^\beta g(\vartheta, P^n z(\vartheta), \dots) - \mathcal{E}^\beta g(\vartheta, z(\vartheta), \dots)\Big\|^2 \\ &\le N_{\mu-\beta}^2 L_g(r) \sum_{i=1}^{p+1} \mathbb{E}\|(P^n - I) z_i\|_\mu^2 \le N_{\mu-\beta}^2 L_g(r)(p+1) \sup_{\vartheta \in [0,T]} \mathbb{E}\|(I - P^n) z(\vartheta)\|_\mu^2,\end{aligned}$$

where each $z_i \in \{z(\vartheta), z(u_1(\vartheta)), \dots, z(u_p(\vartheta))\}$ and we used $u_i(\vartheta) \in [0,T]$ to bound every such term by the single ϑ-independent supremum.

Term $I_5^n - I_5$—worked in full. By the Itô isometry (5) applied to $h_n(s, z(s), \dots) - h(s, z(s), \dots) = h(s, P^n z(s), \dots) - h(s, z(s), \dots)$, exactly as in (29),

$$\begin{aligned}
\mathbb{E}\|I_5^n(\vartheta) - I_5(\vartheta)\|_\mu^2 &= \int_0^\vartheta (\vartheta - s)^{2(\rho-1)} \mathbb{E}\big\|\mathcal{E}^\mu T_\rho(\vartheta - s)\big[h(s, P^n z(s), \dots) - h(s, z(s), \dots)\big]\big\|_{L_2^0}^2 \, ds \\
&\le \frac{B_\mu^2}{\Gamma(1+\rho(1-\mu))^2} \int_0^\vartheta (\vartheta - s)^{2\rho(1-\mu)-2} \, L_h(r)(p+1) \sup_{s\in[0,T]} \mathbb{E}\|(I - P^n)z(s)\|_\mu^2 \, ds \\
&= \frac{B_\mu^2 L_h(r)(p+1)\, T^{2\rho(1-\mu)-1}}{\Gamma(1+\rho(1-\mu))^2\big(2\rho(1-\mu)-1\big)} \sup_{s\in[0,T]} \mathbb{E}\|(I - P^n)z(s)\|_\mu^2,
\end{aligned}$$

using Lemma 1(ii), Assumption 5's Lipschitz bound, and the same Beta-integral evaluation (26) as in the proof of Theorem 1. Terms $I_1^n - I_1$, $I_3^n - I_3$, $I_4^n - I_4$ are bounded analogously (via Lemma 1(i)/Assumption 6 for I_1, and the Cauchy–Schwarz/Beta-integral or finer resolvent-continuity mechanism of I_3, I_4 in Theorem 1 for I_3, I_4), each producing a bound of the form $C_i \sup_{s\in[0,T]} \mathbb{E}\|(I - P^n)z(s)\|_\mu^2$ for a constant $C_i > 0$ independent of n. Substituting all five bounds into (39) and taking the supremum over ϑ,

$$\varepsilon_n^2 = \|F_n z - Fz\|_{T,\mu}^2 \le 5\Big(\sum_{i=1}^5 C_i\Big) \sup_{s\in[0,T]} \mathbb{E}\|(I - P^n)z(s)\|_\mu^2. \tag{40}$$

It remains to show the right-hand side of (40) $\to 0$ as $n \to \infty$. For each *fixed* ϑ, $\mathbb{E}\|(I - P^n)z(\vartheta)\|_\mu^2 \to 0$ as $n \to \infty$: since $\{\chi_i\}$ is a complete orthonormal system (Assumption 1), $P^n \to I$ strongly on Υ_μ, so $(I - P^n)z(\vartheta) \to 0$ in Υ_μ for $\mathbb{P}$-a.e. ω, and dominated convergence (using $\mathbb{E}\|(I - P^n)z(\vartheta)\|_\mu^2 \le 4\mathbb{E}\|z(\vartheta)\|_\mu^2 \le 4R^2$, uniformly in n, as a dominating bound, since $z \in B_R(D_T^\mu, \tilde{\Psi})$) gives $\mathbb{E}\|(I - P^n)z(\vartheta)\|_\mu^2 \to 0$. To upgrade this from pointwise-in-ϑ to uniform-in-$\vartheta \in [0,T]$ convergence (needed for the supremum in (40)), note that $\vartheta \mapsto \mathbb{E}\|(I - P^n)z(\vartheta)\|_\mu^2$ is continuous on the compact interval $[0,T]$ (since $z \in D_T^\mu$ has continuous paths and P^n is a fixed bounded operator) and decreases monotonically in n for each fixed ϑ (as the inclusion of the range of $P^n \subset P^{n+1}$ gives $\big\|(I - P^{n+1})y\big\|_\mu \le \|(I - P^n)y\|_\mu$ for every y); by Dini's theorem, pointwise-monotone convergence of continuous functions on a compact set to a continuous limit (0) is automatically uniform, so $\sup_{\vartheta\in[0,T]} \mathbb{E}\|(I - P^n)z(\vartheta)\|_\mu^2 \to 0$ as $n \to \infty$. By (40), $\varepsilon_n \to 0$, and hence, by (38), $\|z_n - z\|_{T,\mu} \to 0$ as $n \to \infty$; in particular $\{z_n\}$ converges (a fortiori is Cauchy) to the unique mild solution z. □

Remark 8. *Unlike the pairwise Cauchy comparison of z_n against z_k used in ([7], Theorem 3.2)for the deterministic Caputo case—which bounds $\mathbb{E}\big\|(I - P^n)\mathcal{E}^{-\theta}\big\|$ by the spectral-gap rate $\lambda_n^{-(\theta-\mu)}$ and so requires $\alpha_i \to \infty$ for that rate to vanish—the argument above compares each z_n directly to the (already known to exist, by Theorem 1) true mild solution z, and uses only that $P^n \to I$* strongly *on Υ_μ, a property of any complete orthonormal eigenbasis regardless of whether $\{\alpha_i\}$ is bounded or diverges. This is what allows the concrete example of Section The Faedo–Galerkin Scheme, where $\mathcal{E}$ is a* bounded *operator with $\alpha_i \uparrow 1$ (so the classical spectral-gap rate $\lambda_n^{-(\theta-\mu)} \to 0$ is simply false there, since $\lambda_n \not\to \infty$), to still satisfy the hypotheses of Theorems 4 and 5 below. The trade-off is that no explicit convergence* rate *is asserted in general (only $\varepsilon_n \to 0$); when $\alpha_i \to \infty$, one can additionally invoke the quantitative rate $\varepsilon_n = O(\lambda_n^{-(\theta-\mu)})$, recovering the sharper statement of ([7], Theorem 3.2) as a special case.*

The Faedo–Galerkin Scheme

Proposition 1 (Roadmap of the Faedo–Galerkin construction). *The three objects z_n, $\bar{z}_n$, F_n introduced above, and the coordinate system (42) below, are related as follows.*

(a) $F_n : B_R(D_T^\mu, \tilde{\Psi}) \to D_T^\mu$ *of (37) is the* projected solution operator*: it is built exactly like the solution operator F of (8), except that the neutral, forcing, and diffusion maps g, f, h are replaced by their n-th Galerkin truncations g_n, f_n, h_n (arguments projected onto Y^n via P^n before g, f, h are applied). It is* not itself *restricted to take values in Y^n.*

(b) $z_n \in B_R(D_T^\mu, \tilde{\Psi})$*, given by Theorem 3, is the (genuinely* Y*-valued, in general not finite-dimensional) fixed point $F_n z_n = z_n$; it is the solution of an auxiliary problem in which only the* nonlinearities *have been truncated, and $\{z_n\}$ is shown in Theorem 4 to converge to the true mild solution z as $n \to \infty$.*

(c) $\bar{z}_n := P^n z_n \in \mathrm{Y}^n$ *is the additional, genuinely finite-dimensional projection of z_n itself onto* Y^n*;* this *is the Faedo–Galerkin approximation proper, i.e., the object satisfying the projected mild-solution identity (41) and, ultimately, the scalar system (42). Theorem 5 shows $\bar{z}_n \to z$ by combining $z_n \to z$ (Theorem 4) with $P^n \to I$ strongly (triangle inequality, as in the short proof given there); no separate contraction argument for $\bar{z}_n$ is needed.*

(d) Applying P^n to (41) and expanding both $z_n(\vartheta) = \sum_{i\geq 0} \rho_i(\vartheta)\chi_i$ (so that $\bar{z}_n(\vartheta) = P^n z_n(\vartheta) = \sum_{i=0}^n \rho_i(\vartheta)\chi_i$, matching the notation ρ_i^n used below) and $\mathcal{E}\chi_i = \alpha_i \chi_i$ in the common eigenbasis of Assumption 1 converts the Y*-valued identity (41) into $n+1$ coupled* scalar *equations, one per coordinate $i = 0, \ldots, n$, by taking the inner product of (41) with each χ_i and using that P^n (and hence every term of (41)) is supported on* span$\{\chi_0, \ldots, \chi_n\}$*; differentiating this integral identity in the Hilfer sense recovers precisely the finite-dimensional Hilfer stochastic differential system (42), which is therefore the coordinate form of (41) and not a separate approximation.*

Let $\bar{z}_n := P^n z_n$. The Faedo–Galerkin approximation to (1)–(3) is

$$\begin{aligned}\bar{z}_n(\vartheta) = {} & S_{\rho,\sigma}(\vartheta)\big[P^n(z_0 + \Psi(0)) + P^n g(0,\ldots) - P^n\varphi(z)\big] - P^n g\big(\vartheta, z_n(\vartheta),\ldots\big) \\ & - \int_0^\vartheta (\vartheta - s)^{\rho-1} T_\rho(\vartheta - s)\, \mathcal{E} P^n g(s, z_n(s),\ldots)\, ds + \int_0^\vartheta (\vartheta - s)^{\rho-1} T_\rho(\vartheta - s)\, P^n f(s, z_n(s),\ldots)\, ds \\ & + \int_0^\vartheta (\vartheta - s)^{\rho-1} T_\rho(\vartheta - s)\, P^n h(s, z_n(s),\ldots)\, dW(s), \qquad \vartheta \in [0, T],\end{aligned} \tag{41}$$

and $\bar{z}_n(\vartheta) = P^n(z_0 + \tilde{\Psi}(\vartheta))$, $\vartheta \in [-\eta, 0]$. Writing $z(\vartheta) = \sum_{i=0}^\infty \rho_i(\vartheta)\chi_i$, $\bar{z}_n(\vartheta) = \sum_{i=0}^n \rho_i^n(\vartheta)\chi_i$, with $\rho_i(\vartheta) = \langle z(\vartheta), \chi_i\rangle$, $\rho_i^n(\vartheta) = \langle \bar{z}_n(\vartheta), \chi_i\rangle$, substitution into (41) yields the system of scalar stochastic fractional equations, for $i = 0, \ldots, n$,

$${}^H D_{0^+}^{\rho,\sigma} \rho_i^n(\vartheta) + \alpha_i \rho_i^n(\vartheta) = g_i^n(\vartheta, \boldsymbol{\rho}^n(\vartheta), \boldsymbol{\rho}^n(u_1(\vartheta)),\ldots) + f_i^n(\vartheta, \boldsymbol{\rho}^n(\vartheta),\ldots) + \dot{h}_i^n(\vartheta, \boldsymbol{\rho}^n(\vartheta),\ldots), \qquad \rho_i^n(0) = \chi_i^0, \tag{42}$$

where $\boldsymbol{\rho}^n = (\rho_0^n, \ldots, \rho_n^n)$, $g_i^n = \langle g(\vartheta, \sum_i \rho_i^n \chi_i, \ldots), \chi_i\rangle$ and similarly for f_i^n, h_i^n; (42) is a coupled system of finite-dimensional Hilfer fractional stochastic differential equations that can, in principle, be solved/simulated numerically (e.g., via a fractional Euler–Maruyama scheme).

Theorem 5. *Suppose Assumptions 1–7 and conditions (20) and (21) hold, and $\Psi(\vartheta) \in D(\mathcal{E})$, $\vartheta \in [-\eta, 0]$. Then there exist $\bar{z}_n \in B_R(D_T^\mu, \tilde{\Psi})$ satisfying (41) and $z \in B_R(D_T^\mu, \tilde{\Psi})$ satisfying (8) such that*

$$\|\bar{z}_n - z\|_{T,\mu} \longrightarrow 0 \qquad \text{as } n \to \infty.$$

Proof. By Theorem 4, $z_n \to z$ in D_T^μ. Since $P^n \to I$ strongly on Y_μ and $\|P^n\| \leq 1$,

$$\|\bar{z}_n - z\|_{T,\mu} \leq \|P^n z_n - P^n z\|_{T,\mu} + \|P^n z - z\|_{T,\mu} \leq \|z_n - z\|_{T,\mu} + \|(P^n - I)z\|_{T,\mu} \longrightarrow 0,$$

using $\|z_n - z\|_{T,\mu} \to 0$ (Theorem 4) and dominated convergence together with $\|(P^n - I)z(\vartheta)\|_\mu \to 0$ pointwise (Assumption 1) and $\mathbb{E}\|z(\vartheta)\|_\mu^2 \le R^2$ uniformly. □

5. Application: A Stochastically Perturbed Digital Filter Model

We illustrate the abstract results on a Sobolev-type stochastic filter model generalizing the deterministic filter example of ([7], Section 5). A digital filter processes a sampled input signal to produce an output signal; in practice the input, and hence the filter state, is corrupted by measurement or channel noise, which is naturally modeled as an additive Q-Wiener perturbation acting through the diffusion coefficient h. The system is described symbolically by (1)–(3) with g representing the (delayed) phase-shifting network, f the deterministic Fourier/convolution processing block, and h the noise-injection block; Figure 1 depicts the resulting signal flow. The phase-shifted input $g(\vartheta, z(\vartheta), z(u_1(\vartheta)), \ldots, z(u_p(\vartheta)))$ is combined with the neutral Hilfer fractional memory kernel $S_{\rho,\sigma}(\vartheta)$ acting on the nonlocal initial data, the deterministic block $(\vartheta - s)^{\rho-1} T_\rho(\vartheta - s) f$ is integrated against ds over $(0, \vartheta)$, and the stochastic block $(\vartheta - s)^{\rho-1} T_\rho(\vartheta - s) h$ is integrated against the increments $dW(s)$ of the driving Wiener process; the summer network combines all contributions to output $z(\vartheta)$.

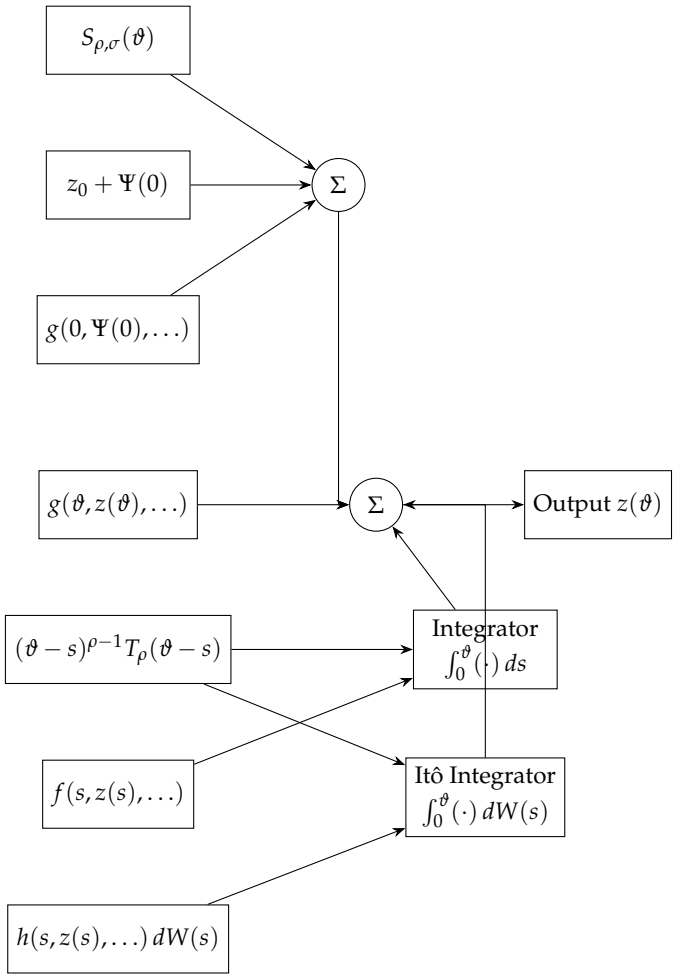

Figure 1. Signal-flow diagram of the stochastically perturbed Sobolev-type Hilfer filter model. The deterministic Fourier/convolution path (f, ordinary integrator) and the noise-injection path (h, Itô integrator) are combined at the summer network together with the neutral memory term g and the nonlocal/initial contribution processed by $S_{\rho,\sigma}(\vartheta)$.

Taking $\sigma = 1$ (Caputo case) and $h \equiv 0$ (no noise) in this model recovers exactly the deterministic filter example of ([7], Section 5), consistent with the reduction noted after Theorem 1.

A Concrete Sobolev-Type Stochastic Heat/Diffusion Example

To make Assumptions 1–7 verifiable in a specific case, consider the spatial domain $\mathcal{O} = (0, \pi) \subset \mathbb{R}$, the Hilbert space $\mathrm{Y} = L^2(0, \pi)$, and the standard eigenbasis $\chi_i(x) = \sqrt{2/\pi}\sin(ix)$, $i = 1, 2, \ldots$, of the Dirichlet Laplacian $-\partial_{xx}^2$ on $H^2(0, \pi) \cap H_0^1(0, \pi)$, with eigenvalues i^2. Let

$$\mathcal{A} = I - \partial_{xx}^2, \qquad \mathcal{B} = -\partial_{xx}^2,$$

both with domain $H^2(0, \pi) \cap H_0^1(0, \pi)$; these are the classical Sobolev-type operators used, e.g., for pseudo-parabolic (Sobolev, Barenblatt–Zheltov–Kochina) equations, and satisfy (C1)–(C3) with $\mathcal{A}^{-1}$ compact and self-adjoint, and with χ_i simultaneously diagonalizing $\mathcal{A}, \mathcal{B}, \mathcal{E}$ as required by Assumption 1. Then $\mathcal{E} = \mathcal{B}\mathcal{A}^{-1}$ has eigenpairs $(\alpha_i, \chi_i) = (i^2/(1 + i^2), \chi_i)$, so $0 < \alpha_i < 1$ for every i and Assumption 1 holds with $\alpha_i \uparrow 1$ as $i \to \infty$—a *bounded*, rather than unbounded, spectrum, which is typical of genuinely Sobolev-type (as opposed to purely parabolic) problems, since (C2) forces $D(\mathcal{A}) \subset D(\mathcal{B})$, which is easiest to arrange when $\mathcal{A}, \mathcal{B}$ have comparable order, giving a *bounded* $\mathcal{E} = \mathcal{B}\mathcal{A}^{-1}$. As discussed after Assumption 1, this is deliberately compatible with the Faedo–Galerkin theory of Section 4 (Remark 8), even though the classical spectral-gap rate $\lambda_n^{-(\theta-\mu)} \to 0$ does *not* apply here (since $\lambda_n \not\to \infty$); only the finite-T bound $N_0 = \sup_{\vartheta\in[0,T]}\|S(\vartheta)\| < \infty$ of (4) is used, which holds automatically for this bounded $\mathcal{E}$.

Consider the stochastically forced, delayed, nonlocal Sobolev–Hilfer initial-boundary value problem, for $x \in (0, \pi)$, $\vartheta \in (0, T]$,

$$\begin{aligned} {}^{H}D_{0^+}^{\rho,\sigma}\Big[(I - \partial_{xx}^2)z(\vartheta, x) + c_0\int_0^\pi k(x, \xi)\, z\big(u_1(\vartheta), \xi\big)\, d\xi\Big] &= -\partial_{xx}^2 z(\vartheta, x) + c_1 \sin\big(z(\vartheta, x)\big) + c_2\, z(\vartheta, x)\, \dot\beta(\vartheta), \\ z(\vartheta, 0) = z(\vartheta, \pi) &= 0, \qquad \vartheta \in [0, T], \\ z(\vartheta, x) &= z_0(x) + \Psi(\vartheta, x), \qquad \vartheta \in [-\eta, 0], \end{aligned} \tag{43}$$

with nonlocal condition $I_{0^+}^{(1-\rho)(1-\sigma)}\big[(I - \partial_{xx}^2)z\big](0^+, x) + \sum_{j=1}^k c_j' z(\vartheta_j, x) = (I - \partial_{xx}^2)[z_0(x) + \Psi(0, x)]$ for fixed $0 < \vartheta_1 < \cdots < \vartheta_k \le T$, $k(\cdot, \cdot) \in L^2((0, \pi)^2)$, and β a standard real Brownian motion so that $W(\vartheta) = \beta(\vartheta)e_0$ (a one-dimensional, hence trivially trace-class, Q-Wiener process with $K = \mathbb{R}$). Here $g_0(\vartheta, z, z(u_1(\vartheta))) = c_0\int_0^\pi k(\cdot, \xi)z(u_1(\vartheta), \xi)\, d\xi =: c_0\,\mathcal{K}z(u_1(\vartheta))$, $f_0(\vartheta, z) = c_1\sin(z)$, $h_0(\vartheta, z) = c_2 z$, $\varphi_0(z) = \sum_j c_j' z(\vartheta_j, \cdot)$ are the "physical" maps of (1), and $\mathcal{K}$ denotes the Hilbert–Schmidt integral operator with kernel k.

By Assumption 3, we must verify Assumptions 2–6 for the $\mathcal{A}^{-1}$-composed maps $g = \mathcal{A}^{-1}g_0 = c_0\mathcal{A}^{-1}\mathcal{K}z(u_1(\vartheta))$, $f = \mathcal{A}^{-1}f_0$, $h = \mathcal{A}^{-1}h_0$, $\varphi = \mathcal{A}^{-1}\varphi_0$—and, per Assumption 2 (using $\mathcal{E}^\beta$, as justified in Remark 3), this means verifying that $\mathcal{E}^\beta g$ (not merely g itself) is Lipschitz. This is straightforward precisely because $\mathcal{E}$ is bounded in this example: $\mathcal{E}^\beta$ has eigenvalues $\alpha_i^\beta \in (0, 1)$ (bounded by 1, since $0 < \alpha_i < 1$), so $\|\mathcal{E}^\beta\| \le 1$ on all of $\mathrm{Y} = L^2(0, \pi)$—no domain restriction, and no separate regularity condition on the kernel k beyond $k \in L^2((0, \pi)^2)$ is needed. Since $\mathcal{A}^{-1}$ is also bounded ($W_1 = \|\mathcal{A}^{-1}\| \le 1$, as $1 + i^2 \ge 1$) and $\mathcal{K}$ is a bounded (Hilbert–Schmidt) operator with $\|\mathcal{K}\| \le \|k\|_{L^2((0,\pi)^2)}$,

$$\mathcal{E}^\beta g = c_0\,\mathcal{E}^\beta\mathcal{A}^{-1}\mathcal{K}\,z(u_1(\vartheta)), \qquad \mathbb{E}\Big\|\mathcal{E}^\beta g(\vartheta_1, z_1, \ldots) - \mathcal{E}^\beta g(\vartheta_2, w_1, \ldots)\Big\|^2 \le c_0^2\|k\|_{L^2((0,\pi)^2)}^2\,\mathbb{E}\|z(u_1(\vartheta_1)) - w(u_1(\vartheta_2))\|^2,$$

verifying Assumption 2 for $\mathcal{E}^\beta g$ (not merely for g) with $L_g(r) = c_0^2\|k\|_{L^2((0,\pi)^2)}^2$, independent of r. As in Remark 3, had we instead tried to hypothesize Lipschitz continuity of $\mathcal{B}^\beta g$ or

$\mathcal{A}^\beta g$ directly, the argument would fail: $\mathcal{A} = I - \partial_{xx}^2$ has eigenvalues $1 + i^2 \to \infty$, so $\mathcal{A}^\beta$, and hence $\mathcal{B}^\beta = \mathcal{E}^\beta \mathcal{A}^\beta$, is an *unbounded* operator for every $\beta > 0$, and no analogous clean bound would be available. This is precisely why Assumption 2 is phrased in terms of $\mathcal{E}^\beta$, not $\mathcal{B}^\beta$. Similarly $f = \mathcal{A}^{-1} f_0$ is bounded ($\|f\| \le W_1 c_1 \sqrt{\pi}$, $W_1 = \left\|\mathcal{A}^{-1}\right\| \le 1$) and globally Lipschitz with constant $\le W_1^2 c_1^2 \le c_1^2$ (since $\mathcal{A}^{-1}$ is a contraction here, $\left\|\mathcal{A}^{-1}\right\| \le 1$, and $|\sin a - \sin b| \le |a - b|$), verifying Assumption 4 with $L_f(r) \equiv c_1^2$; $h = \mathcal{A}^{-1} h_0$ is linear, hence globally Lipschitz on Y_μ with $L_h(r) \le c_2^2$, verifying Assumption 5; and $\varphi = \mathcal{A}^{-1}\varphi_0 = \sum_{j=1}^k c_j' \mathcal{A}^{-1} z(\vartheta_j, \cdot)$ is a finite sum of evaluations of the Y_μ-valued path $z(\cdot) \in D_T^\mu$ at the *fixed times* $\vartheta_j \in [0, T]$, composed with the bounded operator $\mathcal{A}^{-1}$—i.e., for each j, $z \mapsto \mathcal{A}^{-1} z(\vartheta_j, \cdot)$ is the (bounded) coordinate map $D_T^\mu \to \mathrm{Y}_\mu$, $z \mapsto z(\vartheta_j)$, composed with $\mathcal{A}^{-1}$; no spatial embedding is needed (in particular $\mathrm{Y}_\mu \hookrightarrow C(0, \pi)$ plays no role here, since φ is not evaluated at a point in space, only in time). Directly from the triangle inequality and $\left\|\mathcal{A}^{-1}\right\| \le 1$, $\mathbb{E}\|\varphi(z_1) - \varphi(z_2)\|_\mu^2 \le \left(\sum_j |c_j'|\right)^2 \sup_{\vartheta \in [0,T]} \mathbb{E}\|z_1(\vartheta) - z_2(\vartheta)\|_\mu^2$, verifying Assumption 6 with $C_\varphi = \left(\sum_j |c_j'|\right)^2$. Consequently, provided T, c_0, c_1, c_2, c_j' are small enough (or T is small enough relative to the fixed c_0, c_1, c_2, c_j') that (20) and (21) hold. Because all the Lipschitz constants above are global (independent of r), the conditions reduce here to a single explicit smallness inequality in T, whose right-hand side involves only $\rho, \mu, \beta, \sigma, T$ and the coefficients c_0, c_1, c_2, c_j', all of which are freely at the experimenter's disposal.

Remark 9 (An explicit admissible parameter choice)**.** *Take $\rho = 0.9$, $\mu = 0.3$, $\beta = 0.1$, $\sigma = 1$ as in Remark 1, so that $\rho(1 - \mu) = 0.63 > 1/2$; a single delay term ($p = 1$) with $u_1(\vartheta) = q\vartheta$, $0 \le q \le 1$ (e.g., $q = 1/2$), which satisfies Assumption 7; a single nonlocal evaluation time $k = 1$, $\vartheta_1 = T/2$; and coefficients $c_0 = c_1 = c_2 = c_1' = 0.05$, together with a horizon $T = 0.1$. With this choice every Lipschitz/growth constant appearing in (20) and (21) above is proportional to c_0^2, c_1^2, c_2^2, or $(c_1')^2$, i.e., to 0.0025, while the various Gamma-function and T-power prefactors in (20) and (21) are, for these values of ρ, μ, β and $T = 0.1$, all of moderate size (none of the relevant exponents $2\rho(1 - \mu) - 1 = 0.134$, $2\rho(\mu - \beta) = 0.36$ is close to a Gamma-function pole); consequently the left-hand sides of (20) and (21) are of order $c^2 \times T^{(\text{small positive power})} \ll 1$ for $c = 0.05$, and (20) and (21) hold for this choice (and, by continuity, for an open neighborhood of it) with substantial room to spare, e.g., by taking $R = 1$ once $\mathbb{E}\|z_0 + \Psi(0)\|_\mu^2$ and $\mathbb{E}\|\Psi\|_{0,\mu}^2$ are also fixed at order-1 values. This confirms concretely, rather than only qualitatively, that problem (43) satisfies the hypotheses of Theorems 1, 4, and 5.*

Theorem 1 guarantees problem (43) has a unique mild solution $z \in B_R(D_T^\mu, \tilde{\Psi})$, Theorem 4 guarantees the fixed point iterates z_n converge to it in mean square, and Theorem 5 guarantees that its truncated Fourier (Galerkin) approximations $\bar{z}_n(\vartheta, x) = \sum_{i=1}^n \rho_i^n(\vartheta)\chi_i(x)$, obtained by numerically solving the coupled scalar system (42)—here, a system of n coupled real-valued Hilfer stochastic fractional differential equations driven by the single Brownian motion β—converge to z in mean square as $n \to \infty$. This example also illustrates the role of σ: taking $\sigma \to 1$ makes the nonlocal condition above converge to the Caputo-type pointwise condition $(I - \partial_{xx}^2) z(0, x) + \sum_j c_j' z(\vartheta_j, x) = (I - \partial_{xx}^2)[z_0(x) + \Psi(0, x)]$, whereas $\sigma < 1$ allows the initial datum to be prescribed only on a fractional integral (i.e., a time-averaged, memory-weighted version) of the neutral term, which is the physically relevant choice whenever the "initial" measurement of the system is itself obtained through an averaging or filtering process, as is common in the digital-filter interpretation of Figure 1. In terms of the underlying filter, c_0 scales the strength of the delayed feedback/memory coupling encoded by the kernel k, c_1 scales the (saturating, since $|\sin| \le 1$) nonlinear forcing intrinsic to the filter dynamics, and c_2 scales the intensity of the additive channel or measurement noise; the smallness condition of Remark 9 therefore has the concrete reading that a stable, uniquely solvable filter model of this Sobolev type is guaranteed whenever

the feedback strength, nonlinearity, and noise intensity are jointly moderate relative to the observation horizon T, which is the regime of practical interest for a lightly damped digital filter. The advantage of the Hilfer derivative over the Caputo derivative in this model is precisely the extra type parameter σ just discussed: a purely Caputo-based model ($\sigma = 1$ in our formulation) forces the nonlocal/initial condition to be a pointwise statement about $z(0, \cdot)$ itself, whereas a filter whose very first reading is already an average over a short warm-up window—for instance, a running-average or low-pass pre-filter applied before sampling begins—is more faithfully modeled by $\sigma < 1$, where the initial data constrains only $I_{0^+}^{(1-\rho)(1-\sigma)}[\cdot](0^+)$, a fractional-integral (memory-weighted) functional of the neutral term rather than its instantaneous value; the Hilfer formulation therefore strictly enlarges, rather than merely generalizes for its own sake, the class of physically realizable initial measurements this model can represent.

6. Conclusions

We generalized the neutral, nonlocal, Sobolev-type fractional differential equation with finite delay of Kaliraj et al. [7] in two independent directions: replacing the Caputo derivative by the more general Hilfer fractional derivative ${}^{H}D_{0^+}^{\rho,\sigma}$, and adding a stochastic perturbation driven by a Q-Wiener process. Using a probability-density (Mainardi–Wright) representation of the associated Hilfer semigroup-type operators $S_{\rho,\sigma}(\vartheta)$ and $T_{\rho}(\vartheta)$, we derived the stochastic Volterra mild-solution formula, established existence and uniqueness of the mild solution in the mean-square sense via the Banach contraction principle, and constructed a Faedo–Galerkin approximation scheme whose solutions converge in mean square to the unique mild solution. Setting $\sigma = 1$ and $h \equiv 0$ recovers precisely the results of [7], confirming the present work as a strict generalization; we also established, as a genuinely new result with no counterpart in [7], mean-square continuous dependence of the mild solution on the nonlocal and delay data (Theorem 2), and verified the full set of hypotheses on a concrete Sobolev-type stochastic heat/diffusion example with a finite-delay nonlocal condition (Section The Faedo–Galerkin Scheme).

Several directions remain open for future work.

(a) **Local, rather than global, Lipschitz hypotheses.** Assumptions 2–6 are global (or, in the case of L_g, L_f, L_h, uniform over bounded sets with the ball radius r fixed in advance); a genuinely local theory, in which the Lipschitz constants are allowed to blow up as $r \to \infty$ and existence is obtained on a possibly shrinking time interval via a standard continuation/truncation argument, would substantially broaden the class of nonlinearities g, f, h covered, at the cost of a more delicate analysis of the maximal interval of existence.

(b) **Approximate controllability.** A natural next step, in the spirit of [13,24–27], is to append a control term $\mathcal{C}u(\vartheta)$ (with $\mathcal{C}$ a bounded linear control operator and u taking values in a separable control Hilbert space U) to the right-hand side of (1) and to study approximate controllability of the resulting Hilfer stochastic Sobolev-type neutral system with finite delay, using the associated controllability (Gramian) operator together with the fixed point and Faedo–Galerkin machinery developed here.

(c) **Impulsive effects.** Finally, combining the present framework with instantaneous or non-instantaneous impulses at the switching times ϑ_k already present in the weighted space D_T^{μ} (but not otherwise used in the Equation (1) itself) would connect this work back to the impulsive Sobolev-type Caputo theory of [8] and its references, now in the Hilfer stochastic setting.

(d) **Infinite delay and other driving noises.** The finite delay hypothesis (Assumption 7) could be relaxed to infinite delay via a suitable phase space, in the spirit of [14,15]; combined with the present Sobolev-type, finite-delay, Faedo–Galerkin machinery, this

would remove one of the few remaining gaps relative to the existing Hilfer stochastic literature discussed in Section 1. Likewise, replacing the Q-Wiener process W by fractional Brownian motion or a Rosenblatt process, as in [18], would connect the present results to the growing literature on non-Markovian driving noise for Hilfer fractional systems.

(e) **Numerical validation.** The coordinate system (42) underlying the Faedo–Galerkin scheme is, by construction, amenable to numerical simulation via a fractional Euler–Maruyama or similar scheme; a systematic numerical study of the convergence rate of $\bar{z}_n$ to z (complementing the qualitative convergence established in Theorem 5 and the rate discussion of Remark 8) on the concrete filtering model of Section 5 is a natural and, we expect, tractable next step.

Taken together, these directions would extend the present theory of existence, uniqueness, and approximation toward the controllability, long-time behavior, and computational questions that are of most direct interest in the filtering applications that motivate the Sobolev-type model of Section 5.

Author Contributions: Conceptualization, S.R.M. and A.K.; methodology, S.R.M.; formal analysis, S.R.M. and A.K.; investigation, S.R.M. and A.K.; writing—original draft preparation, S.R.M. and A.K.; writing—review and editing, A.K.; supervision, S.R.M. and A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KFU265229].

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Khan, W.A.; Wani, S.A.; Ayman-Mursaleen, M.; Kotecha, K.; Jadhav, P. Extended forms of Legendre-Laguerre-based hybrid polynomials and their characteristics via fractional operator approach. *J. Inequal. Appl.* **2026**, *2026*, 24. [CrossRef]
2. Riyasat, M.; Wani, S.A.; Khan, S.; Shah, S. Operational perspective on degenerate Gould–Hopper–Fubini polynomials using fractional calculus. *Georgian Math. J.* **2026**. [CrossRef]
3. Hilfer, R. (Ed.) *Applications of Fractional Calculus in Physics*; World Scientific: Singapore, 2000.
4. Furati, K.M.; Kassim, M.D.; Tatar, N.E. Existence and uniqueness for a problem involving Hilfer fractional derivative. *Comput. Math. Appl.* **2012**, *64*, 1616–1626. [CrossRef]
5. Gu, H.; Trujillo, J.J. Existence of mild solution for evolution equation with Hilfer fractional derivative. *Appl. Math. Comput.* **2015**, *257*, 344–354. [CrossRef]
6. Debbouche, A.; Nieto, J.J. Sobolev type fractional abstract evolution equations with nonlocal conditions and optimal multi-controls. *Appl. Math. Comput.* **2014**, *245*, 74–85. [CrossRef]
7. Kaliraj, K.; Manjula, M.; Thilakraj, E.; Ravichandran, C.; Nisar, K.S.; El-Ebiary, Y.A.B.; Hourani, A.O. Discussions on Sobolev type Neutral Nonlocal fractional differential equation. *Partial Differ. Equ. Appl. Math.* **2025**, *13*, 101018. [CrossRef]
8. Kaliraj, K.; Manjula, M.; Ravichandran, C.; Nisar, K.S. Results on neutral differential equation of sobolev type with nonlocal conditions. *Chaos Solitons Fractals* **2022**, *158*, 112060. [CrossRef]
9. Prato, G.D.; Zabczyk, J. *Stochastic Equations in Infinite Dimensions*; Cambridge University Press: Cambridge, UK, 1992.
10. Byszewski, L. Theorems about the existence and uniqueness of solutions of a semilinear evolution nonlocal Cauchy problem. *J. Math. Anal. Appl.* **1991**, *162*, 497–505. [CrossRef]
11. Zhou, Y.; Jiao, F. Nonlocal Cauchy problem for fractional evolution equations. *Nonlinear Anal. Real World Appl.* **2010**, *11*, 4465–4475. [CrossRef]
12. Mao, X. *Stochastic Differential Equations and Applications*, 2nd ed.; Horwood Publishing: Chichester, UK, 2007.
13. Sakthivel, R.; Ren, Y.; Mahmudov, N.I. On the approximate controllability of semilinear fractional differential systems. *Comput. Math. Appl.* **2011**, *62*, 1451–1459. [CrossRef]

14. Pradeesh, J.; Vijayakumar, V. On the Asymptotic Stability of Hilfer Fractional Neutral Stochastic Differential Systems with Infinite Delay. *Qual. Theory Dyn. Syst.* **2024**, *23*, 153. [CrossRef]
15. Sivasankar, S.; Udhayakumar, R.; Muthukumaran, V.; Gokul, G.; Al-Omari, S. Existence of Hilfer fractional neutral stochastic differential systems with infinite delay. *Bull. Karaganda Univ. Math. Ser.* **2024**, *113*, 174–193. [CrossRef]
16. Shah, R.; Irshad, N. Existence and uniqueness of solutions of Hilfer fractional neutral impulsive stochastic delayed differential equations with nonlocal conditions. *J. Nonlinear Complex Data Sci.* **2026**, *27*, 43–63. [CrossRef]
17. Raheem, A.; Alamrani, F.M.; Akhtar, J.; Alatawi, A.; Alshaban, E.; Khatoon, A.; Khan, F.A. Study on Controllability for Ψ-Hilfer Fractional Stochastic Differential Equations. *Fractal Fract.* **2024**, *8*, 727. [CrossRef]
18. Lavanya, M.; Vadivoo, B.S.; Nisar, K.S. Controllability Analysis of Neutral Stochastic Differential Equation Using Ψ-Hilfer Fractional Derivative With Rosenblatt Process. *Qual. Theory Dyn. Syst.* **2025**, *24*, 19. [CrossRef]
19. Lv, J.; Yang, M. A Study Concerning to the Existence and Averaging Principle for Ψ-Hilfer Fractional Stochastic Evolution Equations. *Math. Methods Appl. Sci.* **2025**, *48*, 15773–15783. [CrossRef]
20. Dhanush, A.; Vijayakumar, V. Results on the existence and time optimal control results for Hilfer fractional stochastic differential inclusions in Hilbert spaces. *J. Appl. Math. Comput.* **2025**, *71*, 3997–4023. [CrossRef]
21. Gou, H. Study on Sobolev type Hilfer evolution equations with non-instantaneous impulses. *Int. J. Comput. Math.* **2023**, *100*, 1153–1170. [CrossRef]
22. Gou, H.; Li, Y. Extremal mild solutions to Hilfer evolution equations with non-instantaneous impulses and nonlocal conditions. *Fract. Calc. Appl. Anal.* **2023**, *26*, 1145–1185. [CrossRef]
23. Pazy, A. *Semigroups of Linear Operators and Applications to Partial Differential Equations*; Applied Mathematical Sciences; Springer: New York, NY, USA, 1983; p. 44.
24. Ayman-Mursaleen, M.; Alshaban, E.; Nasiruzzaman, M. Approximation to family of α-Bernstein operators using shifted knot properties. *J. Inequal. Appl.* **2025**, *2025*, 107. [CrossRef]
25. Ayman-Mursaleen, M. Controllability of tempered Caputo fractional dynamical systems. *J. Nonlinear Convex Anal.* **2025**, *26*, 2501–2511.
26. Haque, I. Controllability for Ψ-Hilfer fractional Sobolev-type stochastic differential system. *Fixed Point Theory Algorithms Sci. Eng.* **2026**, *2026*, 27. [CrossRef]
27. Ayman-Mursaleen, M.; Saeed, S.; Almohammadi, S.; Arif, K.; Imran, M. A deep neural network model for heat transfer in darcy–forchheimer hybrid nanofluid flow with activation energy. *Sci. Rep.* **2026**, *16*, 8339. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(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.