Discrete and Continuous Dynamical Systems - Series S
Vol. 15, No. 8, August 2022, pp. 1883–1918
doi:10.3934/dcdss.2022012

A UNIFIED FRAMEWORK FOR OPTIMAL CONTROL OF FRACTIONAL IN TIME SUBDIFFUSIVE SEMILINEAR PDES

Harbir Antil

The Center for Mathematics and Artificial Intelligence
Department of Mathematical Sciences
George Mason University
Fairfax, VA 22030, USA

Ciprian G. Gal

Department of Mathematics and Statistics
Florida International University
Miami, FL 33199, USA

Mahamadi Warma

The Center for Mathematics and Artificial Intelligence
Department of Mathematical Sciences
George Mason University
Fairfax, VA 22030, USA

Abstract. We consider optimal control of fractional in time (subdiffusive, i.e., for $0 < \gamma < 1$) semilinear parabolic PDEs associated with various notions of diffusion operators in an unifying fashion. Under general assumptions on the nonlinearity we first show the existence and regularity of solutions to the forward and the associated backward (adjoint) problems. In the second part, we prove existence of optimal controls and characterize the associated first order optimality conditions. Several examples involving fractional in time (and some fractional in space diffusion) equations are described in detail. The most challenging obstacle we overcome is the failure of the semigroup property for the semilinear problem in any scaling of (frequency-domain) Hilbert spaces.

1. Introduction.

Optimization problems constrained by partial differential equations (PDEs) are ubiquitous in science and engineering. Without any specific mention, we will refer to these problems as optimal control problems. See the monographs [34, 27, 7, 28] and references therein for many applications and general results for such problems. In particular, such optimization problems arise in fluid dynamics, superconductivity, phase field modeling, regularized variational inequalities and contact problems, etc. These all are the examples of optimization problems with parabolic semilinear PDEs as constraints. Semilinear optimal control problems are known to be a key testbed for developing new algorithms and/or analysis and there

2020 *Mathematics Subject Classification.* Primary: 49J20, 49K20; Secondary: 35S15,65R20.

Key words and phrases. Generalized Caputo time-fractional derivative, time fractional sudiffusive semilinear equations, optimal solutions, optimality conditions.

The first author is partially supported by NSF grants DMS-2110263, DMS-1913004 and the Air Force Office of Scientific Research under Award NO: FA9550-19-1-0036. The last author is partially supported by the Air Force Office of Scientific Research under Award NO: FA9550-18-1-0242 and the Army Research Office (ARO) under Award NO: W911NF-20-1-0115.

is a significant amount of literature available on this topic [19, 20, 12, 8]. The optimization variable (control variable) either acts in the interior (distributed control), or on the boundary (boundary control), or in the exterior (exterior control). The first two notions of controls are well-known, the exterior control is new, see [36, 6, 11].

The goal of this paper is to develop a unified framework for optimal control problems constrained by fractional in time semilinear PDEs

$$\begin{cases} \partial_t^\gamma u\,(x,t) + Au\,(x,t) = f\,(u\,(x,t)) + \mathbb{B}z\,(x,t)\,, & \text{in } \Omega \times (0,\infty), \\ u\,(\cdot,0) = u_0 \text{ in } \Omega. \end{cases} \tag{1.1}$$

where A is a given linear operator and f is a given nonlinear map which depends on the unknown PDE solution u. Some examples of f where our theory directly applies are the following.

- **Allen-Cahn (phase field) equation:** Here f is a cubic type of nonlinearity $f = -F'$ associated with the double-well potential
$$F(u) = c_1 u^4 - c_2 u^2, \quad \text{for } c_2 > c_1 > 0.$$
- **Subdiffusive Fisher-KPP:** Here f is a logistic term of the form $ru(1 - uK^{-1})$ with $r, K > 0$.
- **Subdiffusive Burger's equation:** Here $f(u) := -u\mathrm{div}(J * u)$, where
$$(J * u)(x) = \int_\Omega J(x-y)u(y)dy, \quad x \in \Omega.$$

We call u the state variable. The control z enters into the problem in a linear fashion, but with the help of operator $\mathbb{B}$, it can be in the interior, on the boundary, or in the exterior. Notice, that this framework not only allows A to be a standard local operator such as $-\mathrm{div}(K\nabla\cdot)$, but also a nonlocal operator such as a fractional Laplacian $(-\Delta)^s$ $(0 < s < 1)$. In addition, the boundary control not only can, be of Dirichlet, Neumann or Robin type, but also of Wentzell type [35]. Notice, that in the standard case of $\gamma = 1$, there are several existing results on boundary control of Dirichlet, Neumann or Robin problems, but there are no existing results on Wentzell type boundary control under the weaker conditions provided in this paper.

A key novelty of this paper is the presence of fractional in time derivative ∂_t^γ in (1.1) in the sense of Caputo (Definition 2.1). Recently, there has been a considerable interest in optimal control of fractional PDEs and ODEs, but most of the results are limited to linear problems or they consider very special scenarios [2, 10, 29]. Besides, low regularity requirements, the interest in fractional time derivative stems from its ability in capturing hereditary effects in materials and anomalous random walks [14, 30]. This hereditary property has also been recently used in designing new deep neural networks [4, 5], and gradient based algorithms [32].

The amount of literature on fractional in time derivative is growing due to many new emerging applications, however many of these results are empirical observations, and the analytical results are still very limited. It is well-known that the results from the classical models, i.e., $\gamma = 1$, do not directly extend to the fractional case $\gamma < 1$, see the monograph [26] and references therein. However, the results in this monograph do not apply to our problem as it focuses on the homogeneous case, i.e., $z = 0$ and does not consider any optimal control problems. Furthermore, problem (1.1) is *ill-posed* (even when $f \equiv 0$) in the sense that the solution flow is **not** strongly continuous (near $t = 0$) in the scale of Hilbert spaces associated with **any** fractional order of the (self-adjoint) operator A (see Remark 3.8). Another technical difficulty is the fact that the formulation of the necessary optimality conditions for the optimal control problem (3.8) requires a rigorous passage of the Caputo-derivative from the forward to the backward variables, a step which must be carefully analized (and, unfortunately, omitted quite often in the current literature on optimal control for subdiffusive parabolic problems). It turns out that the arguments, leading to the neccesary optimality conditions, need an additional condition of the behavior of the time derivative of the control variable near the origin. We emphasize that this is a technical condition which is in fact necessary, and which we believe, it cannot be discarded as it takes into account the solution behavior of (1.1) near the time $t = 0$. Besides, the control-to-state mapping for our problem appears to be (Frechet) differentiable **only** under this hypothesis.

The key novelties of this paper are the following.

- Well posedness of the fractional in time state equation under minimal (**generic**) assumptions on f.
- Well posedness of the linearized state (adjoint) equation.
- Existence of solution to the optimal control problem.
- Lipschitz continuity of the control to state map and its derivative.
- Rigorous derivation of the first order necessary conditions using the notion of **strong**[1] **Caputo derivative** (2.1).
- Applications to (subdiffusive) phase-transition phenomena, (subdiffusive) Fisher-KPP equations and subdiffusive Burger's equation, subject to nonlocal transport.

The remainder of the paper is organized as follows: In section 2 we introduce the basic definitions of fractional derivatives, properties of abstract operator A and collect various other tools that will be needed in the remainder of the paper. Section 3 focuses on the well-posedness of the forward semilinear problem (1.1) under minimal conditions on f and z. This is followed by section 4 where we state the optimal control problem. Here we assume that the final time T is finite and $T < T_{\max}$, for a maximal time $T_{\max} > 0$ where either $T_{\max} = \infty$ (global solution) or $T_{\max} < \infty$, and finite time blow-up may occur in some V_α-norm. Global well-posedness is briefly touched upon in Section 5 for V_1-solutions. With respect to the optimal control problem, we first establish the Lipschitz continuity of the control-to-state map and then existence of solution to the optimal control problem. Next, we show the differentiability of the control-to-state operator followed by Lipschitz continuity of the derivative of the control-to-state map. Subsequently, we establish well-posedness of the linearized state equation and rigorously derive the first order necessary conditions. We present several examples of cost functionals and control problems in section 6 where the abstract theory can be applied. Finally, we provide in Section Appendix 7 the proofs of most of teh technical results stated in Section 3.

2. **The basic functional framework.**

Let Y, Z be two Banach spaces endowed with norms $\|\cdot\|_Y$ and $\|\cdot\|_Z$, respectively. We denote by $Y \hookrightarrow Z$ if $Y \subseteq Z$ and there exists a constant $C > 0$ such that $\|u\|_Z \le C\|u\|_Y$, for $u \in Y \subseteq Z$. In particular, this means that the injection of Y into Z is continuous. In addition, if Y is dense

[1] It remains unclear how this derivation can be performed with a notion of generalized Caputo derivative (2.2), **without** the validity of Proposition 2.3.

in Z, then we denote by $Y \overset{d}{\hookrightarrow} Z$, and finally if the injection is also compact we shall denote it by $Y \overset{c}{\hookrightarrow} Z$. We denote by $\mathcal{L}(Y,Z)$ the space of all (bounded) linear operators from Y to Z. If $Y=Z$, we let $\mathcal{L}(Y,Z)=\mathcal{L}(Y)$. By the dual Y^* of Y, we think of Y^* as the set of all (continuous) linear functionals on Y. When equipped with the operator norm $\|\cdot\|_{Y^*}$, Y^* is also a Banach space. We use throughout the notation $h \lesssim g$ to denote $h \le Cg$, for some constant $C>0$ when the dependance of the constant $C=C(\gamma,s,q,p,...)$ on some physical parameters is not relevant, and so it is suppressed.

We give next the notion of fractional-in-time derivative in the sense of Caputo and Riemann-Liouville. Let $\gamma \in (0,1)$ and define

$$g_\gamma(t) := \begin{cases} \dfrac{t^{\gamma-1}}{\Gamma(\gamma)} & \text{if } t>0, \\ 0 & \text{if } t \le 0, \end{cases}$$

where Γ is the usual Gamma function. Let Y be a Banach space which possesses the Radon-Nikodym property, and let $T>0$.

Definition 2.1 (**Strong Caputo fractional derivative**). *Let $u \in W^{1,1}((0,T);Y)$. The (strong) Caputo fractional derivative of order $\gamma \in (0,1)$ is given by*

$$_C\partial_t^\gamma u(t) := \int_0^t g_{1-\gamma}(t-\tau)\, u^{'}(\tau)\, d\tau = (g_{1-\gamma} * u^{'})(t), \tag{2.1}$$

for all $t \in (0,T]$.

Definition 2.2 (**(Left) Riemann-Liouville fractional derivative**). *Let $u \in C([0,T];Y)$ be such that $g_{1-\gamma} * u \in W^{1,1}((0,T);Y)$. The (left) Riemann-Liouville fractional derivative of order $\gamma \in (0,1)$ is given by*

$$D_t^\gamma u(t) := \frac{d}{dt}\Big(g_{1-\gamma} * u\Big)(t) = \frac{d}{dt}\int_0^t g_{1-\gamma}(t-\tau)\, u(\tau)\, d\tau,$$

for almost all $t \in (0,T)$.

We next set

$$\partial_t^\gamma u := D_t^\gamma (u - u(0)), \tag{2.2}$$

and recall that the right-hand side of (2.2) is usually dubbed in the literature as a **generalized Caputo derivative** (see, e.g., [26]). We observe that the notions of fractional derivatives in (2.2) and (2.1), respectively, are in fact equivalent under the Radon-Nikodym property.

Proposition 2.3. *Let the assumptions of Definition 2.1 be satisfied. Then*

$$_C\partial_t^\gamma u(t) = \partial_t(g_{1-\gamma} * (u - u(0)))(t) = D_t^\gamma (u - u(0))(t),$$

for almost all $t \in (0,T]$, and $g_{1-\gamma} * (u - u(0)) \in W^{1,1}((0,T);Y)$.

Proof. We note that $\|u(0,\cdot)\|_Y < \infty$ since each $u \in W^{1,1}((0,T);Y)$ is also continuous on $[0,T]$ with values in Y. As a consequence of the Radon-Nikodym theorem (see [21]), each function that belongs to $W^{1,1}([0,T];Y)$ is also absolutely continuous on $[0,T]$ (modulo a null set of Lebesgue measure) with values in Y. Then the conclusion follows from a standard result in [22, Theorem 3.1] in the case $Y=\mathbb{R}$ (the proof in the case of a general Banach space Y follows with some, albeit, obvious modifications). ∎

Definition 2.4 (**(Right) Riemann-Liouville fractional derivative**). *The (right) Riemann-Liouville fractional derivative of order $\gamma \in (0,1)$ is defined by*

$$\partial_{t,T}^{\gamma} u(t) = -\frac{d}{dt}(I_{t,T}^{1-\gamma} u)(t), \tag{2.3}$$

where

$$I_{t,T}^{\gamma} u\,(t) := \frac{1}{\Gamma(\gamma)} \int_t^T (\tau - t)^{\gamma-1} u(\tau) d\tau$$

is the right Riemann-Liouville fractional integral of order γ. The left Riemann-Liouville fractional integral of order $\gamma \in (0,1)$ is given by

$$I_{0,t}^{\gamma} u(t) = \frac{1}{\Gamma(\gamma)} \int_0^t (t - \tau)^{\gamma-1} u(\tau) d\tau.$$

From (2.3), we observe that if u is differentiable, then $\partial_{t,T}^{1} u = -\partial_t u$. We will employ the following well-known result (see, e.g., [1]) to determine the corresponding adjoint problem associated with the initial boundary value problem, that we have set up in Section 4.

Proposition 2.5. *Under the assumptions of Proposition 2.3, the following* integration by parts formula *holds:*

$$\int_0^T v\,(t)_C\, \partial_t^{\gamma} u(t) dt = \int_0^T \partial_{t,T}^{\gamma} v(t) u(t) dt + \left[(I_{t,T}^{1-\gamma} v)\,(t)\, u(t) \right]_{t=0}^{t=T}, \tag{2.4}$$

provided that the left and right-hand sides expressions make sense (i.e., the products inside the integrals are well-defined, at least in a duality sense $\langle \cdot, \cdot \rangle_{Y,Y^*}$). Note that if $u\,(0) = (I_{t,T}^{1-\gamma} v)\,(T) = 0$, the bracket term in (2.4) vanishes.

Remark 2.6. The integral expressions in (2.4) are well defined (cf. [31, pg. 76], owing to $\partial_{t,T}^{\gamma} v = I_{t,T}^{1-\gamma}(\partial_t v)$ and ${}_C\partial_t^{\gamma} u = I_{0,t}^{1-\gamma}(\partial_t u)$), for instance, when $u \in W^{1,p}((0,T);Y^*)$ and $v \in W^{1,q}((0,T);Y)$, with $p, q \geq 1$ and $p^{-1} + q^{-1} \leq 2 - \gamma$ ($p \neq 1$ and $q \neq 1$ when $p^{-1} + q^{-1} = 2 - \gamma$).

We recall next the following Gronwall type inequality from [24, Lemma 6.3].

Lemma 2.7. *Let the function $\varphi(t) \geq 0$ be continuous for $0 < t \leq T$. If*

$$\varphi\,(t) \leq C_1 t^{\alpha-1} + C_2 \int_0^t (t-s)^{\beta-1} \varphi\,(s)\, ds,\ 0 < t \leq T,$$

for some constants $C_1, C_2 \geq 0$ and $\alpha, \beta > 0$, then there is a positive constant $C_* = C_*(\alpha, \beta, T, C_2)$ such that[2]

$$\varphi\,(t) \leq C_* C_1 t^{\alpha-1},\ 0 < t \leq T. \tag{2.5}$$

We shall now assume (Ω, g) is a (n-dimensional) Riemannian manifold and g a complete Riemannian metric on Ω, which is at least of Lipschitz class (i.e., $\|g\|_{\mathrm{Lip}} < \infty$).

(HA) In that case, let A be a strictly positive[3] self-adjoint (unbounded) operator in $L^2(\Omega)$ (i.e., $0 \in \rho\,(A)$) whose resolvent $(I + A)^{-1}$ is compact[4] in $L^2(\Omega)$.

[2] A crucial part of this estimate that we shall exploit repeteadly is that the final exponent in (2.5) is independent of β.

[3] The strict positivity of the operator is generally not required, and one can assume instead $A \geq 0$. Indeed, one can replace A by $I + A$ by adding the identity on both sides of equation (3.1). The assumptions on f hold for the modified nonlinearity as well.

[4] $L^2(\Omega)$ is locally compact since Ω is complete.

By the spectral theory, A has its eigenvalues forming a non-decreasing sequence of real numbers $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq \cdots$ satisfying $\lim_{n\to\infty} \lambda_n = \infty$. In addition, the eigenvalues are of finite multiplicity. Let $(\varphi_n)_{n\in\mathbb{N}}$ be the orthonormal basis of eigenfunctions associated with $(\lambda_n)_{n\in\mathbb{N}}$. Then $\varphi_n \in D(A)$ for every $n \in \mathbb{N}$, $(\varphi_n)_{n\in\mathbb{N}}$ satisfies $A\varphi_n = \lambda_n\varphi_n$. We shall denote by $D(A)^\star$ the dual of $D(A)$ with respect to the pivot space $L^2(\Omega)$ so that we have the continuous embedding $D(A) \overset{c}{\hookrightarrow} L^2(\Omega) \overset{c}{\hookrightarrow} D(A)^\star$. Spectral theory also allows us to define the spaces $V_\alpha = D(A^{\alpha/2})$ with norms

$$|u|_\alpha := ||A^{\alpha/2}u||_{L^2(\Omega)}, \text{ for any real } \alpha. \tag{2.6}$$

A simple argument shows that $D(A^{-\alpha/2}) = V_{-\alpha}$, for $\alpha \geq 0$. Indeed, by the standard spectral theory, fractional powers of A can be defined by means of

$$\begin{cases} A^{\alpha/2}u &= \displaystyle\sum_{n=1}^{\infty} \lambda_n^{\alpha/2} \left(u, \varphi_n\right)_{L^2(\Omega)} \varphi_n, \\ D(A^{\alpha/2}) &= \left\{ u \in L^2\left(\Omega\right) : \left\|A^{\alpha/2}u\right\|^2_{L^2(\Omega)} = \displaystyle\sum_{n=1}^{\infty} \lambda_n^{\alpha} \left|\left(u, \varphi_n\right)_{L^2(\Omega)}\right|^2 < \infty \right\}. \end{cases} \tag{2.7}$$

In particular, there holds $V_1 = D(A^{1/2})$ and $V_{-1} = V_1^*$ such that

$$|u|_1 \simeq \left(A^{1/2}u, A^{1/2}u\right)_{L^2(\Omega)}$$

and $|u|_{-1} \simeq \left(A^{-1}u, u\right)_{L^2}$ (in the sense of equivalent norms), respectively. A complete characterization of the spaces V_α and their embedding properties into $L^p\left(\Omega\right)$-spaces ($p \geq 1$) can be found in [3] (see also [25]).

Next, we recall the definition of the Wright type (also sometimes called the Mainardi) function (see [26] and the references therein),

$$\Phi_\gamma(z) := \sum_{n=0}^{\infty} \frac{(-z)^n}{n!\Gamma(-\gamma n + 1 - \gamma)}, \quad 0 < \gamma < 1, \quad z \in \mathbb{C}. \tag{2.8}$$

It is well known that $\Phi_\gamma(t)$ is a probability density function, namely,

$$\Phi_\gamma(t) \geq 0, \quad t > 0; \quad \int_0^\infty \Phi_\gamma(t)dt = 1.$$

Furthermore, $\Phi_\gamma(0) = 1/\Gamma\left(1-\gamma\right)$ and

$$\int_0^\infty t^p \Phi_\gamma(t)dt = \frac{\Gamma(p+1)}{\Gamma(\gamma p + 1)}, \quad p > -1, \quad 0 < \gamma < 1. \tag{2.9}$$

We let $\left(T\left(t\right)\right)_{t\geq 0}$ denote the analytic semigroup on $L^2(\Omega)$ generated by the operator $-A$, and consider an extension of T (which we still denote by T, for the simplicity of notation) on all scales of negative fractional order spaces V_α, $\alpha < 0$. Next, we define two additional operators

$$S_\gamma(t) : V_\alpha \to V_\alpha, \; P_\gamma(t) : V_\alpha \to V_\alpha, \; \alpha \in \left[-2, 2\right],$$

by

$$\begin{cases} S_\gamma(t)v := \displaystyle\int_0^\infty \Phi_\gamma(\tau)T(\tau t^\gamma)v d\tau, \\ P_\gamma(t)v := \gamma t^{\gamma-1} \displaystyle\int_0^\infty \tau\Phi_\gamma(\tau)T(\tau t^\gamma)v d\tau. \end{cases} \tag{2.10}$$

We next recall the definition of the Mittag-Leffler function

$$E_{\alpha,\beta}(z) := \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n+\beta)}, \ \ \alpha>0,\ \beta\in\mathbb{C}, \quad z\in\mathbb{C}.$$

A particular estimate for $E_{\alpha,\beta}$ that we shall often use in the paper (and whenever necessary) is

$$|E_{\alpha,\beta}(x)| \le C_{\alpha,\beta}(1-x)^{-1}, \quad \text{for } x\le 0 \text{ and } \alpha\in(0,1),\ \beta>0. \tag{2.11}$$

It is also well-known that $E_{\alpha,\beta}(z)$ is an entire function, and that both operators in (2.10) can be also cast in terms of these functions (see, for instance, [18, Theorem 4.2]). In particular, by the spectral theory we further have

$$\begin{cases} S_\gamma(t)v = \displaystyle\sum_{n=0}^{\infty} (v,\varphi_n)_{L^2(\Omega)} E_{\gamma,1}(-\lambda_n t^\gamma)\varphi_n, \\ \\ P_\gamma(t)v = \displaystyle\sum_{n=0}^{\infty} (v,\varphi_n)_{L^2(\Omega)} t^{\gamma-1}E_{\gamma,\gamma}(-\lambda_n t^\gamma)\varphi_n. \end{cases} \tag{2.12}$$

Proposition 2.8. *The operator families $\{S_\gamma(t)\}$, $\{P_\gamma(t)\}$ are analytic for $t>0$, and satisfy the following estimates:*

$$|S_\gamma(t)v|_\beta \le C_{\alpha,\beta,\gamma} t^{-\gamma\left(\frac{\beta-\alpha}{2}\right)} |v|_\alpha, \quad -2\le\alpha\le\beta\le 2, \tag{2.13}$$

with $\beta-\alpha\in[0,2]$, and

$$|P_\gamma(t)v|_{\widetilde{\beta}} \le C_{\widetilde{\alpha},\widetilde{\beta},\gamma} t^{\gamma-1-\gamma\left(\frac{\widetilde{\beta}-\widetilde{\alpha}}{2}\right)} |v|_{\widetilde{\alpha}}, \quad -2\le\widetilde{\alpha}\le\widetilde{\beta}\le 2. \tag{2.14}$$

The positive constants $C_{\alpha,\beta,\gamma}, C_{\widetilde{\alpha},\widetilde{\beta},\gamma}$ are independent of t and v, and are bounded as $\gamma\to 1^-$. Finally, the operator S_γ is also a contraction (strongly continuous) mapping from $V_\alpha \to V_\alpha$.

Proof. The analyticity of the semigroup $T(t)=\exp(-tA)$ together with the representation (2.10) implies the analyticity of $S_\gamma(t)$ and $P_\gamma(t)$ for $t>0$. The analyticity of $T(t)$ is also reflected in the inequality

$$|T(t)v|_\beta \le C_{\alpha,\beta} t^{-(\beta-\alpha)/2} |v|_\alpha, \ \text{with } -2\le\alpha\le\beta\le 2. \tag{2.15}$$

Combining (2.15) with the norm $|v|_\alpha$ via (2.6)-(2.7) for all real α, and exploiting the identities (2.9)-(2.10) (or, respectively (2.11)-(2.12)), we easily obtain the estimates (2.13)-(2.14). ☐

3. Well-posedness results for the forward problem.

In the above framework, we can conveniently rewrite the semilinear problem as follows:

$$\begin{cases} \partial_t^\gamma u(x,t) + Au(x,t) = f(u(x,t)) + \mathbb{B}z(x,t), \quad \text{in } \Omega\times(0,\infty), \\ u(\cdot,0) = u_0 \text{ in } \Omega. \end{cases} \tag{3.1}$$

Our main goal in this section is to state sufficiently general conditions on the data (f,z,u_0) for which we can infer the existence of properly-defined solutions[5] for (3.1). Let $T\in(0,\infty)$ and denote by J a time interval of the form $[0,T]$, $[0,T)$ or $[0,\infty)$.

Definition 3.1. *By a mild solution of (3.1) on the interval J, we mean that the measurable function u has the following properties:*

[5] We will also investigate in which sense (3.1) is satisfied by non-regular solutions.

(a) $u \in C(J; V_\alpha)$, *for some* $\alpha \in \mathbb{R}$.

(b) $f(u(\cdot,t)) \in V_\beta$, *for all* $t \in J$, *for some* $\alpha \geq \beta$, $\beta \in \mathbb{R}$.

(c) $u(\cdot,t) = S_\gamma(t)\, u_0 + \int_0^t P_\gamma(t-\tau)\, f(u(\cdot,\tau))\, d\tau + \int_0^t P_\gamma(t-\tau)\, \mathbb{B} z(\cdot,\tau)\, d\tau$, *for all* $t \in J \backslash \{0\}$, *where the integral is an absolutely converging Bochner integral in the space* V_α.

(d) The initial datum u_0 *is assumed in the following sense:*

$$\lim_{t \downarrow 0^+} |u(\cdot,t) - u_0|_\alpha = 0, \tag{3.2}$$

for $u_0 \in V_\alpha$.

Our first goal is to establish the existence of maximally-defined mild solutions under some suitable assumptions on the parameters of the problem. These assumptions are as follows.

(H1) Let $\widetilde{\alpha} \in [-1,0]$ be such that $\mathbb{B} \in \mathcal{L}\left(L^2(D); V_{\widetilde{\alpha}}\right)$, for an arbitrary[6] (Hausdorff, at least) space D. Set $I_\beta := [\beta, \beta+2)$, for $\beta \in \mathbb{R}$ and assume that $u_0 \in V_\alpha$, for some $\alpha \in I_{\widetilde{\alpha}} \cap I_\beta \neq \varnothing$.

(H2) The 'control' function $z \in L^q_{\text{loc}}(\mathbb{R}_+; L^2(D))$, for some $q \in \left(\frac{2}{\gamma(2-\alpha+\widetilde{\alpha})}, \infty\right]$.

(H3) The nonlinear function $f \in C^{0,1}_{\text{loc}}(\mathbb{R})$, $f(0) = 0$, induces a locally Lipschitzian map

$$f : V_\alpha \to V_\beta;$$

namely, for every $R > 0$, with $|u|_\alpha, |v|_\alpha \leq R$, there exists $C_R > 0$ such that

$$|f(u) - f(v)|_\beta \leq C_R\, |u - v|_\alpha, \text{ for all } u, v \in V_\alpha \hookrightarrow V_\beta.$$

Our first result is concerned with the existence and uniqueness of locally-defined mild solutions (see Appendix 7 for the proof).

Lemma 3.2. *(**Local existence**). Assume (HA) and (H1)-(H3). Then there exists a time* $T_* > 0$ *(depending on* u_0*) such that the problem (3.1) possesses a unique mild solution in the sense of Definition 3.1 on the interval* $J = [0, T_*]$.

Our second statement shows that locally-defined mild solutions in V_α can be (uniquely) extended to a larger interval (see Appendix 7 for the proof).

Lemma 3.3. *(**Unique continuation**) Let the assumptions of Lemma 3.2 be satisfied. Then the unique integral solution on* $J = [0, T^\star]$ *of (3.1) can be extended to the interval* $[0, T^\star + \tau]$, *for some* $\tau > 0$, *so that, the extended function is the unique mild solution of (3.1) on* $[0, T^\star + \tau]$ *in the sense of Definition 3.1.*

The following statement is then straightforward on account of the above lemmas.

Theorem 3.4. *Assume (HA) and (H1)-(H3). Problem (3.1) has a unique mild solution on* $J = [0, T_{\max})$ *in the sense of Definition 3.1, where either* $T_{\max} = \infty$ *or* $T_{\max} < \infty$, *and in that case,*

$$\limsup_{t \to T_{\max}^-} |u(t)|_\alpha = \infty.$$

Proof. The proof is standard owing to Lemma 3.2 and Lemma 3.3, respectively, and a contradiction argument (see, e.g., [3, 18]). $\square$

[6] The control region D depends upon the choice of A, the geometry of Ω and the control operator $\mathbb{B}$, as some examples will show in Section 6.

We derive a sufficient condition in order to conclude additional (temporal) regularity for the mild solution (under the same assumptions of Theorem 3.4; see Appendix 7, for a proof).

Theorem 3.5. *Let $u : [0,T] \to V_\alpha$ be a mild solution in the sense of Theorem 3.4 for any $T < T_{\max}$. In addition, assume that $u_0 \in V_{\beta+2} \hookrightarrow V_\alpha$ and $z \in W^{1,1}((0,T); L^2(D))$, provided that*

$$\|\partial_t z(t)\|_{L^2(D)} \leq C t^{\rho-1}, \quad \text{for all } 0 < t \leq T, \tag{3.3}$$

for some $\rho > 0$, and $C > 0$ independent of t and z. Then, for the above mild solution, we have

$$u \in W^{1,1+\xi}((0,T); V_\alpha), \text{ for some } \xi > 0. \tag{3.4}$$

Furthermore, the identity ${}_C\partial_t^\gamma u(t) = \partial_t^\gamma u(t)$ is satisfied for almost all $t \in (0,T)$.

In what follows it is more convenient to set $\widetilde{\alpha} = \beta \in [-1,0]$ and recall that $\alpha \in [\beta, \beta+2)$ (so that assumption (H1) is satisfied). The previous theorems imply the following two (major) statements which conclude the section.

Theorem 3.6. ***(The problem for the generalized Caputo derivative)*** *If (HA), (H1)-(H3) hold, then problem (3.1) has a unique mild solution*

$$u \in C([0,T_{\max}); V_\alpha),$$

for which

$$\partial_t^\gamma u := \partial_t (g_{1-\gamma} * (u-u_0)) \in C([0,T_{\max}); V_{\widetilde{\alpha}-\delta}), \text{ for any } \delta \in \left(\frac{2}{\gamma q}, 2\right]. \tag{3.5}$$

Moreover, the variational identity

$$\langle \partial_t^\gamma u(t) + Au(t), v\rangle_{V_{\widetilde{\alpha}-\delta}, V_{-\widetilde{\alpha}+\delta}} = \langle f(u(t)) + \mathbb{B}z(t), v\rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}}, \tag{3.6}$$

holds for any $v \in V_{-\widetilde{\alpha}+\delta} \subset V_{-\widetilde{\alpha}}$, and for almost all $t \in (0, T_{\max})$.

Proof. The statement (3.5) is a consequence of the proof of Theorem 3.4, since $f(u) \in C([0,T_{\max}); V_{\widetilde{\alpha}}) \subset L^q(0,T_{\max}; V_{\widetilde{\alpha}})$,

$$\begin{aligned}&\left[|A(P_\gamma * \mathbb{B}z)(t)|_{\widetilde{\alpha}-\delta} + |A(P_\gamma * f(u))(t)|_{\widetilde{\alpha}-\delta}\right] \\ &\lesssim t^{\frac{\gamma\delta}{2}-\frac{1}{q}} \left(\|\mathbb{B}z\|_{L^q(0,T;V_{\widetilde{\alpha}})} + \|f(u)\|_{L^q(0,T;V_{\widetilde{\alpha}})}\right),\end{aligned}$$

and

$$|AS_\gamma(t)u_0|_{\widetilde{\alpha}-\delta} \leq |AS_\gamma(t)u_0|_{\widetilde{\alpha}} \lesssim t^{\frac{\gamma}{2}(2+\widetilde{\alpha}-\alpha)} |u_0|_\alpha,$$

for all $T_{\max} > t > 0$. The identity (3.6) then follows from the solution representation in Definition 3.1 and from (3.5). ☐

Corollary 3.7. ***(The problem for the strong Caputo derivative)*** *Let the assumptions of Theorem 3.6 be satisfied, and in addition, assume $u_0 \in V_{\beta+2} \subset V_\alpha$ and (3.3) hold. Then the mild solution of problem (3.1) satisfies $|\partial_t u(t)|_\alpha \lesssim t^{\theta-1}$, and therefore,*

$$u \in W^{1,1+\xi}((0,T); V_\alpha) \cap L^\sigma((0,T); V_{2+\widetilde{\alpha}}), \ {}_C\partial_t^\gamma u \in L^{1+\xi}((0,T); V_{\widetilde{\alpha}}), \tag{3.7}$$

for $\xi = \xi(\theta) \in (0, \frac{\theta}{1-\theta})$, $\sigma := \min(1+\xi, q) > 1$. Moreover, the variational identity

$$\langle \partial_t^\gamma u(t) + Au(t), v\rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}} = \langle f(u(t)) + \mathbb{B}z(t), v\rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}}, \tag{3.8}$$

holds for any $v \in V_{-\widetilde{\alpha}}$, and for almost all $t \in (0, T_{\max})$.

Proof. The proof is a consequence of Theorem 3.5, in view of the assumptions of Theorem 3.6. Indeed, owing to the fact that $\mathbb{B}z \in L^q((0,T);V_{\widetilde{\alpha}})$, $f(u) \in C([0,T];V_{\widetilde{\alpha}})$ and ${}_C\partial_t^\gamma u = \partial_t^\gamma u \in L^{1+\xi}((0,T);V_{\widetilde{\alpha}})$, one may argue, by comparison in (3.6), that $Au \in L^\sigma((0,T);V_{\widetilde{\alpha}})$, which implies that $u \in L^\sigma((0,T);V_{2+\widetilde{\alpha}})$. □

Remark 3.8. We observe the following facts.

(a) When $\widetilde{\alpha} = \beta \in [-1,0]$ and $\alpha \in [\beta, \beta+2)$, we have in Corollary 3.7, $\theta = \frac{\gamma(2-\alpha+\beta)}{2} - \frac{1}{q} > 0$ whenever $q \in (\frac{2}{\gamma(2-\alpha+\beta)}, \infty]$. Moreover, $\sigma = \min(1+\xi, q) = 1+\xi \in (1, \eta(\theta))$ since $\eta(\theta) := \frac{1}{1-\theta} \in (1, q]$.

(b) When $\widetilde{\alpha} \neq \beta$, the explicit value of $\theta > 0$, in terms of α, β, γ and $\widetilde{\alpha}$, can be found in (7.14), namely[7],

$$\theta = \min\left\{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha}) - \frac{1}{q}, \frac{\gamma}{2}(2-\alpha+\beta)\right\}.$$

In this case, the analogue of Corollary 3.7 is

$$\partial_t^\gamma u + Au = f(u) + \mathbb{B}z, \text{ holds in } V_{\min\{\widetilde{\alpha},\beta\}}, \text{ for almost all } t \in (0,T), \tag{3.9}$$

and each solution of (3.9) belongs to

$$W^{1,1+\xi}((0,T);V_\alpha) \cap L^\sigma\left((0,T);V_{2+\min\{\widetilde{\alpha},\beta\}}\right), {}_C\partial_t^\gamma u \in L^{1+\xi}\left((0,T);V_{\min\{\widetilde{\alpha},\beta\}}\right).$$

For additional regularity theory of problem (3.1), the restriction $\widetilde{\alpha} \in [-1,0]$ does not appear to be a necessary condition. More precisely, in the statement of Corollary 3.7, one can assume instead that $\widetilde{\alpha} \in \mathbb{R}_+$, for as long as $\alpha \in I_{\widetilde{\alpha}} \cap I_\beta \neq \varnothing$ (see, e.g., Theorem 5.2 in Section 5). However, the restriction that $\widetilde{\alpha} > 0$ appears necessary for the rigourous justification of the V_1-energy equality for the corresponding subdiffusive problem.

(c) The gap in regularity between the initial datum $u_0 \in V_{\beta+2} \subsetneq V_\alpha$ and the solution $u \in C([0,T];V_\alpha)$ (but[8] with $u \notin C([0,T];V_{\beta+2})$) is not of technical nature. It is due to a complete **failure** of the semigroup property for the solution operator associated with (3.1) (see, e.g., [26, 23], and the references therein). In particular, this means that the fractional in time problem is not well-posed in the classical sense formulated by Hadamard; namely, there **does not exist a (strongly) continuous** flow map $\Phi : u_0 \mapsto u(t)$ in any scale of the operator spaces V_α. This is in contrast to the strong continuity of the flow map for the classical problem when $\gamma = 1$.

4. The optimal control problem.

In view of Corollary 3.7, we set the control space to be

$$Z_{\rho,\infty} := \left\{z \in C([0,T];L^2(D)) : \|\partial_t z(t)\|_{L^2(D)} \lesssim t^{\rho-1}, \text{ a.e. } 0 < t \leq T\right\},$$

for some $T < T_{\max}$ (with a value T which we will fix from now on) and $0 < \rho \leq 1$. Notice[9] that $Z_{\rho,\infty}$ is a closed (bounded) subset of

$$W^{1,1+\lambda}((0,T);L^2(D)) \subset C([0,T];L^2(D)),$$

[7] *The value of θ is independent of $\rho > 0$.*

[8] *Strictly speaking, this is a consequence of the singular behavior of $P_\gamma(t)$ near $t = 0$.*

[9] Due to the embedding, we may immediately take $q = \infty$ in Remark 3.8. It follows that $\sigma = 1 + \xi(\theta) > 1$ and $\theta = \frac{\gamma(2-\alpha+\beta)}{2}$.

for some $\lambda > 0$ depending on ρ (i.e, $(1+\lambda)(\rho-1) > -1$), when we endow it with the norm

$$\|z\|_{Z_{\rho,\infty}} := \|z\|_{C([0,T];L^2(D))} + \sup_{t\in[0,T]} t^{1-\rho} \|\partial_t z(t)\|_{L^2(D)} . \tag{4.1}$$

Thus, the control-to-state operator

$$\mathcal{S} : Z_{\rho,\infty} \to Y_{\theta,\alpha},\ z \longmapsto u =: \mathcal{S}(z)$$

is well-defined as a mapping from $Z_{\rho,\infty}$ into the Banach space

$$Y_{\theta,\alpha} = \left\{ u \in W^{1,1+\xi(\theta)}((0,T);V_\alpha) : |\partial_t u(t)|_\alpha \lesssim t^{\theta-1},\ \text{a.e. } 0 < t \le T \right\}; \tag{4.2}$$

$Y_{\theta,\alpha}$ is endowed with the natural norm (for some $\xi = \xi(\theta) > 0$, depending[10] on $\theta > 0$)

$$\|u\|_{Y_{\theta,\alpha}} := \sup_{t\in[0,T]} |u(t)|_\alpha + \sup_{t\in[0,T]} t^{1-\theta} |\partial_t u(t)|_\alpha .$$

We notice first that $\mathcal{S}$ is Lipschitz continuous from $Z_{\rho,\infty}$ into $C([0,T];V_\alpha)$ (see (4.5)). Secondly, if we impose additional assumptions on f, the state mapping $\mathcal{S}$ satisfies an improved stability estimate (4.6). To this end, let us denote by $\mathcal{U}$ a nonempty, open and bounded subset of $Z_{\rho,\infty}$.

(H4) The nonlinearity $f \in C^{1,1}$ induces a bounded[11] mapping

$$g_u(v) := (\partial_u f(u))\, v : V_\alpha \to V_\beta, \tag{4.3}$$

such that, for every $u_1, u_2 \in V_\alpha$ satisfying $|u_i|_\alpha \le R$ $(i = 1,2)$, there is a constant $C_R > 0$ such that,

$$|g_{u_1}(v) - g_{u_2}(v)|_\beta \le C_R |v|_\alpha |u_1 - u_2|_\alpha . \tag{4.4}$$

Theorem 4.1. *Let the assumptions of Corollary 3.7 be satisfied.*

(i) Then for each $T < T_{\max}$, there exists a constant $K_1 > 0$, depending on T, R, f, such that whenver $z_1, z_2 \in \mathcal{U}$ are given and $u_1, u_2 \in Y_{\theta,\alpha}$ denote the associated solutions of the state system, we have

$$\|u_1 - u_2\|_{C([0,T];V_\alpha)} \le K_1 \|z_1 - z_2\|_{C([0,T];L^2(D))} . \tag{4.5}$$

(ii) If in addition, (H4) holds, then there is a constant $K_2 > 0$ such that

$$\sup_{t\in[0,T]} t^{1-\theta} |(\partial_t u_1 - \partial_t u_2)(t)|_\alpha \le K_2 \|z_1 - z_2\|_{Z_{\rho,\infty}} . \tag{4.6}$$

Proof. Set $u := u_1 - u_2$ and $z := z_1 - z_2$. Then, every mild/weak solution u satisfies on $(0,T) \subset (0,T_{\max})$,

$$u(t) = \int_0^t P_\gamma(t-\tau)\left(f(u_1(\tau)) - f(u_2(\tau))\right) d\tau + \int_0^t P_\gamma(t-\tau)\, \mathbb{B} z(\tau)\, d\tau \tag{4.7}$$

since $u(0) = 0$, for $i = 1,2$, whenever $u_i(0) = u_0$. By Theorem 3.5,

$$\begin{aligned} \partial_t u_i(t) = {} & A P_\gamma(t) u_0 + P_\gamma(t) f(u_0) + P_\gamma(t)\, \mathbb{B} z_i(0) \\ & + \int_0^t P_\gamma(t-\tau)\, \mathbb{B} \partial_t z_i(\tau)\, d\tau + \int_0^t P_\gamma(t-\tau)\, \partial_{u_i} f(u_i(\tau))\, \partial_t u_i(\tau)\, d\tau, \end{aligned} \tag{4.8}$$

[10] Namely, $\xi > 0$ is such that $(1+\xi)(\theta-1) > -1$.

[11] Note that boundedness is a consequence of (4.4).

for almost all $t \in (0,T) \subset (0,T_{\max})$. Then (4.5) follows easily employing once again an extension argument for (4.7) to the whole interval $(0,T)$, via the proofs of Lemma 3.2 and Lemma 3.3. Let us now set $v := \partial_t u_1 - \partial_t u_2$, and notice that

$$\begin{aligned} v(t) = &\int_0^t P_\gamma(t-\tau)\left(g_{u_1}(\partial_t u_1(\tau)) - g_{u_2}(\partial_t u_1(\tau))\right) d\tau + P_\gamma(t)\,\mathbb{B}z(0) \\ &+ \int_0^t P_\gamma(t-\tau)\,\partial_{u_2} f(u_2(\tau))\,v(\tau)\,d\tau + \int_0^t P_\gamma(t-\tau)\,\mathbb{B}\partial_t z(\tau)\,d\tau. \end{aligned} \tag{4.9}$$

The argument in (7.13), given $\widetilde{\alpha} = \beta$ and $\alpha \in [\beta, \beta+2)$, easily yields

$$\begin{aligned} \left| \int_0^t P_\gamma(t-\tau)\,\mathbb{B}\partial_t z(\tau)\,d\tau \right|_\alpha &\lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}+\rho-1} \sup_{t\in[0,T]} t^{1-\rho} \left\|\partial_t z(t)\right\|_{L^2(D)} \\ &\lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}+\rho-1} \left\|z\right\|_{Z_{\rho,\infty}} \\ &\lesssim t^{\theta-1} T^\rho \left\|z\right\|_{Z_{\rho,\infty}} \\ &\lesssim t^{\theta-1} \left\|z\right\|_{Z_{\rho,\infty}} \end{aligned}$$

while, in light of the boundedness of g_u (by (4.4), $|g_u(v)|_\beta \lesssim |v|_\alpha (|u|_\alpha + 1)$, since $0 \in V_\alpha$), it follows that

$$\left| \int_0^t P_\gamma(t-\tau)\,\partial_{u_2} f(u_2(\tau))\,v(\tau)\,d\tau \right|_\alpha \lesssim \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1} |v(\tau)|_\alpha\,d\tau.$$

Moreover, in view of (4.4) and (4.5), we can deduce that

$$\begin{aligned} &\left| \int_0^t P_\gamma(t-\tau)\left(\partial_{u_1} f(u_1(\tau)) - \partial_{u_2} f(u_2(\tau))\right) \partial_t u_1(\tau)\,d\tau \right|_\alpha \\ &\lesssim \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1} |\partial_t u_1(\tau)|_\alpha\,d\tau \left\|u\right\|_{C([0,T];V_\alpha)} \\ &\lesssim \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1} \tau^{\theta-1} d\tau \left\|u_1\right\|_{Y_{\theta,\alpha}} \left\|u\right\|_{C([0,T];V_\alpha)} \\ &\lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}+\theta-1} \left\|z\right\|_{C([0,T];L^2(D))} \\ &\lesssim t^{\theta-1} \left\|z\right\|_{C([0,T];L^2(D))}. \end{aligned}$$

Finally, we have

$$|P_\gamma(t)\,\mathbb{B}z(0)|_\alpha \lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}-1} |\mathbb{B}z(0)|_{\beta=\widetilde{\alpha}} \lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}-1} \left\|z(0)\right\|_{L^2(D)} \lesssim t^{\theta-1} \left\|z\right\|_{Z_{\rho,\infty}}.$$

Collecting the previous estimates, from (4.9) we find that

$$|v(t)|_\alpha \lesssim t^{\theta-1} \left\|z\right\|_{Z_{\rho,\infty}} + \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1} |v(\tau)|_\alpha\,d\tau. \tag{4.10}$$

Application of the Gronwall inequality (see Lemma 2.7) then yields

$$|v(t)|_\alpha \lesssim t^{\theta-1} \left\|z\right\|_{Z_{\rho,\infty}},$$

from which we can immediately infer (4.6). The proof is complete. □

Remark 4.2. Part (ii) of Theorem 4.1 implies that the solution operator $\mathcal{S}$ is (Lipschitz) continuous when viewed as a mapping from $Z_{\rho,\infty}$ into $Y_{\theta,\alpha}$ (with $\theta = \gamma(2-\alpha+\beta)/2 \in (0,1)$). Indeed, (4.6) yields

$$\left\|u_1 - u_2\right\|_{Y_{\theta,\alpha}} \le K_3 \left\|z_1 - z_2\right\|_{Z_{\rho,\infty}}, \tag{4.11}$$

for some $K_3 > 0$, depending only on K_1, K_2.

Now we define the cost functional, i.e., $J(u,z) := J_1(u) + J_2(z)$. Consider D_1 and D_2 to be the effective domains of the (proper) functionals J_1 and J_2, respectively. We let $J_1 : X_1 \to (-\infty, +\infty]$ and $J_2 : X_2 := Z_{\rho,\infty} \to (-\infty, +\infty]$ (with the convention that $J_i(u) = +\infty$, for $u \in X_i \backslash D_i$, $i = 1, 2$), and as a result we can write the reduced minimization problem

$$\min_{z \in Z_{ad}} \mathcal{J}(z) := J_1(\mathcal{S}(z)) + J_2(z) = J(\mathcal{S}(z), z). \tag{4.12}$$

First, we assume there is an admissible set Z_{ad} which is a convex and closed subset of $Z_{\rho,\infty}$. We impose the following (specific) assumptions on J_1 and J_2.

(H5) $J_1 : D_1 := Y_{\rho,\widetilde{\alpha}} \to \mathbb{R}$ is weakly lower semicontinuous (for some $\rho > 0$); $J_2 : D_2 := Z_{ad} \to \mathbb{R}$ is convex, lower-semicontinuous and the level set $\{z \in Z_{ad} : J_2(z) \le \kappa\}$ is bounded for some $\kappa \in \mathbb{R}$.

Theorem 4.3 (**Existence of optimal controls**). *Let the assumptions of Corollary 3.7 hold. Assume in addition that (H5) holds. Then the optimal control problem (4.12) admits a solution.*

Proof. We begin by noticing an infimizing sequence $\{z_n\}_{n\in\mathbb{N}}$ always exists (see [13, pg. 84]). Let $\{z_n\}_{n\in\mathbb{N}}$ be a infimizing sequence, that is, $z_n \in Z_{ad}$ and $u_n = \mathcal{S}(z_n)$, for $n \in \mathbb{N}$, are such that $\mathcal{J}(z_n) \to j$ as $n \to \infty$. Moreover, $u_n(0) = u_{0n} \in V_\alpha$ are such that $u_{0n} \rightharpoonup u_0$ weakly in V_α. Notice that

$$Z_{ad} \subseteq \mathcal{X} := W^{1,1+\lambda}\left((0,T); L^2(D)\right),$$

where $\mathcal{X}$ is reflexive. Since Z_{ad} is a closed and convex subspace of $\mathcal{X}$, Z_{ad} is also reflexive, and therefore, by taking a subsequence if necessary, we may assume that $\{z_n\}_{n\in\mathbb{N}}$ converges weakly in the space $Z_{\rho,\infty}$, to some $z_* \in Z_{ad} \subset Z_{\rho,\infty}$ (since Z_{ad} is weakly compact in the topology of $Z_{\rho,\infty}$; see [13, Proposition 3.2.8 and Theorem 3.2.1]).

Next, we aim to show that the state $\{u_n\}_{n\in\mathbb{N}}$ converges, as $n \to \infty$, to some u_* in a suitable sense, and that (u_*, z_*) satisfies the state equation $u_* = \mathcal{S}(z_*)$. More precisely, z_* becomes the desired optimal control for the problem, owing to the (weak) lower-sequential semicontinuity of the cost functional $\mathcal{J}$. By virtue of the proof of Corollary 3.7, we observe that u_n is bounded uniformly (with respect to $n \in \mathbb{N}$),

$$u_n \in C([0,T]; V_\alpha) \tag{4.13}$$

and

$$u_n \in W^{1,1+\xi}((0,T); V_\alpha) \cap L^{\sigma=1+\xi}((0,T); V_{\widetilde{\alpha}+2}) \overset{c}{\hookrightarrow} L^\sigma((0,T); V_\alpha), \tag{4.14}$$

since $V_{\widetilde{\alpha}+2} \overset{c}{\hookrightarrow} V_\alpha \hookrightarrow V_{\widetilde{\alpha}} = V_\beta$, and $\widetilde{\alpha} \le \alpha < \widetilde{\alpha} + 2$ (owing to $(I + A)^{-1}$ being compact in $L^2(\Omega)$). We recall that the embedding in (4.14) is also compact due to the (standard) Aubin-Lions-Simon compactness lemma. It follows that, as $n \to \infty$,

$$u_n \to u_* \text{ strongly in } L^\sigma((0,T); V_\alpha).$$

Together with (H3), this strong convergence implies that

$$f(u_n) \to f(u_*) \text{ strongly in } L^\sigma((0,T); V_\beta),$$

which is enough to pass to the limit in a standard way, in the sequence of mild solutions

$$u_n = S_\gamma(t) u_{0n} + \int_0^t P_\gamma(t-\tau) f(u_n(\tau)) d\tau + \int_0^t P_\gamma(t-\tau) \mathbb{B} z_n(\tau) d\tau. \tag{4.15}$$

Indeed, while it is easy to pass to the limit in the first and last summands on the right-hand side of (4.15), for the second convolution term we have

$$\left|P_\gamma * (f(u_n) - f(u_*))\right|_{V_\beta} \le t^{(\gamma-1)\frac{\sigma}{\xi}+1} \left\|f(u_n) - f(u_*)\right\|_{L^\sigma(0,T;V_\beta)},$$

for all $0 < \delta \le t \le T$. In particular, we have established the strong convergence in $C((0,T];V_\beta)$, of

$$\int_0^t P_\gamma(t-\tau) f(u_n(\tau))\, d\tau \to \int_0^t P_\gamma(t-\tau) f(u_*(\tau))\, d\tau, \text{ as } n \to \infty$$

for any $T < T_{\max}$. Therefore, there holds in V_β, for almost all $t \in (0,T)$,

$$u_* = S_\gamma(t) u_0 + \int_0^t P_\gamma(t-\tau) f(u_*(\tau))\, d\tau + \int_0^t P_\gamma(t-\tau) \mathbb{B} z_*(\tau)\, d\tau.$$

Due to (4.13), clearly $u_* \in C([0,T];V_\alpha)$; in fact, one may conclude as in the proof of Corollary 3.7 that $u_* \in Y_{\theta,\alpha} \subseteq Y_{\theta,\widetilde{\alpha}}$ (for each $\alpha \ge \widetilde{\alpha}$) is a solution in the sense of Theorem 3.6.

The element z_* is the right candidate for the optimal control. Indeed, z_* is the minimizer. We first notice that since J_2 is convex, proper, and lower-semicontinuous, therefore it is weakly lower-semicontinuous with respect to the X_2-norm topology (see [13, Theorem 3.3.3]). We have

$$\begin{aligned}\inf_{z\in Z_{ad}} \mathcal{J}(z) &= \liminf_{n\to\infty} \mathcal{J}(z_n) \ge \liminf_{n\to\infty} \mathcal{J}_1(\mathcal{S}(z_n)) + \liminf_{n\to\infty} \mathcal{J}_2(z_n)\\ &\ge \mathcal{J}_1(\mathcal{S}(z_*)) + \mathcal{J}_2(z_*) = \mathcal{J}(z_*).\end{aligned}$$

The proof is complete. □

Our next goal is to show differentiability of the control-to-state operator. We begin with another assumption on f.

(H4bis) The nonlinearity $f \in C^{2,1}(\mathbb{R})$ induces (for a fixed $u \in V_\alpha$) a bounded[12] nonlinear form

$$b_u(v,w) := \left(\partial_u^2 f(u)\right) vw : V_\alpha \times V_\alpha \to V_\beta,$$

such that, for every $u_1, u_2 \in V_\alpha$, satisfying $|u_i|_\alpha \le R$ $(i = 1,2)$, there is a constant $C_R > 0$ such that,

$$\left|b_{u_1}(v,w) - b_{u_2}(v,w)\right|_\beta \le C_R |v|_\alpha |w|_\alpha |u_1 - u_2|_\alpha. \tag{4.16}$$

Suppose now that $z_* \in Z_{ad}$ is a local minimizer for the control problem, and let $u_* = \mathcal{S}(z_*)$ be the associated state. We consider, for a fixed $h \in \mathcal{U}$, the linearized system:

$$\partial_t^\gamma \eta(t) = -A\eta(t) + f'(u_*(t))\eta(t) + \mathbb{B}h, \quad \eta(0) = 0, \text{ in } \Omega. \tag{4.17}$$

We now show that problem (4.17) admits for every $h \in Z_{\rho,\infty}$, a unique solution $\eta \in Y_{\theta,\alpha}$ (in the sense of Corollary 3.7), and that the linear mapping

$$\Phi : Z_{\rho,\infty} \to Y_{\theta,\alpha},\ h \mapsto \eta := \eta^h \tag{4.18}$$

is continuous from $Z_{\rho,\infty}$ into $Y_{\theta,\alpha}$. Namely, there is a constant $K_4 > 0$ such that

$$\|\eta\|_{Y_{\theta,\alpha}} \le K_4 \|h\|_{Z_{\rho,\infty}}. \tag{4.19}$$

[12]Note that boundedness of b_u is a consequence of (4.16).

Lemma 4.4. *Let the assumptions of Corollary 3.7 hold, and in addition, assume (H4) and (H4bis). Then the above statement for (4.18) holds. Moreover,* $\eta \in L^{1+\xi}((0,T);V_{2+\widetilde{\alpha}})$ *and* $\partial_t^{\gamma}\eta \in L^{1+\xi}((0,T);V_{\widetilde{\alpha}})$, *for the same value* $\xi > 0$ *defined previously.*

Proof. The existence of a (unique) mild solution follows exactly along the lines of Theorem 3.4, and is based on the formula

$$\eta(t) = \int_0^t P_\gamma(t-\tau) f'(u_*(\tau))\eta(\tau)\,d\tau + \int_0^t P_\gamma(t-\tau)\mathbb{B}h(\tau)\,d\tau.$$

We skip the (basic) details for the sake of brevity, and focus mainly on the stability estimate (4.19). Since η is also differentiable for almost all $t \in (0,T)$, we have

$$\begin{aligned}\partial_t\eta(t) &= \int_0^t P_\gamma(t-\tau) f'(u_*(\tau))\partial_t\eta(\tau)\,d\tau \\ &\quad + \int_0^t P_\gamma(t-\tau) f'(u_*(\tau))\partial_t\eta(\tau)\,d\tau \\ &\quad + \int_0^t P_\gamma(t-\tau)\mathbb{B}\partial_t h(\tau)\,d\tau + P_\gamma(t)\mathbb{B}h(0).\end{aligned}$$

This is due to the fact that if L is continuous at $t=0$ and L is of bounded variation on $(0,T)$, we have that

$$\frac{d}{dt}(P_\gamma * L)(t) = P_\gamma(t)L(0) + (P_\gamma * \partial_t L)(t), \text{ for } t > 0.$$

Following the basic argument developed in the proof of (4.5) (see Theorem 4.1), it is first easy to see that

$$\|\eta\|_{C([0,T];V_\alpha)} \le K_5\|h\|_{C([0,T];L^2(D))}, \tag{4.20}$$

for some $K_5 > 0$ independent of h. Similarly, arguing as in the proof of (4.9)-(4.10), owing to the boundedness of the form $b_u : V_\alpha \times V_\alpha \to V_\beta$, one has

$$|\partial_t\eta(t)|_\alpha \lesssim t^{\theta-1}\|h\|_{Z_{\rho,\infty}} + \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1}|\partial_t\eta(\tau)|_\alpha\,d\tau. \tag{4.21}$$

The application of Gronwall lemma to (4.21), together with (4.20), implies the desired conclusion (4.19). The final regularity on η can be deduced as in the proof of Corollary 3.7 (see also the proof of Proposition 4.6 below). □

Lemma 4.5. *Let the assumptions of Lemma 4.4 hold. Then the following statements hold:*

(i) *Let* $z_*(=z) \in \mathcal{U}$ *be arbitrary. Then the control-to-state mapping* $\mathcal{S} : Z_{\rho,\infty} \to Y_{\theta,\alpha}$ *is (Frechet) differentiable at* z_*, *and the frechet derivative* $d\mathcal{S}(z_*)$ *is given by* $d\mathcal{S}(z_*)(h) = \eta$, *where for any given* $h \in Z_{\rho,\infty}$, *the function* η *denotes the solution of the linearized system (4.17).*

(ii) *The mapping* $d\mathcal{S}(z_*) : \mathcal{U} \to \mathcal{L}(Z_{\rho,\infty}, Y_{\theta,\alpha})$, $z_* \mapsto d\mathcal{S}(z_*)$ *is Lipschitz continuous on* $\mathcal{U}$ *in the following sense: there exists a constant* $K_6 > 0$ *such that for all* $z_1, z_2 \in \mathcal{U}$ *and all* $h \in \mathcal{U}$ *the following estimate holds:*

$$\|d\mathcal{S}(z_1)h - d\mathcal{S}(z_2)h\|_{Y_{\theta,\alpha}} \le K_6\|z_1 - z_2\|_{Z_{\rho,\infty}}\|h\|_{Z_{\rho,\infty}}.$$

Proof. We begin with (i). Let $z \in \mathcal{U}$ be arbitrarily chosen and let $u = \mathcal{S}(z)$ be the associated solution to the state system. Since $\mathcal{U}$ is open in $Z_{\rho,\infty}$, there is $\zeta > 0$ such that for any $h \in Z_{\rho,\infty}$ with $\|h\|_{Z_{\rho,\infty}} \le \zeta$ there holds $z + h \in \mathcal{U}$. In what follow, we also consider solutions $u^h = \mathcal{S}(z+h)$, for such variations $h \in \mathcal{U}$. Next, we let $v^h := u^h - u - \eta^h$, where $\eta = \eta^h$ denotes the unique solution to the linearized system associated with a given $h \in \mathcal{U}$. Our goal is to show that $\|v^h\|_{Y_{\theta,\alpha}} = o(\|h\|_{Z_{\rho,\infty}})$, as $\|h\|_{Z_{\rho,\infty}} \to 0$. To this end, we notice that v^h satisfies

$$v^h(t) = \int_0^t P_\gamma(t-\tau)\Big(f(u^h(\tau)) - f(u(\tau)) - f'(u(\tau))\eta(\tau)\Big)d\tau,$$

for almost all $t \in (0,T) \subset (0,T_{\max})$. By Taylor's theorem [16, Theorem 30.1.3], we have for almost all $t \in (0,T)$,

$$f(u^h(t)) - f(u(t)) - f'(u(t))\eta(t) = f'(u(t))v^h(t) + r^h(t), \tag{4.22}$$

with remainder

$$r^h(t) := \int_0^1 (f'(u(t) + x(u^h-u)(t)) - f'(u(t)))(u^h-u)(t)\,dx.$$

Moreover, $v^h(0) = 0$, $r^h(0) = 0$, and v^h is differentiable a.e. on $(0,T)$, with

$$\begin{aligned}
\partial_t v^h(t) &= \int_0^t P_\gamma(t-\tau)\, b_u(\partial_t u(\tau), v^h(\tau))d\tau + \int_0^t P_\gamma(t-\tau)\, g_u(\partial_t v^h(\tau))d\tau \qquad (4.23)\\
&+ \int_0^1\int_0^t P_\gamma(t-\tau)\left[g_{u+x(u^h-u)}(\partial_t(u^h-u)) - g_u(\partial_t(u^h-u))\right]d\tau dx\\
&+ \int_0^1\int_0^t P_\gamma(t-\tau)\left[b_{u+x(u^h-u)}(\partial_t u, u^h-u) - b_u(\partial_t u, u^h-u)\right]d\tau dx\\
&+ \int_0^1\int_0^t P_\gamma(t-\tau)\, x b_{u+x(u^h-u)}(u^h-u, \partial_t(u^h-u))d\tau dx\\
&=: Q_1 + \ldots + Q_5.
\end{aligned}$$

In view of (4.22), we can also rewrite

$$v^h(t) = \int_0^t P_\gamma(t-\tau)\, f'(u(\tau))v^h(\tau))d\tau + \int_0^t P_\gamma(t-\tau)\, r^h(\tau))d\tau.$$

We deduce

$$\begin{aligned}
\left|v^h(t)\right|_\alpha &\le C_R\int_0^t |P_\gamma(t-\tau)\, f'(u(\tau))v^h(\tau))|_\alpha d\tau\\
&\quad + C_R\int_0^1\int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1} x|u^h(\tau) - u(\tau)|_\alpha d\tau dx\\
&\lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}}||v^h||_{C([0,T];V_\alpha)} + t^{\frac{\gamma(2-\alpha+\beta)}{2}}||u^h-u||^2_{C([0,T];V_\alpha)},
\end{aligned}$$

so that for small enough $T \ll 1$, we obtain

$$||v^h||_{C([0,T];V_\alpha)} \lesssim ||u^h-u||^2_{C([0,T];V_\alpha)} \lesssim \|h\|^2_{Z_{\rho,\infty}}, \text{ by (4.11)}. \tag{4.24}$$

The continuation argument exploited in the proof of Lemma 3.3 yields the same estimate on the whole interval $(0,T)$, for any $T < T_{\max}$. It remains to estimate all Q_i-terms in (4.23). The assumptions (H4)-(H4bis) and (4.24) are mainly exploited in these estimates. We thus find that

$$|Q_1(t)|_\alpha \lesssim \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1}\tau^{\theta-1}d\tau\left(1+\|u\|_{C([0,T];V_\alpha)}\right)\|u\|_{Y_{\Theta,\varepsilon}}\|v^h\|_{C([0,T];V_\alpha)} \tag{4.25}$$

$$\lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}+\theta-1}\|h\|^2_{Z_{\rho,\infty}},$$

and

$$|Q_2(t)|_\alpha \lesssim \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1}\left|\partial_t v^h(\tau)\right|_\alpha d\tau.$$

Analogously, we obtain that

$$\begin{aligned}|Q_4(t)|_\alpha &\lesssim \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1}|\partial_t u(\tau)|_\alpha |(u^h-u)(\tau)|^2_\alpha d\tau\\ &\lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}+\theta-1}\|u^h-u\|^2_{C([0,T];V_\alpha)}\|u\|_{Y_{\Theta,\epsilon}}\\ &\lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}+\theta-1}\|h\|^2_{Z_{\rho,\infty}},\end{aligned}$$

$$\begin{aligned}|Q_5(t)|_\alpha &\lesssim \int_0^t (t-\tau)^{\frac{\gamma(2-\alpha+\beta)}{2}-1}\left|\partial_t(u^h-u)(\tau)\right|_\alpha\left|(u^h-u)(\tau)\right|_\alpha d\tau\\ &\lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}+\theta-1}\|h\|^2_{Z_{\rho,\infty}},\end{aligned}$$

as well as

$$|Q_3(t)|_\alpha \lesssim t^{\frac{\gamma(2-\alpha+\beta)}{2}+\theta-1}\|h\|^2_{Z_{\rho,\infty}}. \tag{4.26}$$

Once again collecting the previous estimates, we obtain by means of Gronwall's lemma that

$$\sup_{t\in[0,T]} t^{1-\theta}|\partial_t v^h(t)|_\alpha \lesssim \|h\|^2_{Z_{\rho,\infty}}. \tag{4.27}$$

Combining (4.24)-(4.27), we finally arrive at the conclusion (i). The proof of (ii) follows in a similar fashion; we leave the details to the interested reader. ☐

It is now straightforward to derive the standard variational inequality that optimal controls must satisfy. However, for the method to be practical, it is critical to identify the adjoint equation. We shall proceed on these two fronts simultaneously. Exploiting first the integration by parts formula (2.4) of Proposition 2.5, we introduce a dual problem, that can be associated with (4.17) whenever $h \equiv 0$, also owing to $\eta(0)=0$,

$$\begin{cases}\partial^\gamma_{t,T}w + Aw &= f'(u_*)w + k, \quad \text{in } Q := (0,T)\times\Omega,\\ I^{1-\gamma}_{t,T}w(T,\cdot) &= 0 \qquad \text{in } \Omega.\end{cases} \tag{4.28}$$

Roughly speaking, one identifies k with $d_u J(\mathcal{S}(z_*), z_*)$, where the function $u_* = \mathcal{S}(z_*)$ is the state associated with a minimizer $z_* \in Z_{ad} \subset Z_{\rho,\infty}$, of (4.12). Note that (4.28) is a (linear) backward in time partial differential equation for the (right) Riemann-Liouville fractional derivative $\partial^\gamma_{t,T}$. We now use the time transformation $t \mapsto T-t := \bar{t}$ in (4.28) to set $p(t) = w(T-t)$ (which also yields that $w(t) = w(T-\bar{t}) = p(\bar{t})$) and $k(t) := g(\bar{t})$. Employing the basic identities

$$(I^{1-\gamma}_{t,T}w)(t) = (I^{1-\gamma}_{0,\bar{t}}p)(\bar{t}), \quad -\frac{d}{dt} = \frac{d}{d\bar{t}}, \tag{4.29}$$

we can then transform (4.28) into a (forward) problem for the left Riemann-Liouville derivative $D^{\gamma}_{T-t} = D^{\gamma}_{\bar{t}}$. In particular, solvability of problem (4.28) turns out to be related to solvability of the following (generic) initial-value problem

$$\begin{cases} D^{\gamma}_{\bar{t}}p + Ap = f^{'}(u_*)\, p + g =: l, \quad \bar{t} \in (0,T), \\ (I^{1-\gamma}_{0,\bar{t}}p)(0) = p_0. \end{cases} \tag{4.30}$$

Here, once again in the context of (4.28) g must be equal to $d_u J(\mathcal{S}(z_*), z_*)$ and $p_0 = 0$. The solvability of the linearized problem (4.30) under suitable conditions on l has also been investigated in detail by Bajlekova [15, Section 4, Theorem 4.16]. However, we prefer to give a more direct proof of the solvability here due to the additional summand $f^{'}(u_*)\, p$. To this end, we recall that if $(I^{1-\gamma}_{0,\bar{t}}p)(0) = p_0$, one has the following integral solution representation for the above linearized problem:

$$p(t) = P_{\gamma}(t)\, p_0 + \int_0^t P_{\gamma}(t-\tau) \Big(f^{'}(u_*(\tau))\, p(\tau) + g(\tau) \Big) d\tau, \tag{4.31}$$

where $u_*(t) \in \mathbb{B}_R$ ($\mathbb{B}_R$ is a ball of radius R, in the corresponding strong topology of $Y_{\theta,\alpha}$; see (4.2)), and we have dropped the bar from $\bar{t}$, for the sake of notational simplicity. We also recall from (4.2) that the Banach space

$$Y_{\rho,\widetilde{\alpha}} = \left\{ k \in C([0,T]; V_{\widetilde{\alpha}}) : |\partial_t k(t)|_{\widetilde{\alpha}} \lesssim t^{\rho-1}, \quad \text{a.e. } 0 < t \leq T \right\},$$

is subject to the (natural) norm

$$\|k\|_{Y_{\rho,\widetilde{\alpha}}} := \|z\|_{C([0,T];V_{\widetilde{\alpha}})} + \sup_{t\in[0,T]} t^{1-\rho} |\partial_t k(t)|_{\widetilde{\alpha}},$$

for some $\rho > 0$. We sketch a proof of the subsequent result in the Appendix (see Section 7).

Proposition 4.6. *Let $u := u_*(t) \in \mathbb{B}_R$, be an optimal solution of (3.1) in the sense of Corollary 3.7, for some $R > 0$ and $t \in (0,T)$ with $T \leq T_{\max}$ ($T_{\max} > 0$ is the maximal existence time for u; see Theorem 3.4).*

(i) If (H4) holds and $k \in L^q((0,T); V_{\widetilde{\alpha}})$, for some $q \in (\frac{2}{\gamma(2-\alpha+\widetilde{\alpha})}, \infty]$, then the linearized problem (4.28) admits a unique mild solution on $(0,T)$. In particular, one has

$$w \in C([0,T]; V_{\alpha}); \; \{\partial^{\gamma}_{t,T} w, Aw\} \in C([0,T]; V_{\widetilde{\alpha}-\delta}), \text{ for } \frac{2}{\gamma q} < \delta \leq 2.$$

The variational equation

$$\langle \partial^{\gamma}_{t,T} w(t) + Aw(t), v \rangle_{V_{\widetilde{\alpha}-\delta}, V_{-\widetilde{\alpha}+\delta}} = \langle f^{'}(u(t))\, w(t) + k(t), v \rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}},$$

is satisfied, for any $v \in V_{-\widetilde{\alpha}+\delta} \subset V_{-\widetilde{\alpha}}$, for almost all $t \in (0,T)$.

(ii) If (H4), (H4bis) hold and additionally $k \in Y_{\rho,\widetilde{\alpha}}$, then the mild solution of (4.28) also satisfies[13] *$w \in Y_{\theta,\alpha}$, and*

$$\langle \partial^{\gamma}_{t,T} w(t) + Aw(t), v \rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}} = \langle f^{'}(u(t))\, w(t) + k(t), v \rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}},$$

for any $v \in V_{-\widetilde{\alpha}}$, for almost all $t \in (0,T)$. Furthermore, $w \in L^{1+\xi}((0,T); V_{2+\widetilde{\alpha}})$, $\partial^{\gamma}_{t,T} w \in L^{1+\xi}((0,T); V_{\widetilde{\alpha}})$, for some $\xi = \xi(\theta) > 0$.

[13] We recall once again that θ is independent of $\rho > 0$.

It follows from the chain rule that the reduced cost functional $\mathcal{J}(z) = J(u,z) = J(\mathcal{S}(z), z)$ is Frï¿½chet differentiable at every $z \in \mathcal{U}$ (provided that $u \mapsto J_1(u)$ and $z \mapsto J_2(z)$ are continuously (Frï¿½chet) differentiable) with Frï¿½chet derivative

$$d\mathcal{J}(z) = d_u J(\mathcal{S}(z), z) \circ d\mathcal{S}(z) + d_z J(\mathcal{S}(z), z). \tag{4.32}$$

Equivalently for any $h \in \mathcal{U}$, we have

$$\begin{aligned}
&\int_0^T (d\mathcal{J}(z(t)), h(t))_{L^2(D)}\, dt \qquad (4.33)\\
&= \int_0^T \langle d_u J(\mathcal{S}(z(t)), z(t)), d\mathcal{S}(z(t))\, h(t)\rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}} dt\\
&+ \int_0^T (d_z J(\mathcal{S}(z(t)), z(t)), h(t))_{L^2(D)}\, dt,
\end{aligned}$$

where we have used the differentiability of $\mathcal{S}$ from Lemma 4.5(i) and the fact that $d\mathcal{S}(z) \in \mathcal{L}(Z_{\rho,\infty}, Y_{\theta,\widetilde{\alpha}})$ because $Y_{\theta,\alpha} \subseteq Y_{\theta,\widetilde{\alpha}}$, for each $a \geq \widetilde{\alpha} \geq -1$. Since $d\mathcal{S}(z)$ is bounded and linear, it follows that its adjoint $d\mathcal{S}(z)^* \in \mathcal{L}(Z_{\rho,\infty}, Y_{\theta,\widetilde{\alpha}})^*$ is well-defined. Consequently, from (4.33) we obtain

$$\begin{aligned}
&\int_0^T (d\mathcal{J}(z(t)), h(t))_{L^2(D)}\, dt \qquad (4.34)\\
&= \int_0^T \big(d\mathcal{S}(z(t))^* d_u J(\mathcal{S}(z(t)), z(t)) + d_z J(\mathcal{S}(z(t)), z(t)), h(t)\big)_{L^2(D)}\, dt.
\end{aligned}$$

Thus to evaluate $d\mathcal{J}$, we need to identify $d\mathcal{S}(z(\cdot))^*$, we do this next.

Lemma 4.7. *Let $(z,u) \in Z_{\rho,\infty} \times Y_{\theta,\alpha}$ solve the state equation in the sense of Theorem 3.6. For a.e. $t \in [0,T]$, the adjoint operator $d\mathcal{S}(z(t))^*\psi(t) : V_{-\widetilde{\alpha}} \to L^2(D)$ is given by*

$$d\mathcal{S}(z(t))^*\psi = \mathbb{B}^* w(t) \in L^2(D).$$

Furthermore, w solves the linear equation

$$\begin{cases} \partial_{t,T}^{\gamma} w + Aw &= f'(u_*)\, w + \psi, \ t \in (0,T),\\ I_{t,T}^{1-\gamma} w(T,\cdot) &= 0. \end{cases} \tag{4.35}$$

Proof. Recall that for each $\widetilde{\alpha} \in [-1,0]$, $\eta = d\mathcal{S}(z)h$ solves (4.17), while $\eta \in Y_{\theta,\alpha} \subset Y_{\theta,\widetilde{\alpha}}$ and

$$\eta \in L^{\sigma}((0,T); V_{2+\widetilde{\alpha}}) \subseteq L^{\sigma}((0,T); V_{-\widetilde{\alpha}}) \text{ and } \partial_t^{\gamma}\eta \in L^{1+\xi}((0,T); V_{\widetilde{\alpha}})$$

(the values σ and ξ are given in Corollary 3.7). For every $\psi \in Y_{\rho,\widetilde{\alpha}}$ and $h \in Z_{\rho,\infty}$, we have that

$$\int_0^T \langle \psi(t), d\mathcal{S}(z(t))h(t)\rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}} dt = \int_0^T (d\mathcal{S}(z)^*\psi(t), h(t))_{L^2(D)}\, dt. \tag{4.36}$$

Testing (4.35) with η solving (4.17), in view of Proposition 4.6(ii) (recall that w is sufficiently smooth), we obtain

$$\begin{aligned}&\int_0^T \langle \psi(t), d\mathcal{S}(z(t))h(t)\rangle_{V_{\widetilde{\alpha}},V_{-\widetilde{\alpha}}} dt\\&=\int_0^T \langle \psi(t), \eta(t)\rangle_{V_{\widetilde{\alpha}},V_{-\widetilde{\alpha}}} dt\\&=\int_0^T \left[\langle \partial_{t,T}^{\gamma} w(t), \eta(t)\rangle_{V_{\widetilde{\alpha}},V_{-\widetilde{\alpha}}} + \langle Aw(t)-f'(u(t))w(t), \eta(t)\rangle_{V_{\widetilde{\alpha}},V_{-\widetilde{\alpha}}}\right] dt.\end{aligned}$$

Applying integration by parts in time (see Proposition 2.5) and using the fact that A can be extended to a (self-adjoint) isomorphism acting from $V_{\widetilde{\alpha}}$ into $V_{-\widetilde{\alpha}}$, we arrive at

$$\int_0^T \langle \psi(t), d\mathcal{S}(z(t))h(t)\rangle_{V_{\widetilde{\alpha}},V_{-\widetilde{\alpha}}} dt = \int_0^T \left\langle w(t), \partial_t^{\gamma}\eta(t)+A\eta(t)-f'(u(t))\eta(t)\right\rangle_{V_{-\widetilde{\alpha}},V_{\widetilde{\alpha}}} dt.$$

Then using (4.17), we immediately obtain

$$\int_0^T \langle \psi(t), d\mathcal{S}(z(t))h(t)\rangle_{V_{\widetilde{\alpha}},V_{-\widetilde{\alpha}}} dt = \int_0^T \langle w(t), \mathbb{B}h(t)\rangle_{V_{-\widetilde{\alpha}},V_{\widetilde{\alpha}}} dt. \tag{4.37}$$

The asserted result then follows from (4.36) and (4.37). $\square$

Finally, we are ready to state the first order necessary optimality conditions.

Theorem 4.8. *Let $z_* \in Z_{ad}$ be a local minimum for (4.12) with $u_* = \mathcal{S}(z_*)$ solving the state equation. If w solves the adjoint equation (4.35) with ψ replaced by $d_u J(\mathcal{S}(z), z)$, then the following necessary optimality conditions hold:*

$$\begin{aligned}&\int_0^T \left(d\mathcal{J}(z_*(t)), z(t)-z_*(t)\right)_{L^2(D)} dt\\&=\int_0^T \left(\mathbb{B}^* w(t) + d_z J(\mathcal{S}(z_*(t)), z_*(t)), z(t)-z_*(t)\right)_{L^2(D)} dt \geq 0\end{aligned} \tag{4.38}$$

for all $z \in Z_{ad}$.

Proof. The proof immediately follows from the convexity of Z_{ad} and the assumed differentiability of $\mathcal{J}$. Towards this end, using Lemma 4.7, we can write (4.34) equivalently as the right-hand side of (4.38), where w solves the adjoint equation (4.35) with ψ replaced by $d_u J(\mathcal{S}(z), z)$. $\square$

5. **A global regularity result for energy solutions.** Our goal in this section is to describe the proper regularity conditions necessary to obtain globally defined bounded solutions, i.e., $T_{\max} = \infty$ (see Corollary 3.7). We begin with a simple energy estimate which is a consequence of the Hardy-Littlewood theorem. To this end, let H be a Hilbert space with its associated inner product $(\cdot,\cdot)$ and norm $|\cdot|_H$, respectively.

Proposition 5.1. *If $u \in W^{1,p}((0,T);H)$ with $p \geq \frac{2}{2-\gamma}$, then the following holds:*

$$\int_0^T \left(\partial_t^{\gamma} u(t), \partial_t u(t)\right) dt \geq \sin\frac{\gamma\pi}{2} \sum_{n=1}^{\infty} \left\| g_{(1-\gamma)/2} * \partial_t u_n \right\|_{L^2(0,T)}^2 \|\psi_n\|_H^2 \geq 0,$$

for any $\gamma \in (0,1)$ and $T > 0$. Here, $u_n(t) := (u(t), \psi_n)_H$ where $(\psi_n)_{n\in\mathbb{N}}$ is an orthogonal basis for H.

Proof. We have in H,

$$\partial_t^\gamma u(t) = \sum_{n=1}^{\infty} (g_{1-\gamma} * \partial_t u_n)(t)\,\psi_n,\ \ \partial_t u(t) = \sum_{n=1}^{\infty} \partial_t u_n(t)\,\psi_n.$$

Clearly, $u_n \in W^{1,p}(0,T)$. It follows that

$$\begin{aligned}\int_0^T (\partial_t^\gamma u(t), \partial_t u(t))\,dt &= \sum_{n=1}^{\infty} \|\psi_n\|_H^2 \int_0^T (g_{1-\gamma} * \partial_t u_n)(t)\,\partial_t u_n(t)\,dt \\ &\geq \sin\frac{\gamma\pi}{2} \sum_{n=1}^{\infty} \left\|g_{(1-\gamma)/2} * \partial_t u_n\right\|_{L^2(0,T)}^2 \|\psi_n\|_H^2\end{aligned}$$

where the last bound follows from [33, Corollary 2.1]. The proof is finished. □

We need a crucial regularity assumption on the operator $\mathbb{B}$ in what follows. For the sake of convenience, we also define $F(u) = \int_0^u f(\tau)\,d\tau$.

(H6) Let $z \in Z_{\rho,\infty}$ and $\mathbb{B} \in \mathcal{L}\left(L^2(D); V_{\widetilde{\alpha}}\right)$ with $u_0 \in V_{\beta+2} \subset V_{\alpha=1}$ such that $F(u_0) \in L^1(\Omega)$. Assume also that (H3) is satisfied with $\alpha = 1$, in *either one* of the following regimes: (i) $0 < \beta < \widetilde{\alpha} \leq 1 = \alpha$; (ii) $0 < \widetilde{\alpha} \leq \beta \leq 1 = \alpha$.

The following global regularity for $\gamma \in (0,1)$ is the main result of the section. Let $\kappa := \beta$, when (i) holds and $\kappa := \widetilde{\alpha}$, if (ii) holds, respectively. Then set $\theta := \frac{\gamma}{2}(1+\kappa)$.

Theorem 5.2. *Let (H6) hold and assume*[14] $F(s) \leq -C_F$, *for all* $s \in \mathbb{R}$, *for some* $C_F \in \mathbb{R}$. *Then problem (3.1) admits a unique (globally-defined) weak solution* $u \in Y_{\theta,1}$. *Namely,*

$$u \in W^{1,\frac{1}{1-\theta}-}((0,T);V_1) \cap L^{\frac{1}{1-\theta}-}((0,T);V_{\kappa+2}),\ {}_C\partial_t^\gamma u = \partial_t^\gamma u \in L^{\frac{1}{1-\theta}-}((0,T);V_\kappa), \tag{5.1}$$

for any (fixed) but otherwise arbitrary $T > 0$.

Proof. Note that by (H6), $1 = \alpha \in I_{\widetilde{\alpha}} \cap I_\beta \neq \varnothing$ and (H2)-(H3) are satisfied since $z \in Z_{\rho,\infty}$. Therefore, as in the proof of Corollary 3.7, the problem admits a unique weak solution satisfying (5.1) for every $0 < T < T_{\max}$ (where $T_{\max} > 0$ is such that either $T_{\max} = \infty$, *or* $\lim_{t\to T_{\max}^-} |u(t)|_1 = \infty$, if $T_{\max} < \infty$). Observe that $u : [0,T] \to V_1$ is absolutely continuous, and notice also that $p := \frac{1}{1-\theta} > \frac{2}{2-\gamma}$ if and only if $\theta > \gamma/2$, which holds since $\beta > 0$ on account[15] of (i) (the case (ii) is similar). Thus, Proposition 5.1 applies with $H = L^2(\Omega)$, and we obtain

$$\begin{aligned}&2(\partial_t^\gamma u(t), \partial_t u(t))_{L^2(\Omega)} + \frac{d}{dt}\left[|u(t)|_1^2 - 2(F(u(t)),1) - 2(\mathbb{B}z(t), u(t))\right] \\ &= -2(\mathbb{B}\partial_t z(t), u(t))_{L^2(\Omega)},\end{aligned} \tag{5.2}$$

for almost all $t \in (0,T)$. The right hand side of (5.2) is bounded in terms of $C|u(t)|_1 t^{\rho-1}$, for some $C > 0$ independent of t, T, owing to the fact that $z \in Z_{\rho,\infty}$ and $\mathbb{B} \in \mathcal{L}\left(L^2(D); V_{-1}\right)$. In particular,

$$2\left|(\mathbb{B}\partial_t z(t), u(t))_{L^2(\Omega)}\right| \lesssim \left(1 + |u(t)|_1^2\right) t^{\rho-1},\ \text{for } 0 < t \leq T$$

[14]Strictly speaking, this assumption can be easily relaxed depending on the application one has in mind for the problem.

[15]This is the only place where one needs to assume $\beta > 0$, in order to derive the regularity $W^{1,p}((0,T);V_1)$, for $p > 2/(2-\gamma)$ and effectively exploit Proposition 5.1.

and

$$2|\left(\mathbb{B}z\left(t\right),u\left(t\right)_{L^2(\Omega)}\right)|\leq C_\delta\left\|z\right\|^2_{Z_{\rho,\infty}}+\delta\left|u\left(t\right)\right|^2_1,\text{ for every }\delta>0. \tag{5.3}$$

Set now

$$E_\gamma\left(t\right):=C_T+\left|u\left(t\right)\right|^2_1-2\left(F\left(u\left(t\right)\right),1\right)_{L^2(\Omega)}-2\left(\mathbb{B}z\left(t\right),u\left(t\right)\right)_{L^2(\Omega)},$$

where $C_T>0$ is sufficiently large (depending clearly on $z\in Z_{\rho,\infty}$) such that $E_\gamma\geq 0$ on $(0,T)$ (this is possible due to (5.3), for a sufficiently small $\delta\ll 1$). Notice also that $E_\gamma\left(0\right)\leq(C_T+|u_0|^2_1+\|F\left(u_0\right)\|_{L^1(\Omega)})$. We immediately deduce from (5.2) that,

$$\partial_t E_\gamma\left(t\right)+2\left(\partial_t^\gamma u\left(t\right),\partial_t u\left(t\right)\right)_{L^2(\Omega)}\leq C_T^{'}\left(1+\left|u\left(t\right)\right|^2_1\right)t^{\rho-1}\leq C_T^{''}E_\gamma\left(t\right)t^{\rho-1}$$

for some $C_T^{'},C_T^{''}>0$. Integrating the foregoing inequality over $(0,T)$, we deduce on account of Gronwall's lemma, that there exists $C_T^{'''}<\infty$, such that

$$C_{T,\delta}(1+\left|u\left(T\right)\right|^2_1)\leq E_\gamma\left(T\right)\leq C_T^{'''}E_\gamma\left(0\right)<\infty, \tag{5.4}$$

for any $T>0$. The energy inequality (5.4) finally yields in view of Theorem 3.4 that $T_{\max}=\infty$. The proof is finished. □

The above proof underlines once again the additional smoothness required of the sources on the right hand side of the equation in order to be able to (rigorously) justify the energy equality[16] for (3.1). This is in contrast to what happens in the classical case where generally much less is required on the sources on the right-hand side (and, which is due to the absence of strongly singular behavior of the solution operator near $t=0$). This result is optimal since $p=2$ in Proposition 5.1 corresponds exactly to the case when $\gamma=1$. In general, $1<p<2$ whenever $0<\gamma<1$, and the value of p diminishes toward the value 1, as γ goes to zero, no matter how smooth the right-hand side turns out to be.[17]

6. Control setting and operator examples.

In this section we give some examples of operators, functionals, and nonlinearities that enter in our framework described in the previous sections.

6.1. Cost functionals and admissible set.

We set, for given functions

$$z_Q\in L^2\left((0,T);L^2\left(\Omega\right)\right),\ z_\Sigma\in L^2\left((0,T);L^2\left(\partial\Omega\right)\right) \tag{6.1}$$

and constants $a_i\geq 0$ (not all identically zero), the cost functional

$$J_1(u)=\frac{a_1}{2}\int_0^T\left\|u\left(\cdot,t\right)-z_Q\left(\cdot,t\right)\right\|^2_{L^2(\Omega)}dt+\frac{a_2}{2}\int_0^T\left\|u\left(\cdot,t\right)-z_\Sigma\left(\cdot,t\right)\right\|^2_{L^2(\partial\Omega)}dt. \tag{6.2}$$

Additionally, we let

$$J_2(z)=\frac{\zeta}{2}\|z\|^2_{L^2((0,T)\times D)}, \tag{6.3}$$

where $D=\Omega$ if the control z lies in the interior of Ω, and $D=\partial\Omega$ in case the control is placed on the boundary $\partial\Omega$. Moreover, $\zeta>0$ is a control regularization parameter. We consider the problem of minimizing the total cost functional $J:=$

[16] Note also that a smoother initial datum $u_0\in V_{1+\varepsilon}$, for some $\varepsilon>0$, is neccesary due to the "loss" of regularity of the solution flow near $t=0$.

[17] A common mistake we had found in scientific journal publications nowadays is the fact that many authors generally assume in their definition of weak or smooth solutions that the $p=2$ regularity can be always achieved for arbitrary $0<\gamma<1$ and when sources are sufficiently smooth.

$J_1 + J_2$, subject to the constraint $z \in Z_{ad}$, where we define the admissible control set (with prescribed singular behavior near $t = 0$) to be

$$\begin{aligned} Z_{ad} := \{ & z \in W^{1,2}((0,T); L^2(D)) \ : \ \|\partial_t z\|_{L^2(D)} \le M t^{\rho - 1}, \\ & \text{and } z_a \le z \le z_b, \text{ a.e. in } (0,T) \times D \}. \end{aligned} \tag{6.4}$$

Here, $z_a, z_b \in L^2((0,T) \times D)$ with $z_a \le z_b$ are given, and we assume that $M > 0$, $\rho > 1/2$. Notice that, Z_{ad} is a closed and convex subset of $Z_{\rho,\infty}$. Moreover, Z_{ad} depicts a generic situation with box constraints z_a and z_b. The hypothesis (H5) is then satisfied by the above J.

Our aim is to formulate necessary optimality conditions for our nonlocal in time (subdiffusive) problem. We introduce $\mathcal{U}$ as a nonempty open subset of $Z_{\rho,\infty}$ which contains Z_{ad}; without loss of generality, we may assume that $\mathcal{U}$ is also open in $W^{1,2}((0,T); L^2(D))$. Recall that the control-to-state mapping $\mathcal{S} : \mathcal{U} \to Y_{\theta,\alpha}$, $\mathcal{S}(z) = u$ is the unique (variational) solution of (3.1) (in the sense of Corollary 3.7). In the context of (6.2)-(6.3), the 'reduced' cost functional $\mathcal{J} : \mathcal{U} \to \mathbb{R}$ is then given by

$$\mathcal{J}(z) = J(u, z) := J_1(\mathcal{S}(z)) + J_2(z).$$

As Z_{ad} is convex, the desired necessary condition for optimality is

$$\langle d\mathcal{J}(z_*), z - z_* \rangle \ge 0, \tag{6.5}$$

for every $z \in Z_{ad}$ (for a proper optimal control $z_* \in Z_{ad}$), provided that $d\mathcal{J}(z_*)$ is well-defined (at least in the Gï¿½teaux sense) in the dual space $\left(W^{1,2}((0,T); L^2(D))\right)^*$. Following Section 4, it turns out that $\mathcal{S}$ is (continuously) Frï¿½chet differentiable at z_* so that the chain rule can be applied. As we had seen in Lemma 4.5, this leads to the linearized problem (4.17) which can then be stated for a generic element $h \in \mathcal{U}$. This, in turn, leads to the fact that the Frechet derivative $d\mathcal{S}(z) \in \mathcal{L}(Z_{\rho,\infty}, Y_{\theta,\alpha})$ exists for (a given generic) $\rho > 1/2$, such that $d\mathcal{S}(z)h = \eta$, where η is the unique (variational) solution of the aforementioned linearized problem. The latter can be described in detail following the statement (ii) of Proposition 4.6, provided that some regularity criteria for $k = k(t) \in Y_{\pi,\widetilde{\alpha}}$ is given for another (generic) parameter $\pi \in (0,1]$ (to be determined below, see (6.7)). We thus can immediately apply the chain rule and exploit the formula (4.32), to find that (6.5) takes on the form

$$a_1 \int_0^T (u_* - z_Q, \eta)_{L^2(\Omega)}\, dt + a_2 \int_0^T (u_* - z_\Sigma, \eta)_{L^2(\partial\Omega)}\, dt + \zeta \int_0^T \int_D z_* h\, dx dt \ge 0, \tag{6.6}$$

for any given $z \in Z_{ad}$, where the function η is the solution of the linearized problem corresponding to $h = z - z_*$. As usual, the final procedure consists in eliminating η in (6.6) by exploiting the 'backward-in-time' solution

$$w \in L^{1+\xi}((0,T); V_{2+\widetilde{\alpha}}),\ \partial_{t,T}^{\gamma} w \in L^{1+\xi}((0,T); V_{\widetilde{\alpha}}),$$

of the corresponding adjoint problem (see (ii) of Proposition 4.6), but now set

$$k := d_u J_1(\mathcal{S}(z_*), z_*) = \begin{pmatrix} a_1 (u_* - z_Q) \\ a_2 (u_* - z_\Sigma) \end{pmatrix} \in Y_{\pi,\widetilde{\alpha}}.$$

Namely, w satisfies

$$\langle \partial_{t,T}^{\gamma} w(t) + Aw(t), v \rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}} = \langle f'(u(t)) w(t) + k(t), v \rangle_{V_{\widetilde{\alpha}}, V_{-\widetilde{\alpha}}},$$

for any $v \in V_{-\widetilde{\alpha}}$, for almost all $t \in (0,T)$. We note that for $z_* \in Z_{ad} \subset Z_{\rho,\infty}$, it follows from Lemma 4.5 that $u_* = \mathcal{S}(z_*) \in Y_{\theta,\alpha} \subseteq Y_{\theta,\widetilde{\alpha}}$ (since $\alpha \geq \widetilde{\alpha}$). Thus, we must choose $\pi := \theta$ and consider further regularity assumptions on the data, i.e.,

$$z_Q, z_\Sigma \in Y_{\theta,\widetilde{\alpha}}. \tag{6.7}$$

This allows[18] to conclude that $k \in Y_{\theta,\widetilde{\alpha}}$; finally, we can eliminate

$$\eta \in L^{1+\xi}\left((0,T);V_{2+\widetilde{\alpha}}\right) \subset L^{1+\xi}\left((0,T);V_{-\widetilde{\alpha}}\right)$$

(where $\partial_t^\gamma \eta \in L^{1+\xi}\left((0,T);V_{\widetilde{\alpha}}\right)$). Indeed, notice that all the operations on the left-hand side of (6.6) are now well-defined in view of (6.1) and (6.7), respectively, and we can perform calculations exactly as in the proof of Lemma 4.7.

We can state a simpler form of the optimality conditions (6.6) in the context of various examples of diffusion operators. We do that next.

6.2. **The fractional Neumann problem for the Laplacian.** For the sake of simplicity, assume that $\Omega \subset \mathbb{R}^n$, $n \geq 1$, has a smooth boundary $\partial\Omega$. In this section denote by $B := -\Delta_{\Omega,N}$ the realization of $(-\Delta)$ in $L^2(\Omega)$ with the zero Neumann boundary condition. Since Ω is assumed to be smooth we have that $D(B) = \{u \in H^2(\Omega) : \partial_\nu u = 0 \text{ on } \partial\Omega\}$. Thus, fractional powers $A := B^s$ of order $s \in [0,1]$ can be defined as usual by the semigroup theory and the following domain characterization holds:

$$X_{2s} := D\left((B+I)^s\right) = \begin{cases} \{u \in H^{2s}(\Omega) : \partial_\nu u = 0 \text{ on } \partial\Omega\}, & s \in (3/4,1] \\ H^{2s}(\Omega), & s \in (0,3/4), \end{cases}$$

(in the case $s = 3/4$, $u \in X_{3/2} \subset H^{3/2}(\Omega)$, and equality does not hold). As usual we equip $D(A)$ with the H^{2s}-norm. For each $s \in (0,1]$, we consider an internally controlled system

$$\begin{cases} {}_C\partial_t^\gamma u + B^s u = f(u(t)) + z, & \text{in } Q := (0,T) \times \Omega, \\ \\ u(0,\cdot) = u_0 & \text{in } \Omega. \end{cases} \tag{6.8}$$

This can be rewritten as the abstract Cauchy problem

$$\begin{cases} {}_C\partial_t^\gamma u(t) + Au(t) = f(u(t)) + \mathbb{B}z(t), & t \in (0,T), \\ u(0) = u_0, \end{cases} \tag{6.9}$$

where $\mathbb{B} = \mathbb{I}$ and $D := \Omega$. Notice that for $\alpha \in (0,2]$, $V_\alpha = D(A^{\alpha/2}) = X_{s\alpha}$. In what follows, we thus let $\widetilde{\alpha} = 0$, $\beta = -1$ and $s \in (0,1]$.

Example 6.1. *One prototype for f is a cubic type of nonlinearity $f = -F'$ associated with the double-well potential*

$$F(u) = c_1 u^4 - c_2 u^2, \ \textit{for } c_2 > c_1 > 0. \tag{6.10}$$

With the choice (6.10), one refers to (6.8) as an internal optimal control problem for the subdiffusive Allen-Cahn (or phase field) type equation. The assumptions (H3), (H4), (H4bis) are satisfied by the cubic nonlinearity provided that $s(3\alpha+1) > \frac{n}{2}$ for $s \in (0,1]$ and $\alpha \in (0,1)$.

[18] By the definition of C_1, C_2, (6.7) implies only additional temporal regularity of the data.

Example 6.2. *When f is a logistic reaction term of the form $ru\left(1-uK^{-1}\right)$ $(r,K>0)$, the associated system (6.9) is an internal control problem for the subdiffusive Fisher-KPP equation. The assumptions (H3), (H4), (H4bis) are satisfied by the Fisher-KPP type logistic source provided that $s\in(0,1]$, $\alpha\in(0,1)$ satisfy the condition $s\left(2\alpha+1\right)>\frac{n}{2}$.*

Example 6.3. *We can also apply our results to the associated optimal control problem for a (subdiffusive) Burger's equation, subject to nonlocal advection or transport. Indeed, let $f\left(u\right):=-u\mathrm{div}(J*u)=-u(\mathcal{G}*u)$, where $\mathcal{G}:=\mathrm{div}(J)$,*

$$\left(J*u\right)\left(x\right)=\int_{\Omega}J\left(x-y\right)u\left(y\right)dy,\ x\in\Omega,$$

for some $J\in W_{loc}^{1,div}\left(\mathbb{R}^n\right):=\left\{J\in L_{loc}^1\left(\mathbb{R}^n\right):\mathcal{G}\in L_{loc}^1\left(\mathbb{R}^n\right)\right\}$. We justify this definition in the one dimensional case, $\Omega=(-L,L)$, $L>0$. Given the Dirac delta function acting δ_x at a point $x\in\Omega$,

$$\delta_x\left[u\right]=\int_{\Omega}u\left(y\right)\delta\left(x-y\right)dy=u\left(x\right),\ x\in\Omega,$$

as a distribution, $\delta_x\in C^k$ and $\partial_x^k\delta_x\left[u\right]=\left(-1\right)^k\delta_x\left[\partial_x u\right]=\left(-1\right)^k\partial_x^k u\left(x\right)$, for every positive integer k. Then, we observe that, for $u\in C_c^1\left(\Omega\right)$ and $v=-\partial_y\left(\delta\left(x-y\right)\right)$, we find that

$$\int_{\Omega}v\left(y\right)u\left(y\right)dy=\int_{\Omega}\partial_y u\left(y\right)\delta\left(x-y\right)dy=\partial_x u\left(x\right).$$

Thus, whenever $u\in C_c^1\left(\Omega\right)$, $\partial_x u\left(x\right)$ occurs as an approximation of the convolution $\mathcal{G}*u$, where $\mathcal{G}=-\partial_y\delta\left(x-y\right)$ (i.e., $f\left(u\right)\approx-u\left(u_x\right)$). Thus, in general we may replace any $\partial_x u\left(x\right)$ with a convolution $\mathcal{G}*u$ to reflect the nonlocal behavior of transport at microscopic levels. The assumptions (H3), (H4)-(H4bis) apply to the nonlocal nonlinearity provided that $s\left(3\alpha+1\right)\geq n$ if $n>2s$; no restrictions are required when $n\leq 2s$.

However, in what follows we will not take a particular choice for $f\left(u\right)$ since many of the technical assumptions (H3), (H4), (H4bis) can be verified directly in applications for such nonlinearities.

Corollary 6.4. *Let $u_0\in V_1=X_s$ and $z\in Z_{\rho,\infty}$, for some $\rho>1/2$. Assume (H3) for some $\alpha\in(0,1)$ and $\beta=-1$. Then (6.8) admits a unique weak solution on $(0,T_{\max})$ such that $u\in Y_{\theta,\alpha}$ with $\theta:=\frac{\gamma}{2}\left(1-\alpha\right)$. The variational equality*

$$\langle{}_C\partial_t^{\gamma}u\left(t\right)+Au\left(t\right)-z,v\rangle_{V_{-1},V_1}=\langle f\left(u\left(t\right)\right),v\rangle_{V_{-1},V_1},$$

holds for any $v\in V_1$, for almost all $t\in(0,T_{\max})$. In particular,

$$u\in W^{1,1+\xi}\left((0,T);V_\alpha\right),\ Au\in L^{1+\xi}\left((0,T);X_s^*\right),\ {}_C\partial_t^{\gamma}u\in L^{1+\xi}\left((0,T);X_s^*\right),$$

for any $T<T_{\max}$, where $\xi\in(0,\frac{\theta}{1-\theta})$.

Proof. By assumption, (H1)-(H2) are automatically satisfied for the operator $\mathbb{B}=\mathbb{I}$ and the initial datum u_0, since $V_1\subset V_\alpha$ and $z\in Z_{\rho,\infty}$. Thus the conclusions of Corollary 3.7 hold. □

The conclusions of Section 4 hold as well provided that f satisfies the corresponding hypotheses (H4)-(H4bis) in that section with a given $\alpha\in(0,1)$ and $\beta=-1$.

In what follows we consider the cost functional J, defined in Section 6, by setting $a_2=0$, $a_1>0$ and $\zeta\geq 0$. Next, the datum z_Q is assumed to belong to $Y_{\theta,0}$, for the same value $\theta = \frac{\gamma}{2}(1-\alpha)$, as in Corollary 6.4. We take the same admissible set Z_{ad}, as defined in (6.4).

Consequently, it follows on account of Theorem 4.8 and the previous considerations of Section 6.1, the following.

Theorem 6.5. *Let $z_* \in Z_{ad}$ be an admissible optimal control and $u_* = \mathcal{S}(z_*)$, the associated state. The necessary optimal condition (6.6) reads*

$$\int_0^T \int_\Omega (w + \zeta z_*)[z - z_*]\, dxdt \geq 0, \text{ for all } z \in Z_{ad}, \tag{6.11}$$

where $T < T_{\max}$.

Remark 6.6. *Notice that since Z_{ad} is a closed convex subset of a Hilbert space, we have in view of [13, Theorem 3.3.5], instead of solving the variational inequality (6.11), we can equivalently find z_* by computing the projection of $-\frac{1}{\zeta}w$ onto the set Z_{ad} with respect to the topology on Z_{ad}. However, this projection maybe challenging to evaluate in general, see for instance [9] for the H^1-case where each projection requires solving a variational inequality itself.*

6.3. **The nonhomogeneous Wentzell-Robin problem for the Laplacian.** Assume that Ω has a Lipschitz continuous boundary. Let $\beta \in L^\infty(\partial\Omega)$ be such that $\beta(x) \geq \beta_0 > 0$ for σ-a.e. $x \in \partial\Omega$ and for some $\beta_0 \in \mathbb{R}$, $\delta \in \{0,1\}$ and

$$\mathbb{H}^{1,\delta}(\overline{\Omega}) := \Big\{U = (u, u|_{\partial\Omega}) :\ u \in H^1(\Omega) \text{ and } \delta u|_{\partial\Omega} \in H^1(\partial\Omega)\Big\},$$

be endowed with the norm

$$\|u\|_{\mathbb{H}^{1,\delta}(\overline{\Omega})} = \begin{cases} \left(\|u\|^2_{H^1(\Omega)} + \|u\|^2_{H^1(\partial\Omega)}\right)^{\frac{1}{2}} & \text{if } \delta = 1 \\ \left(\|u\|^2_{H^1(\Omega)} + \|u\|^2_{H^{\frac{1}{2}}(\partial\Omega)}\right)^{\frac{1}{2}} & \text{if } \delta = 0. \end{cases}$$

Then

$$\mathbb{H}^{1,0}(\overline{\Omega}) \hookrightarrow L^q(\Omega) \times L^q(\partial\Omega), \tag{6.12}$$

with

$$1 \leq q \leq \frac{2(n-1)}{n-2} \text{ if } n > 2 \text{ and } 1 \leq q < \infty \text{ if } n \leq 2, \tag{6.13}$$

and

$$\mathbb{H}^{1,1}(\overline{\Omega}) \hookrightarrow L^q(\Omega) \times L^q(\partial\Omega), \tag{6.14}$$

with

$$1 \leq q \leq \frac{2n}{n-2} \text{ if } n > 2 \text{ and } 1 \leq q < \infty \text{ if } n \leq 2. \tag{6.15}$$

Let $\mathcal{E}_{\delta,W}$ with $D(\mathcal{E}_{\delta,W}) := \mathbb{H}^{1,\delta}(\overline{\Omega})$ be given by

$$\mathcal{E}_{\delta,W}(U,V) := \int_\Omega \nabla u \cdot \nabla v dx + \delta \int_{\partial\Omega} \nabla_\Gamma u \cdot \nabla_\Gamma v d\sigma + \int_{\partial\Omega} \beta(x) u v d\sigma. \tag{6.16}$$

Let $\Delta_{\delta,W}$ be the self-adjoint operator in $L^2(\Omega) \times L^2(\partial\Omega)$ associated with $\mathcal{E}_{\delta,W}$. That is,

$$\begin{cases} D(\Delta_{\delta,W}) = \Big\{U := (u, u|_\Gamma) \in \mathbb{H}^{1,\delta}(\overline{\Omega}) :\ \exists\, F := (f,g) \in L^2(\Omega) \times L^2(\partial\Omega), \\ \qquad\qquad \mathcal{E}_{\delta,W}(U,V) = (f,v)_{L^2(\Omega)} + (g,v)_{L^2(\partial\Omega)}\ \forall\, V := (v, v|_{\partial\Omega}) \in \mathbb{H}^{1,\delta}(\overline{\Omega})\Big\} \\ \Delta_{\delta,W} U = F. \end{cases} \tag{6.17}$$

Then $\Delta_{\delta,W}$ is a realization in $L^2(\Omega)\times L^2(\partial\Omega)$ of $\Big(-\Delta,-\Delta_\Gamma\Big)$ with the generalized Wentzell boundary conditions. More precisely, using an integration by parts arguments, we have that

$$D(\Delta_{\delta,W})=\Big\{(u,u|_\Gamma)\in\mathbb{H}^{1,\delta}(\overline{\Omega}):\Delta u\in L^2(\Omega)\text{ and }\\ -\delta\Delta_\Gamma(u|_{\partial\Omega})+\partial_\nu u+\beta(u|_{\partial\Omega})\in L^2(\partial\Omega)\Big\},$$

and

$$\Delta_{\delta,W}(u,u|_\Gamma)=\Big(-\Delta u,-\delta\Delta_\Gamma(u|_{\partial\Omega})+\partial_\nu u+\beta(u|_{\partial\Omega})\Big).$$

We notice that for $1\le q\le\infty$, the space $L^q(\Omega)\times L^q(\partial\Omega)$ endowed with the norm

$$\|(f,g)\|_{L^q(\Omega)\times L^q(\partial\Omega)}=\begin{cases}\Big(\|f\|^q_{L^q(\Omega)}+\|g\|^q_{L^q(\partial\Omega)}\Big)^{1/q} & \text{if } 1\le q<\infty,\\ \max\{\|f\|_{L^\infty(\Omega)},\|g\|_{L^\infty(\partial\Omega)}\} & \text{if } q=\infty,\end{cases}$$

can be identified with $L^q(\overline{\Omega},\mu)$ where the measure μ on $\overline{\Omega}$ is defined for a measurable set $A\subset\overline{\Omega}$ by $\mu(A)=|\Omega\cap A|+\sigma(\partial\Omega\cap A)$. In addition, we have that the embedding $\mathbb{H}^{1,\delta}(\overline{\Omega})\hookrightarrow L^2(\overline{\Omega},\mu)$ is compact.

For $\overline{\delta}\in\{0,1\}$, let us consider the following semilinear problem:

$$\begin{cases}{}_C\partial_t^\gamma u-\Delta u=f(u(t,x)), & \text{in } Q:=(0,T)\times\Omega,\\ \overline{\delta}_C\partial_t^\gamma u|_{\partial\Omega}-\delta\Delta_\Gamma(u|_{\partial\Omega})+\partial_\nu u+\beta(u|_{\partial\Omega})=z, & \text{in } \Gamma:=(0,T)\times\partial\Omega),\\ u(0,\cdot)=(u_0,v_0) & \text{in } \overline{\Omega}.\end{cases}\tag{6.18}$$

The system (6.18) can be written as the following abstract Cauchy problem

$$\begin{cases}\mathbb{K}_{\overline{\delta}}\left({}_C\partial_t^\gamma U\right)-\Delta_{\delta,W}U=(f(u),z) & \text{in } (0,T)\times(\Omega\times\partial\Omega)\\ U(0,\cdot)=(u_0,v_0) & \text{in } \overline{\Omega},\end{cases}\tag{6.19}$$

where we have identified ${}_C\partial_t^\gamma U$ with $({}_C\partial_t^\gamma u|_\Omega,{}_C\partial_t^\gamma u|_{\partial\Omega})$, and set

$$\mathbb{K}_{\overline{\delta}}:=\begin{pmatrix}1&0\\0&\overline{\delta}\end{pmatrix}.$$

Clearly, the system (6.19) can be rewritten as the abstract Cauchy problem (3.1) when $D:=\partial\Omega$, $A=-\Delta_{\delta,W}$, and

$$\mathbb{B}=\begin{pmatrix}0&0\\0&1\end{pmatrix}\in\mathcal{L}(\{0\}\times L^2(D),L^2(\overline{\Omega},\mu))$$

is a fixed control operator (and so, in that case $\widetilde{\alpha}=0$). The statement of Corollary 3.7 then applies with $\alpha\in(0,1)$, $\theta=\frac{\gamma}{2}(1-\alpha)$, $\widetilde{\alpha}=0$, $\beta=-1$ and $(u_0,v_0)\in\mathbb{H}^{1,\delta}(\overline{\Omega})=V_1$, provided that $(f(u),0)$ is locally Lipschitz in the sense of (H3). All the results of Section 4 are satisfied as well provided that all the corresponding assumptions (H4)-(H4bis) are applied[19] to the vector $(f(u),0)$ (see Examples 6.1, 6.2, 6.3 with $s=1$).

With this setup in mind, we take $a_1=0$, $a_2>0$, $\zeta\ge0$ and consider the datum $z_\Sigma\in Y_{\theta,0}$ (as in Section 6.1). Since $\mathbb{B}=\mathbb{B}^*$, we conclude with the following.

---

[19]For instance, for $f(u)=ru(1-u/K)$, these assumptions are satisfied provided that $(2\alpha+1)>\frac{n}{2}$, a condition which only restricts the value of $\alpha\in(0,1)$ in higher space dimensions $n\ge3$.

Theorem 6.7. *Let $z_* \in Z_{ad}$ be an admissible optimal control and $u_* = \mathcal{S}(z_*)$, the associated state. The necessary optimal condition (6.6) for the problem (6.18) reads*

$$\int_0^T \int_{\partial\Omega} \left(w_{|\partial\Omega} + \zeta z_*\right)[z - z_*]\, d\sigma dt \geq 0, \text{ for all } z \in Z_{ad},$$

where $T < T_{\max}$.

Notice that the comment from Remark 6.6, also applies to Theorem 6.7.

7. **Appendix.** For the sake of completeness, we list here the proofs of some statements that appear in Section 3.

Proof of Lemma 3.2. We employ the application of the contraction mapping principle to the mapping

$$\Psi(u)(t) := S_\gamma(t) u_0 + \int_0^t P_\gamma(t-\tau) f(u(\tau))\, d\tau + \int_0^t P_\gamma(t-\tau)\mathbb{B}z(\tau)\, d\tau$$

in the ball

$$B_T := \left\{u \in C([0,T]; V_\alpha) : \|u\|_{C([0,T];V_\alpha)} \leq R\right\}.$$

We first show that $\Psi : B_{T_*} \to B_{T_*}$, for some $T_* > 0$ and (any sufficiently large) $R > 0$. Indeed, on account of the estimates of Proposition 2.8, we have

$$\begin{aligned} |\Psi(u)(t)|_\alpha &\leq |S_\gamma(t) u_0|_\alpha + \int_0^t |P_\gamma(t-\tau) f(u(\tau))|_\alpha\, d\tau \\ &\quad + \int_0^t |P_\gamma(t-\tau)\mathbb{B}z(\tau)|_\alpha\, d\tau \\ &\lesssim |u_0|_\alpha + \int_0^t (t-\tau)^{\frac{\gamma}{2}(2-\alpha+\beta)-1} |f(u(\tau))|_\beta\, d\tau \\ &\quad + \int_0^t (t-\tau)^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1} |\mathbb{B}z(\tau)|_{\widetilde{\alpha}}\, d\tau \\ &\lesssim |u_0|_\alpha + C_R t^{\frac{\gamma}{2}(2-\alpha+\beta)} \|u\|_{C([0,T];V_\alpha)} + t^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/p} \|z\|_{L^q((0,T_*);L^2(D))} \end{aligned} \tag{7.1}$$

owing to the fact that $\mathbb{B} \in \mathcal{L}\left(L^2(D), V_{\widetilde{\alpha}}\right)$ and $\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha}) - 1 + 1/p > 0$. The preceeding estimate then yields the claim that $\Phi : B_{T_*} \to B_{T_*}$ since $|\Psi(u)(t)|_\alpha \leq R$, for all $0 \leq t \leq T_*$, provided that $T_*, R > 0$ are such that $R \gtrsim 2|u_0|_\alpha$ and

$$C_R T_*^{\frac{\gamma}{2}(2-\alpha+\beta)} R + T_*^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/p} \|z\|_{L^q((0,T_*);L^2(D))} \lesssim |u_0|_\alpha.$$

On the other hand, we can choose a much smaller $T_* > 0$, such that the mapping $\Psi : B_{T_*} \to B_{T_*}$ is a contraction. Indeed, for $u, v \in B_{T_*}$, by the same standard argument, we have

$$\begin{aligned} |\Psi(u)(t) - \Psi(v)(t)|_\alpha &\leq \int_0^t |P_\gamma(t-\tau)(f(u(\tau)) - f(v(\tau)))|_\alpha\, d\tau \\ &\lesssim C_R t^{\frac{\gamma}{2}(2-\alpha+\beta)} \|u - v\|_{C([0,T_*];V_\alpha)} \\ &\lesssim C_R T_*^{\frac{\gamma}{2}(2-\alpha+\beta)} \|u - v\|_{C([0,T_*];V_\alpha)}, \end{aligned} \tag{7.2}$$

provided that T_* is small enough such that $C_R T_*^{\frac{\gamma}{2}(2-\alpha+\beta)} \lesssim 1/2$. Henceforth, the existence of a unique fixed point $u \in B_{T_*}$ for the mapping Ψ is an immediate consequence of the contraction mapping principle.

Next, we verify the sense in which the initial datum $u(0) = u_0$ is satisfied. Recall that S_γ is strongly continuous as a mapping from $V_\alpha \to V_\alpha$. Exploiting the argument of (7.1), we find that

$$\begin{aligned}|u(t) - u_0|_\alpha &\leq |u(t) - S_\gamma(t) u_0|_\alpha + |S_\gamma(t) u_0 - u_0|_\alpha \\ &\leq |S_\gamma(t) u_0 - u_0|_\alpha + \int_0^t |P_\gamma(t-\tau) f(u(\tau))|_\alpha \, d\tau \\ &+ \int_0^t |P_\gamma(t-\tau) \mathbb{B} z(\tau)|_\alpha \, d\tau \\ &\lesssim |S_\gamma(t) u_0 - u_0|_\alpha + C_R t^{\frac{\gamma}{2}(2-\alpha+\beta)} R \\ &+ t^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/p} \|z\|_{L^q((0,T_*);L^2(D))}\end{aligned} \tag{7.3}$$

Therefore, the claim in (3.2) follows by passing to the limit as $t \downarrow 0^+$ in (7.3). The proof is finished. $\square$

Proof of Lemma 3.3. Let $T^\star$ be the time from Lemma 3.2. Fix $\tau > 0$ and consider the space

$$\begin{aligned}\mathbb{K} := \Big\{ v \in C([0, T^\star + \tau]; V_\alpha) \colon v(\cdot, t) = u(\cdot, t) \qquad &\forall\, t \in [0, T^\star], \\ |v(\cdot, t) - u(\cdot, T^\star)|_\alpha \leq R, \ \ \forall\, t \in [T^\star, T^\star + \tau] \Big\}&.\end{aligned}$$

Define the mapping Ψ on $\mathbb{K}$ by

$$\Psi(v)(t) = S_\gamma(t) u_0 + \int_0^t P_\gamma(t-s) f(v(s)) \, ds + \int_0^t P_\gamma(t-\tau) \mathbb{B} z(\tau) \, d\tau.$$

Note that $\mathbb{K}$ when endowed with the norm of $C([0, T^\star + \tau]; V_\alpha)$ is a closed subspace of $C([0, T^\star + \tau]; V_\alpha)$. We show that Ψ has a fixed point in $\mathbb{K}$.

We first show that Ψ maps $\mathbb{K}$ into $\mathbb{K}$. Indeed, let $v \in \mathbb{K}$.

If $t \in [0, T^\star]$, then $v(\cdot, t) = u(\cdot, t)$. Hence $\Psi(v)(t) = \Psi(u)(t) = u(\cdot, t)$ and there is nothing to prove. If $t \in [T^\star, T^\star + \tau]$, then

$$\begin{aligned}&|\Psi(v)(t) - u(T^\star)|_\alpha \\ &\leq |S_\gamma(t) u_0 - S_\gamma(T^\star) u_0|_\alpha \\ &+ \left| \int_0^t P_\gamma(t-s) f(v(s)) \, ds - \int_0^{T_*} P_\gamma(T_* - s) f(u(s)) \, ds \right|_\alpha \\ &+ \int_{T_*}^t |P_\gamma(t-s) \mathbb{B} z(\tau)|_\alpha \, ds \\ \leq\, &|S_\gamma(t) u_0 - S_\gamma(T^\star) u_0|_\alpha \\ &+ \int_0^{T_*} |(P_\gamma(t-s) - P_\gamma(T_* - s)) f(u(s))|_\alpha \, ds \\ &+ \int_{T_*}^t |P_\gamma(t-s) f(v(s))|_\alpha \, ds + \int_{T_*}^t |P_\gamma(t-s) \mathbb{B} z(\tau)|_\alpha \, ds \\ &=: Q_1 + Q_2 + Q_3 + Q_4.\end{aligned}$$

Since for every $T \geq 0$, the mapping $t \mapsto S_\gamma(t) u_0$ belongs to $C([0, T], V_\alpha)$, we can choose $\tau > 0$ small such that for $t \in [T^\star, T^\star + \tau]$, we have

$$Q_1 = |S_\gamma(t) u_0 - S_\gamma(T^\star) u_0|_\alpha \leq \frac{R}{4}. \tag{7.4}$$

Proceeding as in the proof of Lemma 3.2 we can choose $\tau > 0$ small such that for $t \in [T^\star, T^\star + \tau]$, we have

$$Q_3 = \int_{T_*}^{t} \left| P_\gamma (t-s) f(v(s)) \right|_\alpha ds \lesssim C_R \tau^{\frac{\gamma}{2}(2-\alpha+\beta)} R \le \frac{R}{4}, \tag{7.5}$$

$$Q_4 \lesssim \tau^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/p} \left\| z \right\|_{L^q((T_*,T_*+\tau);L^2(D))} \le \frac{R}{4}. \tag{7.6}$$

We next write in view of (2.10),

$$\begin{aligned} Q_2 &= \int_0^{T_*} \left| (P_\gamma (t-s) - P_\gamma (T_* - s)) f(u(s)) \right|_\alpha ds \qquad (7.7)\\ &= \int_0^{T_*} \left| \left(\gamma (t-s)^{\gamma-1} - \gamma (T_*-s)^{\gamma-1} \right) \int_0^\infty \tau \Phi_\gamma (\tau) T(\tau (t-s)^\gamma) d\tau w(s) \right|_\alpha ds \\ &+ \int_0^{T_*} \left| \gamma (T_*-s)^{\gamma-1} \int_0^\infty \tau \Phi_\gamma (\tau) (T(\tau (t-s)^\gamma) - T(\tau (T_*-s)^\gamma)) d\tau w(s) \right|_\alpha ds \\ &=: Q_{21} + Q_{22}. \end{aligned}$$

Noting that $w(s) = f(u(s)) \in V_\beta$ for $s \in [0, T_*]$, and recalling the semigroup estimate (2.15) for $T(t) := \exp(-At)$, we have

$$\left| \left(\gamma (t-s)^{\gamma-1} - \gamma (T_*-s)^{\gamma-1} \right) \int_0^\infty \tau \Phi_\gamma (\tau) T(\tau (t-s)^\gamma) d\tau w(s) \right|_\alpha \to 0$$

in the limit as $t \to T_*$, and there exists a constant $C_R > 0$ (where $|u(s)|_\alpha \le R$, $s \in [0, T_*]$) such that

$$\begin{aligned} &\left| \left(\gamma (t-s)^{\gamma-1} - \gamma (T_*-s)^{\gamma-1} \right) \int_0^\infty \tau \Phi_\gamma (\tau) T(\tau (t-s)^\gamma) d\tau w(s) \right|_\alpha \qquad (7.8)\\ &\lesssim C_R (T_*-s)^{\gamma-1} (t-s)^{-\frac{\gamma}{2}(\alpha-\beta)} \lesssim C_R (T_*-s)^{\frac{\gamma}{2}(2-\alpha+\beta)-1}. \end{aligned}$$

Thus by the Lebesgue Dominated Convergence Theorem[20], we can choose $\tau > 0$ small enough such that for $t \in [T^\star, T^\star + \tau]$, there holds $Q_{21} \le \frac{R}{8}$. A similar argument yields that $Q_{22} \le \frac{R}{8}$ on the interval $[T_*, T_* + \tau]$, for a sufficiently small $\tau > 0$. Henceforth, all the foregoing estimates imply that $|\Psi(v)(t) - u(T^\star)|_\alpha \le R$, for all $t \in [T^\star, T^\star + \tau]$. We have shown that Ψ maps $\mathbb{K}$ into $\mathbb{K}$.

The second step is to show that Ψ is a contraction on $\mathbb{K}$. Let $v, w \in \mathbb{K}$. Then

$$\Psi(v)(t) - \Psi(w)(t) = \int_0^t P_\gamma (t-s) (f(v(s)) - f(w(s))) ds.$$

Once again if $t \in [0, T^\star]$, then

$$|\Psi(v)(t) - \Psi(w)(t)|_\alpha = 0 \le \frac{1}{2} \|v(t) - w(t)\|_{C([0,T_*];V_\alpha)},$$

[20]Note that the right-hand side of (7.8) is integrable.

owing to the fact that $v = w = u$ on $[0, T_*]$. On the other hand, if $t \in [T^\star, T^\star + \tau]$, then proceeding as in the proof of Lemma 3.2, we find

$$\begin{aligned} |\Psi(v)(t) - \Psi(w)(t)|_\alpha &\leq \int_{T_*}^{t} |P_\gamma (t-s) (f(v(s)) - f(w(s)))|_\alpha \, ds \\ &\lesssim C_R \tau^{\frac{\gamma}{2}(2-\alpha+\beta)} \|v(t) - w(t)\|_{C([T_*, T_*+\tau]; V_\alpha)} \\ &\leq \frac{1}{2} \|v(t) - w(t)\|_{C([T_*, T_*+\tau]; V_\alpha)}, \end{aligned} \tag{7.9}$$

provided that $\tau > 0$ is small enough. We deduce once again that Ψ is a contraction on $\mathbb{K}$ so that it has a unique fixed point v on $\mathbb{K}$. The proof is finished. □

Proof of Theorem 3.5. Let $\delta \in (0, T/2)$ be an arbitrarily small number and consider the right-difference

$$Z(t,h) := h^{-1} (u(t+h) - u(t)), \text{ for } h \in (0, \delta] \text{ and } \delta < t \leq T.$$

Notice that $Z(t-h,h)$ coincides with the left-difference. We will derive a uniform estimate for $Z(t,h)$, while we leave the details of the uniform estimate for $Z(t-h,h)$ to the interested reader. For every mild (continuous) solution $u(t) \in V_\alpha$, the continuous function $Z(t,h)$ satisfies

$$\begin{aligned} Z(t,h) &:= h^{-1} (S_\gamma(t+h) - S_\gamma(t)) u_0 \\ &\quad + h^{-1} \int_t^{t+h} P_\gamma(s) f(u(t+h-s)) \, ds \\ &\quad + h^{-1} \int_0^t P_\gamma(t-s) (f(u(s+h)) - f(u(s))) \, ds \\ &\quad + h^{-1} \int_t^{t+h} P_\gamma(s) \mathbb{B} z(t+h-s) \, ds \\ &\quad + h^{-1} \int_0^t P_\gamma(s) \mathbb{B} (z(t+s+h) - z(t-s)) \, ds \\ &\quad =: Z_1 + Z_2 + Z_3 + Z_4 + Z_5. \end{aligned} \tag{7.10}$$

We have $|Z_1(t,h)|_\alpha \to 0$, as $h \downarrow 0^+$ uniformly in $t \in [\delta, T]$ since $S_\gamma(t)$ is analytic for $t \geq \delta > 0$ and $S_\gamma^{'}(t) u_0 = P_\gamma(t) A u_0$, $u_0 \in D(A)$. Hence, we can pick a suffiiciently small $h_0 \leq \delta$ such that, for all $0 < h \leq h_0$ and $\delta \leq t \leq T$,

$$|Z_1(t,h)|_\alpha \leq 1 + |P_\gamma(t) A u_0|_\alpha \lesssim 1 + t^{\gamma/2(2-\alpha+\beta)-1} |A u_0|_\beta \leq C_T t^{\gamma/2(2-\alpha+\beta)-1},$$

for some positive constant C_T that depends clearly on $u_0 \in V_{\beta+2}$ but is independent of t, h and δ. Next, on account of Proposition 2.8, we estimate

$$\begin{aligned} |Z_2(t,h)|_\alpha &\lesssim h^{-1} \int_t^{t+h} s^{\frac{\gamma}{2}(2-\alpha+\beta)-1} |f(u(t+h-s))|_\beta \, ds \\ &\lesssim C_R h^{-1} \left((t+h)^{\frac{\gamma}{2}(2-\alpha+\beta)} - t^{\frac{\gamma}{2}(2-\alpha+\beta)} \right) \\ &\leq C C_R t^{\frac{\gamma}{2}(2-\alpha+\beta)-1}, \end{aligned} \tag{7.11}$$

owing to the fact that $(t+h)^r - t^r \leq h t^{r-1}$ for all $h, t > 0$, and any $r > 0$. Here (and below), the constant $C > 0$ is independent of h, t and δ. Moreover, since f is a (locally) Lipschitz mapping from $V_\alpha \to V_\beta$, we bound

$$\begin{aligned} |Z_3(t,h)|_\alpha &\lesssim h^{-1}\int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)-1} |f(u(s+h)) - f(u(s))|_\beta \, ds \\ &\leq CC_R \int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)-1} |Z(s,h)|_\alpha \, ds, \end{aligned} \tag{7.12}$$

for $0 < h \leq \delta$, and for all $\delta \leq t \leq T$. Similarly, we deduce in a similar fashion to (7.1), that

$$\begin{aligned} |Z_4(t,h)|_\alpha &\lesssim h^{-1}\int_t^{t+h} s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1} |\mathbb{B}z(t+h-s)|_{\widetilde{\alpha}} \, ds \\ &\lesssim h^{-1} \|z\|_{L^q((0,T);L^2(D))} \left[(t+h)^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/p} - t^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/p}\right] \\ &\leq C \|z\|_{L^q((0,T);L^2(D))} t^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-2+1/p}, \end{aligned}$$

recalling that $\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha}) - 1 + 1/p > 0$, with $1/p + 1/q = 1$. Finally, we find that

$$\begin{aligned} |Z_5(t,h)|_\alpha &\lesssim h^{-1}\int_0^t s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1} |\mathbb{B}(z(t+s+h) - z(t-s))|_{\widetilde{\alpha}} \, ds \\ &\lesssim \int_0^t s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1} \left| \frac{z(t-s+h) - z(t-s)}{h} \right|_0 ds \\ &\lesssim \int_0^t s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1} \|\partial_t z(t-s)\|_{L^2} \, ds \\ &\lesssim \int_0^t s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1} (t-s)^{\rho-1} \, ds \\ &\lesssim t^{\left(\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})+\rho\right)-1}, \end{aligned} \tag{7.13}$$

by assumption (3.3). Collecting the uniform estimates for the right and left differences, one arrives at the following two inequalities:

$$|Z(t,h)|_\alpha \leq C_{T,R} t^{\theta-1} + C_R \int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)-1} |Z(s,h)|_\alpha \, ds,$$

$$|Z(t-h,h)|_\alpha \leq C_{T,R} t^{\theta-1} + C_R \int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)-1} |Z(s-h,h)|_\alpha \, ds,$$

where, by (H1),

$$\theta := \min\left\{ \frac{\gamma}{2}(2-\alpha+\widetilde{\alpha}) - \frac{1}{q}, \frac{\gamma}{2}(2-\alpha+\beta), \ \frac{\gamma}{2}(2-\alpha+\widetilde{\alpha}) + \rho \right\} > 0. \tag{7.14}$$

We can now apply the Gronwall Lemma 2.7 to infer the existence of a constant $C > 0$, independent of h, δ and t, such that

$$|Z(t,h)|_\alpha \leq Ct^{\theta-1}, \ \ |Z(t-h,h)|_\alpha \leq Ct^{\theta-1}, \tag{7.15}$$

for $0 < h \leq h_0 \leq \delta$, and for all $t \in [\delta, T]$. We can now pass to the limit as $h \downarrow 0^+$ in the limsup and liminf sense in (7.15), to deduce that both lower Dini derivatives $\partial_+ u(t)$, $\partial_- u(t)$ and both upper Dini derivatives $\partial^+ u(t)$, $\partial^- u(t)$ are bounded (as functions with values in V_α) by $Ct^{\theta-1}$ for all $\delta \leq t \leq T$. Since $\delta > 0$ was arbitrary, all four Dini derivatives are bounded (and thus finite) in the range for $0 < t \leq T$. By application of the celebrated theorem of Denjoy-Young-Saks (see [17, Chapter IV, Theorem 4.4]), the continuous integral solution $u : [0,T] \to V_\alpha$ is differentiable for almost all $0 < t \leq T$, and that all four Dini derivatives are equal to $\partial_t u(t)$ on the set $t \in (0,T] \setminus E$ (where E is a null set of Lebesegue measure; in fact E is a set of first category, see [17, Chapter IV, Theorem 4.7]). In particular, this yields the fact that

$$\left|\partial_t u\left(t\right)\right|_{\alpha} \leq Ct^{\theta-1},$$

for allmost all $0 < t \leq T$, and the regularity (3.4) follows. Finally, the last conclusion of the thereom is an immediate consequence of Proposition 2.3. The proof of the theorem is finished. □

Proof of Proposition 4.6. (i) Owing to (4.31), we can work with the transformed equation (4.30) for the couple (p,g). Briefly, since $p_0 = 0$, the contraction mapping principle can be applied to the operator

$$\Upsilon\left(p\right)\left(t\right) := \int_0^t P_\gamma\left(t-\tau\right) f^{'}\left(u\left(\tau\right)\right) p\left(\tau\right) d\tau + \int_0^t P_\gamma\left(t-\tau\right) g\left(\tau\right) d\tau$$

in the ball

$$Q_{T_*} := \left\{ p \in C\left(\left[0,T_*\right]; V_\alpha\right) : \left\|p\right\|_{C([0,T];V_\alpha)} \leq M \right\}.$$

It turns out that one can choose a suffiiciently small time $T_* \leq T$, and a sufficiently large $M > 0$, where

$$C_R T_*^{\frac{\gamma}{2}(2-\alpha+\beta)} \leq \frac{1}{2},\ 2T_*^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/\overline{q}} \left\|g\right\|_{L^q((0,T_*);V_{\widetilde{\alpha}})} \leq M,$$

($\overline{q}$ is conjugate to q) so that Υ is a strict contraction on Q_{T_*}. Indeed, owing to (H4), one has for each $t \in (0,T_*)$,

$$\left|\Upsilon\left(p\right)\left(t\right)\right|_{\alpha} \leq C_R t^{\frac{\gamma}{2}(2-\alpha+\beta)} \left\|p\right\|_{C([0,T];V_\alpha)} + t^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/\overline{p}} \left\|g\right\|_{L^q((0,T_*);V_{\widetilde{\alpha}})}$$

and

$$\left|\Upsilon\left(p_1\right)\left(t\right) - \Upsilon\left(p_2\right)\left(t\right)\right|_{\alpha} \leq C_R t^{\frac{\gamma}{2}(2-\alpha+\beta)} \left\|p_1 - p_2\right\|_{C([0,T_*];V_\alpha)}.$$

As in the proof of Theorem 3.4, the locally defined solution p can be extended to the whole interval $(0,T)$, for as long as an optimal solution $u := u_*$ exists on $(0,T)$. Moreover, owing to (H4), one has $\xi := f^{'}\left(u\right) p \in C([0,T];V_\beta)$, which in turn implies

$$\left[\left|A\left(P_\gamma * g\right)\left(t\right)\right|_{\widetilde{\alpha}-\delta} + \left|A\left(P_\gamma * \xi\right)\left(t\right)\right|_{\widetilde{\alpha}-\delta}\right] \lesssim t^{\frac{\gamma\delta}{2}-\frac{1}{q}} \left(\left\|g\right\|_{L^q((0,T);V_{\widetilde{\alpha}})} + \left\|\xi\right\|_{L^q((0,T);V_{\widetilde{\alpha}})}\right)$$

since $g \in L^q\left((0,T);V_{\widetilde{\alpha}}\right)$ and $\widetilde{\alpha} = \beta$ (this latter condition is chosen for the sake of simplicity; see Section 3). The rest follows analogously to the proof of Theorem 3.2.

For the proof of (ii), as in the proof of Theorem 3.5, let

$$V\left(t,h\right) := h^{-1}\left(p\left(t+h\right) - p\left(t\right)\right),\ \text{for } h > 0 \text{ and } 0 < t \leq T.$$

For every mild continuous solution $p\left(t\right) \in V_\alpha$, $V\left(t,h\right)$ satisfies

$$\begin{aligned} V\left(t,h\right) :=\ & h^{-1}\int_t^{t+h} P_\gamma\left(s\right)\xi\left(t+h-s\right) ds \\ & + h^{-1}\int_0^t P_\gamma\left(t-s\right)\left(\xi\left(s+h\right) - \xi\left(s\right)\right) ds \\ & + h^{-1}\int_t^{t+h} P_\gamma\left(s\right) g\left(t+h-s\right) ds \\ & + h^{-1}\int_0^t P_\gamma\left(s\right)\left(g\left(t+s+h\right) - g\left(t-s\right)\right) ds \\ & =: V_1 + V_2 + V_3 + V_4. \end{aligned}$$

As before, we estimate by employing (H4)-(H4bis),

$$\begin{aligned}
|V_1(t,h)|_\alpha &\lesssim h^{-1}\int_t^{t+h} s^{\frac{\gamma}{2}(2-\alpha+\beta)-1}|\xi(t+h-s)|_\beta\, ds\\
&\le C_{R,M}h^{-1}\left((t+h)^{\frac{\gamma}{2}(2-\alpha+\beta)}-t^{\frac{\gamma}{2}(2-\alpha+\beta)}\right)\\
&\le C_{R,M}t^{\frac{\gamma}{2}(2-\alpha+\beta)-1}.
\end{aligned}$$

and

$$\begin{aligned}
|V_2(t,h)|_\alpha &\lesssim \int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)-1}\left|f'(u(s+h))V(s,h)\right|_\beta ds\\
&+h^{-1}\int_0^t P_\gamma(t-s)(f'(u(s+h))-f'(u(s)))p(s)\,ds\\
&\le C_R\int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)}|Z(s,h)|_\alpha|p(s)|_\alpha\, ds\\
&+C_R\int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)-1}|V(s,h)|_\alpha\, ds\\
&\lesssim \int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)}s^{\theta-1}ds+\int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)}|V(s,h)|_\alpha\, ds\\
&\lesssim t^{\frac{\gamma}{2}(2-\alpha+\beta)+\theta-1}+\int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)}|V(s,h)|_\alpha\, ds,
\end{aligned}$$

as well as,

$$\begin{aligned}
|V_3(t,h)|_\alpha &\lesssim h^{-1}\int_t^{t+h} s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1}|g(t+h-s)|_{\widetilde{\alpha}}\, ds\\
&\lesssim h^{-1}\|g\|_{L^q((0,T);V_\sim)}[(t+h)^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/\overline{q}}-t^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/\overline{q}}]\\
&\le C\|g\|_{L^q((0,T);V_\sim)}t^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-2+1/\overline{q}},
\end{aligned}$$

recalling that $\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1+1/\overline{q}>0$, with $1/\overline{q}+1/q=1$. Finally, we find that

$$\begin{aligned}
|V_4(t,h)|_\alpha &\lesssim h^{-1}\int_0^t s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1}|g(t+s+h)-g(t-s)|_{\widetilde{\alpha}}\, ds\\
&\lesssim \int_0^t s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1}|\partial_t g(t-s)|_{\widetilde{\alpha}}\, ds\\
&\lesssim \int_0^t s^{\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})-1}(t-s)^{\rho-1}\, ds\\
&\lesssim t^{(\frac{\gamma}{2}(2-\alpha+\widetilde{\alpha})+\rho)-1},
\end{aligned}$$

since $g\in\widetilde{Z}_{\rho,\infty}$. Collecting these estimates, for the same value $\theta>0$ as defined earlier, we obtain

$$|V(t,h)|_\alpha \lesssim t^{\theta-1}+\int_0^t (t-s)^{\frac{\gamma}{2}(2-\alpha+\beta)-1}|V(s,h)|_\alpha\, ds.$$

This implies by application of the Gronwall Lemma 2.7,

$$|V(t,h)|_\alpha \le Ct^{\theta-1}.$$

A similar estimate applies to the left-difference $V(t-h,h)$. The constant on the right hand side is independent of h, so one can pass to the limit once again as $h \downarrow 0^+$. The proof is finished. $\square$

REFERENCES

[1] O. P. Agrawal, Fractional variational calculus in terms of Riesz fractional derivatives, *J. Phys. A*, **40** (2007), 6287–6303.

[2] O. P. Agrawal, A general finite element formulation for fractional variational problems, *J. Math. Anal. Appl.*, **337** (2008), 1–12.

[3] E. Alvarez, C. G. Gal, V. Keyantuo and M. Warma, Well-posedness results for a class of semi-linear super-diffusive equations, *Nonlinear Anal.*, **181** (2019), 24–61.

[4] H. Antil, H. C. Elman, A. Onwunta and D. Verma, Novel deep neural networks for solving bayesian statistical inverse problems, *arXiv preprint arXiv:2102.03974*, 2021.

[5] H. Antil, R. Khatri, R. Lö hner and D. Verma, Fractional deep neural network via constrained optimization, *Machine Learning: Science and Technology*, **2** (2020), 015003.

[6] H. Antil, R. Khatri and M. Warma, External optimal control of nonlocal PDEs, *Inverse Problems*, **35** (2019), 084003, 35 pp.

[7] H. Antil, D. P. Kouri, M.-D. Lacasse and D. Ridzal, *Frontiers in PDE-Constrained Optimization*, Springer, New York, 2018.

[8] H. Antil, D. P. Kouri and D. Ridzal, ALESQP: An augmented Lagrangian equality-constrained sqp method for optimization with general constraints, Submitted to SIOPT, 2021.

[9] H. Antil, R. Nochetto and P. Venegas, Controlling the Kelvin force: Basic strategies and applications to magnetic drug targeting, *Optim. Eng.*, **19** (2018), 559–589.

[10] H. Antil, E. Otárola and A. J. Salgado, A space-time fractional optimal control problem: Analysis and discretization, *SIAM J. Control Optim.*, **54** (2016), 1295–1328.

[11] H. Antil, D. Verma and M. Warma, External optimal control of fractional parabolic PDEs, *ESAIM Control Optim. Calc. Var.*, **26** (2020), Paper No. 20, 33 pp.

[12] H. Antil and M. Warma, Optimal control of fractional semilinear PDEs, *ESAIM Control Optim. Calc. Var.* , **26** (2020), Paper No. 5, 30pp.

[13] H. Attouch, G. Buttazzo and G. Michaille, *Variational Analysis in Sobolev and BV Spaces*, 2^{nd} edition, Mathematical Optimization Society, Philadelphia, PA, 2014.

[14] E. Barkai, R. Metzler and J. Klafter, From continuous time random walks to the fractional F okker-Planck equation, *Phys. Rev. E*, **61** (2000), 132–138.

[15] E. G. Bazhlekova, *Fractional Evolution Equations in Banach Spaces*, Technische Universiteit Eindhoven, Eindhoven, 2001. Eindhoven University of Technology, Eindhoven, 2001.

[16] P. Blanchard and E. Bruning, *Distributions, Hilbert Space Operators, Variational Methods, and Applications in Quantum Physics*, volume 69 of *Progress in Mathematical Physics*. 2^{nd} edition, Birkhäuser, 2015.

[17] A. Bruckner, *Differentiation of Real Functions*, Lecture Notes in Mathematics 659, Springer, Berlin, 1978.

[18] P. M. d. Carvalho Neto, *Fractional Differential Equations: A Novel Study of Local and Global Solutions in Banach Spaces*, PhD thesis, Universidade de S ão Paulo, 2013.

[19] E. Casas, Boundary control of semilinear elliptic equations with pointwise state constraints, *SIAM J. Control Optim.*, **31** (1993), 993–1006.

[20] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, *SIAM J. Control Optim.*, **35** (1997), 1297–1327.

[21] J. Diestel and J. J. Uhl, Jr, *Vector Measures*, American Mathematical Society, Providence, R. I., 1977.

[22] K. Diethelm, *The Analysis of Fractional Differential Equations*, Springer-Verlag, Berlin, 2010.

[23] S. Elagan, On the invalidity of semigroup property for the mittag-leffler function with two parameters, *J. Egyptian Math. Soc.*, **24** (2016), 200–203.

[24] C. M. Elliott and S. Larsson, Error estimates with smooth and nonsmooth data for a finite element method for the cahn-hilliard equation, *Math. Comp.*, **58** (1992), 603–630.

[25] C. G. Gal and M. Warma, On some degenerate non-local parabolic equation associated with the fractional p-Laplacian, *Dyn. Partial Differ. Equ.*, **14** (2017), 47–77.

[26] C. G. Gal and M. Warma, *Fractional-in-Time Semilinear Parabolic Equations and Applications*, Mathématiques & Applications (Berlin) [Mathematics & Applications], 84. Springer, Cham, 2020.

[27] M. Hinze, R. Pinnau, M. Ulbrich and S. Ulbrich, *Optimization with PDE Cnstraints*, Mathematical Modelling: Theory and Applications, 23. Springer, New York, 2009.

[28] K. Ito and K. Kunisch, *Lagrange Multiplier Approach to Variational Problems and Applications*, Advances in Design and Control, 15. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[29] B. Jin, B. Li and Z. Zhou, Pointwise-in-time error estimates for an optimal control problem with subdiffusion constraint, *IMA J. Numer. Anal.*, **40** (2020), 377–404.

[30] R. Metzler and J. Klafter, The restaurant at the end of the random walk: Recent developments in the description of anomalous transport by fractional dynamics, *J. Phys.*, **37** (2004), 161–208.

[31] S. G. Samko, A. A. Kilbas and O. I. Marichev, *Fractional Integrals and Derivatives*, Gordon and Breach Science Publishers, Yverdon, 1993. Theory and applications, Edited and with a foreword by S. M. Nikolskiĭ, Translated from the 1987 Russian original, Revised by the authors.

[32] Y. Shin, J. Darbon and G. E. Karniadakis, A caputo fractional derivative-based algorithm for optimization, *arXiv preprint arXiv:2104.02259*, 2021.

[33] T. Tang, H. Yu and T. Zhou, On energy dissipation theory and numerical stability for time-fractional phase-field equations, *SIAM J. Sci. Comput.*, **41** (2019), A3757–A3778.

[34] F. Tröltzsch, *Optimal Control of Partial Differential Equations*, Theory, methods and applications. Translated from the 2005 German original by Jürgen Sprekels. Graduate Studies in Mathematics, 112. American Mathematical Society, Providence, RI, 2010.

[35] M. Warma, The Robin and Wentzell-R obin Laplacians on Lipschitz domains, *Semigroup Forum*, **73** (2006), 10–30.

[36] M. Warma, Approximate controllability from the exterior of space-time fractional diffusive equations, *SIAM J. Control Optim.*, **57** (2019), 2037–2063.

Received for publication October 2021; early access January 2022.

E-mail address: hantil@gmu.edu
E-mail address: cgal@fiu.edu
E-mail address: mwarma@gmu.edu