FRACTIONAL FLOWS DRIVEN BY SUBDIFFERENTIALS IN HILBERT SPACES

GORO AKAGI

Dedicated to Professor Mitsuharu Ôtani on the occasion of his 70th birthday

Abstract. This paper presents an abstract theory on well-posedness for time-fractional evolution equations governed by subdifferential operators in Hilbert spaces. A proof relies on a regularization argument based on maximal monotonicity of time-fractional differential operators as well as energy estimates based on a *nonlocal chain-rule formula* for subdifferentials. Moreover, it will be extended to a Lipschitz perturbation problem. These abstract results will be also applied to time-fractional nonlinear PDEs such as time-fractional porous medium, fast diffusion, p-Laplace parabolic, Allen-Cahn equations.

1. Introduction

1.1. **Time-fractional derivatives and PDEs.** A notion of *fractional derivative* already appeared in a letter of Leibniz to l'Hôpital in 1695 and afterword several notions of fractional derivative were proposed by Riemann, Liouville, Riesz, Caputo and so on. In particular, notions of fractional derivative were also employed during the last decade or two to improve physical models to cover various phenomena beyond the scope of classical theories in Physics. Among those, time-fractional derivatives are particularly attracting much attention in the study of *anomalous diffusion*, in which the MSD (Mean-Squared Displacement) $\langle (x(t)-x(0))^2\rangle$ of randomly moving particles exhibits a nonlinear growth in time t, and hence, the diffusion coefficient (= MSD$/t$) cannot be determined as a constant. In [25] (see also [35]), Fokker-Planck equations including time-fractional derivatives are derived from the so-called CTRW (Continuous-Time Random Walk, see [37]), which is a stochastic process and in which each jump-length Δx and waiting-time Δt of each particle are randomly determined by certain probability density functions (e.g., Brownian motion is reproduced by determining Δx and Δt by means of normal and exponential distributions). For instance, if Δx and Δt are determined by a normal distribution and a power distribution (of power α), respectively, then

2010 *Mathematics Subject Classification. Primary*: 35R11; *Secondary*: 34K37, 47J35.

Key words and phrases. Evolution equation, Riemann-Liouville and Caputo fractional derivatives, subdifferential.

Acknowledgment. GA is supported by JSPS KAKENHI Grant Number JP18K18715, JP16H03946, JP16K05199, JP17H01095 and by the Alexander von Humboldt Foundation and by the Carl Friedrich von Siemens Foundation.

evolution of the (density) distribution $u(x,t)$ of particles is described in terms of a time-fractional partial differential equation (PDE for short),

$$\partial_t^\alpha \left[u(x,\cdot) - u_0(x)\right](t) - \Delta u(x,t) = 0,$$

where u_0 is an initial distribution and ∂_t^α is the *Riemann-Liouville fractional derivative* defined by

$$\partial_t^\alpha f(t) = {}_0D_t^\alpha f(t) := \frac{1}{\Gamma(1-\alpha)} \frac{\mathrm{d}}{\mathrm{d}t} \int_0^t (t-s)^{-\alpha} f(s)\,\mathrm{d}s, \quad t > 0, \quad 0 < \alpha < 1.$$

Here we also remark that $\partial_t^\alpha[u(x,\cdot) - u_0(x)]$ coincides with the Caputo derivative ${}^C D_t^\alpha u(x,\cdot)$ for smooth $u(x,\cdot)$ (see Remark 3.1 below). Therefore such linear PDEs involving time-fractional derivatives have been studied vigorously so far (see, e.g., [43, 28], [20, 33, 34] and references therein).

1.2. **Nonlinear PDEs involving time-fractional derivatives.** Time-fractional PDEs have been studied mostly based on linear technique such as Laplace and Fourier transforms; indeed, they are two of the few effective methods to handle fractional derivatives; however, they are not always effective for nonlinear problems (e.g., degenerate and singular diffusion equations). On the other hand, studies on time-fractional PDEs are recently extending to nonlinear problems. Here, let us give a couple of typical examples.

The dynamics of fluids in unsaturated non-swelling soils is often described by means of the *Richards equation*,

$$\partial_t u(x,t) = \operatorname{div}\left(C(u)\nabla u(x,t)\right),$$

where $u = u(x,t)$ denotes the local volume fraction of water and $C(u) \geq 0$ is the diffusivity depending on u. The Richards equation is derived from a conservation law as well as Darcy's law, and the nonlinearity residing in the soil-water diffusivity $C(u)$ arises from *water retention curves*, which characterize each medium (e.g., soil, sand, clay, silt) and are determined by experiments. On the other hand, anomalous diffusion is also observed in experiments of moisture dispersion in building materials (see [48, 31, 18, 45]). However, a nonlinear scaling $\langle x^2 \rangle \sim t^\alpha$ of moisture dispersion into building materials is not reproduced by means of the Richards equation. In [19, 3], to bridge a gap, the following *time-fractional Richards equation* is introduced:

$$\partial_t^\alpha \left[u(x,\cdot) - u_0(x)\right](t) = \operatorname{div}\left(C(u)\nabla u(x,t)\right),$$

where $0 < \alpha < 1$ and $u_0 := u|_{t=0}$.

Moreover, in [40, 41, 42], *time-fractional porous medium equations*, where $C(u)$ is a power of u, are also considered. In these papers, numerical schemes for nonlinear PDEs involving time-fractional derivatives are proposed and applied to the equations. Furthermore, long-time behaviors of solutions for such time-fractional nonlinear PDEs are investigated in [52, 50, 17, 1], where existence and regularity of solutions are simply assumed. Theoretical analysis on nonlinear PDEs involving time-fractional derivatives has just begun and has not yet been fully pursued, and there still remain open questions on fundamental issues such as well-posedness for fractional variants of degenerate or singular parabolic PDEs (e.g., time-fractional porous medium equation). The main purpose of this paper is to present a general theory which is concerned with an abstract Cauchy problem for nonlocal gradient flows and guarantees well-posedness for a wide class of time-fractional (possibly, degenerate or singular) parabolic PDEs.

Finally, we also refer the reader to [38, 49] for an approach based on *viscosity solutions* to time-fractional nonlinear PDEs.

1.3. **Abstract setting.** In this paper, we shall present an abstract theory on time-fractional evolution equations, which include only most significant common features of time-fractional PDEs mentioned above. To this end, let us first give an abstract setting and formulate a problem. We refer the reader to Section 6 below, concerning how to apply the following abstract theory to concrete time-fractional PDEs.

Let H be a real Hilbert space and let $\varphi : H \to [0,+\infty]$ be a proper (i.e., $\varphi \not\equiv +\infty$) lower semicontinuous convex functional with *effective domain* $D(\varphi) := \{w \in H \colon \varphi(w) < +\infty\}$. Here, φ is supposed to be non-negative. However, it is not restrictive in view of affine boundedness from below for each lower semicontinuous convex functionals (see, e.g., [4, Proposition]). We shall discuss existence, uniqueness and continuous dependence (on prescribed data) of (strong) solutions $u : (0,T) \to H$ to the equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u - u_0)\right](t) + \partial\varphi(u(t)) \ni f(t) \ \text{ in } H, \quad 0 < t < T, \tag{1.1}$$

where $T > 0$, $u_0 \in H$ and $f : (0,T) \to H$ are given, the convolution $k * (u - u_0)$ with a kernel $k \in L^1_{\mathrm{loc}}([0,\infty))$ is defined by

$$(k * w)(t) := \int_0^t k(t-s)w(s)\,\mathrm{d}s \quad \text{ for } \ w \in L^1_{\mathrm{loc}}([0,\infty); H), \ t > 0$$

and $\partial\varphi : H \to 2^H$ is the *subdifferential operator* of φ, that is, for $w \in D(\varphi)$,

$$\partial\varphi(w) := \{\xi \in H \colon \varphi(z) - \varphi(w) \geq (\xi, z - w)_H \ \text{ for all } z \in H\}, \tag{1.2}$$

with domain $D(\partial\varphi) := \{w \in D(\varphi) \colon \partial\varphi(w) \neq \emptyset\}$. It is well known that every subdifferential operator is maximal monotone in H (see [8]). Throughout this paper, keeping Riemann-Liouville derivative in our mind, we assume that

(K): The kernel $k \in L^1_{\mathrm{loc}}([0,+\infty))$ is nonnegative and nonincreasing. Moreover, there exists a nonnegative and nonincreasing kernel $\ell \in L^1_{\mathrm{loc}}([0,+\infty))$ such that

$$k * \ell \equiv 1.$$

A typical example of (k,ℓ) satisfying (K) is the Riemann-Liouville kernel,

$$k_\beta(t) = \frac{t^{-\beta}}{\Gamma(1-\beta)}, \quad \ell_\beta(t) = \frac{t^{\beta-1}}{\Gamma(\beta)}, \quad 0 < \beta < 1. \tag{1.3}$$

Then the nonlocal derivative $(\mathrm{d}/\mathrm{d}t)[k_\beta * (u - u_0)]$ coincides with the *Riemann-Liouville derivative* $\partial_t^\beta(u - u_0)$ of $u - u_0$ of the order β (see also Remark 3.2).

1.4. **Related results.** By convolving (1.1) with ℓ and by using (K), equation (1.1) is reduced into a *nonlinear Volterra equation*,

$$u(t) + \big(\ell * \xi\big)(t) = (\ell * f)(t) + u_0, \quad \xi(t) \in \partial\varphi(u(t)), \quad 0 < t < T.$$

Nonlinear Volterra equations (in Hilbert and Banach spaces) were already studied in 1960's (see, e.g., [36]) in the following general form,

$$u(t) + \big(a * \xi\big)(t) = g(t), \quad \xi(t) \in A(u(t)), \quad 0 < t < T, \tag{1.4}$$

where $A : X \to X$ is a nonlinear operator in a Hilbert or Banach space X, $g : (0,T) \to X$ is given and $a : [0,+\infty) \to [0,+\infty]$ is a kernel function. In [7, 6, 5], Barbu studied an abstract nonlinear Volterra equation (1.4), which arises in the study of mechanical systems with memory effects, under assumptions that a is of class $W^{1,1}_{\mathrm{loc}}([0,+\infty))$ (in particular, a is differentiable and finite at $t = 0$) and positive, $g \in W^{1,2}(0,T;X)$ and $A = \partial\varphi$ in a Hilbert space X. Due to the regularity of the kernel, by differentiating both sides of (1.4) in time, we have

$$u'(t) + a(0)\xi(t) + \big(a' * \xi\big)(t) = g'(t), \quad \xi(t) \in A(u(t)), \quad 0 < t < T,$$

which can be regarded as a nonlocal (in time) perturbation problem of a (local) nonlinear evolution equation. We also refer the reader to [15, 22, 32, 27, 11, 2].

Concerning singular kernels, Clément and Nohel [12, Theorem 3.1] studied (1.4) for *completely positive kernels* $a \in L^1_{\mathrm{loc}}([0,+\infty))$ and proved existence of a *generalized solution*, that is, a weak limit of certain class of approximate solutions. Moreover, the abstract theory is also applied to a couple of nonlocal (in time) PDEs (see also [13]). The literature [23] may be also related to the present paper (see also [26, Theorems 2 and 3]). Indeed, Theorem 1 of [23] is concerned with existence and uniqueness of strong solutions for (1.4) under (K) and some assumptions, which particularly require g of (1.4) is sufficiently smooth, e.g., $g \in W^{1,1}_{\mathrm{loc}}([0,+\infty);X)$ and $g' \in BV_{\mathrm{loc}}([0,+\infty);X)$. Moreover, evolution equations including nonlocal derivatives (e.g., Riemann-Liouville derivative) are also studied in [10, 24, 14] by finding out that nonlocal differential operators are m-accretive in Bochner spaces under (K). On the other hand, most of existence results are established for generalized solutions and, to the best of author's knowledge, there had been no result corresponding to strong (in time) solutions for a long while.

In [56], under a Gelfand triplet setting,

$$V \hookrightarrow H \equiv H^* \hookrightarrow V^*,$$

where V and V^* are a Hilbert space and its dual space, respectively, and H is a pivot Hilbert space, Zacher proved existence and uniqueness of strong (in time) solutions to the abstract equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\big([k * (u - u_0)]\,(t), v\big)_H + b(t, u(t), v) = \langle f(t), v\rangle_V \ \text{ for all } v \in V,$$

where $(\cdot,\cdot)_H$ and $\langle\cdot,\cdot\rangle_V$ stand for an inner product of H and a duality pairing between V and V^*, respectively, k is a completely positive kernel, $f : (0,T) \to V^*$ and $u_0 \in H$ are given and $b(t,\cdot,\cdot)$ is a (time-dependent) bounded coercive bilinear form defined on V, by employing two important devices: m-accretivity of nonlocal differential operators and nonlocal energy identities (see, e.g., Lemma 3.3 below). In [51], convergence of solutions of nonlocal (in time) gradient flows in Euclidean spaces $\mathbb{R}^d$ to equilibria is studied under (K) and additional assumptions (which may exclude Riemann-Liouville and Caputo derivatives). We also refer the reader to [55], [30], [52, 50, 53] on various properties of solutions such as boundedness, regularity, decaying property, blow-up phenomena and so on. In [47], a nonlinear variant of the above equation is studied and further extended to stochastic PDEs.

1.5. **Construction of the paper.** Section 2 presents main results. We shall first give a definition of strong solutions to (1.1) and then state a theorem on existence and uniqueness of strong solutions to (1.1) and continuous dependence on prescribed data along with regularity of strong solutions and energy inequalities for $u_0 \in D(\varphi)$ (see Theorem 2.3). It can be regarded as a fractional variant of the celebrated Brézis-Kōmura theory for gradient flows in Hilbert spaces (see [8] and [29]). Moreover, we shall give a proposition on initial condition (see Proposition 2.5). Indeed, it is delicate in which sense initial condition is fulfilled by strong solutions to (1.1); it still holds true in a classical sense under a certain additional assumption on ℓ, which corresponds to the case that the order β of the Riemann-Liouville fractional derivative is greater than $1/2$. Let us emphasize that, even for $\beta \leq 1/2$, a strong solution is uniquely determined by each initial datum according to Theorem 2.3, and therefore, there is a one-to-one correspondence between initial data and strong solutions. A further smoothing property of strong solutions for u_0 belonging to the closure of $D(\varphi)$ in H will be also discussed under some additional assumptions of kernels in Theorem 2.8. Furthermore, we shall also give a corollary on a contraction property of the solution operator for (1.1) under some additional assumptions (see Corollary 2.7).

Section 3 contains materials related to nonlocal time-derivatives which will be used later. In §3.1, we arrange some preliminary facts on nonlocal time-derivatives. In particular, we shall recall maximal monotonicity of nonlocal derivatives $(\mathrm{d}/\mathrm{d}t)[k * u]$ under the assumption (K) as well as some well-known and useful facts for energy methods. In §3.2, we present a chain-rule for convex functionals and nonlocal derivatives, which will play a crucial role to prove the main result. Subsection 3.3 is concerned with maximality of the sum of two maximal monotone operators, that is, a standard differential operator and a nonlocal one. Indeed, the maximality of the sum of two maximal monotone operators is not always obvious and one needs in particular to pay careful attention for it, when the domains of two operators have no interior point. This part will be also used to construct approximate solutions for (1.1) (see §4.1) and to derive a priori estimates for them (see §4.2).

Section 4 is concerned with a proof of the main result (i.e., Theorem 2.3). It consists of several steps. We first assume $f \in W^{1,2}(0,T;H)$ to derive additional regularity of strong solutions and energy inequalities. Then we introduce approximate problems,

$$\lambda \frac{\mathrm{d}u_\lambda}{\mathrm{d}t}(t) + \frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u_\lambda - u_0)\right](t) + \partial\varphi_\lambda(u_\lambda(t)) \ni f(t) \quad \text{in H}, \quad 0 < t < T,$$

where $\lambda \in (0,1)$, $\partial\varphi_\lambda$ denotes the Yosida approximation of $\partial\varphi$, along with the initial condition $u_\lambda(0) = u_0$ in a classical sense. Existence of strong solutions for the approximate problems will be proved based on maximal monotone operator theory and some facts developed in §3.3. Moreover, in §4.2, some a priori estimates will be established by applying materials developed in §3.1 and 3.3. In §4.3, a limiting procedure is discussed by proving that (u_λ) forms a Cauchy sequence and by employing standard techniques such as demiclosedness of maximal monotone operators. Thus existence of a strong solution for (1.1) will be proved for $f \in W^{1,2}(0,T;H)$. In §4.5, we shall prove continuous dependence of strong solutions on prescribed data. In §4.6, we shall set about proving existence of solutions to (1.1) for $f \in L^2(0,T;H)$. Here the nonlocal chain-rule developed in §3.2 will play a crucial role. Furthermore, §4.7 is devoted to a proof of Corollary 2.7. Finally, we shall give a proof for Theorem 2.8 in §4.8.

In Section 5, we shall treat a Lipschitz perturbation problem,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[k*(u-u_0)\right](t) + \partial\varphi(u(t)) + F(t,u(t)) \ni f(t) \quad \text{in } H, \quad 0<t<T,$$

where $F = F(t,w)$ is a mapping from $(0,T) \times H$ into H, measurable in t and Lipschitz continuous in w. Indeed, the Lipschitz perturbation theory enhances applicability of the abstract theory to (time-fractional) nonlinear PDEs and also extends the main result to abstract fractional gradient flows for the so-called λ-*convex* functionals, which may not be convex for $\lambda < 0$. The proof relies on a classical contraction argument. However, the choice of a certain function space and the proof of a contraction property of an associated mapping are somewhat delicate due to the nonlocal nature of the equation.

Section 6 concerns applications of the preceding abstract results to simple but typical examples of nonlinear PDEs including time-fractional derivatives. More precisely, we shall apply the main results in §2 to time-fractional p-Laplace parabolic equations in §6.1 and time-fractional porous medium and fast diffusion equations in §6.2. Furthermore, the Lipschitz perturbation theory developed in §5 will be applied to a time-fractional Allen-Cahn equation in §6.3.

2. Main results

This section is devoted to presenting an abstract theory for (1.1). Throughout this paper, we are concerned with *strong solutions* of (1.1) in the following sense:

Definition 2.1 (Strong solution of (1.1)). *A function $u \in L^2(0,T;H)$ is called a* strong solution *of* (1.1) *on* $[0,T]$ *if the following conditions are all satisfied*:

(i) *It holds that*

$$k*(u-u_0) \in W^{1,2}(0,T;H), \quad [k*(u-u_0)](0) = 0,$$
$$u(t) \in D(\partial\varphi) \quad \text{for a.e. } t \in (0,T).$$

(ii) *There exists $\xi \in L^2(0,T;H)$ such that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[k*(u-u_0)\right](t) + \xi(t) = f(t) \quad \text{and} \quad \xi(t) \in \partial\varphi(u(t)) \quad \text{in } H \tag{2.1}$$

for a.e. $t \in (0,T)$.

REMARK 2.2 (Further integrability of strong solutions). (i) The definition of strong solution entails

$$k * \|u-u_0\|_H^2 \in L^\infty(0,T), \quad k(\cdot)\|u(\cdot)-u_0\|_H^2 \in L^1(0,T) \tag{2.2}$$

(see §A in Appendix). Furthermore, it also follows immediately that

$$\varphi(u(\cdot)) \in L^1(0,T)$$

from the definition of subdifferential.

(ii) Let k be the Riemann-Liouville kernel (of the order β) defined by (1.3). Then the following regularity is directly derived from the definition above:

$$u \in \begin{cases} L^{\frac{2}{1-2\beta},\infty}(0,T;H) & \text{if } \beta \in (0,1/2), \\ \cap_{p<+\infty} L^{p,\infty}(0,T;H) & \text{if } \beta = 1/2, \end{cases} \tag{2.3}$$

where $L^{p,\infty}$ stands for the weak L^p-space. As for the case $\beta > 1/2$, we shall obtain $u \in C([0,T];H)$ in Proposition 2.5 below.

Our main result reads,

THEOREM 2.3 (Well-posedness of (1.1)). *Assume that* (K) *is satisfied. For any* $T>0$, $f \in L^2(0,T;H)$ *and* $u_0 \in D(\varphi)$, *the Cauchy problem* (1.1) *admits a unique strong solution* $u \in L^2(0,T;H)$ *on* $[0,T]$. *Moreover, there exists a non-increasing function* $\mathcal{F} : [0,T] \to (-\infty, 0]$ *satisfying* $\mathcal{F}(0)=0$ *such that*

$$\left\| \frac{\mathrm{d}}{\mathrm{d}t}\left[k*(u-u_0)\right](t) \right\|_H^2 + \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(t) \le 0 \quad \textit{for a.e. } t \in (0,T) \tag{2.4}$$

and

$$\mathcal{F}(t) = \Big[k * \big(\varphi(u(\cdot)) - \varphi(u_0)\big)\Big](t) - \int_0^t \left(f(\tau), \frac{\mathrm{d}}{\mathrm{d}t}\left[k*(u-u_0)\right](\tau) \right)_H \mathrm{d}\tau \tag{2.5}$$

for a.e. $t \in (0,T)$.

In addition, if $f \in W^{1,2}(0,T;H)$, *it is then satisfied that*

$$u \in L^\infty(0,T;H), \quad \ell * \left\| \frac{\mathrm{d}}{\mathrm{d}t}[k*(u-u_0)](\cdot) \right\|_H^2 \in L^\infty(0,T), \quad \varphi(u(\cdot)) \in L^\infty(0,T), \tag{2.6}$$

and the following energy inequality holds:

$$\begin{aligned} &\frac{1}{2}\left(\ell * \left\| \frac{\mathrm{d}}{\mathrm{d}t}[k*(u-u_0)](\cdot) \right\|_H^2 \right)(t) + \varphi(u(t)) \\ &\le \varphi(u_0) + (f(t),u(t))_H - (f(0),u_0)_H - \int_0^t (f'(\tau),u(\tau))_H \,\mathrm{d}\tau \end{aligned} \tag{2.7}$$

for a.e. $t \in (0,T)$. *Moreover,* (2.5) *holds for all* $t \in [0,T]$.

The unique strong solution continuously depends on prescribed data in the following sense: *let* u_1 *and* u_2 *be strong solutions on* $[0,T]$ *to* (1.1) *with* (u_0,f) *replaced* *by* $(u_{0,1}, f_1), (u_{0,2}, f_2) \in H \times L^2(0,T;H)$, *respectively. Then* $w := u_1 - u_2$ *and* $w_0 := u_{0,1} - u_{0,2}$ *fulfill*

$$\int_0^T \|w(t) - w_0\|_H^2 \,\mathrm{d}t \le C \left(\|f_1 - f_2\|_{L^2(0,T;H)}^2 + \|w_0\|_H^2 \right) \tag{2.8}$$

for some constant C *depending only on* $\|k\|_{L^1(0,T)}$ *and* $\|\ell\|_{L^1(0,T)}$.

Remark 2.4 (Energy inequality). (i) The regularity (2.6) and the energy inequality (2.7) do not follow from Definition 2.1. Indeed, it can be obtained by formally testing (1.1) with u'; however, u is not supposed to be of class $W^{1,2}(0,T;H)$ in the definition. Hence, (2.7) can be regarded as an extra regularity of strong solutions under the assumption $f \in W^{1,2}(0,T;H)$.

(ii) If $f \equiv 0$, then the inequality (2.4) can be regarded as a fractional variant of Lyapunov property (in particular, the non-increase of the energy $t \mapsto \varphi(u(t))$ of classical gradient flows (i.e., the case $\alpha = 1$). In particular, if k is given by (1.3), then the inequality can be formally regarded as

$$\left\| \partial_t^\beta (u - u_0) \right\|_H^2 + \partial_t^\beta \left[\varphi(u(\cdot)) - \varphi(u_0) \right] \le 0 \quad \text{for} \ \ t > 0.$$

However, it does not imply the non-increase of the energy $t \mapsto \varphi(u(t))$ due to the nonlocal nature of the fractional derivative. Indeed, it is known that functions are not always monotone on some interval, even though their fractional derivatives of some order $\beta \in (0,1)$ are signed on the interval (see [16, Example 2.1]). On the other hand, the function $t \mapsto [k * (\varphi(u(\cdot)) - \varphi(u_0))](t)$ is non-increasing, provided that $f \equiv 0$.

(iii) As for the classical case $\alpha = 1$, the following simpler relation holds true:

$$\|u_1(t) - u_2(t)\|_H \le \|u_{0,1} - u_{0,2}\|_H + \int_0^t \|f_1(\tau) - f_2(\tau)\|_H \,\mathrm{d}\tau \ \ \text{for all} \ \ t > 0,$$

which particularly assures that (1.1) (with $\alpha = 1$) generates a contraction semigroup (see [8, Lemma 3.1]). Corollary 2.7 below is a counterpart for fractional gradient flows. Here we also remark in advance that $\log \ell_\beta$ is convex and $\ell_\beta \in L^{1/(1-\beta),\infty}(0,+\infty)$, when (k_β, ℓ_β) is given as in (1.3).

Let us next remark that the initial condition, $u(0) = u_0$, holds true in a classical sense, when the conjugate kernel ℓ belongs to $L^2(0,T)$.

Proposition 2.5 (Initial condition in a classical sense). *In addition to* (K)*, assume that* $\ell \in L^2(0,T)$. *Then for any* $f \in L^2(0,T;H)$ *and* $u_0 \in D(\varphi)$ *the unique strong solution has a continuous representative,* $u \in C([0,T];H)$. *Moreover,* $u(t) \to u_0$ *strongly in* H *as* $t \to 0_+$.

Proof. Convolve (2.1) with ℓ. Then

$$\left(\ell * \frac{\mathrm{d}}{\mathrm{d}t} \left[k * (u - u_0) \right] \right)(t) + \left[\ell * (\xi - f) \right](t) = 0, \quad \xi(t) \in \partial\varphi(u(t))$$

for a.e. $t \in (0,T)$. Note by $[k * (u-u_0)](0) = 0$ that

$$\ell * \frac{\mathrm{d}}{\mathrm{d}t}[k*(u-u_0)] = \frac{\mathrm{d}}{\mathrm{d}t}[\ell * k * (u-u_0)] = \frac{\mathrm{d}}{\mathrm{d}t}[1*(u-u_0)] = u-u_0$$

for a.e. $t \in (0,T)$. By $\ell \in L^2(0,T)$ and $(\mathrm{d}/\mathrm{d}t)[k*(u-u_0)] \in L^2(0,T;H)$, we find that $u-u_0$ has a continuous representative with values in H on $[0,T]$, and then, we denote it by $u-u_0$ again. Thus we find that

$$\begin{aligned}\|u(t)-u_0\|_H &= \|[\ell * (\xi - f)](t)\|_H \\ &\le \int_0^t \ell(s)\|\xi(t-s) - f(t-s)\|_H \,\mathrm{d}s \le \|\ell\|_{L^2(0,t)}\|\xi - f\|_{\mathcal{H}}\end{aligned}$$

for any $t \in [0,T]$. Therefore we conclude that $u(t) \to u_0$ strongly in H as $t \to 0_+$. □

Remark 2.6 (Classical initial condition for fractional derivatives). As for the Riemann-Liouville kernel $k_\beta(t) = t^{-\beta}/\Gamma(1-\beta)$, the conjugate kernel $\ell_\beta(t) = t^{\beta-1}/\Gamma(\beta)$ is of class $L^2(0,T)$ for any $T>0$ if and only if $\beta > 1/2$.

As in [11], we have

Corollary 2.7 (Contraction property). *In addition to* (K), *assume that*

$$\ell > 0 \quad \textit{and} \quad \log \ell(t) \ \textit{is convex.} \tag{2.9}$$

For $i=1,2$, let $u_{0,i}$ and f_i fulfill the same assumptions as in Theorem 2.3. Let u_i be the unique strong solution on $[0,T]$ of (1.1) *with u_0 and f replaced by $u_{0,i}$ and f_i, respectively. Then it holds that*

$$\|u_1(t) - u_2(t)\|_H \le \|u_{0,1} - u_{0,2}\|_H + \big(\ell * \|f_1 - f_2\|_H\big)(t) \tag{2.10}$$

for a.e. $t \in (0,T)$. In particular, if $\ell \in L^2(0,T)$, then (1.1) *generates a non-expansive (in H) mapping $S(t) : D(\varphi) \to D(\varphi)$, $u_0 \mapsto u(t)$ for $t \in [0,T]$, where $u(\cdot)$ stands for the unique solution to* (1.1) *with the initial datum u_0.*

Finally, let us consider the case $u_0 \in \overline{D(\varphi)}^H$, where one can no longer expect that every strong solution u of (1.1) lies on the domain of the nonlocal derivative (more precisely, $k*(u-u_0)$ does not belong to $W^{1,2}(0,T;H)$). Therefore we need to modify the notion of strong solutions defined by Definition 2.1. The following theorem is concerned with a further smoothing effect.

Theorem 2.8 (Smoothing effect for $u_0 \in \overline{D(\varphi)}^H$). *In addition to* (K), *suppose that*

$$\textit{the function } t \mapsto t\left(\int_0^t \ell(s)\,\mathrm{d}s\right)^{-2} \ \textit{belongs to } L^1(0,T). \tag{2.11}$$

Then for any $f \in L^2(0,T;H)$ *and* $u_0 \in \overline{D(\varphi)}^H$, *there exists a function* $u \in L^2(0,T;H)$ *such that*

$$k * \|u-u_0\|_H^2 \in L^\infty(0,T), \quad k(\cdot)\|u(\cdot)-u_0\|_H^2 \in L^1(0,T),$$
$$\varphi(u(\cdot)) \in L^1(0,T), \quad t^{1/2}\frac{\mathrm{d}}{\mathrm{d}t}[k*(u-u_0)] \in L^2(0,T;H),$$
$$u(t) \in D(\partial\varphi) \ \text{ for a.e. } t \in (0,T),$$
$$[k*(u-u_0)](t) \to 0 \ \text{ strongly in } H \ \text{ as } \ t \to 0_+,$$

and there exists $\xi \in L^2_{\mathrm{loc}}((0,T];H)$ *such that* $t^{1/2}\xi \in L^2(0,T;H)$ *and* (2.1) *holds true. Furthermore, any two solutions* u_1, u_2 *constructed in this theorem satisfy* (2.8). *In addition, if* $\ell \in L^2(0,T)$ *and* (2.9) *is satisfied, then* u *belongs to* $C([0,T];H)$ *and* $u(0) = u_0$. *Moreover,* (2.10) *holds for all* $t \in [0,T]$.

3. Nonlocal time-differential operators

In §3.1, we shall arrange preliminary facts from a functional analytic theory for nonlocal time-differential operators. In §3.2, a chain-rule for convex functionals and nonlocal derivatives will be provided. In §3.3, we shall give a proposition on the maximal monotonicity of the sum of the standard differential operator and nonlocal one, which will play a crucial role to prove main results.

3.1. **Preliminaries.** Let $T > 0$ and $p \in [1,+\infty]$ be fixed and let X be a Banach space. We first recall the (ordinary) time-differential operator $\mathcal{A} : D(\mathcal{A}) \subset L^p(0,T;X) \to L^p(0,T;X)$ defined by

$$D(\mathcal{A}) = \left\{ w \in W^{1,p}(0,T;X) \colon w(0) = 0 \right\} \quad \text{and} \quad \mathcal{A}(w) := \frac{\mathrm{d}w}{\mathrm{d}t} \quad \text{for } \ w \in D(\mathcal{A}).$$

Then it is well known that $\mathcal{A}$ is linear and m-accretive in $L^p(0,T;X)$.

Let us next define a *nonlocal time-differential operator* $\mathcal{B} : D(\mathcal{B}) \subset L^p(0,T;X) \to L^p(0,T;X)$ by

$$D(\mathcal{B}) = \{ w \in L^p(0,T;X) \colon k * w \in D(\mathcal{A}) \} \tag{3.1}$$

and

$$\mathcal{B}(w) := \mathcal{A}\,(k*w) = \frac{\mathrm{d}}{\mathrm{d}t}(k*w) \quad \text{for } \ w \in D(\mathcal{B}). \tag{3.2}$$

Then we note that

$$D(\mathcal{A}) \subset D(\mathcal{B}).$$

Indeed, for any $u \in D(\mathcal{A})$, we find that $(k*u)(0) = 0$, since $k \in L^1(0,T)$ and $u \in W^{1,p}(0,T;X) \subset L^\infty(0,T;X)$. Moreover, we infer that $k * u \in W^{1,p}(0,T;X)$ by $u(0) = 0$ and $u \in W^{1,p}(0,T;X)$. Hence $u \in D(\mathcal{B})$.

Remark 3.1 (Riemann-Liouville and Caputo derivatives). Assume that (K) is satisfied. For $u \in W^{1,2}(0,T;H)$ taking an initial datum $u(0) = u_0 \in H$, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[k*(u-u_0)\right](t) = \left(k * \frac{\mathrm{d}u}{\mathrm{d}t}\right)(t), \quad 0 < t < T. \tag{3.3}$$

Hence the restriction of $\mathcal{B}$ onto $D(\mathcal{A})$ coincides with the operator $\mathcal{C}$ defined on $D(\mathcal{C}) := D(\mathcal{A})$ by

$$\mathcal{C}(u) := k * \frac{\mathrm{d}u}{\mathrm{d}t} \quad \text{for} \ \ u \in D(\mathcal{C}).$$

In particular, if k is given by (1.3), then $\mathcal{B}$ and $\mathcal{C}$ correspond to the Riemann-Liouville and Caputo differential operators, respectively.

We recall that $\mathcal{B}$ is also m-accretive in $L^p(0,T;X)$ under the assumption (K) (see, e.g., [10], [24], [14], [51], [56]). Then for $n \in \mathbb{N}$ the resolvent $\mathcal{J}_n : L^p(0,T;X) \to D(\mathcal{B})$ and the Yosida approximation $\mathcal{B}_n : L^p(0,T;X) \to L^p(0,T;X)$ of $\mathcal{B}$ are given by

$$\mathcal{J}_n(w) := \left(I + \frac{1}{n}\mathcal{B}\right)^{-1}(w), \quad \mathcal{B}_n(w) := n(I - \mathcal{J}_n)(w) = \mathcal{B}(\mathcal{J}_n w) = \frac{\mathrm{d}}{\mathrm{d}t}(k_n * w),$$

where I stands for the identity mapping (we shall use the same letter for identity mappings defined on any spaces unless any confusion may arise). Moreover, $k_n \in W^{1,1}(0,T)$ is a nonincreasing and nonnegative kernel given by $k_n = ns_n$, where s_n is a unique solution of the Volterra equation,

$$s_n + n(\ell * s_n) = 1 \ \ \text{in} \ (0,+\infty).$$

Hence k_n depends only on ℓ and n; in particular, it is independent of the choices of X and p. Then a general theory for (linear) m-accretive operators (see e.g. [39], [4]) ensures that

$$\mathcal{B}_n(w) \to \mathcal{B}(w) \quad \text{strongly in } L^p(0,T;X) \ \ \text{as} \ \ n \to \infty, \tag{3.4}$$

provided that $w \in D(\mathcal{B})$. Indeed, $\mathcal{B}$ is densely defined in $L^p(0,T;X)$, and hence, for $w \in D(\mathcal{B})$, we deduce that $\mathcal{B}_n(w) = \mathcal{B}(\mathcal{J}_n(w)) = \mathcal{J}_n(\mathcal{B}(w)) \to \mathcal{B}(w)$ strongly in $L^p(0,T;X)$ as $n \to +\infty$. Moreover, (3.4) particularly implies that

$$k_n \to k \quad \text{strongly in } L^1(0,T) \ \ \text{as} \ \ n \to \infty,$$

by setting $w \equiv 1$, $p = 1$ and $X = \mathbb{R}$.

Remark 3.2 (Class $\mathcal{K}^1(\alpha,\theta)$). Let us further define a class of kernels (see [44] and [51, §2] for more details): $h \in L^1_{\mathrm{loc}}([0,+\infty))$ is said to be of class $\mathcal{K}^1(\alpha,\theta)$ for some $\alpha \geq 0$ and $\theta > 0$ if the following conditions hold true:

- h is of *subexponential growth*, i.e., $\int_0^\infty e^{-\varepsilon t}|h(t)|\,\mathrm{d}t < +\infty$ for any $\varepsilon > 0$;
- h is 1-*regular*, i.e., there exists a constant $c > 0$ such that $|\lambda \hat{h}'(\lambda)| \leq c|\hat{h}(\lambda)|$ for all $\operatorname{Re}\lambda > 0$;
- h is θ-*sectorial*, i.e., $|\arg(\hat{h})(\lambda)| \leq \theta$ for all $\operatorname{Re}\lambda > 0$;
- it holds that

$$\limsup_{\lambda\to+\infty} |\hat{h}(\lambda)|\lambda^\alpha < +\infty, \quad \liminf_{\lambda\to+\infty} |\hat{h}(\lambda)|\lambda^\alpha > 0, \quad \liminf_{\lambda\to 0} |\hat{h}(\lambda)| > 0.$$

Let us give a couple of remarks:

(i) Suppose that $h \in \mathcal{K}^1(\alpha,\theta)$ for some $\alpha \in (0,1)$ and $\theta \in (0,\pi)$. If $w \in L^2(0,T;H)$, $h * w \in W^{1,2}(0,T;H)$ and $(h * w)(0) = 0$, then $h * \|w\|_H^2 \in W^{1,1}(0,T)$ and $(h * \|w\|_H^2)(0) = 0$ (see §2 of [51]).

(ii) Any Riemann-Liouville kernel $k(s) = s^{-\beta}/\Gamma(1-\beta)$, $0 < \beta < 1$, is of class $\mathcal{K}^1(\beta, \beta\pi/2)$ (see Example 2.1 of [51]).

(iii) Let X be a Banach space such that the Hilbert transform is bounded in $L^p(\mathbb{R}; X)$ for some $p \in (1, +\infty)$ (in particular, any Hilbert space satisfy the property). If k is of class $\mathcal{K}^1(\alpha, \theta)$ for some $\alpha \in (0,1)$ and $\theta \in (0, \pi)$, then the domain $D(\mathcal{B})$ of $\mathcal{B}$ coincides with the space $H_0^{\alpha,p}(0,T;X) = \{u|_{[0,+\infty)} \colon u \in H^{\alpha,p}(\mathbb{R};X) \text{ and } \operatorname{supp} u \subset [0,+\infty)\}$, where $H^{\alpha,p}(\mathbb{R};X)$ is the so-called *Bessel potential space*, i.e., $u \in H^{\alpha,p}(\mathbb{R};X)$ if and only if $u \in L^p(\mathbb{R};X)$ and there exists $g \in L^p(\mathbb{R};X)$ whose Fourier transform $\widehat{g}(\rho)$ coincides with $|\rho|^\alpha \widehat{u}(\rho)$ (here $\widehat{u}$ denotes the Fourier transform of u). We refer the reader to [54, Corollary 2.1] and [56, Corollary 3.1] for more details.

Let us recall the following lemma, which will be frequently used later.

Lemma 3.3 (See Lemma 2.1 of [56]). *Let H be a Hilbert space. For any $h \in W^{1,1}(0,T)$ and $u \in L^2(0,T;H)$, it holds that*

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}(h * u)(t), u(t)\right)_H = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(h * \|u\|_H^2\right)(t) + \frac{1}{2}h(t)\|u(t)\|_H^2 - \frac{1}{2}\int_0^t h'(s)\|u(t) - u(t-s)\|_H^2 \, \mathrm{d}s \tag{3.5}$$

for a.e. $t \in (0,T)$. Here each term of the right-hand side belongs to $L^1(0,T)$. In particular, if $h' \le 0$, then

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}(h * u)(t), u(t)\right)_H \ge \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(h * \|u\|_H^2\right)(t) + \frac{1}{2}h(t)\|u(t)\|_H^2 \quad \textit{for a.e. } t \in (0,T).$$

Here we emphasize that the identity above in (3.5) holds only for *absolutely continuous* kernels. Hence one cannot directly apply it to k, ℓ satisfying only (K).

3.2. **Chain-rule for convex functionals and nonlocal derivatives.** The following proposition provides a chain-rule for convex functionals and nonlocal derivatives (cf. Lemma 2.2. of [50]). Here we denote by ψ^* the *convex conjugate* (or *Legendre transform*) of a convex functional $\psi : H \to [0, +\infty]$, that is,

$$\psi^*(g) := \sup_{v \in H}\left[(g, v)_H - \psi(v)\right] \quad \text{for } g \in H. \tag{3.6}$$

Then ψ^* is proper, lower semicontinuous and convex, provided that so is ψ.

Proposition 3.4 (Nonlocal chain-rule for convex functionals). *Let $h \in W^{1,1}(0,T)$ and $\psi : H \to [0,+\infty]$ be a proper (i.e., $\psi \not\equiv +\infty$) convex functional. Let $u \in L^1(0,T;H)$ be such that $\psi(u(\cdot)) \in L^1(0,T)$. Then for each $t \in (0,T)$ satisfying*

$u(t) \in D(\partial\psi)$ and for any $u_0 \in H$ and $g \in \partial\psi(u(t))$, it holds that

$$\begin{aligned}&\left(\frac{\mathrm{d}}{\mathrm{d}t}[h*(u-u_0)](t), g\right)_H \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\big[h*\psi(u(\cdot))\big](t) + h(t)\big[(u(t)-u_0,g)_H - \psi(u(t))\big] \\ &\quad + \int_0^t h'(\tau)\big[(u(t-\tau)-u(t),g)_H + \psi(u(t)) - \psi(u(t-\tau))\big]\,\mathrm{d}\tau.\end{aligned}$$

Moreover, assume that $h' \le 0$. Then one has

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}[h*(u-u_0)](t), g\right)_H \ge \frac{\mathrm{d}}{\mathrm{d}t}\big[h*\psi(u(\cdot))\big](t) + h(t)\left[\psi^*(g) - (u_0,g)_H\right].$$

In addition, if $u_0 \in D(\psi)$ and $h \ge 0$, then

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}[h*(u-u_0)](t), g\right)_H \ge \frac{\mathrm{d}}{\mathrm{d}t}\big[h*(\psi(u(\cdot))-\psi(u_0))\big](t).$$

Proof. Let $u \in L^1(0,T;H)$ be such that $\psi(u(\cdot)) \in L^1(0,T)$ and let $t \in (0,T)$ be such that $u(t) \in D(\partial\varphi)$. Then we note that the function $t \mapsto [h*\psi(u(\cdot))](t)$ belongs to $W^{1,1}(0,T)$. By straightforward computation, one finds that, for any $u_0 \in H$ and $g \in \partial\varphi(u(t))$,

$$\begin{aligned}&\left(\frac{\mathrm{d}}{\mathrm{d}t}[h*(u-u_0)](t), g\right)_H \\ &= h(0)\,(u(t)-u_0,g)_H + \int_0^t h'(t-s)\,(u(s)-u_0,g)_H\,\mathrm{d}s \\ &= h(0)\psi(u(t)) + h(0)\big[(u(t)-u_0,g)_H - \psi(u(t))\big] \\ &\quad + \int_0^t h'(t-s)\psi(u(s))\,\mathrm{d}s + \int_0^t h'(t-s)\big[(u(s)-u_0,g)_H - \psi(u(s))\big]\,\mathrm{d}s \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\big[h*\psi(u(\cdot))\big](t) - \int_0^t h'(\tau)\,\mathrm{d}\tau\big[(u(t)-u_0,g)_H - \psi(u(t))\big] \\ &\quad + h(t)\big[(u(t)-u_0,g)_H - \psi(u(t))\big] \\ &\quad + \int_0^t h'(\tau)\big[(u(t-\tau)-u_0,g)_H - \psi(u(t-\tau))\big]\,\mathrm{d}\tau \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\big[h*\psi(u(\cdot))\big](t) + h(t)\big[(u(t)-u_0,g)_H - \psi(u(t))\big] \\ &\quad + \int_0^t h'(\tau)\big[(u(t-\tau)-u(t),g)_H + \psi(u(t)) - \psi(u(t-\tau))\big]\,\mathrm{d}\tau.\end{aligned}$$

Here we note by definition of $\partial\psi$ that

$$(u(t-\tau)-u(t),g)_H + \psi(u(t)) - \psi(u(t-\tau)) \le 0.$$

On the other hand, we note by the Fenchel-Moreau identity that

$$(u(t),g)_H - \psi(u(t)) = \psi^*(g) \quad \text{if and only if} \quad g \in \partial\psi(u(t))$$

In addition, assume that $u_0 \in D(\psi)$ and $h \geq 0$. It then follows that

$$(u(t) - u_0, g)_H - \psi(u(t)) \geq -\psi(u_0),$$

whence follows

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}t}[h * (u - u_0)](t), g\right)_H &\geq \frac{\mathrm{d}}{\mathrm{d}t}\big[h * \psi(u(\cdot))\big] - \psi(u_0)h(t)\\ &\geq \frac{\mathrm{d}}{\mathrm{d}t}\Big[h * \big(\psi(u(\cdot)) - \psi(u_0)\big)\Big](t).\end{aligned}$$

This completes the proof. □

Lemma 3.3 also follows from Proposition 3.4 by setting $\psi(w) = (1/2)\|w\|_H^2$.

3.3. **Maximality of $\mathcal{A}+\mathcal{B}$.** Throughout this subsection, we set X to be a Hilbert space H and $p = 2$ and simply write $\mathcal{H} = L^2(0, T; H)$. We prove the following:

Proposition 3.5. *Under the assumption* (K)*, the sum $\mathcal{A}+\mathcal{B}$ is maximal monotone in $\mathcal{H} \times \mathcal{H}$.*

The maximality of the sum $\mathcal{A} + \mathcal{B}$ has already been proved in a more direct way (see, e.g., [10]). Here, we shall give an alternative proof based on a sufficient condition for the maximality of the sum of two maximal monotone operators. Moreover, some part of the following argument will be also used to derive energy estimates later.

Proof of Proposition 3.5. We first show that

$$(\mathcal{A}(u), \mathcal{B}(u))_{\mathcal{H}} = \left(\frac{\mathrm{d}u}{\mathrm{d}t}, \frac{\mathrm{d}}{\mathrm{d}t}(k * u)\right)_{\mathcal{H}} \geq 0 \tag{3.7}$$

for each $u \in D(\mathcal{A})$, that is,

$$u,\ k * u \in W^{1,2}(0, T; H),\ u(0) = 0 \ \text{ and } (k * u)(0) = 0.$$

Here we note by (K) that

$$\mathcal{A}(u) = \frac{\mathrm{d}u}{\mathrm{d}t} = \frac{\mathrm{d}^2}{\mathrm{d}t^2}(\ell * k * u) = \frac{\mathrm{d}}{\mathrm{d}t}\left[\ell * \frac{\mathrm{d}}{\mathrm{d}t}(k * u)\right]. \tag{3.8}$$

Set $v := \mathcal{B}(u) = (\mathrm{d}/\mathrm{d}t)(k * u) \in \mathcal{H}$. Then it is satisfied that

$$\ell * v = u \in D(\mathcal{A}).$$

Hence v belongs to the domain of the operator $\mathcal{B}_\ell$ which is defined by (3.1) and (3.2) with the kernel k replaced by ℓ. Then by (K) there exists $\ell_n \in W^{1,1}(0, T)$ such that $(B_\ell)_n = (\mathrm{d}/\mathrm{d}t)(\ell_n * \cdot\,)$ and $\ell_n \to \ell$ strongly in $L^1(0, T)$. By Lemma 3.3, we observe that

$$\begin{aligned}(\mathcal{A}(u)(t), \mathcal{B}(u)(t))_H &= \left(\frac{\mathrm{d}}{\mathrm{d}t}(\ell * v)(t), v(t)\right)_H\\ &\geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\ell_n * \|v\|_H^2\right)(t) + \frac{1}{2}\ell_n(t)\|v(t)\|_H^2 + h_n(t),\end{aligned} \tag{3.9}$$

where $h_n \in L^1(0,T)$ is given by

$$h_n(t) := \left(\frac{\mathrm{d}}{\mathrm{d}t}(\ell * v)(t) - \frac{\mathrm{d}}{\mathrm{d}t}(\ell_n * v)(t), v(t) \right)_H .$$

Here we note that $h_n \to 0$ strongly in $L^1(0,T)$. Indeed, since $\mathcal{B}_\ell$ is linear maximal monotone in $\mathcal{H}$, we find by $v \in D(\mathcal{B}_\ell)$ that

$$(\mathcal{B}_\ell)_n(v) = \frac{\mathrm{d}}{\mathrm{d}t}(\ell_n * v) \to \mathcal{B}_\ell(v) = \frac{\mathrm{d}}{\mathrm{d}t}(\ell * v) \quad \text{strongly in } \mathcal{H},$$

which implies that $h_n \to 0$ strongly in $L^1(0,T)$. Integrating both sides of (3.9) over $(0,T)$, one has

$$\begin{aligned}\int_0^T (\mathcal{A}(u)(t), \mathcal{B}(u)(t))_H \,\mathrm{d}t &\geq \frac{1}{2}\left(\ell_n * \|v\|_H^2\right)(T) + \int_0^T h_n(t)\,\mathrm{d}t \\ &\geq \int_0^T h_n(t)\,\mathrm{d}t \to 0 \quad \text{as } \ n \to \infty,\end{aligned}$$

where we also used the nonnegativity of ℓ_n and the fact that

$$(\ell_n * \|v\|_H^2)(0) = 0.$$

Indeed, we note that $\ell_n * \|v\|_H^2$ is continuous by $\ell_n \in W^{1,1}(0,T)$ and $v \in \mathcal{H}$. Then

$$\begin{aligned}\left|(\ell_n * \|v\|_H^2)(t)\right| &\leq \int_0^t |\ell_n(t-s)| \|v(s)\|_H^2 \,\mathrm{d}s \\ &\leq \sup_{\tau \in [0,t]} |\ell_n(\tau)| \int_0^t \|v(s)\|_H^2 \,\mathrm{d}s \to 0 \quad \text{as } \ t \to 0_+.\end{aligned}$$

Let $\mathcal{J}_n$ denote the resolvent of $\mathcal{B}$, that is, $\mathcal{J}_n := (I + \mathcal{B}/n)^{-1} : \mathcal{H} \to D(\mathcal{B})$. We next claim that

$$\mathcal{J}_n(D(\mathcal{A})) \subset D(\mathcal{A}) \quad \text{for any } \ n \in \mathbb{N}.$$

Let $u \in D(\mathcal{A})$. Then $u_n := \mathcal{J}_n u$ satisfies

$$u_n = u - \frac{1}{n}\mathcal{B}_n(u) \in W^{1,2}(0,T;H).$$

Indeed, we find by $u \in W^{1,2}(0,T;H)$ and $k_n \in W^{1,1}(0,T)$ that

$$\mathcal{B}_n(u) = \frac{\mathrm{d}}{\mathrm{d}t}(k_n * u) = k_n(0)u + k_n' * u \in W^{1,2}(0,T;H).$$

Furthermore, we observe that $u_n(0) = u(0) - \frac{1}{n}\mathcal{B}_n(u)(0)$ and that

$$\begin{aligned}\|\mathcal{B}_n(u)(t)\|_H &\leq k_n(0)\|u(t)\|_H + \int_0^t |k_n'(t-s)| \|u(s)\|_H \,\mathrm{d}s \\ &\leq k_n(0)\|u(t)\|_H + \sup_{\tau \in [0,t]} \|u(\tau)\|_H \int_0^t |k_n'(\sigma)|\,\mathrm{d}\sigma \to 0\end{aligned}$$

as $t \to 0_+$ by $k_n' \in L^1(0,T)$ and $u \in D(\mathcal{A})$. Therefore $\mathcal{J}_n u = u_n \in D(\mathcal{A})$ if $u \in D(\mathcal{A})$.

The rest of the proof runs as in the classical literature (see [8, Chap.II, §9 and Chap.IV, §4]). However, for the reader's convenience, we give the details of the proof. Let us claim that

$$(\mathcal{A}(u), \mathcal{B}_n(u))_{\mathcal{H}} \geq 0 \quad \text{for all } \ u \in D(\mathcal{A}) \ \text{ and } \ n \in \mathbb{N}. \tag{3.10}$$

Indeed, since $\mathcal{J}_n u \in D(\mathcal{A})$ by $u \in D(\mathcal{A})$, we observe by the last claim that

$$\begin{aligned}(\mathcal{A}(u), \mathcal{B}_n(u))_{\mathcal{H}} &= (\mathcal{A}(\mathcal{J}_n u), \mathcal{B}_n(u))_{\mathcal{H}} + (\mathcal{A}(u) - \mathcal{A}(\mathcal{J}_n u), \mathcal{B}_n(u))_{\mathcal{H}} \\ &= (\mathcal{A}(\mathcal{J}_n u), \mathcal{B}(\mathcal{J}_n u))_{\mathcal{H}} + n\,(\mathcal{A}(u) - \mathcal{A}(\mathcal{J}_n u), u - \mathcal{J}_n u)_{\mathcal{H}} \overset{(3.7)}{\geq} 0.\end{aligned}$$

Thus (3.10) follows. Now, we are in position to show the maximality of $\mathcal{A} + \mathcal{B}$, which is equivalent to the surjectivity of $I + \mathcal{A} + \mathcal{B}$ on $\mathcal{H}$. Let $f \in \mathcal{H}$ be arbitrarily given. Since $\mathcal{B}_n$ is monotone and Lipschitz continuous in $\mathcal{H}$, the sum $\mathcal{A} + \mathcal{B}_n$ turns out to be maximal monotone in $\mathcal{H}$ (see, e.g., [8, Lemma 2.4]). Thus one can take $u_n \in D(\mathcal{A})$ such that

$$u_n + \mathcal{A}(u_n) + \mathcal{B}_n(u_n) = f \quad \text{in } \mathcal{H}. \tag{3.11}$$

Test both sides by u_n and employ the monotonicity of $\mathcal{A}$ and $\mathcal{B}$ to get

$$\|u_n\|_{\mathcal{H}} \leq \|f\|_{\mathcal{H}},$$

which yields, up to a (not relabeled) subsequence,

$$u_n \to u \quad \text{weakly in } \mathcal{H}.$$

Multiplying the both sides of (3.11) by $\mathcal{A}(u_n)$, we have

$$(u_n, \mathcal{A}(u_n))_{\mathcal{H}} + \|\mathcal{A}(u_n)\|_{\mathcal{H}}^2 + (\mathcal{B}_n(u_n), \mathcal{A}(u_n))_{\mathcal{H}} = (f, \mathcal{A}(u_n))_{\mathcal{H}},$$

which along with (3.10) implies

$$\|\mathcal{A}(u_n)\|_{\mathcal{H}} \leq \|f\|_{\mathcal{H}}.$$

Multiply the both sides by $\mathcal{B}_n(u_n)$. It also follows by (3.10) that

$$\|\mathcal{B}_n(u_n)\|_{\mathcal{H}} \leq \|f\|_{\mathcal{H}}.$$

Moreover, since the graphs of $\mathcal{A}$ and $\mathcal{B}$ are weakly closed in $\mathcal{H}$ (by the weak closedness of linear maximal monotone operators), we assure that

$$\begin{aligned}\mathcal{A}(u_n) &\to \mathcal{A}(u) && \text{weakly in } \mathcal{H},\\ \mathcal{J}_n u_n &\to u && \text{weakly in } \mathcal{H},\\ \mathcal{B}_n(u_n) &\to \mathcal{B}(u) && \text{weakly in } \mathcal{H}.\end{aligned}$$

Indeed, we note that $\mathcal{J}_n u_n = u_n - (1/n)\mathcal{B}_n(u_n)$, which derives the second assertion. Moreover, the third assertion follows from $\mathcal{B}_n(u_n) = \mathcal{B}(\mathcal{J}_n u_n)$. Hence, the limit u fulfills $u + \mathcal{A}(u) + \mathcal{B}(u) = f$. Consequently, $\mathcal{A} + \mathcal{B}$ is maximal monotone in $\mathcal{H}$. $\square$

We also obtain the following corollary, which will be used later to derive a priori estimates.

COROLLARY 3.6. *Under the assumption* (K), *it holds that*

$$\begin{aligned}\int_s^t (\mathcal{A}(u)(\tau), \mathcal{B}(u)(\tau))_H \, \mathrm{d}\tau \\ \geq \frac{1}{2}\left(\ell * \|B(u)(\cdot)\|_H^2\right)(t) - \frac{1}{2}\left(\ell * \|B(u)(\cdot)\|_H^2\right)(s) \\ + \frac{1}{2}\int_s^t \ell(\tau)\|\mathcal{B}(u)(\tau)\|_H^2 \, \mathrm{d}\tau \end{aligned} \tag{3.12}$$

for a.e. $s, t \in (0, T)$ *with* $s < t$ *and* $u \in D(\mathcal{A})$. *In particular,* $\ell(\cdot)\|\mathcal{B}(u)(\cdot)\|_H^2$ *is integrable in* $(0, T)$, *and moreover,*

$$\begin{aligned}\int_0^t (\mathcal{A}(u)(\tau), \mathcal{B}(u)(\tau))_H \, \mathrm{d}\tau \\ \geq \frac{1}{2}\left(\ell * \|B(u)(\cdot)\|_H^2\right)(t) + \frac{1}{2}\int_0^t \ell(\tau)\|\mathcal{B}(u)(\tau)\|_H^2 \, \mathrm{d}\tau \end{aligned} \tag{3.13}$$

for all $t \in [0, T]$ *and* $u \in D(\mathcal{A})$. *Furthermore, for all* $u \in D(\mathcal{A})$, *the function* $t \mapsto (\ell * \|\mathcal{B}(u)(\cdot)\|_H^2)(t)$ *is differentiable a.e. in* $(0, T)$, *and hence, it holds that*

$$(\mathcal{A}(u)(t), \mathcal{B}(u)(t))_H \geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\ell * \|\mathcal{B}(u)(\cdot)\|_H^2\right)(t) + \frac{1}{2}\ell(t)\|\mathcal{B}(u)(t)\|_H^2 \tag{3.14}$$

for a.e. $t \in (0, T)$. *In addition, suppose that* ℓ *is of class* $\mathcal{K}^1(\alpha, \theta)$ *for some* $\alpha \in (0, 1)$ *and* $\theta \in (0, \pi)$. *Then the function* $t \mapsto (\ell * \|\mathcal{B}(u)(\cdot)\|_H^2)(t)$ *belongs to* $W^{1,1}(0, T)$ *(hence it is absolutely continuous on* $[0, T]$*) and vanishes at* $t = 0$.

Proof. Recall (3.9) in the proof of Proposition 3.5. Integrate both sides of (3.9) over (s, t), $0 < s < t < T$, to observe that

$$\begin{aligned}\int_s^t (\mathcal{A}(u)(\tau), \mathcal{B}(u)(\tau))_H \, \mathrm{d}\tau \\ \geq \frac{1}{2}\left(\ell_n * \|\mathcal{B}(u)(\cdot)\|_H^2\right)(t) - \frac{1}{2}\left(\ell_n * \|\mathcal{B}(u)(\cdot)\|_H^2\right)(s) + \frac{1}{2}\int_s^t \ell_n(\tau)\|\mathcal{B}(u)(\tau)\|_H^2 \, \mathrm{d}\tau \\ - \int_0^T |h_n(\tau)| \, \mathrm{d}\tau.\end{aligned}$$

Here we note that $\left(\ell_n * \|\mathcal{B}(u)(\cdot)\|_H^2\right)(s) = 0$ if $s = 0$. Since $\ell_n \to \ell$ strongly in $L^1(0, T)$ and $\mathcal{B}(u) \in D(\mathcal{B}_\ell) \subset L^2(0, T; H)$, by Fatou's lemma, we have $\ell(\cdot)\|\mathcal{B}(u)(\cdot)\|_H^2 \in L^1(0, T)$ and

$$\begin{aligned}\int_s^t (\mathcal{A}(u)(\tau), \mathcal{B}(u)(\tau))_H \, \mathrm{d}\tau \\ \geq \frac{1}{2}\left(\ell * \|\mathcal{B}(u)(\cdot)\|_H^2\right)(t) - \frac{1}{2}\left(\ell * \|\mathcal{B}(u)(\cdot)\|_H^2\right)(s) + \frac{1}{2}\int_s^t \ell(\tau)\|\mathcal{B}(u)(\tau)\|_H^2 \, \mathrm{d}\tau\end{aligned}$$

for a.e. $0 < s < t < T$ (if $s = 0$, then the second term of the right-hand side can be neglected). It also implies that the function $t \mapsto (\ell * \|\mathcal{B}(u)(\cdot)\|_H^2)(t)$ is differentiable a.e. in $(0, T)$. Therefore dividing both sides by $t - s$ and taking a limit as $s \to t - 0$, we obtain (3.14) for a.e. $t \in (0, T)$.

If $\ell \in \mathcal{K}^1(\alpha,\theta)$ and $v := \mathcal{B}(u) \in D(\mathcal{B}_\ell)$, we deduce by [51, Proposition 2.1] (see also Remark 3.2) that the function $t \mapsto (\ell * \|v\|_H^2)(t)$ belongs to $W^{1,1}(0,T)$ and vanishes at $t = 0$. □

4. Proofs of main results

This section is devoted to proving main results stated in §2. We first assume that

$$f \in W^{1,2}(0,T;H) \quad \text{and} \quad u_0 \in D(\varphi),$$

which will be always assumed until the end of §4.4. Furthermore, we also write $\mathcal{H} = L^2(0,T;H)$ and use the same notation $\mathcal{A}$ and $\mathcal{B}$ as in §3.

4.1. Approximate problems.

For $\lambda \in (0,1)$, we consider the following approximate problems:

$$(\lambda\mathcal{A} + \mathcal{B})(u_\lambda - u_0) + \partial\Phi_\lambda(u_\lambda) \ni f \quad \text{in } \mathcal{H}, \tag{4.1}$$

where $\Phi_\lambda : \mathcal{H} \to [0,+\infty)$ is defined by

$$\Phi_\lambda(w) := \int_0^T \varphi_\lambda(w(t))\,\mathrm{d}t \quad \text{for} \quad w \in \mathcal{H}$$

and $\varphi_\lambda : H \to [0,+\infty)$ stands for the Moreau-Yosida regularization of φ defined by

$$\varphi_\lambda(w) := \min_{z\in H}\left(\frac{1}{2\lambda}\|w - z\|_H^2 + \varphi(z)\right) \quad \text{for} \quad w \in H.$$

Let us recall that the minimum of the above is attained by $J_\lambda w$, where $J_\lambda := (I + \lambda\partial\varphi)^{-1}$, and moreover, φ_λ is Fréchet differentiable in H and its derivative coincides with the Yosida approximation of $\partial\varphi$ (see, e.g., [8, Proposition 2.11], for more details). So we denote by $\partial\varphi_\lambda$ the derivative of φ_λ as well as the Yosida approximation of $\partial\varphi$.

As in Proposition 3.5, one can check that the sum $\lambda\mathcal{A}+\mathcal{B}$ is maximal monotone in $\mathcal{H}$. Moreover, the (translated) operator $w \mapsto \partial\Phi_\lambda(w+u_0)$ is maximal monotone and $D(\partial\Phi_\lambda) = \mathcal{H}$ (see, e.g., [8, Proposition 2.16]). Then the sum $\lambda\mathcal{A}+\mathcal{B}+\partial\Phi_\lambda(\,\cdot\,+u_0)$ turns out to be maximal monotone in $\mathcal{H}$. Furthermore, it is also surjective in $\mathcal{H}$, since $\mathcal{A}$ is coercive in $\mathcal{H}$ (see, e.g., [8, Chap.II, §5], [4]); indeed, for any $\varepsilon > 0$, one can take $C_\varepsilon > 0$ such that, for all $w \in D(\mathcal{A})$,

$$\begin{aligned}\frac{1}{2}\|w(t)\|_H^2 &= \frac{1}{2}\int_0^t \frac{\mathrm{d}}{\mathrm{d}t}\|w(s)\|_H^2\,\mathrm{d}s\\ &= \int_0^t (w'(s), w(s))_H\,\mathrm{d}s \le \varepsilon t \sup_{s\in[0,t]}\|w(s)\|_H^2 + C_\varepsilon\int_0^t \|w'(s)\|_H^2\,\mathrm{d}s,\end{aligned}$$

which implies

$$\sup_{t\in[0,T]}\|w(t)\|_H^2 \le C\int_0^T \|w'(s)\|_H^2\,\mathrm{d}s \quad \text{for all} \quad w \in D(\mathcal{A})$$

(in particular, $\mathcal{A}$ is coercive in $\mathcal{H}$). Thus for each $\lambda > 0$ and $f \in \mathcal{H}$, we obtain a unique solution $u_\lambda \in W^{1,2}(0,T;H)$ of (4.1) such that $u_\lambda - u_0 \in D(\mathcal{A})$.

Here it is noteworthy that the regularization term $\lambda\mathcal{A}$ is used not only for deriving the coercivity (indeed, $\mathcal{B}$ is also coercive in $\mathcal{H}$) but also for guaranteeing a regularity of approximate solutions, i.e., $u_\lambda \in W^{1,2}(0,T;H)$. Then $\mathcal{B}(u_\lambda - u_0)$ coincides with $\mathcal{C}(u_\lambda)$, and hence, one can employ fine properties of both nonlocal derivatives.

4.2. **A priori estimates.** We next establish a priori estimates. Fix $v_0 \in D(\varphi) \neq \emptyset$.[1] Let us first test (4.1) by $u_\lambda - v_0 \in D(\mathcal{A})$. Then

$$\begin{aligned}&\frac{\lambda}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_\lambda(t) - u_0\|_H^2 + \left(\frac{\mathrm{d}}{\mathrm{d}t}[k * (u_\lambda - u_0)](t), u_\lambda(t) - u_0\right)_H + \varphi_\lambda(u_\lambda(t))\\
&\leq \varphi_\lambda(v_0) + (f(t), u_\lambda(t) - v_0)_H + (\lambda\mathcal{A}(u_\lambda - u_0) + \mathcal{B}(u_\lambda - u_0), v_0 - u_0)_H .\end{aligned}$$

Here by Lemma 3.3, we note that

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}t}[k * (u_\lambda - u_0)](t), u_\lambda(t) - u_0\right)_H &= \left(\frac{\mathrm{d}}{\mathrm{d}t}[k_n * (u_\lambda - u_0)](t), u_\lambda(t) - u_0\right)_H + h_n(t)\\
&\geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(k_n * \|u_\lambda - u_0\|_H^2\right)(t) + h_n(t),\end{aligned}$$

where $h_n(t)$ is given by

$$h_n(t) := \left(\frac{\mathrm{d}}{\mathrm{d}t}\left[(k - k_n) * (u_\lambda - u_0)\right](t), u_\lambda(t) - u_0\right)_H .$$

Since $k_n \to k$ strongly in $L^1(0,T)$, $t \mapsto \|u_\lambda(t) - u_0\|_H^2 \in W^{1,1}(0,T)$ and $\|u_\lambda(0) - u_0\|_H^2 = 0$, one can deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t}(k_n * \|u_\lambda - u_0\|_H^2) \to \frac{\mathrm{d}}{\mathrm{d}t}(k * \|u_\lambda - u_0\|_H^2) \quad \text{strongly in } L^1(0,T),$$

and moreover, as in the proof of Proposition 3.5,

$$h_n \to 0 \quad \text{strongly in } L^1(0,T) \quad \text{as } n \to +\infty.$$

Thus, passing to the limit as $n \to \infty$, we obtain

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u_\lambda - u_0)\right](t), u_\lambda(t) - u_0\right)_H \geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(k * \|u_\lambda - u_0\|_H^2\right)(t)$$

for a.e. $t \in (0,T)$. Moreover, we see that

$$\begin{aligned}&(\lambda\mathcal{A}(u_\lambda - u_0) + \mathcal{B}(u_\lambda - u_0), v_0 - u_0)_H\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\big(\lambda(u_\lambda(t) - u_0) + [k * (u_\lambda - u_0)](t), v_0 - u_0\big)_H.\end{aligned}$$

[1]For simplicity, one may also take $v_0 = u_0$

It follows that

$$\begin{aligned}
&\frac{\lambda}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_\lambda(t)-u_0\|_H^2+\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(k*\|u_\lambda-u_0\|_H^2\right)(t)+\varphi_\lambda(u_\lambda(t))\\
&\le \varphi_\lambda(v_0)+(f(t),u_\lambda(t)-v_0)_H\\
&\quad+\frac{\mathrm{d}}{\mathrm{d}t}\big(\lambda(u_\lambda(t)-u_0)+[k*(u_\lambda-u_0)](t),v_0-u_0\big)_H
\end{aligned}\tag{4.2}$$

for a.e. $t\in(0,T)$. Integrating both sides of (4.2) over $(0,t)$ and using $\|u_\lambda(0)-u_0\|_H^2=(k*\|u_\lambda-u_0\|_H^2)(0)=0$ and $u_\lambda(0)-u_0=[k*(u_\lambda-u_0)](0)=0$, we assure that

$$\begin{aligned}
&\frac{\lambda}{2}\|u_\lambda(t)-u_0\|_H^2+\frac{1}{2}\left(k*\|u_\lambda-u_0\|_H^2\right)(t)+\int_0^t\varphi_\lambda(u_\lambda(\tau))\,\mathrm{d}\tau\\
&\le T\varphi(v_0)+\int_0^t\|f(\tau)\|_H\|u_\lambda(\tau)-v_0\|_H\,\mathrm{d}\tau\\
&\quad+\big(\lambda(u_\lambda(t)-u_0)+[k*(u_\lambda-u_0)](t),v_0-u_0\big)_H
\end{aligned}\tag{4.3}$$

for all $t\in[0,T]$. Here we further note that

$$\begin{aligned}
&\big(\lambda(u_\lambda(t)-u_0)+[k*(u_\lambda-u_0)](t),v_0-u_0\big)_H\\
&\le\frac{\lambda}{4}\|u_\lambda(t)-u_0\|_H^2+\frac{1}{4}\left(k*\|u_\lambda-u_0\|_H^2\right)(t)+\left(1+\|k\|_{L^1(0,T)}\right)\|v_0-u_0\|_H^2.
\end{aligned}$$

Hence it particularly follows that

$$\frac{1}{4}\left(k*\|u_\lambda-u_0\|_H^2\right)(t)\le\int_0^t\|f(\tau)\|_H\|u_\lambda(\tau)-u_0\|_H\,\mathrm{d}\tau+C.$$

Convolving both sides with ℓ, we infer that

$$\begin{aligned}
\frac{1}{4}\int_0^T\|u_\lambda(t)-u_0\|_H^2\,\mathrm{d}t&\le\left(\ell*\int_0^t\|f(\tau)\|_H\|u_\lambda(\tau)-u_0\|_H\,\mathrm{d}\tau\right)(t)+C\\
&\le\|\ell\|_{L^1(0,T)}\|f\|_{L^2(0,T;H)}\|u_\lambda-u_0\|_{L^2(0,T;H)}+C,
\end{aligned}$$

which yields

$$\frac{1}{8}\int_0^T\|u_\lambda(t)-u_0\|_H^2\,\mathrm{d}t\le2\|\ell\|_{L^1(0,T)}^2\|f\|_{L^2(0,T;H)}^2+C.$$

Therefore, recalling (4.3) again, we obtain

$$\begin{aligned}
&\lambda\sup_{t\in[0,T]}\|u_\lambda(t)-u_0\|_H^2+\sup_{t\in[0,T]}\left(k*\|u_\lambda-u_0\|_H^2\right)(t)\\
&\quad+\int_0^T\varphi_\lambda(u_\lambda(\tau))\,\mathrm{d}\tau+\int_0^T\|u_\lambda(\tau)-u_0\|_H^2\,\mathrm{d}\tau\le C,
\end{aligned}\tag{4.4}$$

where C depends on T, $\|f\|_{\mathcal{H}}$, $\|k\|_{L^1(0,T)}$, $\|\ell\|_{L^1(0,T)}$, $\|u_0\|_H$ and $\varphi(v_0)$, $\|v_0\|_H$, but it is independent of λ.

We next test (4.1) by $u_\lambda'(t) = \mathcal{A}(u_\lambda - u_0)(t)$ and integrate both sides over $(0,t)$. Then we see by (3.13) of Corollary 3.6 that

$$\begin{aligned}
&\lambda \int_0^t \|\mathcal{A}(u_\lambda - u_0)(\tau)\|_H^2 \, \mathrm{d}\tau + \frac{1}{2}\left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t) + \varphi_\lambda(u_\lambda(t)) \\
&\le \varphi_\lambda(u_0) + \int_0^t \left(f(\tau), u_\lambda'(\tau)\right)_H \, \mathrm{d}\tau \\
&\le \varphi(u_0) + \int_0^t \frac{\mathrm{d}}{\mathrm{d}\tau}\left(f(\tau), u_\lambda(\tau)\right)_H \, \mathrm{d}\tau - \int_0^t \left(f'(\tau), u_\lambda(\tau)\right)_H \, \mathrm{d}\tau
\end{aligned} \tag{4.5}$$

for all $t \in [0,T]$. Here we also used the chain-rule for subdifferentials (see [8, Lemma 3.3]),

$$\left(\partial\Phi_\lambda(u_\lambda)(t), \mathcal{A}(u_\lambda - u_0)(t)\right)_H = \left(\partial\varphi_\lambda(u_\lambda(t)), u_\lambda'(t)\right)_H = \frac{\mathrm{d}}{\mathrm{d}t}\varphi_\lambda(u_\lambda(t)).$$

Moreover, note by $u_\lambda - u_0 \in D(\mathcal{A}) \subset D(\mathcal{B})$ that

$$\begin{aligned}
\|u_\lambda(t) - u_0\|_H &= \left\| \frac{\mathrm{d}}{\mathrm{d}t}\left[\ell * k * (u_\lambda - u_0)\right](t) \right\|_H \\
&= \left\| \int_0^t \ell(t-s) \left(\frac{\mathrm{d}}{\mathrm{d}t}[k * (u_\lambda - u_0)] \right)(s) \, \mathrm{d}s \right\|_H \\
&\le \|\ell\|_{L^1(0,t)}^{1/2} \sqrt{\left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t)},
\end{aligned}$$

which implies

$$\|\ell\|_{L^1(0,T)}^{-1} \|u_\lambda(t) - t\|_H^2 \le \left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t).$$

In particular, it follows that

$$\begin{aligned}
&\lambda \int_0^t \|\mathcal{A}(u_\lambda - u_0)(\tau)\|_H^2 \, \mathrm{d}\tau + \frac{1}{4}\|\ell\|_{L^1(0,T)}^{-1} \|u_\lambda(t) - t\|_H^2 \\
&\quad + \frac{1}{4}\left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t) + \varphi_\lambda(u_\lambda(t)) \\
&\le \varphi(u_0) + \left(f(t), u_\lambda(t)\right)_H - \left(f(0), u_0\right)_H - \int_0^t \left(f'(\tau), u_\lambda(\tau)\right)_H \, \mathrm{d}\tau.
\end{aligned}$$

By simple calculation along with (4.4) and the fact that $f \in W^{1,2}(0,T;H) \hookrightarrow C([0,T];H)$, we obtain

$$\begin{aligned}
&\lambda \int_0^t \|\mathcal{A}(u_\lambda - u_0)(\tau)\|_H^2 \, \mathrm{d}\tau + \|u_\lambda(t) - u_0\|_H^2 \\
&\quad + \left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t) + \varphi_\lambda(u_\lambda(t)) \le C
\end{aligned} \tag{4.6}$$

for any $t \in [0,T]$. Convolving both sides with k and recalling that $\ell * k = 1$, we infer that

$$\int_0^T \|\mathcal{B}(u_\lambda - u_0)(\tau)\|_H^2 \, \mathrm{d}\tau \le C. \tag{4.7}$$

By comparison of terms in (4.1), it follows that

$$\|\partial \Phi_\lambda(u_\lambda)\|_{\mathcal{H}}^2 = \int_0^T \|\partial \varphi_\lambda(u_\lambda(\tau))\|_H^2 \, \mathrm{d}\tau \leq C. \tag{4.8}$$

4.3. **Convergence of approximate solutions.** From the preceding a priori estimates (4.4), (4.6)–(4.8), we assure, up to a (not relabeled) subsequence, that

$$\begin{aligned} u_\lambda &\to u && \text{weakly in } \mathcal{H}, \\ \partial \Phi_\lambda(u_\lambda) &\to \xi && \text{weakly in } \mathcal{H}, \\ \lambda \mathcal{A}(u_\lambda - u_0) &\to 0 && \text{strongly in } \mathcal{H}, \\ \mathcal{B}(u_\lambda - u_0) &\to \mathcal{B}(u - u_0) && \text{weakly in } \mathcal{H}, \end{aligned} \tag{4.9}$$

which yields $\mathcal{B}(u-u_0) + \xi = f$. Here we used the weak closedness of the linear maximal monotone operator $\mathcal{B}$ in $\mathcal{H}$ to identify the limit of $\mathcal{B}(u_\lambda - u_0)$.

We next show that (u_λ) forms a Cauchy sequence in $\mathcal{H}$. Let u_λ and u_μ be solutions to (4.1) with parameters λ and μ, respectively, and set $w = u_\lambda - u_\mu \in D(\mathcal{A})$. By subtracting equations,

$$\frac{\mathrm{d}}{\mathrm{d}t} (k * w)(t) + \partial \varphi_\lambda(u_\lambda(t)) - \partial \varphi_\mu(u_\mu(t)) = \mu u_\mu'(t) - \lambda u_\lambda'(t).$$

Multiply both sides by $w(t)$ and apply the so-called *Kōmura's trick* (see, e.g., [8, p.56]) to deal with the term $(\partial\varphi_\lambda(u_\lambda(t)) - \partial\varphi_\mu(u_\mu(t)), u_\lambda(t) - u_\mu(t))_H$. Then it follows that

$$\begin{aligned} \left(\frac{\mathrm{d}}{\mathrm{d}t} (k * w)(t), w(t) \right)_H &\leq \left(\mu \|u_\mu'(t)\|_H + \lambda \|u_\lambda'(t)\|_H \right) \|w(t)\|_H \\ &\quad + \frac{\lambda + \mu}{4} \left(\|\partial\varphi_\lambda(u_\lambda(t))\|_H^2 + \|\partial\varphi_\mu(u_\mu(t))\|_H^2 \right). \end{aligned}$$

Moreover, note that

$$\begin{aligned} \left(\frac{\mathrm{d}}{\mathrm{d}t} (k * w)(t), w(t) \right)_H &= \left(\frac{\mathrm{d}}{\mathrm{d}t} (k_n * w)(t), w(t) \right)_H + \hat{h}_n \\ &\geq \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \left(k_n * \|w\|_H^2 \right)(t) + \hat{h}_n, \end{aligned}$$

where $\hat{h}_n$ is given by

$$\hat{h}_n(\cdot) := \left(\frac{\mathrm{d}}{\mathrm{d}t} [(k - k_n) * w](\cdot), w(\cdot) \right)_H \to 0 \quad \text{strongly in } L^1(0,T).$$

Here we used the fact that $w = u_\lambda - u_\mu \in D(\mathcal{A})$. Moreover, note that

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(k_n * \|w\|_H^2 \right) \to \frac{\mathrm{d}}{\mathrm{d}t} \left(k * \|w\|_H^2 \right) \quad \text{strongly in } L^1(0,T)$$

due to the fact that $\|w\|_H^2 \in W^{1,1}(0,T)$ and $\|w(0)\|_H^2 = 0$. Hence combining all these facts and letting $n \to \infty$, one deduces that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(k * \|w\|_H^2\right)(t) \leq \left(\mu\|u_\mu'(t)\|_H + \lambda\|u_\lambda'(t)\|_H\right)\|w(t)\|_H + \frac{\lambda+\mu}{4}\left(\|\partial\varphi_\lambda(u_\lambda(t))\|_H^2 + \|\partial\varphi_\mu(u_\mu(t))\|_H^2\right)$$

for a.e. $t \in (0,T)$. Integrating both sides over $(0,t)$ and employing $(k * \|w\|_H^2)(0) = 0$, we deduce that

$$\frac{1}{2}\left(k * \|w\|_H^2\right)(t) \leq \left(\mu\|u_\mu'\|_{\mathcal{H}} + \lambda\|u_\lambda'\|_{\mathcal{H}}\right)\|w\|_{\mathcal{H}} + \frac{\lambda+\mu}{4}\left(\|\partial\Phi_\lambda(u_\lambda)\|_{\mathcal{H}}^2 + \|\partial\Phi_\mu(u_\mu)\|_{\mathcal{H}}^2\right) \text{ for all } t \in [0,T],$$

which together with (K), (4.4), (4.6) and (4.8) implies

$$\|w\|_{\mathcal{H}}^2 \leq C\left(\sqrt{\mu} + \sqrt{\lambda}\right) + C(\lambda+\mu) \to 0$$

as $\lambda, \mu \to 0$. Therefore (u_λ) forms a Cauchy sequence in $\mathcal{H}$. Thus we conclude that

$$u_\lambda \to u \quad \text{strongly in } \mathcal{H}, \tag{4.10}$$

which along with (4.8) implies

$$J_\lambda u_\lambda \to u \quad \text{strongly in } \mathcal{H}. \tag{4.11}$$

Due to the demiclosedness of maximal monotone operators, we infer that

$$u \in D(\partial\Phi) \text{ and } \xi \in \partial\Phi(u),$$

where $\Phi : \mathcal{H} \to [0,+\infty]$ is defined by

$$\Phi(w) = \begin{cases} \int_0^T \varphi(w(t))\,\mathrm{d}t & \text{if } \varphi(w(\cdot)) \in L^1(0,T), \\ +\infty & \text{otherwise} \end{cases} \quad \text{for } w \in \mathcal{H},$$

and hence, by [8, Proposition 2.16],

$$u(t) \in D(\partial\varphi) \text{ and } \xi(t) \in \partial\varphi(u(t)) \text{ for a.e. } t \in (0,T).$$

Consequently, the limit u satisfies the relation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u-u_0)\right](t) + \partial\varphi(u(t)) \ni f(t) \text{ in } H, \quad 0 < t < T.$$

4.4. **Energy inequality.** We next prove (2.7) by recalling (4.5). By (4.10) and (4.11), one can further take a (not relabeled) subsequence of (u_λ) and a set $I \subset (0,T)$ satisfying $|(0,T) \setminus I| = 0$ such that

$$u_\lambda(t) \to u(t) \quad \text{strongly in } H,$$
$$J_\lambda u_\lambda(t) \to u(t) \quad \text{strongly in } H$$

for all $t \in I$. Hence we have

$$\begin{aligned}
&\frac{1}{2}\left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t) + \varphi_\lambda(u_\lambda(t)) \\
&\le \varphi(u_0) + (f(t), u_\lambda(t))_H - (f(0), u_0)_H - \int_0^t \left(f'(\tau), u_\lambda(\tau)\right)_H \,\mathrm{d}\tau \\
&\to \varphi(u_0) + (f(t), u(t))_H - (f(0), u_0)_H - \int_0^t \left(f'(\tau), u(\tau)\right)_H \,\mathrm{d}\tau
\end{aligned} \tag{4.12}$$

for all $t \in I$. From the lower semicontinuity of φ, it follows that

$$\liminf_{\lambda \to 0} \varphi_\lambda(u_\lambda(t)) \ge \liminf_{\lambda \to 0} \varphi(J_\lambda u_\lambda(t)) \ge \varphi(u(t)) \quad \text{for all } \ t \in I.$$

Moreover, estimate the first term of the left-hand side from below as follows:

$$\begin{aligned}
\liminf_{\lambda \to 0} \left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t) &= \liminf_{\lambda \to 0} \int_0^t \ell(t-\tau)\|\mathcal{B}(u_\lambda - u_0)(\tau)\|_H^2 \,\mathrm{d}\tau \\
&\ge \int_0^t \ell(t-\tau)\|\mathcal{B}(u - u_0)(\tau)\|_H^2 \,\mathrm{d}\tau \qquad (4.13) \\
&= \left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t)
\end{aligned}$$

for all $t \in I$. Indeed, for each $t \in I$, we find by (4.12) that

$$\int_0^t \left\|\sqrt{\ell(t-\tau)}\,\mathcal{B}(u_\lambda - u_0)(\tau)\right\|_H^2 \mathrm{d}\tau = \left(\ell * \|\mathcal{B}(u_\lambda - u_0)(\cdot)\|_H^2\right)(t) \le C,$$

whence follows, up to a subsequence (which may depend on t and will be not relabeled), that

$$\sqrt{\ell(t-\cdot)}\,\mathcal{B}(u_\lambda - u_0) \to \zeta \quad \text{weakly in } L^2(0,t;H)$$

for some $\zeta \in L^2(0,t;H)$. We next identify the limit ζ. For any $z \in C_c^\infty((0,t);H)$, let us take $\delta > 0$ such that $\operatorname{supp} z \subset (\delta, t-\delta)$ and observe that

$$\begin{aligned}
&\int_0^t \left(\sqrt{\ell(t-\tau)}\,\mathcal{B}(u_\lambda - u_0)(\tau), z(\tau)\right)_H \mathrm{d}\tau \\
&= \int_\delta^{t-\delta} \left(\sqrt{\ell(t-\tau)}\,\mathcal{B}(u_\lambda - u_0)(\tau), z(\tau)\right)_H \mathrm{d}\tau \\
&= \int_\delta^{t-\delta} \left(\mathcal{B}(u_\lambda - u_0)(\tau), \sqrt{\ell(t-\tau)}\,z(\tau)\right)_H \mathrm{d}\tau \\
&\overset{(4.9)}{\to} \int_\delta^{t-\delta} \left(\mathcal{B}(u - u_0)(\tau), \sqrt{\ell(t-\tau)}\,z(\tau)\right)_H \mathrm{d}\tau \\
&= \int_0^t \left(\sqrt{\ell(t-\tau)}\,\mathcal{B}(u - u_0)(\tau), z(\tau)\right)_H \mathrm{d}\tau.
\end{aligned}$$

Here we used the fact that $0 < \ell(t-\tau) \le \ell(\delta) < +\infty$ for $\tau \in (\delta, t-\delta)$. Thus we deduce that $\zeta = \sqrt{\ell(t-\cdot)}\,\mathcal{B}(u - u_0)$. Therefore, the weak lower-semicontinuity of norm yields the inequality in (4.13). Combining all these facts, we derive (2.7). Furthermore, repeating a similar argument to (4.6), one can also verify (2.6).

4.5. **Uniqueness and continuous dependence on initial data.** In this subsection, we shall prove the uniqueness of solutions for (1.1). The uniqueness of strong solutions has more of a significance for fractional gradient flows. Indeed, in Definition 2.1, it is still not clear whether each solution satisfies the initial condition $u(0) = u_0$ (or $u(0_+) = u_0$) in a classical sense (on the other hand, under some additional integrability of ℓ, one can check it. See Proposition 2.5). However, the uniqueness ensures that each solution is uniquely determined by specifying an initial datum as in (1.1).

Assume that $f_1, f_2 \in L^2(0,T;H)$ and $u_{0,1}, u_{0,2} \in H$, and moreover, u_1 and u_2 are strong solutions of (1.1) with (u_0, f) replaced by $(u_{0,1}, f_1)$ and $(u_{0,2}, f_2)$, respectively, in the sense of Definition 2.1. Set $w := u_1 - u_2$ and $w_0 := u_{0,1} - u_{0,2}$. Then we observe that $w - w_0 \in D(\mathcal{B})$. Subtracting equations, we see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[k * (w - w_0)\right](t) + \partial\varphi(u_1(t)) - \partial\varphi(u_2(t)) \ni f_1(t) - f_2(t) \quad \text{in } H$$

for a.e. $t \in (0,T)$. Multiplying both sides by w and using the monotonicity of $\partial\varphi$, one can derive

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(k_n * \|w - w_0\|_H^2\right)(t)\\
&\leq (f_1(t) - f_2(t), w(t))_H - \left(\frac{\mathrm{d}}{\mathrm{d}t}\left[(k - k_n) * (w - w_0)\right](t), w(t) - w_0\right)_H\\
&\quad - \left(\frac{\mathrm{d}}{\mathrm{d}t}\left[k * (w - w_0)\right](t), w_0\right)_H
\end{aligned}$$

for a.e. $t \in (0,T)$. The integration of both sides over $(0,t)$ implies

$$\begin{aligned}
&\frac{1}{2}\left(k_n * \|w - w_0\|_H^2\right)(t)\\
&\leq \int_0^t \|f_1(\tau) - f_2(\tau)\|_H \|w(\tau) - w_0\|_H \,\mathrm{d}\tau + \left\|\frac{\mathrm{d}}{\mathrm{d}t}\left[(k - k_n) * (w - w_0)\right]\right\|_{\mathcal{H}} \|w - w_0\|_{\mathcal{H}}\\
&\quad - ([k * (w - w_0)](t), w_0)_H + \|f_1 - f_2\|_{L^1(0,t;H)} \|w_0\|_H
\end{aligned}$$

for $t \in [0,T]$. Here we used the facts that $[k * (w - w_0)](0) = 0$ by definition and $\left(k_n * \|w(\cdot) - w_0\|_H^2\right)(0) = 0$ by $k_n \in W^{1,1}(0,T) \subset L^\infty(0,T)$ and $\|w(\cdot) - w_0\|_H^2 \in L^1(0,T)$. Recalling that $k_n \to k$ strongly in $L^1(0,T)$ and $\mathcal{B}_n(w - w_0) \to \mathcal{B}(w - w_0)$

strongly in $\mathcal{H}$ by $w - w_0 \in D(\mathcal{B})$ and using Fatou's lemma, one obtains

$$\begin{aligned}
&\frac{1}{2}\left(k * \|w - w_0\|_H^2\right)(t) \\
&\le \int_0^t \|f_1(\tau) - f_2(\tau)\|_H \|w(\tau) - w_0\|_H \,\mathrm{d}\tau + \|[k * (w - w_0)](t)\|_H \, \|w_0\|_H \\
&\quad + \|f_1 - f_2\|_{L^1(0,t;H)} \|w_0\|_H \\
&\le \left(\int_0^t \|f_1(\tau) - f_2(\tau)\|_H^2 \,\mathrm{d}\tau\right)^{1/2} \left(\int_0^t \|w(\tau) - w_0\|_H^2 \,\mathrm{d}\tau\right)^{1/2} \\
&\quad + \frac{1}{4}\left(k * \|w - w_0\|_H^2\right)(t) + \|k\|_{L^1(0,t)} \|w_0\|_H^2 + \|f_1 - f_2\|_{L^1(0,t;H)} \|w_0\|_H
\end{aligned}$$

for a.e. $t \in (0,T)$. Hence the convolution of both sides with ℓ further yields that

$$\begin{aligned}
\frac{1}{8}\int_0^T \|w(t) - w_0\|_H^2 \,\mathrm{d}t &\le 2\|\ell\|_{L^1(0,T)}^2 \|f_1 - f_2\|_{\mathcal{H}}^2 \\
&\quad + \|\ell\|_{L^1(0,T)} \left(\|k\|_{L^1(0,T)} \|w_0\|_H^2 + \|f_1 - f_2\|_{L^1(0,T;H)} \|w_0\|_H\right),
\end{aligned}$$

whence follows (2.8). In particular, if $u_{0,1} = u_{0,2}$ and $f_1 = f_2$, it then follows that

$$\int_0^T \|w(t)\|_H^2 \,\mathrm{d}t = 0,$$

which implies $w \equiv 0$, i.e., $u_1 \equiv u_2$. This completes the proof of uniqueness.

4.6. **Existence of a strong solution for $f \in L^2(0,T;H)$.** In this subsection, we shall discuss existence of strong solutions to (1.1) for $f \in L^2(0,T;H)$ and $u_0 \in D(\varphi)$. Then one can take a sequence (f_n) in $W^{1,2}(0,T;H)$ such that

$$f_n \to f \quad \text{strongly in } \mathcal{H}. \tag{4.14}$$

Let u_n be the unique strong solution to (1.1) with f replaced by f_n. Repeating a similar argument as before, one can derive

$$\sup_{t \in [0,T]} \left(k * \|u_n - u_0\|_H^2\right)(t) + \int_0^T \varphi(u_n(\tau)) \,\mathrm{d}\tau + \int_0^T \|u_n(\tau) - u_0\|_H^2 \,\mathrm{d}\tau \le C. \tag{4.15}$$

Now, let us test (1.1) by $\mathcal{B}_m(u_n - u_0) = (\mathrm{d}/\mathrm{d}t)[k_m * (u - u_0)]$. Then by Proposition 3.4, since $u_0 \in D(\varphi)$, $k_m' \le 0$ and $k_m \ge 0$, it follows that

$$\begin{aligned}
&\left(\frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u_n - u_0)\right](t), \frac{\mathrm{d}}{\mathrm{d}t}\left[k_m * (u_n - u_0)\right](t)\right)_H + \frac{\mathrm{d}}{\mathrm{d}t}\big[k_m * (\varphi(u_n(\cdot)) - \varphi(u_0))\big](t) \\
&\qquad \le \left(f_n(t), \frac{\mathrm{d}}{\mathrm{d}t}\left[k_m * (u_n - u_0)\right](t)\right)_H.
\end{aligned} \tag{4.16}$$

Here we note that $(\mathrm{d}/\mathrm{d}t)[k_m * \varphi(u_0)](t) = k_m(t)\varphi(u_0)$. Integrating both sides of (4.16) over $(0,t)$, we find that

$$\begin{aligned}&\int_0^t \left(\frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u_n - u_0)\right](\tau), \frac{\mathrm{d}}{\mathrm{d}t}\left[k_m * (u_n - u_0)\right](\tau) \right)_H \mathrm{d}\tau + \big[k_m * \varphi(u_n(\cdot))\big](t)\\ &\quad \le \int_0^t \left(f_n(\tau), \frac{\mathrm{d}}{\mathrm{d}t}\left[k_m * (u_n - u_0)\right](\tau) \right)_H \mathrm{d}\tau + \varphi(u_0) \int_0^t k_m(\tau)\,\mathrm{d}\tau.\end{aligned}$$

Letting $m \to +\infty$ and recalling that $k_m \to k$ strongly in $L^1(0,T)$, we deduce that

$$\begin{aligned}&\int_0^t \left\| \frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u_n - u_0)\right](\tau) \right\|_H^2 \mathrm{d}\tau + \big[k * \varphi(u_n(\cdot))\big](t)\\ &\quad \le \int_0^t \left(f_n(\tau), \frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u_n - u_0)\right](\tau) \right)_H \mathrm{d}\tau + \varphi(u_0) \int_0^t k(\tau)\,\mathrm{d}\tau \qquad (4.17)\end{aligned}$$

for all $t \in [0,T]$. Here we used the fact that $\mathcal{B}_m(u_n - u_0) \to \mathcal{B}(u_n - u_0)$ strongly in $\mathcal{H}$ and $k_m * \varphi(u_n(\cdot)) \to k * \varphi(u_n(\cdot))$ strongly in $C([0,T])$ as $m \to +\infty$ by $u_n - u_0 \in D(\mathcal{B})$ and $\varphi(u_n(\cdot)) \in L^\infty(0,T)$, respectively. By comparison of each term in (1.1), it also follows that

$$\int_0^T \|\xi_n(\tau)\|_H^2 \,\mathrm{d}\tau \le C,$$

where $\xi_n(t)$ is a section of $\partial\varphi(u_n(t))$ as in (2.1). Thus we obtain

$$\begin{aligned} u_n &\to u && \text{weakly in } \mathcal{H},\\ \mathcal{B}(u_n - u_0) &\to \mathcal{B}(u - u_0) && \text{weakly in } \mathcal{H},\\ \xi_n &\to \xi && \text{weakly in } \mathcal{H}, \end{aligned}$$

which along with (4.15) ensures $\varphi(u(\cdot)) \in L^1(0,T)$. Furthermore, by (2.8), we infer that

$$\|u_n - u_m\|_{\mathcal{H}} \le C\|f_n - f_m\|_{\mathcal{H}} \to 0$$

as $n, m \to +\infty$. Thus we observe that (u_n) forms a Cauchy sequence in $\mathcal{H}$, and therefore,

$$u_n \to u \quad \text{strongly in } \mathcal{H}.$$

Hence by virtue of the demiclosedness of maximal monotone operators, we conclude that $u(t) \in D(\partial\varphi)$ and $\xi(t) \in \partial\varphi(u(t))$ for a.e. $t \in (0,T)$, and therefore, u solves (1.1).

Finally, let us prove (2.4) along with (2.5). Recalling (4.16) with u_n and f_n replaced by u and f and integrating both sides over (s,t), we infer that the function

$$\mathcal{F}_m(t) := \Big[k_m * \big(\varphi(u(\cdot)) - \varphi(u_0)\big)\Big](t) - \int_0^t \left(f(\tau), \mathcal{B}_m(u - u_0)(\tau)\right)_H \mathrm{d}\tau$$

is nonincreasing and $\mathcal{F}_m(0) = 0$. Then by Helly's lemma, one can take a nonincreasing function $\mathcal{F} : [0,T] \to [-\infty, 0]$ such that

$$\mathcal{F}_m(t) \to \mathcal{F}(t) \quad \text{for all } \ t \in [0,T].$$

Then $\mathcal{F}$ is differentiable a.e. in $(0,T)$, and moreover,

$$\int_s^t \|\mathcal{B}(u-u_0)(\tau)\|_H^2 \, \mathrm{d}\tau + \mathcal{F}(t) \le \mathcal{F}(s) \quad \text{for all } \ 0 \le s < t \le T,$$

which together with the arbitrariness of s,t implies

$$\|\mathcal{B}(u-u_0)(t)\|_H^2 + \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(t) \le 0 \quad \text{for a.e. } t \in (0,T).$$

On the other hand, by $\varphi(u(\cdot)) \in L^1(0,T)$ and $u - u_0 \in D(\mathcal{B})$, one can observe that

$$\mathcal{F}_m(t) \to \Big[k * \big(\varphi(u(\cdot)) - \varphi(u_0)\big)\Big](t) - \int_0^t (f(\tau), \mathcal{B}(u-u_0)(\tau))_H \, \mathrm{d}\tau \tag{4.18}$$

for a.e. $t \in (0,T)$. Thus $\mathcal{F}$ is finite on $[0,T]$ and (2.4) and (2.5) follow. As for the case that $f \in W^{1,2}(0,T;H)$, since $\varphi(u(\cdot))$ belongs to $L^\infty(0,T)$, it follows that $k_m * \varphi(u(\cdot))) \to k * \varphi(u(\cdot))$ strongly in $C([0,T])$ by using (K), and hence, (4.18) holds for each $t \in [0,T]$. This completes the proof of Theorem 2.3. □

4.7. **Proof of Corollary 2.7.** Let (f_n) be a sequence in $C^\infty([0,T];H)$ such that

$$f_n \to f \quad \text{strongly in } \mathcal{H} \tag{4.19}$$

as $n \to +\infty$. Moreover, let u_λ be strong solutions on $[0,T]$ of the approximate problems (4.1) with f replaced by f_n. As in [11], convolve (4.1) with ℓ and rewrite it by (K) as a Volterra equation,

$$u_\lambda + \lambda^{-1}\ell * u_\lambda = u_0 + \lambda^{-1}\ell * J_\lambda u_\lambda + \lambda(\lambda^{-1}\ell) * (f_n - \lambda u_\lambda'), \quad J_\lambda := (I + \lambda\partial\varphi)^{-1}$$

whose solution u_λ can be also represented in terms of the resolvent kernel $r \in L^1_{\mathrm{loc}}([0,+\infty))$ of $\lambda^{-1}\ell$ and $s \in W^{1,1}_{\mathrm{loc}}([0,+\infty))$ satisfying

$$r + (\lambda^{-1}\ell) * r = \lambda^{-1}\ell, \quad s + (\lambda^{-1}\ell) * s = 1$$

as

$$u_\lambda = s u_0 + r * [J_\lambda u_\lambda + \lambda (f_n - \lambda u_\lambda')] . \tag{4.20}$$

Here we recall by (2.9) that $r \ge 0$ and $s \ge 0$ a.e. in $(0,+\infty)$ for any $\lambda > 0$ (see [11]). Let us denote by $u_{i,\lambda}$ the approximate solution corresponding to $u_{0,i}$ and $f_{i,n}$ for $i = 1,2$. The subtraction of equation (4.20) for $u_{1,\lambda}, u_{2,\lambda}$ implies

$$\begin{aligned}
&\|u_{1,\lambda}(t) - u_{2,\lambda}(t)\|_H \\
&\le s\|u_{0,1} - u_{0,2}\|_H + (r * \|u_{1,\lambda} - u_{2,\lambda}\|_H)(t) + \lambda (r * \|f_{1,n} - f_{2,n}\|_H)(t) \\
&\quad + \lambda^2 \left(r * \|u_{1,\lambda}' - u_{2,\lambda}'\|_H\right)(t)
\end{aligned}$$

for all $t \in [0,T]$. Here we used the non-expansivity of the resolvent J_λ in H. Convolving the above with $\lambda^{-1}\ell$ and adding it to the original one, we deduce that

$$\begin{aligned}
&\|u_{1,\lambda}(t) - u_{2,\lambda}(t)\|_H + \lambda^{-1}\left(\ell * \|u_{1,\lambda} - u_{2,\lambda}\|_H\right)(t) \\
&\le (s + \lambda^{-1}\ell * s)\|u_{0,1} - u_{0,2}\|_H + \left[\left(r + \lambda^{-1}\ell * r\right) * \|u_{1,\lambda} - u_{2,\lambda}\|_H\right](t) \\
&\quad + \lambda\left[\left(r + \lambda^{-1}\ell * r\right) * \|f_{1,n} - f_{2,n}\|_H\right](t) \\
&\quad + \lambda^2\left[\left(r + \lambda^{-1}\ell * r\right) * \|u'_{1,\lambda} - u'_{2,\lambda}\|_H\right](t) \\
&\le \|u_{0,1} - u_{0,2}\|_H + \lambda^{-1}\left(\ell * \|u_{1,\lambda} - u_{2,\lambda}\|_H\right)(t) + \left(\ell * \|f_{1,n} - f_{2,n}\|_H\right)(t) \\
&\quad + \lambda\left(\ell * \|u'_{1,\lambda} - u'_{2,\lambda}\|_H\right)(t).
\end{aligned}$$

As we have seen so far, by letting $\lambda \to +0$, it holds that $u_{i,\lambda} \to u_{i,n}$ and $\lambda u'_{i,\lambda} \to 0$ strongly in $\mathcal{H}$ where $u_{i,n}$ is the unique solution to (1.1) with u_0 and f replaced by $u_{0,i}$ and $f_{i,n}$ for $i = 1,2$. Thus noting that $\ell * \|u'_{1,\lambda} - u'_{2,\lambda}\|_H$ is bounded in $L^2(0,T)$, we obtain

$$\|u_{1,n}(t) - u_{2,n}(t)\|_H \le \|u_{0,1} - u_{0,2}\|_H + \left(\ell * \|f_{1,n} - f_{2,n}\|_H\right)(t)$$

for a.e. $t \in (0,T)$ and all $n \in \mathbb{N}$. Furthermore, by virtue of (4.19), one can check that $\ell * \|f_{1,n} - f_{2,n}\|_H \to \ell * \|f_1 - f_2\|_H$ strongly in $L^2(0,T)$. As in §4.6, for $i = 1,2$, we can verify that

$$u_{i,n} \to u_i \quad \text{strongly in } \mathcal{H}$$

where u_i is the unique strong solution of (1.1) with $u_0 = u_{0,i}$ and $f = f_i$ on $[0,T]$, as $n \to +\infty$. Thus we obtain (2.10) for a.e. $t \in (0,T)$. In particular, if $\ell \in L^2(0,T)$, then we deduce by Proposition 2.5 that $u \in C([0,T];H)$. Thus (2.10) holds for all $t \in [0,T]$, and the family of operators $S(t) : u_0 \in D(\varphi) \mapsto u(t) \in D(\varphi)$, where $u(\cdot)$ is the unique strong solution to (1.1) with the initial datum u_0, for $t \in [0,T]$ forms a non-expansive solution operator in H. □

4.8. **Proof of Theorem 2.8.** This subsection is devoted to a proof of Theorem 2.8. One can take sequences (f_n) and $(u_{0,n})$ in $W^{1,2}(0,T;H)$ and $D(\varphi)$, respectively, such that

$$f_n \to f \quad \text{strongly in } L^2(0,T;H) \ \text{ and } \ u_{0,n} \to u_0 \ \text{ strongly in } H \tag{4.21}$$

as $n \to +\infty$. Let u_n be the unique solution to (1.1) with f and u_0 replaced by f_n and $u_{0,n}$, respectively. We also denote by $\xi_n(t)$ the section of $\partial\varphi(u_n(t))$ as in (2.1). Fix $v_0 \in D(\varphi)$. Testing (1.1) with f_n and $u_{0,n}$ by $u_n - v_0$, one can derive

$$\begin{aligned}
&\frac{1}{2}\left(k * \|u_n - u_{0,n}\|_H^2\right)(t) + \frac{1}{2}\int_0^t k(\tau)\|u_n(\tau) - u_{0,n}\|_H^2 \,\mathrm{d}\tau + \int_0^t \varphi(u_n(\tau))\,\mathrm{d}\tau \\
&\le t\varphi(v_0) + \int_0^t \left(f_n(\tau), u_n(\tau) - v_0\right)_H \,\mathrm{d}\tau - \left([k * (u_n - u_{0,n})](t), u_{0,n} - v_0\right)_H
\end{aligned} \tag{4.22}$$

for all $t \in [0,T]$. As in §4.2, one can verify

$$\sup_{t\in[0,T]}\left(k * \|u_n - u_{0,n}\|_H^2\right)(t) + \int_0^T \varphi(u_n(\tau))\,\mathrm{d}\tau + \int_0^T \|u_n(\tau) - u_{0,n}\|_H^2 \,\mathrm{d}\tau \le C. \tag{4.23}$$

Multiply (1.1) with f_n and $u_{0,n}$ by $\mathcal{B}_m(u_n - u_{0,n})$ and employ Proposition 3.4 to observe that

$$\begin{aligned}&\big(\mathcal{B}(u_n - u_{0,n})(t), \mathcal{B}_m(u_n - u_{0,n})(t)\big)_H + \frac{\mathrm{d}}{\mathrm{d}t}\big[k_m * \varphi(u_n(\cdot))\big](t)\\&\le \big(f_n(t), \mathcal{B}_m(u_n - u_{0,n})(t)\big)_H - k_m(t)\left[(u_n(t) - u_{0,n}, \xi_n(t))_H - \varphi(u_n(t))\right]\\&\le \big(f_n(t), \mathcal{B}_m(u_n - u_{0,n})(t)\big)_H - k_m(t)\varphi^*(\xi_n(t)) + k_m(t)\|u_{0,n}\|_H\|\xi_n(t)\|_H.\end{aligned}$$

Here we employed the Fenchel-Moreau identity, $\varphi^*(\xi_n(t)) = (u_n(t), \xi_n(t))_H - \varphi(u_n(t))$. Moreover, we recall that φ^* is affinely bounded from below (see, e.g., [4]), i.e., there exists a constant $c_0 \ge 0$ such that $\varphi^*(\xi) \ge -c_0(\|\xi\|_H + 1)$ for any $\xi \in H$. Multiplying both sides by t and integrating both sides over $(0, t)$, we deduce that

$$\begin{aligned}&\int_0^t \tau\big(\mathcal{B}(u_n - u_{0,n})(\tau), \mathcal{B}_m(u_n - u_{0,n})(\tau)\big)_H \,\mathrm{d}\tau + t\big[k_m * \varphi(u_n(\cdot))\big](t)\\&\le \int_0^t \big[k_m * \varphi(u_n(\cdot))\big](\tau)\,\mathrm{d}\tau + \int_0^t \tau\big(f_n(\tau), \mathcal{B}_m(u_n - u_{0,n})(\tau)\big)_H \,\mathrm{d}\tau\\&\quad + c_0T\int_0^t k_m(\tau)\,\mathrm{d}\tau + (\|u_{0,n}\|_H + c_0)\int_0^t \tau k_m(\tau)\|\xi_n(\tau)\|_H\,\mathrm{d}\tau. \end{aligned}\tag{4.24}$$

Here, recalling the definition of k_m, that is, $k_m(t) + m(\ell * k_m)(t) = m$, one finds by (K) that

$$k_m(t) \le \left(\frac{1}{m} + \int_0^t \ell(s)\,\mathrm{d}s\right)^{-1},$$

whence follows

$$tk_m(t)^2 \le t\left(\int_0^t \ell(s)\,\mathrm{d}s\right)^{-2} =: h(t) \in L^1(0,T)$$

by assumption. Hence by passing to the limit in (4.24) as $m \to +\infty$, we obtain

$$\begin{aligned}&\int_0^t \tau\left\|\mathcal{B}(u_n - u_{0,n})(\tau)\right\|_H^2 \,\mathrm{d}\tau + t\big[k * \varphi(u_n(\cdot))\big](t)\\&\le \int_0^t \big[k * \varphi(u_n(\cdot))\big](\tau)\,\mathrm{d}\tau + \int_0^t \tau\big(f_n(\tau), \mathcal{B}(u_n - u_{0,n})(\tau)\big)_H \,\mathrm{d}\tau\\&\quad + c_0T\|k\|_{L^1(0,T)} + (\|u_{0,n}\|_H + c_0)\,\|h\|_{L^1(0,T)}^{1/2}\left(\int_0^t \tau\|\xi_n(\tau)\|_H^2\,\mathrm{d}\tau\right)^{1/2}\end{aligned}$$

for all $t \in [0,T]$. Thus by (1.1), it follows that

$$\begin{aligned}&\frac{1}{2}\int_0^t \tau\left\|\mathcal{B}(u_n - u_{0,n})(\tau)\right\|_H^2 \,\mathrm{d}\tau + t\big[k * \varphi(u_n(\cdot))\big](t)\\&\le \|k\|_{L^1(0,T)}\int_0^T \varphi(u_n(\cdot))\,\mathrm{d}\tau + c_0T\|k\|_{L^1(0,T)}\\&\quad + C\left[\int_0^t \tau\|f_n(\tau)\|_H^2\,\mathrm{d}\tau + (\|u_{0,n}\|_H + c_0)^2\,\|h\|_{L^1(0,T)}\right],\end{aligned}\tag{4.25}$$

which along with (4.23) and boundedness of f_n and $u_{0,n}$ in $\mathcal{H}$ and H, respectively, yields

$$\int_0^T \tau \,\|\mathcal{B}(u_n - u_{0,n})(\tau)\|_H^2 \,\mathrm{d}\tau + \sup_{t\in[0,T]} \big(t\big[k * \varphi(u_n(\cdot))\big](t)\big) \le C. \tag{4.26}$$

By (1.1) again, it further follows that

$$\int_0^T \tau\|\xi_n(\tau)\|_H^2 \,\mathrm{d}\tau \le C.$$

Thus we obtain

$$\begin{aligned} t^{1/2}\mathcal{B}(u_n - u_{0,n}) &\to \bar\eta \quad \text{weakly in } \mathcal{H}, \\ t^{1/2}\xi_n &\to \bar\xi \quad \text{weakly in } \mathcal{H} \end{aligned} \tag{4.27}$$

for some $\bar\eta, \bar\xi \in \mathcal{H}$. On the other hand, by virtue of (2.8), (u_n) forms a Cauchy sequence in $\mathcal{H}$, and therefore,

$$u_n \to u \quad \text{strongly in } \mathcal{H}. \tag{4.28}$$

Hence by virtue of the demiclosedness of maximal monotone operators, as in §4.3, we can conclude that $u(t) \in D(\partial\varphi)$ and $t^{-1/2}\bar\xi(t) \in \partial\varphi(u(t))$ for a.e. $t \in (0,T)$. To be more precise, we may replace $(0,T)$ by (δ,T) for each $\delta \in (0,T)$ throughout the argument of §4.3 and note by (4.27) that $\xi_n \to t^{-1/2}\bar\xi$ weakly in $L^2(\delta,T;H)$. Then one can identify the limit $\xi(t)$ for a.e. $t \in (\delta,T)$ and finally employ the arbitrariness of δ.

Now, we observe that, for any $w \in C_c^\infty(0,T)$,

$$\begin{aligned} \int_0^T \mathcal{B}(u_n - u_{0,n})(t)w(t)\,\mathrm{d}t &= \int_0^T t^{1/2}\mathcal{B}(u_n - u_{0,n})(t)t^{-1/2}w(t)\,\mathrm{d}t \\ &\to \int_0^T \bar\eta(t)t^{-1/2}w(t)\,\mathrm{d}t \quad \text{weakly in } H. \end{aligned}$$

On the other hand,

$$\begin{aligned} \int_0^T \mathcal{B}(u_n - u_{0,n})(t)w(t)\,\mathrm{d}t &= -\int_0^T \left[k * (u_n - u_{0,n})\right](t)w'(t)\,\mathrm{d}t \\ &\to -\int_0^T \left[k * (u - u_0)\right](t)w'(t)\,\mathrm{d}t \quad \text{strongly in } H. \end{aligned}$$

Thus

$$t^{-1/2}\bar\eta = \frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u - u_0)\right]$$

in the distributional sense. In particular, it follows that $\bar\eta(t) = t^{1/2}(\mathrm{d}/\mathrm{d}t)[k * (u - u_0)](t)$ for a.e. $t \in (0,T)$.

We next improve the regularity of the limit u. To this end, let us recall (4.22) again and pass to the limit as $n \to +\infty$. We here note that

$$\begin{aligned}
&\left\| \left(k * \|u_n - u_{0,n}\|_H^2\right)(t) - \left(k * \|u - u_0\|_H^2\right)(t) \right\|_{L^1(0,T)} \\
&\leq \|k\|_{L^1(0,T)} \int_0^t \left| \|u_n(\tau) - u_{0,n}\|_H^2 - \|u(\tau) - u_0\|_H^2 \right| \, \mathrm{d}\tau \\
&\leq \|k\|_{L^1(0,T)} \|u_n - u - u_{0,n} + u_0\|_{\mathcal{H}} \|u_n + u - u_{0,n} - u_0\|_{\mathcal{H}} \to 0,
\end{aligned}$$

which implies, up to a subsequence, that

$$\left(k * \|u_n - u_{0,n}\|_H^2\right)(t) \to \left(k * \|u - u_0\|_H^2\right)(t) \quad \text{ for a.e. } t \in (0,T).$$

Hence by virtue of (4.28) and Fatou's lemma, we deduce that

$$\begin{aligned}
&\frac{1}{2}\left(k * \|u - u_0\|_H^2\right)(t) + \frac{1}{2}\int_0^t k(\tau)\|u(\tau) - u_0\|_H^2 \, \mathrm{d}\tau + \int_0^t \varphi(u(\tau)) \, \mathrm{d}\tau \\
&\leq t\varphi(v_0) + \int_0^t (f(\tau), u(\tau) - v_0)_H \, \mathrm{d}\tau - ([k * (u - u_0)](t), u_0 - v_0)_H
\end{aligned}$$

for a.e. $t \in (0,T)$. In particular, it also holds that $k * \|u - u_0\|_H^2 \in L^\infty(0,T)$, $k(\cdot)\|u(\cdot) - u_0\|_H^2 \in L^1(0,T)$ and $\varphi(u(\cdot)) \in L^1(\Omega)$. Furthermore, we observe that

$$([k * (u - u_0)](t), u_0 - v_0)_H \leq \|k\|_{L^1(0,t)}^{1/2} (k * \|u - u_0\|_H^2)(t)^{1/2} \|u_0 - v_0\|_H.$$

Thus one can take a set $I \subset (0,T)$ such that $|(0,T) \setminus I| = 0$ and

$$\begin{aligned}
&\frac{1}{4}\left(k * \|u - u_0\|_H^2\right)(t) + \frac{1}{2}\int_0^t k(\tau)\|u(\tau) - u_0\|_H^2 \, \mathrm{d}\tau + \int_0^t \varphi(u(\tau)) \, \mathrm{d}\tau \\
&\leq t\varphi(v_0) + \int_0^t (f(\tau), u(\tau) - v_0)_H \, \mathrm{d}\tau + \|k\|_{L^1(0,t)} \|u_0 - v_0\|_H^2
\end{aligned}$$

for all $t \in I$. Therefore for any sequence (s_n) in I converging to 0_+, it follows that $(k * \|u - u_0\|_H^2)(s_n) \to 0$, whence follows that $[k * (u - u_0)](s_n) \to 0$ strongly in H. On the other hand, since $h(t) := t^{1/2}[k * (u - u_0)](t)$ belongs to $W^{1,r}(0,T;H)$ for any $r \in (1,2)$, we observe that

$$\begin{aligned}
\left\| s^{1/2}[k * (u - u_0)](s) - t^{1/2}[k * (u - u_0)](t) \right\|_H &\leq \int_t^s \left\| \frac{\mathrm{d}h}{\mathrm{d}\tau}(\tau) \right\|_H \mathrm{d}\tau \\
&\leq \|\mathrm{d}h/\mathrm{d}\tau\|_{L^r(0,T;H)} |s - t|^{1/r'},
\end{aligned}$$

where $r' = r/(r-1) > 2$, for any $0 \leq t < s \leq T$. Now, let (t_n) be a sequence in $(0,T)$ such that $t_n \to 0_+$. Then one can take a sequence (s_n) in I such that

$$0 \leq s_n - t_n < t_n^{r'},$$

and hence,

$$\left\| (s_n/t_n)^{1/2}[k * (u - u_0)](s_n) - [k * (u - u_0)](t_n) \right\|_H \leq \|\mathrm{d}h/\mathrm{d}t\|_{L^r(0,T;H)} t_n^{1/2}.$$

Accordingly, we deduce that $k * (u - u_0)(t_n) \to 0$ strongly in H.

Finally, if $\ell \in L^2(0,T)$ and (2.9) is fulfilled, then thanks to Corollary 2.7 we have

$$\sup_{t\in[0,T]} \|u_n(t) - u_m(t)\|_H \le \|u_{0,n} - u_{0,m}\|_H + \|\ell\|_{L^2(0,T)}\|f_n - f_m\|_{\mathcal{H}} \to 0$$

as $n, m \to +\infty$. Consequently, u_n converges to u strongly in $C([0,T];H)$ as $n \to +\infty$, and moreover, u satisfies the initial condition $u(0) = u_0$ in the classical sense. □

5. Lipschitz perturbation problem

This section is devoted to discussing well-posedness of the following Lipschitz perturbation problem for (1.1):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[k * (u - u_0)\right](t) + \partial\varphi(u(t)) + F(t, u(t)) \ni f(t) \quad \text{in } H, \quad 0 < t < T, \tag{5.1}$$

where $F : [0,T] \times H \to H$ satisfies:

(L): The function $t \mapsto F(t,w)$ is strongly measurable on $(0,T)$ with values in H for each $w \in H$, and moreover, there exists $L \ge 0$ such that

$$\|F(t,u) - F(t,v)\|_H \le L\|u - v\|_H$$

for all $u, v \in H$ and for a.e. $t \in (0,T)$. There exists $u_* \in L^2(0,T;H)$ such that the function $t \mapsto F(t, u_*(t))$ belongs to $L^2(0,T;H)$.

Then it follows that

$$\|F(t,u)\|_H \le L\|u\|_H + \rho(t) \quad \text{for all } u \in H \text{ and a.e. } t \in (0,T),$$

where $\rho(\cdot) := L\|u_*(\cdot)\|_H + \|F(\cdot, u_*(\cdot))\|_H \in L^2(0,T)$, and hence, the mapping $u \mapsto F(\cdot, u(\cdot))$ turns out to be Lipschitz continuous in $L^2(0,T;H)$. Our result reads,

Theorem 5.1 (Lipschitz perturbation). *Assume that* (K) *and* (L) *are satisfied. Then for any $f \in L^2(0,T;H)$ and $u_0 \in D(\varphi)$, the Cauchy problem* (5.1) *admits a unique strong solution $u \in L^2(0,T;H)$ on $[0,T]$.*

Before proving the theorem above, we remark that a standard extension strategy (i.e., a local (in time) solution is first constructed and then extended onto an arbitrary interval) cannot be applied to construct a global (in time) solution for (5.1). Indeed, the concatenation of two solutions may not solve the equation due to the nonlocal nature of (5.1). Here we shall set up a weighted function space in order to construct a global solution directly.

Proof. Let $\mu > 1$ be a number which will be determined later. Set $\mathcal{X} := \{w \in L^2(0,T;H) \colon \|w\|_{\mathcal{X}} < +\infty\}$, where $\|\cdot\|_{\mathcal{X}}$ is given by

$$\|w\|_{\mathcal{X}}^2 := \operatorname*{ess\,sup}_{t\in(0,T)} \left[\mathrm{e}^{-\mu t}\left(k * \|w\|_H^2\right)(t)\right] \quad \text{for } w \in \mathcal{X}.$$

Then, we find that $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ is a Banach space (see Appendix C). Now, define a mapping $\mathcal{S} : \mathcal{X} \to \mathcal{X}$ by

$$\mathcal{S}(v) = u \quad \text{for} \quad v \in \mathcal{X},$$

where u stands for the unique solution of the Cauchy problem

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[k*(u-u_0)\right](t)+\partial\varphi(u(t))\ni f(t)-F(t,v(t)) \ \text{ in } H, \quad 0<t<T. \tag{5.2}$$

By (L), we infer that $F(\cdot,v(\cdot))\in L^2(0,T;H)$ for $v\in L^2(0,T;H)$. By Theorem 2.3, (5.2) admits a unique solution $u\in\mathcal{X}$ on $[0,T]$ for each $v\in\mathcal{X}$, and therefore, $\mathcal{S}:\mathcal{X}\to\mathcal{X}$ is well defined. We next prove that $\mathcal{S}$ is a contraction mapping in $\mathcal{X}$ by choosing $\mu>0$ large enough. To this end, let $v_1,v_2\in\mathcal{X}$ and set $u_1=\mathcal{S}(v_1)$ and $u_2=\mathcal{S}(v_2)$. Subtracting equations and testing it by $u_1-u_2\in D(\mathcal{B})$, we deduce that

$$\frac{1}{2}\left(k*\|u_1-u_2\|_H^2\right)(t)\le L\int_0^t \|v_1(\tau)-v_2(\tau)\|_H\|u_1(\tau)-u_2(\tau)\|_H\,\mathrm{d}\tau$$

for all $t\in[0,T]$. Now, let us recall by (K) that

$$\int_0^t h(\tau)\,\mathrm{d}\tau=(\ell*k*h)(t) \ \text{ for } \ h\in L^1(0,t).$$

Moreover, note that

$$(\ell*k*h)(t)=\mathrm{e}^{\mu t}\int_0^t \ell(t-\tau)\mathrm{e}^{-\mu(t-\tau)}\mathrm{e}^{-\mu\tau}(k*h)(\tau)\,\mathrm{d}\tau.$$

Then

$$\begin{aligned}
&\left(k*\|u_1-u_2\|_H^2\right)(t)\\
&\le 2L\mathrm{e}^{\mu t}\int_0^t \ell(t-\tau)\mathrm{e}^{-\mu(t-\tau)}\mathrm{e}^{-\mu\tau}\left[k*(\|v_1-v_2\|_H\|u_1-u_2\|_H)\right](\tau)\,\mathrm{d}\tau\\
&\le 2L\mathrm{e}^{\mu t}\int_0^t \ell(t-\tau)\mathrm{e}^{-\mu(t-\tau)}\mathrm{e}^{-\mu\tau}\left(k*\|v_1-v_2\|_H^2\right)^{1/2}(\tau)\left(k*\|u_1-u_2\|_H^2\right)^{1/2}(\tau)\,\mathrm{d}\tau
\end{aligned}$$

for all $t\in[0,T]$. Here we used the fact by Hölder's inequality that

$$h*(FG)\le(h*F^2)^{1/2}(h*G^2)^{1/2}$$

for nonnegative functions $h\in L^1(0,t)$ and $F,G\in L^2(0,t)$. Therefore by Young's inequality, we deduce that

$$\begin{aligned}
&\sup_{t\in[0,T]}\left[\mathrm{e}^{-\mu t}\left(k*\|u_1-u_2\|_H^2\right)(t)\right]\\
&\le 2L\left(\int_0^T \ell(\tau)e^{-\mu\tau}\,\mathrm{d}\tau\right)\operatorname*{ess\,sup}_{\tau\in(0,T)}\left[\mathrm{e}^{-\mu\tau}\left(k*\|v_1-v_2\|_H^2\right)^{1/2}(\tau)\left(k*\|u_1-u_2\|_H^2\right)^{1/2}(\tau)\right]\\
&\le 2La_\mu\left(\operatorname*{ess\,sup}_{\tau\in(0,T)}\left[\mathrm{e}^{-\mu\tau}\left(k*\|v_1-v_2\|_H^2\right)(\tau)\right]\right)^{1/2}\left(\sup_{\tau\in[0,T]}\left[\mathrm{e}^{-\mu\tau}\left(k*\|u_1-u_2\|_H^2\right)(\tau)\right]\right)^{1/2},
\end{aligned}$$

where a_μ is a positive constant given by

$$a_\mu:=\int_0^T \ell(\tau)\mathrm{e}^{-\mu\tau}\,\mathrm{d}\tau\to 0_+ \quad \text{as} \ \ \mu\to+\infty.$$

Thus

$$\sup_{\tau\in[0,T]}\left[\mathrm{e}^{-\mu\tau}\left(k*\|u_1-u_2\|_H^2\right)(\tau)\right]\le 4L^2a_\mu^2\operatorname*{ess\,sup}_{\tau\in(0,T)}\left[\mathrm{e}^{-\mu\tau}\left(k*\|v_1-v_2\|_H^2\right)(\tau)\right].$$

Now, choosing $\mu>1$ large enough, we find that

$$4L^2a_\mu^2<1.$$

Therefore $\mathcal{S}:\mathcal{X}\to\mathcal{X}$ turns out to be a contraction mapping. Thanks to Banach's contraction mapping principle, we conclude that $\mathcal{S}$ possesses a unique fixed point u, that is, $\mathcal{S}(u)=u$. Equivalently, u solves (5.1) on $[0,T]$. The uniqueness also follows from the contraction property. □

The argument of the proof above can be also applied to derive a continuous dependence of strong solutions for (5.1) on prescribed data. Indeed, let $u_{0,1},u_{0,2}\in D(\varphi)$ and let $f_1,f_2\in L^2(0,T;H)$. We denote by u_1,u_2 the unique strong solutions of (5.1) with $u_0=u_{0,1},u_{0,2}$ and $f=f_1,f_2$, respectively. As in the proof of Theorem 5.1, setting $w:=u_1-u_2$, $w_0:=u_{0,1}-u_{0,2}$ and $g:=f_1-f_2$, subtracting equations for u_1 and u_2 and multiplying it by w, for any $K>0$, one can verify that

$$\begin{aligned}&\frac{1}{2}\left(k*\|w-w_0\|_H^2\right)(t)\\&\le\frac{1}{2K}\int_0^t\|g(\tau)\|_H^2\,\mathrm{d}\tau+\left(\frac{K}{2}+L\right)\int_0^t\|w(\tau)\|_H^2\,\mathrm{d}\tau-\Big(\left[k*(w-w_0)\right](t),w_0\Big)_H\end{aligned}$$

for all $t\in[0,T]$. Moreover, we have

$$\begin{aligned}&\left(k*\|w-w_0\|_H^2\right)(t)\\&\le K^{-1}\int_0^t\|g(\tau)\|_H^2\,\mathrm{d}\tau+(K+2L)\,\mathrm{e}^{\mu t}\int_0^t\ell(t-\tau)\mathrm{e}^{-\mu(t-\tau)}\mathrm{e}^{-\mu\tau}\left(k*\|w\|_H^2\right)(\tau)\,\mathrm{d}\tau\\&\quad+2\|k\|_{L^1(0,T)}^{1/2}\|w_0\|_H(k*\|w\|_H^2)^{1/2}(t)+2\|k\|_{L^1(0,T)}\|w_0\|_H^2\\&\le K^{-1}\int_0^t\|g(\tau)\|_H^2\,\mathrm{d}\tau+(K+2L)\,\mathrm{e}^{\mu t}a_\mu\sup_{\tau\in[0,T]}\left[\mathrm{e}^{-\mu\tau}\left(k*\|w\|_H^2\right)(\tau)\right]\\&\quad+2\|k\|_{L^1(0,T)}^{1/2}\|w_0\|_H(k*\|w\|_H^2)^{1/2}(t)+2\|k\|_{L^1(0,T)}\|w_0\|_H^2,\end{aligned}$$

which implies

$$\begin{aligned}&\frac{1}{2}\left(k*\|w\|_H^2\right)(t)\\&\le K^{-1}\int_0^t\|g(\tau)\|_H^2\,\mathrm{d}\tau+(K+2L)\,\mathrm{e}^{\mu t}a_\mu\sup_{\tau\in[0,T]}\left[\mathrm{e}^{-\mu\tau}\left(k*\|w\|_H^2\right)(\tau)\right]\\&\quad+2\|k\|_{L^1(0,T)}^{1/2}\|w_0\|_H(k*\|w\|_H^2)^{1/2}(t)+3\|k\|_{L^1(0,T)}\|w_0\|_H^2.\end{aligned}$$

Therefore choosing μ so large that $(K+2L)a_\mu < 1/4$, we deduce that

$$\begin{aligned}
&\frac{1}{4}\sup_{\tau\in[0,T]}\left[\mathrm{e}^{-\mu\tau}\left(k*\|w\|_H^2\right)(\tau)\right]\\
&\le K^{-1}\int_0^T \|g(\tau)\|_H^2\,\mathrm{d}\tau + 2\|k\|_{L^1(0,T)}^{1/2}\|w_0\|_H \sup_{\tau\in[0,T]}\left[\mathrm{e}^{-\mu\tau}\left(k*\|w\|_H^2\right)(\tau)\right]^{1/2}\\
&\quad + 3\|k\|_{L^1(0,T)}\|w_0\|_H^2, \qquad (5.3)
\end{aligned}$$

which ensures the continuous dependence of strong solutions u for (5.1) on initial data u_0 and f.

Remark 5.2 (Smoothing effect). As in Therefore 2.8, under (2.11), one can verify that, for any $f \in L^2(0,T;H)$ and $u_0 \in \overline{D(\varphi)}^H$, the Cauchy problem (5.1) possesses a function $u \in L^2(0,T;H)$ satisfying similar conditions to Theorem 2.8. In addition, if $\ell \in L^2(0,T)$ and (2.9) is fulfilled, then $u \in C([0,T];H)$ and $u(0) = u_0$. Indeed, let $(u_{0,n})$ be a sequence in $D(\varphi)$ such that $u_{0,n} \to u_0$ strongly in H. For $n,m \in \mathbb{N}$, we denote by u_n, u_m the unique strong solutions of (5.1) with initial data $u_{0,n}, u_{0,m}$, respectively. By (5.3), we obtain

$$\sup_{\tau\in[0,T]}\left[\mathrm{e}^{-\mu\tau}\left(k*\|u_n-u_m\|_H^2\right)(\tau)\right] \to 0 \quad \text{as } \ n,m\to+\infty,$$

which also yields that (u_n) forms a Cauchy sequence in $L^2(0,T;H)$. Therefore $u_n \to u$ strongly in $L^2(0,T;H)$. The rest of proof runs as in the proof of Theorem 2.8, since $F(\cdot,u_n(\cdot)) \to F(\cdot,u(\cdot))$ strongly in $L^2(0,T;H)$.

6. Applications to fractional PDEs

In this section, we shall apply the preceding abstract results to time-fractional nonlinear parabolic equations. For simplicity, we shall treat only Dirichlet problems, however, one can also apply the results to other boundary conditions, e.g., Neumann, Robin and nonlinear boundary conditions such as

$$-\partial_\nu u \in \beta(u) \ \text{ on } \partial\Omega,$$

where $\partial\Omega$ is the boundary of a domain $\Omega \subset \mathbb{R}^N$, ∂_ν stands for the normal derivative and β is a maximal monotone graph in $\mathbb{R}$ and which is derived from a Stefan-Boltzmann radiation law. On the other hand, one need to pay careful attention to consider other boundary conditions (e.g., Neumann problem) for time-fractional porous medium / fast diffusion equations (see §6.2).

Let Ω be a domain of $\mathbb{R}^N$ and let $T>0$. We denote by ∂_t^β the Riemann-Liouville time-fractional derivative defined by

$$\partial_t^\beta w(x,t) = \partial_t\left[k_\beta * w(x,\cdot)\right](t) = \partial_t\left(\int_0^t k_\beta(t-s)w(x,s)\,\mathrm{d}s\right),$$

where k_β is a kernel given by (1.3), $0<\beta<1$ and k_β fulfills (K) with the conjugate kernel ℓ_β given by (1.3). Moreover, it is known that k_β is of class $\mathcal{K}^1(\beta,\pi/2)$ (see [51]). Furthermore, (2.11) also holds for any $0<\beta<1$. Subsections 6.1 and 6.2 are devoted to *time-fractional p-Laplace parabolic equations* and *time-fractional* *porous medium and fast diffusion equations*, respectively. These time-fractional nonlinear problems have already been treated in [50], where optimal decay estimates are provided without ensuring existence of solutions (see also [52]). Furthermore, in Subsection 6.3 we shall also handle a time-fractional variant of the so-called Allen-Cahn equation, which is an L^2-gradient flow of a (possibly) nonconvex free energy functional and describes various phase-separation phenomena. We also emphasize that there are much more applications of the preceding abstract results to nonlinear PDEs with time-fractional derivatives, as the classical Brézis-Kōmura theory does to those with the standard time-derivative.

6.1. **Time-fractional p-Laplace parabolic equation.** This subsection is concerned with time-fractional parabolic equations involving the so-called *p-Laplacian* given by

$$\Delta_p u := \operatorname{div}\left(|\nabla u|^{p-2}\nabla u\right), \quad 1 < p < +\infty.$$

Let us consider the Cauchy-Dirichlet problem,

$$\partial_t^\beta (u - u_0) - \Delta_p u = f \quad \text{in } \Omega \times (0,T), \tag{6.1}$$

$$u = 0 \quad \text{in } \partial\Omega \times (0,T), \tag{6.2}$$

where $1 < p < \infty$, $0 < \beta < 1$ and $u_0 = u_0(x)$ and $f = f(x,t)$ are given. In order to apply Theorems 2.3 and 2.8, we set $H = L^2(\Omega)$ and

$$\varphi(w) = \begin{cases} \frac{1}{p}\int_\Omega |\nabla w(x)|^p \, \mathrm{d}x & \text{if } w \in W_0^{1,p}(\Omega) \cap L^2(\Omega), \\ +\infty & \text{otherwise.} \end{cases} \tag{6.3}$$

Then φ is proper, lower semicontinuous and convex in H with $D(\varphi) = W_0^{1,p}(\Omega) \cap L^2(\Omega)$, and moreover, $\partial\varphi(w)$ coincides with $-\Delta_p w$ equipped with the homogeneous Dirichlet boundary condition for $w \in D(\partial\varphi) = \{w \in W_0^{1,p}(\Omega) \cap L^2(\Omega) \colon -\Delta_p w \in L^2(\Omega)\}$. Hence the Cauchy-Dirichlet problem (6.1), (6.2) is reduced to the evolution equation (1.1). Therefore Theorems 2.3 and 2.8 ensure

Theorem 6.1. *For any $f \in L^2(0,T;L^2(\Omega))$ and $u_0 \in W_0^{1,p}(\Omega) \cap L^2(\Omega)$, the Cauchy-Dirichlet problem* (6.1), (6.2) *admits a unique $L^2(\Omega)$ solution $u = u(x,t)$ such that*

$$u \in \begin{cases} L^{2/(1-2\beta),\infty}(0,T;L^2(\Omega)) & \text{if } \beta \in (0,1/2), \\ \cap_{p<\infty} L^{p,\infty}(0,T;L^2(\Omega)) & \text{if } \beta = 1/2, \\ C([0,T];L^2(\Omega)) & \text{if } \beta \in (1/2,1), \end{cases} \tag{6.4}$$

$$u \in L^p(0,T;W_0^{1,p}(\Omega)), \quad u - u_0 \in H_0^{\beta,2}(0,T;L^2(\Omega)), \tag{6.5}$$

$$\Delta_p u \in L^2(0,T;L^2(\Omega)). \tag{6.6}$$

For any $u_0 \in L^2(\Omega)$, there exists a function $u \in L^2(0,T;L^2(\Omega))$ solving (6.1), (6.2) *almost everywhere such that $u \in L^p(0,T;W_0^{1,p}(\Omega))$, $t^{1/2}\Delta_p u \in L^2(0,T;L^2(\Omega))$, $t^{1/2}\partial_t^\beta(u-u_0) \in L^2(0,T;L^2(\Omega))$ and $[k_\beta * (u-u_0)](t) \to 0$ strongly in $L^2(\Omega)$ as $t \to 0_+$. In addition, if $\beta > 1/2$, then $u \in C([0,T];L^2(\Omega))$ and $u(+0) = u_0$ in $L^2(\Omega)$.*

6.2. **Time-fractional porous medium and fast diffusion equations.** In this subsection, we shall deal with time-fractional porous medium and fast diffusion equations,

$$\partial_t^\beta(u-u_0) - \Delta\left(|u|^{m-2}u\right) = f \quad \text{in } \Omega\times(0,T), \tag{6.7}$$

$$u = 0 \quad \text{in } \partial\Omega\times(0,T), \tag{6.8}$$

where $1 < m < \infty$, $0 < \beta < 1$ and $u_0 = u_0(x)$ and $f = f(x,t)$ are given. Throughout this section, we assume that the Poincaré inequality holds (e.g., Ω is bounded). Then the Laplace operator $-\Delta : H^1_0(\Omega) \to H^{-1}(\Omega)$ is invertible. We set

$$H = H^{-1}(\Omega), \quad \varphi(w) = \begin{cases} \frac{1}{m}\int_\Omega |w(x)|^m \,\mathrm{d}x & \text{if } w \in L^m(\Omega)\cap H^{-1}(\Omega),\\ +\infty & \text{otherwise}\end{cases}$$

equipped with the inner product $(u,v)_H := \langle u, (-\Delta)^{-1}v\rangle_{H^1_0(\Omega)}$ for $u,v \in H$. Then $D(\varphi) = L^m(\Omega)\cap H^{-1}(\Omega)$, and moreover, for any $w \in D(\partial\varphi) = \{w \in L^m(\Omega)\cap H^{-1}(\Omega) \colon |w|^{m-2}w \in H^1_0(\Omega)\}$, it holds that

$$f = \partial\varphi(w) \quad \text{if and only if} \quad (-\Delta)^{-1}f = |w|^{m-2}w.$$

Thus a weak form of (6.7), (6.8) is rewritten as (1.1) (see [9] for more details). Hence by virtue of Theorems 2.3 and 2.8, one obtains

Theorem 6.2. *For any $f \in L^2(0,T;H^{-1}(\Omega))$ and $u_0 \in L^m(\Omega)\cap H^{-1}(\Omega)$, the Dirichlet problem (6.7), (6.8) admits a unique $H^{-1}(\Omega)$ solution $u = u(x,t)$ such that*

$$u \in \begin{cases} L^{2/(1-2\beta),\infty}(0,T;H^{-1}(\Omega)) & \text{if } \beta \in (0,1/2),\\ \cap_{p<\infty} L^{p,\infty}(0,T;H^{-1}(\Omega)) & \text{if } \beta = 1/2,\\ C([0,T];H^{-1}(\Omega)) & \text{if } \beta\in(1/2,1),\end{cases} \tag{6.9}$$

$$u \in L^m(0,T;L^m(\Omega)), \quad |u|^{m-2}u \in L^2(0,T;H^1_0(\Omega)), \tag{6.10}$$

$$u - u_0 \in H^{\beta,2}_0(0,T;H^{-1}(\Omega)). \tag{6.11}$$

For any $u_0 \in H^{-1}(\Omega)$, there exists a function $u \in L^2(0,T;H^{-1}(\Omega))$ solving (6.7), (6.8) almost everywhere such that $u \in L^m(0,T;L^m(\Omega))$, $t^{1/2}|u|^{m-2}u \in L^2(0,T;H^1_0(\Omega))$, $t^{1/2}\partial_t^\beta(u-u_0) \in L^2(0,T;H^{-1}(\Omega))$ and $[k_\beta * (u-u_0)](t) \to 0$ strongly in $H^{-1}(\Omega)$ as $t \to 0_+$. In addition, if $\beta > 1/2$, then $u \in C([0,T];H^{-1}(\Omega))$ and $u(+0) = u_0$ in $H^{-1}(\Omega)$.

Furthermore, we can also apply the abstract results to fractional variants of Stefan problem as well as obstacle problem.

6.3. **Time-fractional Allen-Cahn equation.** In this subsection, we shall consider a time-fractional variant of the Allen-Cahn equation,

$$\partial_t^\beta(u-u_0) - \Delta u + W'(u) = f \quad \text{in } \Omega\times(0,T), \tag{6.12}$$

$$u = 0 \quad \text{in } \partial\Omega\times(0,T), \tag{6.13}$$

where $W : \mathbb{R} \to \mathbb{R}$ is a double-well potential satisfying a λ-*convexity*,

$$W \in C^2(\mathbb{R}), \quad W''(s) \geq \lambda \ \text{ for all } \ s \in \mathbb{R} \tag{6.14}$$

for some $\lambda \in \mathbb{R}$. One can assume $W(0) = 0$ without any loss of generality. Here we shall treat only the case $\lambda < 0$, where the double-well potential may not be convex. Then one can rewrite W into the difference of two convex parts,

$$W(s) = \left(W(s) - \frac{\lambda}{2} s^2 \right) - \left(-\frac{\lambda}{2} s^2 \right).$$

Here we note that the first-term of the right-hand side is convex. Hence we set $H = L^2(\Omega)$ and

$$\varphi(w) = \begin{cases} \frac{1}{2} \int_\Omega |\nabla w(x)|^2 \,\mathrm{d}x + \int_\Omega \left(W(w(x)) - \frac{\lambda}{2} |w(x)|^2 \right) \,\mathrm{d}x & \text{if } w \in D(\varphi), \\ +\infty & \text{otherwise,} \end{cases} \tag{6.15}$$

where $D(\varphi) := \{ w \in H^1_0(\Omega) \colon W(w(\cdot)) \in L^1(\Omega) \}$. Then $\partial \varphi(w)$ coincides with $-\Delta w + W'(w) - \lambda w$ for $w \in D(\partial\varphi) = \{ w \in D(\varphi) \cap H^2(\Omega) \colon W'(w(\cdot)) \in L^2(\Omega) \}$. Moreover, set

$$F(t, w) := \lambda w \ \text{ for } w \in H.$$

Then the Cauchy-Dirichlet problem (6.12), (6.13) are reduced to (5.1), and therefore, by Theorem 5.1 and Remark 5.2 we assure that

Theorem 6.3. *For any $f \in L^2(0,T; L^2(\Omega))$ and $u_0 \in H^1_0(\Omega)$ satisfying $W(u_0(\cdot)) \in L^1(\Omega)$, the Dirichlet problem* (6.12), (6.13) *admits a unique $L^2(\Omega)$ solution $u = u(x,t)$ such that*

$$u \in \begin{cases} L^{2/(1-2\beta),\infty}(0,T; L^2(\Omega)) & \text{if } \ \beta \in (0, 1/2), \\ \cap_{p<\infty} L^{p,\infty}(0,T; L^2(\Omega)) & \text{if } \ \beta = 1/2, \\ C([0,T]; L^2(\Omega)) & \text{if } \ \beta \in (1/2, 1), \end{cases} \tag{6.16}$$

$$u \in L^2(0,T; H^1_0(\Omega) \cap H^2(\Omega)), \quad u - u_0 \in H^{\beta,2}_0(0,T; L^2(\Omega)). \tag{6.17}$$

For any $u_0 \in L^2(\Omega)$, there exists a function $u \in L^2(0,T; L^2(\Omega))$ solving (6.12), (6.13) *almost everywhere such that $u \in L^2(0,T; H^1_0(\Omega))$, $W(u) \in L^1(\Omega \times (0,T))$, $t^{1/2}\Delta u$, $t^{1/2} W'(u) \in L^2(0,T; L^2(\Omega))$, $t^{1/2}\partial_t^\beta (u - u_0) \in L^2(0,T; L^2(\Omega))$ and $[k_\beta * (u - u_0)](t) \to 0$ strongly in $L^2(\Omega)$ as $t \to 0_+$. In addition, if $\beta > 1/2$, then $u \in C([0,T]; L^2(\Omega))$ and $u(+0) = u_0$ in $L^2(\Omega)$.*

Appendix A. Proof of (2.2)

We shall verify (2.2). By Lemma 3.3, we derive that

$$\begin{aligned} &\left(\frac{\mathrm{d}}{\mathrm{d}t} \left[k * (u - u_0) \right](t), u(t) - u_0 \right)_H \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}t} \left[k_n * (u - u_0) \right](t), u(t) - u_0 \right)_H + h_n(t) \\ &\geq \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \left(k_n * \|u - u_0\|_H^2 \right)(t) + \frac{1}{2} k_n(t) \|u(t) - u_0\|_H^2 + h_n(t), \end{aligned} \tag{A.1}$$

where $h_n \in L^1(0,T)$ is given by

$$h_n(t) := \left(\frac{\mathrm{d}}{\mathrm{d}t} \left[(k - k_n) * (u - u_0) \right](t), u(t) - u_0 \right)_H .$$

As $u - u_0$ lies on $D(\mathcal{B})$, it follows that $h_n \to 0$ strongly in $L^1(0,T)$. Hence integrating both sides of (A.1) over $(0,t)$ and passing to the limit as $n \to +\infty$, one deduces by Fatou's lemma that

$$\begin{aligned} &\frac{1}{2} \left(k * \|u - u_0\|_H^2 \right)(t) + \frac{1}{2} \int_0^t k(\tau) \|u(\tau) - u_0\|_H^2 \, \mathrm{d}\tau \\ &\leq \int_0^t \left(\frac{\mathrm{d}}{\mathrm{d}\tau} \left[k * (u - u_0) \right](\tau), u(\tau) - u_0 \right)_H \mathrm{d}\tau \end{aligned}$$

for all $t \in [0,T]$. Thus (2.2) follows.

Appendix B. Proof of (2.3)

Let k be given by (1.3). We start with the case $\beta \in (0, 1/2)$. Convolve both sides of (2.1) with ℓ. Then it follows by (K) that

$$u - u_0 = \ell * (f - \xi), \quad \xi \in \partial\varphi(u(\cdot)).$$

Note that $F := f - \xi \in L^2(0,T;H)$ by Definition 2.1 and $\ell(t) = t^{\beta-1}/\Gamma(\beta)$ belongs to $L^{\frac{1}{1-\beta},\infty}(0,T)$ (hence, their zero extensions $\bar{F}$ and $\bar{\ell}$ outside $[0,T]$ belong to $L^2(\mathbb{R};H)$ and $L^{\frac{1}{1-\beta},\infty}(\mathbb{R})$, respectively). By weak Young's inequality (see, e.g., [46, §IX.3] and [21, Theorem 1.2.13]), one obtains

$$\begin{aligned} \|u - u_0\|_{L^{\frac{2}{1-2\beta},\infty}(0,T;H)} &= \|\ell * F\|_{L^{\frac{2}{1-2\beta},\infty}(0,T;H)} \\ &= \|\bar{\ell} * \bar{F}\|_{L^{\frac{2}{1-2\beta},\infty}(\mathbb{R};H)} \\ &\leq C_\beta \|\bar{\ell}\|_{L^{\frac{1}{1-\beta},\infty}(\mathbb{R})} \|\bar{F}\|_{L^2(\mathbb{R};H)} = C_\beta \|\ell\|_{L^{\frac{1}{1-\beta},\infty}(0,T)} \|F\|_{L^2(0,T;H)} \end{aligned}$$

for some $C_\beta > 0$. Thus it in particular yields $u \in L^{\frac{2}{1-2\beta},\infty}(0,T;H)$. Concerning the case $\beta = 1/2$, since $\ell \in L^{2,\infty}(0,T)$, according to the argument above, we have $u \in L^{p,\infty}(0,T;H)$ for any $p \in [1,\infty)$.

Appendix C. Completeness of $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$

Let (w_n) be a Cauchy sequence in $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$. Then we observe that

$$\|w_n - w_m\|_{\mathcal{X}}^2 \geq \mathrm{e}^{-\mu t} \left(k * \|w_n(\cdot) - w_m(\cdot)\|_H^2 \right)(t) \quad \text{for a.e. } \ t \in (0,T).$$

Convolving both sides with ℓ, we have

$$(\ell * \mathrm{e}^{\mu t})(t) \|w_n - w_m\|_{\mathcal{X}}^2 \geq \int_0^t \|w_n(s) - w_m(s)\|_H^2 \, \mathrm{d}s \quad \text{for all } \ t \in [0,T].$$

Hence (w_n) is a Cauchy sequence in $L^2(0,T;H)$ as well, and therefore,

$$w_n \to w \quad \text{strongly in } L^2(0,T;H). \tag{C.1}$$

On the other hand, note that

$$\|w_n - w_m\|_{\mathcal{X}}^2 = \operatorname*{ess\,sup}_{t\in(0,T)} \left[\mathrm{e}^{-\mu t} \int_0^t k(t-s) \|w_n(s) - w_m(s)\|_H^2 \, \mathrm{d}s \right]$$

$$\geq \mathrm{e}^{-\mu t} \int_0^t \left\| \sqrt{k(t-s)} w_n(s) - \sqrt{k(t-s)} w_m(s) \right\|_H^2 \mathrm{d}s,$$

which implies that $\sqrt{k(t-\cdot)} w_n(\cdot) \to f$ strongly in $L^2(0,t;H)$ for some $f \in L^2(0,t;H)$. By (C.1), we obtain $f(s) = \sqrt{k(t-s)} w(s)$ for a.e. $s \in (0,t)$. Thus we deduce that

$$\sqrt{k(t-\cdot)} w_n(\cdot) \to \sqrt{k(t-\cdot)} w(\cdot) \quad \text{strongly in } L^2(0,t;H) \tag{C.2}$$

for a.e. $t \in (0,T)$.

Since (w_n) is a Cauchy sequence in $\mathcal{X}$, for any $\varepsilon > 0$ there exists $N_0 \in \mathbb{N}$ such that, for all $n, m \in \mathbb{N}$, if $n, m \geq N_0$, then

$$\operatorname*{ess\,sup}_{t\in(0,T)} \left[\mathrm{e}^{-\mu t} \int_0^t k(t-s) \|w_n(s) - w_m(s)\|_H^2 \, \mathrm{d}s \right] < \varepsilon/2.$$

Hence letting $m \to +\infty$ and employing (C.2), we have, for all $n \geq N_0$,

$$\mathrm{e}^{-\mu t} \int_0^t k(t-s) \|w_n(s) - w(s)\|_H^2 \, \mathrm{d}s \leq \varepsilon/2 \quad \text{for a.e. } t \in (0,T),$$

which implies

$$\operatorname*{ess\,sup}_{t\in(0,T)} \left[\mathrm{e}^{-\mu t} \int_0^t k(t-s) \|w_n(s) - w(s)\|_H^2 \, \mathrm{d}s \right] < \varepsilon.$$

Thus we have proved that $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ is complete.

References

[1] Affili, E., Valdinoci, E., Decay estimates for evolution equations with classical and fractional time-derivatives, J. Differential Equations, to appear.

[2] Aizicovici, S., An abstract doubly nonlinear Volterra integral equation, Funkcialaj Ekvacioj **36** (1993), 479–497.

[3] de Azevedo, E.N., de Sousa, P.L., de Souza, R.E., Engelsberg, M., Miranda, Mirla de N. do N., Silva, M.A., Concentration-dependent diffusivity and anomalous diffusion: A magnetic resonance imaging study of water ingress in porous zeolite, Phys. Rev. E **73** (2006) 011204.

[4] Barbu, V., *Nonlinear Semigroups and Differential Equations in Banach spaces*, Noordhoff, 1976.

[5] Barbu, V., Existence for nonlinear Volterra equations in Hilbert spaces, SIAM J. Math. Anal. **10** (1979), 552–569.

[6] Barbu, V., On a nonlinear Volterra integral equation on a Hilbert space, SIAM J. Math. Anal. **8** (1977), 346–355.

[7] Barbu, V., Nonlinear Volterra equations in a Hilbert space, SIAM J. Math. Anal. **6** (1975), 728–741.

[8] Brézis, H., *Operateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, Math Studies, Vol.5 North-Holland, Amsterdam/New York, 1973.

[9] Brézis, H., *Monotonicity methods in Hilbert spaces and some applications to non-linear partial differential equations*, Contributions to Nonlinear Functional Analysis, ed. Zarantonello, E., Academic Press, New York-London, 1971, pp.101–156.

[10] Clément, Ph., On abstract Volterra equations in Banach spaces with completely positive kernels, Infinite-dimensional systems (Retzhof, 1983), Lecture Notes in Math., vol.1076, Springer, Berlin, 1984, pp.32–40.

[11] Clément, Ph., On abstract Volterra equations with kernels having a positive resolvent, Israel J. Math. **36** (1980), 193–200.

[12] Clément, Ph., Nohel, J.A., Asymptotic behavior of solutions of nonlinear Volterra equations with completely positive kernels, SIAM J. Math. Anal. **12** (1981), 514–535.

[13] Clément, Ph., Nohel, J.A., Abstract linear and nonlinear Volterra equations preserving positivity, SIAM J. Math. Anal. **10** (1979), 365–388.

[14] Clément, Ph., Prüss, J., Completely positive measures and Feller semigroups, Math. Ann. **287** (1990), 73–105.

[15] Crandall, M.G., Nohel, J.A., An abstract functional differential equation and a related nonlinear Volterra equation, Israel J. Math. **29** (1978), 313–328.

[16] Diethelm, K., Monotonicity of functions and sign changes of their Caputo derivatives, Fract. Calc. Appl. Anal. **19** (2016), 561–566.

[17] Dipierro, S., Valdinoci, E., Vespri, V., Decay estimates for evolutionary equations with fractional time-diffusion, `arXiv:1707.08278 [math.AP]`.

[18] Abd El-Ghany El-Abd, Milczarek J.J., Neutron radiography study of water absorption in porous building materials: anomalous diffusion analysis, J. Phys. D: Appl. Phys. **37** (2004), 2305–2313.

[19] Gerolymatou, E., Vardoulakis, I., Hilfer, R., Modelling infiltration by means of a nonlinear fractional diffusion model, J. Physics D: Appl. Phys. **39** (2006), 4104–4110.

[20] Gorenflo, R., Luchko, Y., Yamamoto, M., Time-fractional diffusion equation in the fractional Sobolev spaces, Fract. Calc. Appl. Anal. **18** (2015), 799–820.

[21] Grafakos, L., *Classical Fourier analysis, third edition*, Graduate Texts in Mathematics, vol.249, Springer, New York, 2014.

[22] Gripenberg, G., An existence result for a nonlinear Volterra integral equation in a Hilbert space, SIAM J. Math. Anal. **9** (1978), 793–805.

[23] Gripenberg, G., An abstract nonlinear Volterra equation, Israel J. Math. **34** (1979), 198–212.

[24] Gripenberg, G., Volterra integro-differential equations with accretive nonlinearity, J. Differential Equations **60** (1985), 57–79.

[25] Hilfer, R., Anton, L., Fractional master equations and fractal time random walks, Phys. Rev. E. **51** (1995), R848–R851.

[26] Kiffe, T. A perturbation of an abstract Volterra equation, SIAM J. Math. Anal. **11** (1980), 1036–1046.

[27] Kiffe, T., Stecher, M., An abstract Volterra integral equation in a reflexive Banach space, J. Differential Equations **34** (1979), 303–325.

[28] Kilbas, A.A., Srivastava, H.M., Trujillo, J.J., *Theory and applications of fractional differential equations*, North-Holland Mathematics Studies, vol.204. Elsevier Science B.V., Amsterdam, 2006.

[29] Kōmura, Y., Nonlinear semi-groups in Hilbert space, J. Math. Soc. Japan **19** (1967), 493–507.

[30] Kemppainen, J., Siljander, J., Vergara, V., Zacher, R., Decay estimates for time-fractional and other non-local in time subdiffusion equations in $\mathbb{R}^d$, Math. Ann. **366** (2016), 941–979.

[31] Küntz, M., Lavallée, P., Experimental evidence and theoretical analysis of anomalous diffusion during water infiltration in porous building materials, J. Physics D: Appl. Phys. **34** (2001), 2547–54.

[32] Londen, S.-O., Staffans, O.J., A note on Volterra equations in a Hilbert space, Proc. Amer. Math. Soc. **70** (1978), 57–62.

[33] Luchko, Y., Yamamoto, M., General time-fractional diffusion equation: some uniqueness and existence results for the initial-boundary-value problems, Fract. Calc. Appl. Anal. **19** (2016), 1425–1434.

[34] Luchko, Y., Yamamoto, M., A survey on the recent results regarding maximum principles for the time-fractional diffusion equations, Frontiers in fractional calculus, pp.33–69, Curr. Dev. Math. Sci., vol.1, Bentham Sci. Publ., Sharjah, 2018.

[35] Metzler, R., Klafter, J., The random walk's guide to anomalous diffusion: a fractional dynamics approach, Physics Reports **339** (2000), 1–77.

[36] Miller, R.K., On Volterra integral equations with nonnegative integrable resolvents, J. Math. Anal. Appl. **22** (1968), 319–340.

[37] Montroll, E.W., Weiss, G.H., Random Walks on Lattices. II, J. Math. Phys. **6** (1965), 167–181.

[38] Giga, Y., Namba, T., Well-posedness of Hamilton-Jacobi equations with Caputo's time-fractional derivative, Commun. Partial Differential Equations **42** (2017), 1088–1120.

[39] Pazy, A., *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol.44. Springer-Verlag, New York, 1983.

[40] Płociniczak, Ł., Analytical studies of a time-fractional porous medium equation. Derivation, approximation and applications, Commun. Nonlinear Sci. Numer. Simul. **24** (2015), 169–183.

[41] Płociniczak, Ł., Approximation of the Erdélyi-Kober operator with application to the time-fractional porous medium equation, SIAM J. Appl. Math. **74** (2014), 1219–1237.

[42] Płociniczak, Okrasińska, H., Approximate self-similar solutions to a nonlinear diffusion equation with time-fractional derivative, Phys. D **261** (2013), 85–91.

[43] Podlubny, I., *Fractional differential equations, An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*, Mathematics in Science and Engineering, vol.198. Academic Press, Inc., San Diego, CA, 1999.

[44] Prüss, J., *Evolutionary integral equations and applications*, Monographs in Mathematics, vol.87. Birkhäuser Verlag, Basel, 1993.

[45] Ramos, N.M.M., Delgado, João M.P.Q., de Freitas, V.P., Anomalous diffusion during water absorption in porous building materials – experimental evidence, Defect and Diffusion Forum **273-276** (2008), 156–161.

[46] Reed, M., Simon, B., Methods of modern mathematical physics II, Fourier analysis, self-adjointness, Academic Press, New York-London, 1975.

[47] W. Liu, M. Röckner, J.L. da Silva, Quasi-linear (stochastic) partial differential equations with time-fractional derivatives, SIAM J. Math. Anal. **50** (2018), 2588–2607.

[48] Taylor, S.C., Hoff, W.D., Wilson, M., Green, K.M., Anomalous water transport properties of Portland and blended cement-based materials, J. Materials Science Letters **18**(1999), 1925–1927.

[49] Topp, E., Yangari, M., Existence and uniqueness for parabolic problems with Caputo time derivative, J. Differential Equations **262** (2017), 6018–6046.

[50] Vergara, V., Zacher, R., Optimal decay estimates for time-fractional and other nonlocal subdiffusion equations via energy methods, SIAM J. Math. Anal. **47** (2015), 210–239.

[51] Vergara, V. and Zacher, R., Lyapunov functions and convergence to steady state for differential equations of fractional order, Math. Z. **259** (2008), 287–309.

[52] Vergara, V. and Zacher, R., A priori bounds for degenerate and singular evolutionary partial integro-differential equations, Nonlinear Anal. **73** (2010), 3572–3585.

[53] Vergara, V. and Zacher, R., Stability, instability, and blowup for time fractional and other nonlocal in time semilinear subdiffusion equations, J. Evol. Equ. **17** (2017), 599–626.

[54] Zacher, R., Maximal regularity of type L_p for abstract parabolic Volterra equations, J. Evol. Equ. **5** (2005), 79–103.

[55] Zacher, R., Boundedness of weak solutions to evolutionary partial integro-differential equations with discontinuous coefficients, J. Math. Anal. Appl. **348** (2008), 137–149.

[56] Zacher, R., Weak solutions of abstract evolutionary integro-differential equations in Hilbert spaces, Funkcial. Ekvac. **52** (2009), 1–18.

Mathematical Institute, Tohoku University, Aoba, Sendai 980-8578, Japan.
Email address: goro.akagi@tohoku.ac.jp