



*Research article***A functional-analytic framework for nonlinear adaptive memory: hierarchical kernels, state-dependent sensitivity, and memory-dependent functionals****Jiahao Jiang***

School of Mathematics, Southwest Jiaotong University, Chengdu 610031, China

* **Correspondence:** Email: jiahao.jiang@swjtu.edu.cn.

Abstract: In this work, we developed a systematic framework for nonlinear adaptive memory within nonlinear functional analysis, where the influence of past events depends on both elapsed time and the state values along a trajectory. The framework comprises three hierarchical layers. First, memory kernels were classified into mathematically admissible, regular (uniformly bounded, normalized, Lipschitz), and generalized (bounded variation, possibly sign-changing) classes. Second, adaptive sensitivity functions $\Lambda(s, f(s))$ were introduced, satisfying natural conditions; a concrete construction based on historical deviation accumulation interpolated continuously between instantaneous response and history-dependent sensitivity, with an explicit Lipschitz estimate $\|\Lambda_f - \Lambda_g\|_\infty \leq L_\Lambda \|f - g\|_\infty$. Third, an adaptive memory-dependent functional

$$S_{\kappa, \Lambda}(f) = \sup_{t \in I} \left(|f(t)| + \int_0^t \Lambda(s, f(s)) \kappa(t-s) |f(s)| ds \right)$$

and the associated set $\mathcal{M}_{\kappa, \Lambda}(I) = \{f : S_{\kappa, \Lambda}(f) < \infty\}$ were constructed. Fundamental properties of the framework were established, including absolute convergence, measurability, uniform boundedness, positive definiteness, and comparison with the classical supremum norm. It was shown that $C(I) \subset \mathcal{M}_{\kappa, \Lambda}(I)$ strictly, with discontinuous functions (e.g., indicator functions of subintervals) belonging to the set, capturing abrupt signal changes such as on-off switching in nonlinear systems. When the maximum of $|f|$ was attained in $(0, T]$, a strict inequality $S_{\kappa, \Lambda}(f) > \|f\|_\infty$ was proved, demonstrating the nontrivial contribution of the memory component. The resulting framework could be viewed as a mathematical description of nonlinear adaptive phenomena, including habituation, state-dependent weighting, and selective retention, within a rigorous functional-analytic setting. By separating temporal weighting from state-dependent modulation, the construction offered a modular methodology applicable to adaptive phenomena in fields such as neuroscience, adaptive control, and machine learning, where memory is time-dependent and content-sensitive, and may be of interest across other mathematical disciplines, as a step toward a broader understanding of nonlinear phenomena from a mathematical perspective.

Keywords: nonlinear functional analysis; adaptive memory; adaptive sensitivity function; adaptive memory-dependent functional; applications in mathematical disciplines

Mathematics Subject Classification: 46N99, 46T99, 68T05

1. Introduction

The influence of past events is not just about time; it also depends on what happened. A strong event may leave a different trace than a weak one, even if both happened equally long ago. This paper attempts to develop a mathematical framework for such adaptive memory, where past events are weighted not only by when they occurred but also by their size. The framework introduces a new way to measure signals that combines their present strength with their adaptively weighted history. This work may offer a theoretical foundation for understanding habituation, selective retention, and other memory-based adaptation phenomena in nonlinear science, from the perspective of mathematical methods for nonlinear systems, and could be of relevance to fields such as neuroscience, control theory, and machine learning.

Memory, the persistence of past events in shaping present behavior, is a pervasive feature across scientific disciplines. In physical systems, heat conduction with memory effects modifies the classical diffusion paradigm, leading to anomalous large-time asymptotics that depend critically on the space-time scale under consideration [1, 2]. In material science, nonclassical diffusion equations with memory terms on time-dependent spaces exhibit complex dynamic behavior, including the existence of global attractors under general assumptions on the memory kernel [3]. In structural mechanics, Timoshenko systems incorporating memory terms display a dissipative structure whose spectral analysis yields optimal decay estimates and global existence results in critical Sobolev spaces [4]. Extensions to thermoelastic systems with time-varying delays further reveal global existence, asymptotic behavior, and uniform attractors [5]. In stochastic analysis, semilinear Volterra equations with multiplicative noise and non-analytic kernels exhibit parabolic character, with well-posedness and regularity established under Lipschitz-type conditions [6]. A related class of nonlinear Volterra-Fredholm integro-differential equations has been addressed using evolutionary computational intelligence techniques, where artificial neural networks provide accurate modeling of the integral structure [7]. In optimal control, a posteriori error estimates for finite-element discretizations of nonlinear parabolic integro-differential control problems enable adaptive multi-mesh schemes [8].

Beyond deterministic and stochastic evolution equations, memory concepts appear prominently in neural network theory. Neutral-type hybrid bidirectional associative memory networks with time-varying delays have been analyzed for robust stability, yielding delay-derivative-dependent conditions for equilibrium [9]. Hybrid fractional differential equations combined with artificial neural networks have been employed to model tuberculosis transmission dynamics, demonstrating the practical utility of memory-based modeling in epidemiology [10]. In time series analysis, spectral methods for multifractional long-range dependent functional time series characterize memory through operator-valued spectral densities, with weak-consistent estimation of long-memory operators under Gaussian assumptions [11]. In social dynamics, memory-based reduced modeling of opinion spreading, grounded in the Mori-Zwanzig projection formalism, shows that inclusion of memory

terms significantly improves prediction quality across network topologies [12]. In a related direction, further inequalities for the Wills functional of convex bodies have been established, including Brunn-Minkowski and Rogers-Shephard-type inequalities, as well as a new proof of McMullen's inequality [13]. In fluid mechanics, incompressible flows achieve exponential self-similar mixing of passive tracers, a phenomenon intimately connected to the decay of memory in transported quantities [14]. In geometry, norms on the cohomology of hyperbolic 3-manifolds relate purely topological Thurston norms to more geometric harmonic norms, with explicit constants depending on volume and injectivity radius [15].

Foundational to many of these analyses is the decomposition of functions or operators into components of definite sign or monotonicity. Classical Jordan decomposition theorems for signed measures, functions of bounded variation, and operators in Banach spaces provide essential tools for handling memory kernels that may change sign or exhibit irregular behavior [16, 17]. Further decomposition results and their extensions to maps of finite total variation and to operators in Banach spaces are established in [18–20]. In parallel, inequalities for continuous Archimedean t-norms generalize classical Minkowski-type estimates and offer a functional framework relevant to the analysis of nonlinear integral operators [21]. These ideas extend to fuzzy settings, where interval-valued fuzzy structures with t-norms provide computational tools for defuzzification in fuzzy control [22]. In approximation theory, artificial neural networks with uniform-norm-based loss functions offer alternative training paradigms when data are limited or class sizes are imbalanced [23].

Adaptive algorithms have been extensively developed for solving inverse problems and fixed-point problems. Moreover, Adaptive Gauss-Newton methods for nonlinear equations use residual bounds and quadratic regularization to achieve global convergence without prior knowledge of hyperparameters [24]. Split common fixed-point problems for demicontractive operators have been solved using adaptive algorithms that do not require operator norms, with strong convergence established under suitable conditions [25]. Inertial self-adaptive algorithms with two different inertial factors approximate minimum-norm solutions of split feasibility problems in Banach spaces, employing step-size selection without prior norm knowledge [26]. Multiple-sets split equality problems have been addressed with iterative algorithms whose split self-adaptive step sizes are computed directly from the iterative procedure [27]. Split fixed-point problems for multi-valued demi-contractive mappings admit self-adaptive algorithms with strong convergence guarantees [28]. Limited-memory bundle methods for difference-of-convex optimization enable sparse pairwise kernel learning with non-smooth loss functions, balancing prediction accuracy with sparsity requirements [29]. Maximum-norm a posteriori error bounds for extrapolated upwind schemes applied to singularly perturbed convection-diffusion problems provide robust estimators that guide adaptive mesh generation [30]. Memory-efficient combinatorial attacks on small LWE keys address the challenge of limited memory in cryptographic contexts, outperforming previous approaches when memory is constrained [31]. Physics-based active learning strategies iteratively sample training points based on physical query rules, constructing surrogate models that are accurate in industrially relevant regions of the parameter space [32]. Uniform convergence of multigrid methods for elliptic quasi-variational inequalities has been established in the maximum norm, with applications to impulse control problems [33]. Weak-duality-based adaptive finite element methods for PDE-constrained optimization with pointwise gradient state constraints derive residual-based a posteriori error estimators without requiring constraint qualifications [34]. Contextual inverse multiobjective

optimization recovers multiple objective functions and preferences from observed context-decision pairs, accommodating different norm-based scalarizations [35]. Inverse moment bounds for sample autocovariance matrices based on detrended time series provide mean-square error bounds for finite predictor coefficients under short- or long-memory dependence [36]. Probabilistic adaptive widths of multivariate Sobolev spaces equipped with Gaussian measures determine asymptotic values for approximation in L^q norms [37]. Extensions of radial basis functions address large-scale learning problems where different variables play distinct roles, with applications in artificial intelligence [38]. Bio-inspired neural network architectures modeled on the cerebellum offer avenues for embodied intelligence, where adaptive neural networks learn through physical interaction [39]. Theories of attribute implications in multi-adjoint concept lattices with hedges provide logical foundations for knowledge representation and decision-support systems [40]. The relationship between large deviation rate functions and Kullback-Leibler divergence informs the interpretation of neural estimation of mutual information [41]. Integration of privacy-enhancing technologies into explainable artificial intelligence mitigates attribute inference attacks on feature-based explanations, reducing attack success while preserving model utility [42]. Computationally enhanced projection methods for symmetric Sylvester and Lyapunov matrix equations compute residual norms at reduced cost, making Krylov strategies competitive with more recent approaches [43]. Perturbation analysis of singular values in concatenated matrices extends classical Weyl inequalities, providing stability bounds for low-rank approximations with applications in signal processing and data-driven modeling [44]. Generalizations of fractional calculus have been systematically investigated to study anomalous stochastic processes, leading to fractional Itô calculus and generalized Fokker-Planck equations that describe underdamped and overdamped stochastic dynamics [45]. Adaptive frequency evolution decomposition combined with improved fluctuation-based dispersion entropy has been proposed for muscle fatigue characterization using surface electromyography signals, demonstrating the utility of adaptive signal decomposition in biomedical applications [46]. A mathematical theory of adaptive memory has been developed for stochastic processes whose local regularity adapts dynamically in response to their own state, introducing time-varying and responsive fractional Brownian motion with rigorous analysis of covariance structure, pathwise regularity, and attention-like mechanisms [47]. A two-parameter memory-weighted velocity operator has been introduced and analyzed, providing a framework for describing rates of change in systems with time-varying, power-law memory, with weighted pointwise estimates revealing compensation mechanisms between independent memory weightings [48]. In a related direction, a semiclassical functional calculus on nilpotent Lie groups and their compact nilmanifolds has been developed, showing that this calculus includes the functional calculus of suitable subelliptic operators; moreover, Weyl laws for these operators are obtained, with examples including sub-Laplacians in horizontal divergence form perturbed by a potential and their generalizations [49]. Furthermore, asymptotic weighted approximation of convex functions by piecewise affine convex functions has been studied, where the asymptotic expressions are formulated in terms of a functional that extends the notion of affine surface area to the functional setting as the number of facets tends to infinity [50].

The above-mentioned works provide important mathematical descriptions of memory phenomena in their respective application contexts, and a number of them adopt fixed kernels based on time lags. Furthermore, in real-world systems, such as neural adaptation, gain scheduling, and content-addressable memory, the influence of a past event also depends on its magnitude or novelty at the

time it occurred. This suggests that, in addition to time dependence, state dependence may be a relevant factor in the design of memory weights. In this context, we attempt to develop a systematic functional-analytic framework in which the memory weight depends not only on the time lag, but also on the instantaneous state of the trajectory through an adaptive sensitivity function, thereby achieving a twofold dependence on both time and state. This design enables the memory mechanism to reflect more flexibly the characteristics of the signal, and may offer a possible mathematical foundation for describing nonlinear adaptive phenomena such as habituation and selective retention. This twofold dependence on time and state constitutes the point of departure for the present framework, which is organized into three hierarchical layers and will be developed in detail in the following sections.

In this work, we build upon this broad landscape of memory-related models and adaptive algorithms. The framework is developed through three successive layers: A hierarchical classification of memory kernels, a class of adaptive sensitivity functions that modulate the kernel weights according to the state of the trajectory, and a memory-dependent functional with its associated function set that integrates instantaneous magnitude and adaptively weighted historical accumulation. First, a classification of memory kernels is developed in Section 2: Starting from mathematically admissible kernels satisfying minimal integrability and measurability conditions (Definition 2.1), regular admissible kernels are introduced with uniform boundedness, normalization, and Lipschitz continuity (Definition 2.2); generalized admissible kernels further relax these requirements to allow sign changes, bounded variation, and non-unit total weight (Definition 2.3). Stationary memory kernels of the form $K_\kappa(t, s) = \kappa(t - s)$ are adopted as a foundational premise (Subsection 2.4), and prototypical examples, exponential kernel (Example 2.1), power-law kernel (Example 2.2), and finite-memory kernel (Example 2.3), illustrate the scope of the classification. Essential integral estimates for regular and generalized kernels are established in Lemma 2.1 and Proposition 2.1.

Second, adaptive sensitivity functions

$$\Lambda: I \times \mathbb{R} \rightarrow [0, \infty)$$

are introduced in Section 3 (Definition 3.1), satisfying the conditions of uniform boundedness, Lipschitz continuity in the state variable, measurability in time, and positivity at the zero state. These functions modulate the memory weight according to the state value $f(s)$ along the trajectory, yielding a composite weight $\Lambda(s, f(s))\kappa(t - s)$. A concrete construction based on historical deviation accumulation (Example 3.1) demonstrates that the framework accommodates genuinely history-dependent sensitivity, interpolating continuously between purely instantaneous response ($\beta_0 = 0$) and operator-level historical feedback ($\beta_0 > 0$). Theorem 3.1 establishes fundamental properties of the resulting function Λ_f : uniform boundedness (P1), a Lipschitz-type estimate in the supremum norm $\|\Lambda_f - \Lambda_g\|_\infty \leq L_\Lambda \|f - g\|_\infty$ (P2), continuity in time (P3), and positivity at zero (P4). Corollary 3.1 verifies that the purely instantaneous case reduces to a function $\Lambda \in \mathcal{A}(I)$.

Third, the adaptive memory-dependent functional $S_{\kappa, \Lambda}(f)$ is constructed in Section 4 via a supremum operation (Definition 4.2), integrating the instantaneous magnitude $|f(t)|$ with the adaptively weighted historical accumulation $\mathcal{J}_f(t)$ (Definition 4.1). Lemma 4.1 establishes absolute convergence, measurability, and uniform boundedness of the auxiliary functions $\mathcal{J}_f$ and $\mathcal{M}_f$. Lemma 4.2 verifies existence, controlled boundedness, positive definiteness, and comparison with the classical supremum norm for $S_{\kappa, \Lambda}(f)$. Theorem 4.1 shows that when the maximum of $|f|$ is attained in $(0, T]$, the inequality

is strict: $S_{\kappa,\Lambda}(f) > \|f\|_\infty$. The associated function set

$$\mathcal{M}_{\kappa,\Lambda}(I) = \{f : S_{\kappa,\Lambda}(f) < \infty\}$$

is introduced in Definition 4.3. Theorem 4.2 proves that $C(I) \subset \mathcal{M}_{\kappa,\Lambda}(I)$ and establishes the two-sided estimate

$$\|f\|_\infty \leq S_{\kappa,\Lambda}(f) \leq (1 + \Lambda_\infty \kappa_\infty T) \|f\|_\infty.$$

Proposition 4.1 shows that the inclusion mapping is linear, satisfies a controlled estimate, and is Lipschitz continuous. Proposition 4.2 demonstrates that the set also contains certain discontinuous functions, such as the indicator function of a subinterval, reflecting the framework's flexibility to accommodate signals with jump discontinuities arising in nonlinear phenomena like abrupt environmental changes or on-off switching.

The structure of the paper is as follows: In Section 2, we develop the hierarchical classification of memory kernels, introducing mathematically admissible, regular, and generalized kernels, together with stationary representations, prototypical examples, and essential integral estimates. In Section 3, we introduce adaptive sensitivity functions, establish their axiomatic properties, and present a concrete construction based on historical deviation accumulation, including detailed verification of uniform boundedness, Lipschitz-type estimates, continuity, and positivity at zero. In Section 4, we construct the adaptive memory-dependent functional $S_{\kappa,\Lambda}(f)$ and the associated function set $\mathcal{M}_{\kappa,\Lambda}(I)$, proving well-posedness, fundamental properties, and embedding results, including strict comparison when the maximizer lies in $(0, T]$.

Main notation and symbols

For the reader's convenience, we list below the main notation used throughout the paper. All symbols are defined in context at their first occurrence. Table 1 summarizes the main notation and symbols used throughout the paper.

Table 1. Main notation and symbols used throughout the paper.

Symbol	Description	First defined in
I	fixed time interval $[0, T]$	Section 2.1
Ω	triangular region $\{(t, s) : 0 \leq s \leq t \leq T\}$	Section 2.1
$\mathcal{K}_{\text{math}}$	class of mathematically admissible kernels	Definition 2.1
$\mathcal{K}_{\text{reg}}$	class of regular admissible kernels	Definition 2.2
$\mathcal{K}_{\text{gen}}$	class of generalized admissible kernels	Definition 2.3
$\Lambda(s, x)$	adaptive sensitivity function	Definition 3.1
$\Lambda_f(s)$	trajectory-dependent sensitivity function	Example 3.1
$\mathcal{J}_f(t)$	adaptive weighted cumulative function	Definition 4.1
$\mathcal{M}_f(t)$	instantaneous-memory hybrid function	Definition 4.1
$S_{\kappa,\Lambda}(f)$	adaptive memory-dependent functional	Definition 4.2
$\mathcal{M}_{\kappa,\Lambda}(I)$	adaptive memory-dependent sensitivity set	Definition 4.3

2. A hierarchical framework for memory kernels: concepts and estimates

We begin with the first layer of the framework: the classification of memory kernels. This part establishes the basic objects on which the subsequent layers will be built. Our purpose of this section is to introduce a self-consistent axiomatic framework for describing how past states are weighted in their influence on the present. The central object is the memory kernel, a function that quantifies the relative importance of a historical instant s when examining the system at a later time t . Kernels of this type appear across a broad spectrum of mathematical models: in heat equations with memory [1, 2], Timoshenko systems of memory type [4], nonclassical diffusion equations with memory [3], Volterra equations and Volterra-Fredholm integro-differential equations [6, 7], a posteriori error estimates for finite-element discretizations of parabolic integro-differential optimal control problems [8], neural networks with time-varying delays [9], hybrid fractional differential equations [10], long-range dependent time series [11], the development of a mathematical theory for adaptive memory [47], and the study of a two-parameter memory-weighted velocity operator [48].

A key challenge that emerges from these applications is that memory is rarely a static, time-invariant weighting of the past; instead, it often exhibits intrinsically nonlinear characteristics. Systems may adapt their responsiveness based on the intensity or frequency of past stimuli (a phenomenon known as habituation), or they may assign different weights to past events depending on the state values encountered along the trajectory; what we term state-dependent weighting. Our axiomatic framework addresses this class of phenomena by constructing a hierarchical classification of kernels that, in Sections 3 and 4, will be combined with state-dependent sensitivity functions to yield a theory of adaptive memory that accommodates such nonlinear features. This framework aims to capture the essential mathematical structures shared by the models cited above while remaining self-contained. In doing so, it provides a unified language for describing memory phenomena that are inherently nonlinear, adaptation, habituation, state-dependent weighting, within a rigorous functional-analytic setting.

We proceed by first imposing basic mathematical requirements that guarantee the well-definedness of the relevant expressions (see Definition 2.1); subsequently, to provide the necessary analytical tools for the development of the functional-analytic theory that follows, in particular, for the construction of adaptive sensitivity functions in Section 3 (see Definition 3.1) and adaptive memory sets in Section 4 (see Definitions 4.2 and 4.3), we introduce additional regularity conditions (see Definition 2.2) that will be employed throughout the main body of this work.

2.1. Mathematically admissible kernels: minimal requirements and basic properties

Throughout this work, we fix a finite time horizon $T > 0$ and denote by $I := [0, T]$ the corresponding compact time interval, which will be used throughout without further mention. Memory effects, understood as the dependence of a system's present state on its past history, will be described by functions defined on the triangular region

$$\Omega := \{(t, s) \in I \times I \mid 0 \leq s \leq t\}, \quad (2.1)$$

which reflects the fundamental causal principle that the past may influence the present, but not vice versa.

Definition 2.1 (Mathematically admissible kernel). A function $\kappa: I \rightarrow [0, \infty)$ is said to be a mathematically admissible kernel if it fulfills the following three elementary requirements:

(M1) **Non-negativity:** $\kappa(\tau) \geq 0$ for every $\tau \in I$;

(M2) **Integrability:** $\kappa \in L^1(I)$, i.e.

$$\|\kappa\|_{L^1(I)} := \int_0^T \kappa(\tau) d\tau < \infty;$$

(M3) **Measurability:** κ is Lebesgue measurable on I .

The collection of all mathematically admissible kernels will be denoted by $\mathcal{K}_{\text{math}}$.

Remark 2.1. Conditions (M1)–(M3) constitute the basic hypotheses required to ensure that all integral expressions involving κ appearing in this work are mathematically well-defined. Non-negativity guarantees that κ admits an interpretation as a weight, while integrability and measurability are standard prerequisites for the Lebesgue integration theory. At this stage, no further regularity, such as continuity or boundedness, is imposed; these will be added later as the theory develops and stronger properties become necessary. From a modeling perspective, non-negativity encodes the natural requirement that past events should not exert negative influence (i.e., no “inverse memory”), while integrability ensures that the total weight of past experience remains finite, a natural condition for any physically plausible memory mechanism.

Remark 2.2 (Stationarity as a foundational premise). For the purposes of this work, we choose to take stationary kernels as the starting point for the theoretical development. Specifically, for any $(t, s) \in \Omega$, we write

$$K_\kappa(t, s) := \kappa(t - s), \quad (2.2)$$

which expresses that the weight attributed to a historical state at time s is determined exclusively by the elapsed time $\tau := t - s \in [0, T]$. This stationarity assumption is adopted as a foundational premise for the theory; it captures the essential feature that the influence of past events evolves with the passage of time, while providing a clear structure for the subsequent analysis. A formal definition of stationary memory kernels will be given in subsection 2.4, where their role in the overall framework is further elaborated.

2.2. Regular admissible kernels: enhanced regularity conditions

The basic conditions (M1)–(M3) introduced in Definition 2.1 guarantee that the integral expressions involving the kernel are mathematically well-defined. For the development of a deeper functional-analytic theory, and to model memory mechanisms where the weighting of past events varies in a sufficiently regular, non-erratic manner, stronger regularity properties will be imposed. These additional conditions ensure that the kernel exhibits the kind of smooth, bounded behavior expected in many natural memory processes, from biological adaptation to material relaxation, while providing the analytical tools needed for a rigorous functional-analytic treatment. In this subsection, we therefore introduce additional requirements that will be in force throughout the remainder of this work.

Definition 2.2 (Regular admissible kernel). A mathematically admissible kernel $\kappa \in \mathcal{K}_{\text{math}}$ is called a regular admissible kernel if it satisfies the following three enhanced conditions:

(R1) **Uniform boundedness:** There exist constants $0 < m_\kappa \leq M_\kappa < \infty$ such that

$$m_\kappa \leq \kappa(\tau) \leq M_\kappa \quad \text{for all } \tau \in I;$$

(R2) **Normalization:**

$$\int_0^T \kappa(\tau) d\tau = 1;$$

(R3) **Lipschitz continuity:** There exists a constant $L_\kappa > 0$ such that

$$|\kappa(\tau_1) - \kappa(\tau_2)| \leq L_\kappa |\tau_1 - \tau_2| \quad \text{for all } \tau_1, \tau_2 \in I.$$

The collection of all regular admissible kernels will be denoted by $\mathcal{K}_{\text{reg}}$.

Remark 2.3 (Mathematical and physical interpretation of the regularity conditions). The regularity conditions (R1)–(R3) carry multiple layers of significance, from a mathematical and from a modeling perspective.

(R1) Uniform boundedness excludes pathological situations where the kernel would grow without bound or become arbitrarily small, thereby keeping the memory weights within a controllable range and avoiding mathematical singularities.

(R2) Normalization endows κ with a statistical interpretation as a probability density function on the time-lag interval $[0, T]$. Consequently, the integral $\int_0^t \kappa(t-s) ds$ can be viewed as the “cumulative memory intensity” up to time t . Normalization also makes the total memory strength comparable across different kernels.

(R3) Lipschitz continuity provides sufficient smoothness to ensure that the kernel evolves gradually with the time lag. This not only facilitates subsequent differential calculus involving κ , but also aligns with the physical intuition that the weighting of past events evolves in a continuous manner.

Together, these conditions single out $\mathcal{K}_{\text{reg}}$ as a class of kernels that not only possess favorable mathematical properties but also capture the essential qualitative features of many empirically observed memory processes: a finite total memory budget (normalization), a controllable range of influence (uniform boundedness), and a gradual, non-abrupt change in weighting with the passage of time (Lipschitz continuity). The uniform lower bound condition $m_\kappa > 0$ in (R1) ensures the strict positivity of integrals involving κ on subintervals of positive length, providing a solid analytical foundation for the regular kernel class. Within this setting, the generalized kernel class $\mathcal{K}_{\text{gen}}$ introduced in Definition 2.3 further relaxes this restriction, enabling kernels that vanish at endpoints or decay asymptotically, thereby covering a broader range of physically and biologically relevant memory profiles.

2.3. Generalized admissible kernels: relaxed regularity conditions

The class $\mathcal{K}_{\text{reg}}$ introduced in Subsection 2.2 possesses excellent analytical properties and provides a solid foundation for the initial development of our theoretical framework. To further enhance the flexibility and applicability of our framework, we now introduce a broader class of kernels with weaker regularity requirements, thereby accommodating memory phenomena that may involve sign changes, jump discontinuities, or non-unit total weight, features that arise naturally in contexts such as inhibitory neural feedback, abrupt system resets, or adaptive gain control.

Definition 2.3 (Generalized admissible kernel). A function $\kappa : I \rightarrow \mathbb{R}$ is called a generalized admissible kernel if it satisfies the following three conditions:

(G1) **Essential boundedness:** $\kappa \in L^\infty(I)$, i.e., there exists a constant $C_\kappa > 0$ such that

$$\operatorname{ess\,sup}_{\tau \in I} |\kappa(\tau)| \leq C_\kappa,$$

where $\operatorname{ess\,sup}$ denotes the essential supremum. This condition enables the kernel to exceed the bound C_κ on a set of measure zero, offering greater flexibility in modeling applications.

(G2) **Integrability and non-degeneracy:** $\kappa \in L^1(I)$, i.e.,

$$\|\kappa\|_{L^1(I)} := \int_0^T |\kappa(\tau)| d\tau < \infty,$$

and there exists a constant $\delta_\kappa > 0$ such that

$$\left| \int_0^T \kappa(\tau) d\tau \right| \geq \delta_\kappa.$$

The latter inequality ensures that the total “cumulative weight” of the kernel is non-zero, excluding the trivial case where positive and negative contributions cancel out and the overall memory effect vanishes.

(G3) **Bounded variation:** κ is a function of bounded variation on $I = [0, T]$. More precisely, the total variation of κ on I is defined as

$$\operatorname{Var}_I(\kappa) := \sup_{\mathcal{P}} \sum_{i=1}^n |\kappa(\tau_i) - \kappa(\tau_{i-1})|,$$

where the supremum is taken over all partitions $\mathcal{P} = \{\tau_0, \tau_1, \dots, \tau_n\}$ of I satisfying

$$0 = \tau_0 < \tau_1 < \dots < \tau_n = T.$$

We require $\operatorname{Var}_I(\kappa) < \infty$, i.e., $\kappa \in BV(I)$.

The collection of all generalized admissible kernels will be denoted by $\mathcal{K}_{\text{gen}}$.

Remark 2.4 (Interpretation of the generalized conditions). We now elaborate on the mathematical content and modelling significance of each condition.

- **Essential boundedness (G1).** For a Lebesgue measurable function $f : E \rightarrow \mathbb{R} \cup \{\pm\infty\}$, the essential supremum is defined as

$$\operatorname{ess\,sup}_{x \in E} f(x) := \inf \{M \in \mathbb{R} \cup \{+\infty\} : f(x) \leq M \text{ for almost every } x \in E\},$$

or equivalently,

$$\operatorname{ess\,sup}_{x \in E} f(x) = \inf \{M \in \mathbb{R} : \mu(\{x \in E : f(x) > M\}) = 0\},$$

with the convention that the infimum over an empty set is $+\infty$ (hence $\operatorname{ess\,sup} f = +\infty$ when the set $\{M \in \mathbb{R} : \mu(\{x \in E : f(x) > M\}) = 0\}$ is empty). This notion relaxes the pointwise supremum by ignoring exceptional values on null sets, which is often more natural in applications.

- **Non-degeneracy condition (G2).** The requirement

$$\left| \int_0^T \kappa(\tau) d\tau \right| \geq \delta_\kappa > 0$$

is essential. If $\int_0^T \kappa(\tau) d\tau = 0$, the memory term $\int_0^t \kappa(t-s)u(s) ds$ could reduce to a functional where positive and negative contributions cancel, potentially leading to a diminished or vanishing overall memory effect. The constant δ_κ excludes such degenerate cases, ensuring that the kernel retains a nontrivial cumulative influence.

- **Bounded variation (G3).** Functions of bounded variation possess several key properties that are relevant in the context of memory kernels: They are differentiable almost everywhere, admit a Jordan decomposition as the difference of two monotone nondecreasing functions [17–19], and have at most countably many discontinuities, all of which are of the first kind (jump discontinuities). In the context of memory kernels, the bounded variation condition allows for jump-like changes, capturing abrupt adjustments or resets in the memory mechanism, while ensuring that the total oscillatory magnitude $\text{Var}_{[0,T]}(\kappa)$ remains finite, thereby guaranteeing the stability and controllability of the weighted historical accumulation.

Remark 2.5 (Distinctive features and role of the generalized kernel class). The generalized class $\mathcal{K}_{\text{gen}}$ is designed to extend modeling capabilities in three principal directions:

- **Sign flexibility:** Negative values of κ are permitted, enabling the description of inhibitory effects or negative feedback mechanisms arising from past states.
- **Relaxed regularity:** Kernels of bounded variation are admissible; global Lipschitz continuity is not imposed. Hence, kernels with (at most countably many) jump discontinuities are permitted, which is suitable for capturing abrupt changes, resets, or piecewise constant memory weights.
- **Optional normalization:** The integral $\int_0^T \kappa(\tau) d\tau$ may take any non-zero value, with its magnitude controlled by the constant δ_κ . This provides modeling freedom for systems where the overall memory intensity may vary or be adaptively regulated.

The generalized class $\mathcal{K}_{\text{gen}}$ introduced here serves as an interface for potential extensions: Under additional technical conditions, many of the results obtained for $\mathcal{K}_{\text{reg}}$ are expected to carry over to $\mathcal{K}_{\text{gen}}$, thereby covering an even broader spectrum of memory phenomena and application scenarios.

2.4. Stationary memory kernels and their representation

In many physical, biological, and engineering systems [1–3], the manner in which memory weighting varies frequently depends upon the time elapsed since a past event, as opposed to the absolute moment at which that event occurred. This characteristic, referred to as time-shift stationarity, offers a natural and widely adopted simplification for modeling memory processes. Related studies in structural mechanics, adaptive algorithms, and time series analysis [4, 26, 36], as well as in bio-inspired neural networks and adaptive signal decomposition [39, 46], further illustrate the broad relevance of memory-related modeling. On the basis of the kernel classes introduced in the preceding subsections, we now provide a formal definition of stationary memory kernels.

Definition 2.4 (Stationary memory kernel). Let $\kappa : I \rightarrow \mathbb{R}$ be a mathematically admissible kernel in the sense of Definition 2.1 (more specifically, belonging to $\mathcal{K}_{\text{reg}}$ or $\mathcal{K}_{\text{gen}}$). The associated stationary

memory kernel $K_\kappa : \Omega \rightarrow \mathbb{R}$ is defined by

$$K_\kappa(t, s) := \kappa(t - s), \quad 0 \leq s \leq t \leq T. \quad (2.3)$$

Consequently, the weight attributed to a historical state is determined entirely by the time lag

$$\tau := t - s \in [0, T],$$

and does not depend on the particular historical instant s or the current moment t .

Remark 2.6 (Interpretation of the stationarity assumption). The relation

$$K_\kappa(t, s) = \kappa(t - s)$$

encapsulates two fundamental aspects of stationary memory:

- (1) **Time-translation invariance:** The memory characteristics of the system remain unchanged under shifts of the time axis; translating the timeline by any amount leaves the memory profile unaltered.
- (2) **Dependence on relative time:** The weighting factor is determined exclusively by the elapsed time $\tau = t - s$ since the historical event, and is independent of the specific moment s at which the event occurred. This reflects the principle that the influence of a past event is governed solely by how much time has passed, without reference to absolute calendar time.

The stationarity assumption is adopted throughout this work as a foundational premise for the theoretical development. It captures the characteristic feature that the influence of past events varies with the passage of time, while providing a clear and coherent structure for the analysis that follows.

Remark 2.7 (Notational convention). Unless explicitly stated otherwise, the term “kernel” shall refer to a stationary memory kernel

$$K_\kappa(t, s) = \kappa(t - s)$$

generated by some $\kappa \in \mathcal{K}_{\text{reg}}$ according to Definition 2.4. Depending on the context, we shall employ interchangeably the binary notation $K_\kappa(t, s)$ and the univariate notation $\kappa(\tau)$, as they are essentially equivalent under the stationarity hypothesis. All subsequent developments are conducted within this stationary framework, thereby ensuring clarity and consistency throughout the presentation. For regular admissible kernels $\kappa \in \mathcal{K}_{\text{reg}}$, the range of K_κ is restricted to $[0, \infty)$; for generalized admissible kernels $\kappa \in \mathcal{K}_{\text{gen}}$, the range is allowed to be $\mathbb{R}$ in accordance with the possible sign changes of κ .

2.5. Typical examples of kernel functions

To render the abstract kernel classes introduced above more concrete and to demonstrate that they encompass classical and physically important memory phenomena, in this subsection, we present three prototypical examples that are widely encountered in applications. These examples not only validate the reasonableness of Definitions 2.2 and 2.3, but also furnish concrete objects for the theoretical analysis that follows.

Example 2.1 (Exponential kernel in the regular class). A representative example belonging to the regular admissible kernel class $\mathcal{K}_{\text{reg}}$ is provided by the exponential kernel. Let $\alpha > 0$ be a rate parameter and define

$$\kappa_\alpha^{\text{reg}}(\tau) := \frac{\alpha e^{-\alpha\tau}}{1 - e^{-\alpha T}}, \quad \tau \in I. \quad (2.4)$$

This kernel possesses the following properties:

- **Basic mathematical properties:** It satisfies all the conditions (M1)–(M3) in Definition 2.1 for mathematically admissible kernels. Thus, $\kappa_\alpha^{\text{reg}} \geq 0$ and is continuous, and its integral is given by

$$\int_0^T \kappa_\alpha^{\text{reg}}(\tau) d\tau = \frac{\alpha}{1 - e^{-\alpha T}} \int_0^T e^{-\alpha\tau} d\tau = 1.$$

- **Verification of regularity:** This kernel belongs to the regular class $\mathcal{K}_{\text{reg}}$. First, for any $\tau \in I$,

$$0 < \frac{\alpha e^{-\alpha T}}{1 - e^{-\alpha T}} \leq \kappa_\alpha^{\text{reg}}(\tau) \leq \frac{\alpha}{1 - e^{-\alpha T}} < \infty,$$

which confirms the uniform boundedness condition (R1). The normalization condition (R2) follows directly from the integral evaluation above. For Lipschitz continuity (R3), we compute the derivative

$$\frac{d}{d\tau} \kappa_\alpha^{\text{reg}}(\tau) = -\frac{\alpha^2 e^{-\alpha\tau}}{1 - e^{-\alpha T}},$$

whose absolute value is bounded on I by

$$\frac{\alpha^2}{1 - e^{-\alpha T}};$$

consequently, $\kappa_\alpha^{\text{reg}}$ is Lipschitz continuous with constant

$$L_\kappa = \frac{\alpha^2}{1 - e^{-\alpha T}}.$$

- **Physical interpretation:** This kernel describes the classical “exponential” weighting pattern. The memory weight varies exponentially with the elapsed time τ , and parameter α controls the rate of this variation: Larger values of α correspond to a more rapid decrease, giving prominence to recent history, while smaller α imply a more gradual variation, allowing more distant events to retain some influence. This model is a standard choice for describing short-term memory behavior.

Example 2.2 (Power-law kernel in the generalized class). As a concrete instance of the generalized admissible kernel class $\mathcal{K}_{\text{gen}}$ (see Definition 2.3), let $\gamma \in (0, 1)$ and $\varepsilon > 0$, and define

$$\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) := \frac{1-\gamma}{T^{1-\gamma}} (T - \tau + \varepsilon)^{-\gamma}, \quad \tau \in I. \quad (2.5)$$

This kernel exhibits the following properties:

- **Essential boundedness (G1):** Since $\varepsilon > 0$, the denominator satisfies $(T - \tau + \varepsilon) \geq \varepsilon > 0$ for all $\tau \in I$. Consequently,

$$0 < \frac{1-\gamma}{T^{1-\gamma}} (T + \varepsilon)^{-\gamma} \leq \kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) \leq \frac{1-\gamma}{T^{1-\gamma}} \varepsilon^{-\gamma} < \infty,$$

which shows that $\kappa_{\gamma,\varepsilon}^{\text{gen}}$ is uniformly bounded on I . More precisely,

$$\text{ess sup}_{\tau \in I} |\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau)| \leq \frac{1-\gamma}{T^{1-\gamma}} \varepsilon^{-\gamma}.$$

Taking

$$C_\kappa := \frac{1-\gamma}{T^{1-\gamma}} \varepsilon^{-\gamma},$$

we have

$$\operatorname{ess\,sup}_{\tau \in I} |\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau)| \leq C_\kappa < \infty,$$

confirming that $\kappa_{\gamma,\varepsilon}^{\text{gen}} \in L^\infty(I)$ and thus satisfies condition (G1).

- **Integrability and non-degeneracy (G2):** Computing the integral explicitly,

$$\begin{aligned} \int_0^T \kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) d\tau &= \frac{1-\gamma}{T^{1-\gamma}} \int_0^T (T-\tau+\varepsilon)^{-\gamma} d\tau \\ &= \frac{1-\gamma}{T^{1-\gamma}} \left[-\frac{(T-\tau+\varepsilon)^{1-\gamma}}{1-\gamma} \right]_{\tau=0}^{\tau=T} \\ &= \frac{(T+\varepsilon)^{1-\gamma} - \varepsilon^{1-\gamma}}{T^{1-\gamma}}. \end{aligned}$$

Since $\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) > 0$ for all $\tau \in I$, we have $|\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau)| = \kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau)$, and therefore

$$\|\kappa_{\gamma,\varepsilon}^{\text{gen}}\|_{L^1(I)} = \int_0^T \kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) d\tau = \frac{(T+\varepsilon)^{1-\gamma} - \varepsilon^{1-\gamma}}{T^{1-\gamma}} < \infty,$$

which verifies the integrability condition $\kappa_{\gamma,\varepsilon}^{\text{gen}} \in L^1(I)$. Moreover,

$$\left| \int_0^T \kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) d\tau \right| = \frac{(T+\varepsilon)^{1-\gamma} - \varepsilon^{1-\gamma}}{T^{1-\gamma}} > 0.$$

Taking

$$\delta_\kappa := \frac{(T+\varepsilon)^{1-\gamma} - \varepsilon^{1-\gamma}}{T^{1-\gamma}},$$

we obtain

$$\left| \int_0^T \kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) d\tau \right| = \delta_\kappa > 0,$$

thereby satisfying the non-degeneracy condition in (G2). A closer examination of this integral reveals an interesting feature: For any fixed $\varepsilon > 0$, one has

$$\int_0^T \kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) d\tau < 1.$$

To verify this inequality, consider the auxiliary function $f : [0, \infty) \rightarrow \mathbb{R}$ defined by

$$f(z) := (1+z)^{1-\gamma} - z^{1-\gamma}, \quad z \geq 0.$$

With the change of variable $z = \varepsilon/T \geq 0$, the integral becomes

$$\int_0^T \kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) d\tau = f\left(\frac{\varepsilon}{T}\right).$$

Proof of the inequality. The function f is continuous on $[0, \infty)$ and differentiable on $(0, \infty)$. Its derivative is given by

$$f'(z) = (1 - \gamma) [(1 + z)^{-\gamma} - z^{-\gamma}], \quad z > 0.$$

Since $\gamma \in (0, 1)$, the map $t \mapsto t^{-\gamma}$ is strictly decreasing on $(0, \infty)$; consequently, for any $z > 0$,

$$(1 + z)^{-\gamma} < z^{-\gamma},$$

which implies $f'(z) < 0$ for all $z > 0$. Hence, f is strictly decreasing on $(0, \infty)$. Now, fix an arbitrary $\varepsilon > 0$ and set $z_0 = \varepsilon/T > 0$. Applying the Mean Value Theorem on the interval $[0, z_0]$, there exists $\xi \in (0, z_0)$ such that

$$f(z_0) - f(0) = f'(\xi)z_0.$$

Noting that $f(0) = 1$ and that $f'(\xi) < 0$, we obtain

$$f(z_0) = 1 + f'(\xi)z_0 < 1.$$

Substituting back yields

$$\int_0^T \kappa_{\gamma, \varepsilon}^{\text{gen}}(\tau) d\tau = f\left(\frac{\varepsilon}{T}\right) < 1.$$

- **Bounded variation (G3):** The function $\kappa_{\gamma, \varepsilon}^{\text{gen}}$ is monotonically increasing on I . Indeed, for any $0 \leq \tau_1 < \tau_2 \leq T$, we have $(T - \tau_1 + \varepsilon) > (T - \tau_2 + \varepsilon)$, and since the exponent $-\gamma$ is negative, it follows that

$$(T - \tau_1 + \varepsilon)^{-\gamma} < (T - \tau_2 + \varepsilon)^{-\gamma};$$

consequently $\kappa_{\gamma, \varepsilon}^{\text{gen}}(\tau_1) < \kappa_{\gamma, \varepsilon}^{\text{gen}}(\tau_2)$. By monotonicity, for any partition $\mathcal{P} : 0 = \tau_0 < \tau_1 < \dots < \tau_n = T$ of I , the sum of absolute differences reduces to the difference of the endpoint values:

$$\sum_{i=1}^n |\kappa_{\gamma, \varepsilon}^{\text{gen}}(\tau_i) - \kappa_{\gamma, \varepsilon}^{\text{gen}}(\tau_{i-1})| = \sum_{i=1}^n [\kappa_{\gamma, \varepsilon}^{\text{gen}}(\tau_i) - \kappa_{\gamma, \varepsilon}^{\text{gen}}(\tau_{i-1})] = \kappa_{\gamma, \varepsilon}^{\text{gen}}(T) - \kappa_{\gamma, \varepsilon}^{\text{gen}}(0).$$

This value is independent of the particular partition $\mathcal{P}$, and therefore the total variation is given by

$$\text{Var}_{[0, T]}(\kappa_{\gamma, \varepsilon}^{\text{gen}}) = \sup_{\mathcal{P}} (\kappa_{\gamma, \varepsilon}^{\text{gen}}(T) - \kappa_{\gamma, \varepsilon}^{\text{gen}}(0)) = \kappa_{\gamma, \varepsilon}^{\text{gen}}(T) - \kappa_{\gamma, \varepsilon}^{\text{gen}}(0) < \infty,$$

which establishes (G3).

- **Relation to the regular class:** It is instructive to examine how $\kappa_{\gamma, \varepsilon}^{\text{gen}}$ relates to the regular class $\mathcal{K}_{\text{reg}}$. The kernel satisfies two of the three conditions defining $\mathcal{K}_{\text{reg}}$, namely uniform boundedness (R1) and Lipschitz continuity (R3).

– **Uniform boundedness (R1):** As already shown in the verification of (G1), for any $\tau \in I$,

$$0 < \frac{1 - \gamma}{T^{1-\gamma}} (T + \varepsilon)^{-\gamma} \leq \kappa_{\gamma, \varepsilon}^{\text{gen}}(\tau) \leq \frac{1 - \gamma}{T^{1-\gamma}} \varepsilon^{-\gamma} < \infty,$$

which confirms (R1).

– **Lipschitz continuity (R3):** Differentiating $\kappa_{\gamma,\varepsilon}^{\text{gen}}$ yields

$$\frac{d}{d\tau}\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) = \frac{\gamma(1-\gamma)}{T^{1-\gamma}}(T-\tau+\varepsilon)^{-\gamma-1},$$

and this derivative is uniformly bounded on I ; indeed,

$$\sup_{\tau \in I} \left| \frac{d}{d\tau}\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau) \right| \leq \frac{\gamma(1-\gamma)}{T^{1-\gamma}} \varepsilon^{-\gamma-1} < \infty.$$

By the Mean Value Theorem, this implies that $\kappa_{\gamma,\varepsilon}^{\text{gen}}$ is Lipschitz continuous on I with Lipschitz constant

$$L_{\kappa} = \frac{\gamma(1-\gamma)}{T^{1-\gamma}} \varepsilon^{-\gamma-1},$$

thereby satisfying condition (R3).

As observed in the verification of (G2) above, the integral of $\kappa_{\gamma,\varepsilon}^{\text{gen}}$ is strictly less than 1 for any $\varepsilon > 0$; consequently, the normalization condition (R2) is not satisfied. Hence, $\kappa_{\gamma,\varepsilon}^{\text{gen}}$ does not belong to $\mathcal{K}_{\text{reg}}$. (When $\varepsilon = 0$, the integral equals 1, but in that case $\kappa_{\gamma,0}^{\text{gen}}$ becomes unbounded at $\tau = T$, violating (R1) and therefore also lies outside $\mathcal{K}_{\text{reg}}$.) This observation illustrates the purpose of the generalized class $\mathcal{K}_{\text{gen}}$: it accommodates physically reasonable kernels that possess the core analytical properties (boundedness and regularity) of the regular class, yet are not normalized (or have adjustable total weight) and therefore reside outside $\mathcal{K}_{\text{reg}}$.

- **Physical interpretation:** For sufficiently small regularization parameter $\varepsilon > 0$, this kernel describes a weighting pattern that evolves slowly with the time lag, following the power law

$$(T - \tau + \varepsilon)^{-\gamma}.$$

The parameter $\gamma \in (0, 1)$ controls the rate of this evolution. Consider two historical instants $0 \leq \tau_1 < \tau_2 \leq T$; their weight ratio is

$$\frac{\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau_2)}{\kappa_{\gamma,\varepsilon}^{\text{gen}}(\tau_1)} = \left(\frac{T - \tau_1 + \varepsilon}{T - \tau_2 + \varepsilon} \right)^{\gamma}.$$

Since $T - \tau_1 + \varepsilon > T - \tau_2 + \varepsilon$, this ratio exceeds 1, indicating that the weight varies with the elapsed time in a manner that assigns greater influence to more distant events. For fixed τ_1, τ_2 and ε , the ratio as a function of γ , namely

$$\gamma \mapsto \left(\frac{T - \tau_1 + \varepsilon}{T - \tau_2 + \varepsilon} \right)^{\gamma},$$

is strictly increasing on $(0, 1)$ (because the base is greater than 1). Thus, smaller γ yield a ratio closer to 1, corresponding to a more gradual change and allowing more distant events to retain noticeable influence, while larger γ produce a steeper variation, further amplifying the influence of more distant events. Parameter $\varepsilon > 0$ serves a dual purpose: Mathematically, it eliminates the singularity at $\tau = T$, ensuring that the kernel fully satisfies the technical requirements of the generalized class; physically, it may be interpreted as a small adjustment to the weighting of the very recent past (where τ is close to T), reflecting a finite resolution in the system's perception of the present moment.

Thus, $\kappa_{\gamma,\varepsilon}^{\text{gen}} \in \mathcal{K}_{\text{gen}}$, providing a typical and physically meaningful example of the generalized kernel class.

Example 2.3 (Finite-memory kernel in the generalized class). Another illustrative example, which belongs to the generalized class $\mathcal{K}_{\text{gen}}$ (see Definition 2.3) but not to the regular class $\mathcal{K}_{\text{reg}}$ (see Definition 2.2) (except in the limiting case $\Delta = T$), is provided by the finite-memory kernel. Let $\Delta \in (0, T]$ be a parameter representing the length of the memory window, and define

$$\kappa_{\Delta}^{\text{fm}}(\tau) := \begin{cases} \frac{1}{\Delta}, & 0 \leq \tau \leq \Delta, \\ 0, & \Delta < \tau \leq T. \end{cases} \quad (2.6)$$

This kernel exhibits the following features:

- **Basic mathematical properties:** It satisfies the basic conditions (M1)–(M3) required for mathematically admissible kernels. The function is clearly non-negative, measurable, and its integral evaluates to

$$\int_0^T \kappa_{\Delta}^{\text{fm}}(\tau) d\tau = \int_0^{\Delta} \frac{1}{\Delta} d\tau = 1.$$

- **Relation to the regular class:** The kernel $\kappa_{\Delta}^{\text{fm}}$ fulfills two of the three conditions characterizing $\mathcal{K}_{\text{reg}}$, but fails to satisfy the Lipschitz continuity requirement (R3) when $\Delta < T$.

- **Uniform boundedness (R1):** For every $\tau \in I$, one has

$$0 \leq \kappa_{\Delta}^{\text{fm}}(\tau) \leq 1/\Delta,$$

so (R1) holds.

- **Normalization (R2):** As shown above,

$$\int_0^T \kappa_{\Delta}^{\text{fm}}(\tau) d\tau = 1,$$

hence (R2) is satisfied.

- **Lipschitz continuity (R3):** When $\Delta < T$, the kernel possesses a jump discontinuity at $\tau = \Delta$. Considering the difference quotient

$$\frac{|\kappa_{\Delta}^{\text{fm}}(\Delta + \epsilon) - \kappa_{\Delta}^{\text{fm}}(\Delta - \epsilon)|}{2\epsilon} = \frac{\frac{1}{\Delta}}{2\epsilon} = \frac{1}{2\Delta\epsilon},$$

which becomes unbounded as $\epsilon \rightarrow 0^+$, one sees that $\kappa_{\Delta}^{\text{fm}}$ cannot be Lipschitz continuous on I ; consequently (R3) is not fulfilled (in the limiting case $\Delta = T$, the kernel reduces to the constant function $1/T$ on $[0, T]$, which is Lipschitz continuous and actually belongs to $\mathcal{K}_{\text{reg}}$).

- **Membership in the generalized class:** Despite the lack of Lipschitz continuity, $\kappa_{\Delta}^{\text{fm}}$ meets all three conditions defining the generalized class $\mathcal{K}_{\text{gen}}$.

- **Essential boundedness (G1):** Clearly

$$\text{ess sup}_{\tau \in I} |\kappa_{\Delta}^{\text{fm}}(\tau)| = 1/\Delta < \infty.$$

- **Integrability and non-degeneracy (G2):** One has

$$\int_0^T |\kappa_{\Delta}^{\text{fm}}(\tau)| d\tau = 1,$$

$$\left| \int_0^T \kappa_{\Delta}^{\text{fm}}(\tau) d\tau \right| = 1,$$

so (G2) is satisfied (with $\delta_{\kappa} = 1$).

- **Bounded variation (G3):** The function $\kappa_{\Delta}^{\text{fm}}$ is piecewise constant with a single jump of height $1/\Delta$ at $\tau = \Delta$. For any partition $\mathcal{P}$ of I , the sum of absolute differences cannot exceed this jump height, and by taking partitions that include points arbitrarily close to Δ from both sides, the total variation is seen to be exactly $1/\Delta$. Hence, $\text{Var}_{[0,T]}(\kappa_{\Delta}^{\text{fm}}) = 1/\Delta < \infty$, which establishes (G3).

Therefore $\kappa_{\Delta}^{\text{fm}} \in \mathcal{K}_{\text{gen}}$.

- **Physical interpretation:** This kernel models a system with a strict finite memory horizon: Only the most recent Δ time units are retained, while everything older is completely forgotten. Such a weighting pattern arises naturally in digital systems, devices with limited storage capacity, or sliding-window filters used in signal processing and real-time applications. Parameter Δ governs the memory capacity: As $\Delta \rightarrow 0^+$, the memory window shrinks to zero, so that in the limit, the system retains information only about the present instant, effectively approaching a memoryless regime; when $\Delta = T$, the kernel is constant on the whole interval, representing uniform weighting of the entire past; for intermediate values $\Delta \in (0, T)$, the system possesses a finite but non-trivial memory.

Remark 2.8 (Significance and selection of the examples). The three examples presented above illustrate fundamental memory patterns that are both physically relevant and mathematically tractable. The exponential kernel $\kappa_{\alpha}^{\text{exp}}$ (Example 2.1) belongs to the regular class $\mathcal{K}_{\text{reg}}$ and captures short-memory behavior, where the weighting assigned to past events diminishes with increasing time lag. The power-law kernel $\kappa_{\gamma,\varepsilon}^{\text{gen}}$ (Example 2.2) represents a contrasting long-memory pattern: due to its increasing nature as a function of the time lag, it assigns greater weight to more distant past events, thereby complementing the short-memory scenario. The finite-memory kernel $\kappa_{\Delta}^{\text{fm}}$ (Example 2.3) belongs to the generalized class $\mathcal{K}_{\text{gen}}$ and illustrates how kernels with discontinuities can still be accommodated within the framework while retaining essential analytical properties such as bounded variation. Together, these examples demonstrate the flexibility of the proposed kernel classes in capturing a range of memory phenomena.

2.6. Essential integral estimates for admissible kernels

Integral estimates for kernel functions play a fundamental role in the analysis of memory-dependent operators. In this subsection, we establish a set of basic integral control inequalities. The estimates are formulated primarily for regular admissible kernels $\kappa \in \mathcal{K}_{\text{reg}}$ (see Definition 2.2); their favorable properties, in particular boundedness and normalization, lead to particularly concise forms.

Lemma 2.1 (Integral control estimates for regular kernels). *Let $\kappa \in \mathcal{K}_{\text{reg}}$ be a regular admissible kernel in the sense of Definition 2.2. Throughout, I denotes the fixed time interval $[0, T]$. Then the following estimates hold:*

- (1) **Cumulative weight estimate:** For every $t \in I$,

$$\int_0^t \kappa(t-s) ds \leq 1. \quad (2.7)$$

In particular, the normalization condition $\int_0^T \kappa(\tau) d\tau = 1$ is satisfied.

(2) **Weighted supremum estimate:** For every $f \in C(I)$ and every $t \in I$,

$$\int_0^t \kappa(t-s)|f(s)| ds \leq \|f\|_\infty. \quad (2.8)$$

Here, $\|f\|_\infty := \sup_{s \in I} |f(s)|$ denotes the usual supremum norm on $C(I)$.

(3) **Integral control of differences:** For every $f, g \in C(I)$ and every $t \in I$,

$$\left| \int_0^t \kappa(t-s)(f(s)-g(s)) ds \right| \leq \|f-g\|_\infty. \quad (2.9)$$

Proof. (1) By the change of variables $\tau = t-s$, one obtains

$$\int_0^t \kappa(t-s) ds = \int_0^t \kappa(\tau) d\tau.$$

The non-negativity of κ (condition (R1) in Definition 2.2) together with the normalization condition (R2) yields

$$\int_0^t \kappa(\tau) d\tau \leq \int_0^T \kappa(\tau) d\tau = 1.$$

(2) From $\kappa \geq 0$ and the definition of the supremum norm,

$$\int_0^t \kappa(t-s)|f(s)| ds \leq \sup_{s \in [0,t]} |f(s)| \int_0^t \kappa(t-s) ds \leq \|f\|_\infty \int_0^t \kappa(\tau) d\tau.$$

Applying estimate (1) gives $\|f\|_\infty \cdot 1 = \|f\|_\infty$.

(3) Applying estimate (2) with f replaced by $f-g$ yields

$$\left| \int_0^t \kappa(t-s)(f(s)-g(s)) ds \right| \leq \int_0^t \kappa(t-s)|f(s)-g(s)| ds \leq \|f-g\|_\infty.$$

The first inequality follows from the triangle inequality for integrals together with the non-negativity of κ . $\square$

Proposition 2.1 (Integral estimates for generalized kernels). *Let $\kappa \in \mathcal{K}_{\text{gen}}$ be a generalized admissible kernel as introduced in Definition 2.3. Then the following estimates hold:*

(1) **Non-negative case:** If $\kappa(\tau) \geq 0$ for almost every $\tau \in I$, then for every $f \in C(I)$ and every $t \in I$,

$$\int_0^t \kappa(t-s)|f(s)| ds \leq \|f\|_\infty \int_0^T \kappa(\tau) d\tau. \quad (2.10)$$

(2) **General case (possibly sign-changing):** For every $f, g \in C(I)$ and every $t \in I$,

$$\left| \int_0^t \kappa(t-s)f(s) ds \right| \leq \|f\|_\infty \int_0^T |\kappa(\tau)| d\tau, \quad (2.11)$$

$$\left| \int_0^t \kappa(t-s)(f(s)-g(s)) ds \right| \leq \|f-g\|_\infty \int_0^T |\kappa(\tau)| d\tau. \quad (2.12)$$

Proof. We prove each case separately.

Proof of (1). Assume $\kappa(\tau) \geq 0$ for almost every $\tau \in I$. Then $|\kappa(\tau)| = \kappa(\tau)$ almost everywhere. Using the non-negativity of κ and the definition of the supremum norm,

$$\int_0^t \kappa(t-s) |f(s)| ds \leq \sup_{s \in [0,t]} |f(s)| \int_0^t \kappa(t-s) ds \leq \|f\|_\infty \int_0^t \kappa(\tau) d\tau.$$

Since $\kappa \geq 0$, one has

$$\int_0^t \kappa(\tau) d\tau \leq \int_0^T \kappa(\tau) d\tau,$$

which yields the desired estimate.

Proof of (2). For arbitrary $\kappa \in \mathcal{K}_{\text{gen}}$, condition (G2) in Definition 2.3 guarantees $\kappa \in L^1(I)$ with

$$\int_0^T |\kappa(\tau)| d\tau < \infty.$$

For any $t \in I$, the triangle inequality for integrals gives

$$\left| \int_0^t \kappa(t-s) f(s) ds \right| \leq \int_0^t |\kappa(t-s)| |f(s)| ds \leq \|f\|_\infty \int_0^t |\kappa(t-s)| ds.$$

By the change of variables $\tau = t - s$, the last integral equals

$$\int_0^t |\kappa(\tau)| d\tau \leq \int_0^T |\kappa(\tau)| d\tau,$$

establishing (2.11). Estimate (2.12) follows by applying the same argument to $f - g$ in place of f . $\square$

Remark 2.9. The estimates established in this subsection provide basic tools for handling integral expressions involving kernel functions. Lemma 2.1 presents three fundamental estimates for regular kernels $\kappa \in \mathcal{K}_{\text{reg}}$, relying on their non-negativity, boundedness, and normalization. Proposition 2.1 extends these estimates to the broader class $\mathcal{K}_{\text{gen}}$, where the L^1 norm $\int_0^T |\kappa(\tau)| d\tau$ naturally appears in place of the constant 1. In the non-negative case, the factor reduces to $\int_0^T \kappa(\tau) d\tau$, consistent with the structure of the regular estimates. Together, these results illustrate how the integral estimates adapt to the regularity properties of the kernel, offering flexibility for the regular and generalized settings considered in this work.

3. Adaptive sensitivity functions: definition, construction, and basic properties

We now turn to the second layer of the framework: the adaptive sensitivity functions, which modulate the kernel weights according to the state of the trajectory. This layer introduces a state-dependent mechanism that works in tandem with the kernel hierarchy established in the previous section. The kernel-theoretic framework developed in Section 2 provides a hierarchical classification of memory kernels based on their regularity and integrability properties, capturing the temporal weighting patterns through stationary kernels of the form $K_\kappa(t, s) = \kappa(t - s)$. This framework, with its emphasis on the temporal dimension of memory, characterizes the weight assigned to a past state $f(s)$ as a function

$\kappa(t - s)$ of the elapsed time. The adaptive sensitivity function introduced below retains this temporal weighting as a foundational layer, while incorporating an additional mechanism that modulates the weight according to the state value $f(s)$. Specifically, we augment the kernel κ with a sensitivity factor $\Lambda(s, f(s))$, yielding a composite memory weight of the form $\Lambda(s, f(s))\kappa(t - s)$. This construction gives mathematical expression to two key capabilities: The capacity for selective retention of historically salient events, and the capacity for differentiated assessment based on state-dependent importance.

3.1. Mathematical framework of adaptive sensitivity functions

The memory formalism developed in the preceding section establishes a foundational temporal weighting mechanism, wherein the influence of a past state $f(s)$ is mediated through the kernel κ via $K_\kappa(t, s) = \kappa(t - s)$. This construction captures the key idea that the weight attributed to a historical instant depends naturally on the time elapsed. The adaptive framework proposed below retains this temporal weighting as a core component, while extending it to incorporate sensitivity to the state values. Specifically, we introduce a modulation factor $\Lambda(s, f(s))$ that works in tandem with $\kappa(t - s)$, yielding a composite weight $\Lambda(s, f(s))\kappa(t - s)$. The following definition formalizes this extended object.

Definition 3.1 (Adaptive sensitivity function). Let $I = [0, T]$ denote the fixed time horizon. A function $\Lambda : I \times \mathbb{R} \rightarrow [0, \infty)$ is termed an adaptive sensitivity function if it satisfies the following four conditions: (AS1) **Uniform boundedness:** There exist constants $0 < \lambda_{\min} \leq \lambda_{\max} < \infty$, such that

$$\lambda_{\min} \leq \Lambda(s, x) \leq \lambda_{\max} \quad \text{for all } (s, x) \in I \times \mathbb{R}.$$

This condition precludes unbounded amplification and unphysical vanishing to zero, thereby maintaining the regularity essential for subsequent analytical developments.

(AS2) **Lipschitz continuity with respect to the state variable:** There exists a constant $L_\Lambda > 0$, such that

$$|\Lambda(s, x) - \Lambda(s, y)| \leq L_\Lambda |x - y| \quad \text{for all } s \in I, x, y \in \mathbb{R}.$$

The Lipschitz requirement guarantees that sensitivity adjustments occur gradually as the state varies, precluding abrupt jumps that would otherwise complicate physical interpretation and mathematical treatment.

(AS3) **Measurability with respect to the temporal variable:** For every fixed $x \in \mathbb{R}$, the mapping $s \mapsto \Lambda(s, x)$ is Lebesgue measurable on I . This measurability requirement is a fundamental prerequisite for the Lebesgue integration theory, ensuring that expressions, such as

$$\int_0^t \Lambda(s, f(s))\kappa(t - s)|f(s)| ds$$

are mathematically well-defined.

(AS4) **Positivity at the zero state:** There exists a constant $\varrho > 0$, such that

$$\Lambda(s, 0) \geq \varrho \quad \text{for all } s \in I.$$

This condition excludes the degenerate case $\Lambda(s, 0) = 0$, which could compromise the distinctive status of the zero function in the functional-analytic framework.

The collection of all adaptive sensitivity functions satisfying conditions (AS1)–(AS4) will be denoted by $\mathcal{A}(I)$.

In this definition, the variable $x \in \mathbb{R}$ denotes a generic state value; when such a function is applied to a specific system trajectory $f \in C(I)$, we write $\Lambda(s, f(s))$, where $f(s) \in \mathbb{R}$ is the value of f at time s . This is precisely the form appearing in expressions such as

$$\int_0^t \Lambda(s, f(s)) \kappa(t-s) |f(s)| ds,$$

which will be central to the construction of the adaptive memory-dependent functional in the following subsection (see Definition 4.2).

Remark 3.1 (Systematic rationale underlying the sensitivity conditions). The four conditions encapsulated in Definition 3.1 constitute a carefully balanced axiomatic system that reconciles mathematical tractability with modeling fidelity:

- **Complementary roles of (AS1) and (AS4):** While condition (AS1) provides global bounds controlling the extreme excursions of Λ , condition (AS4) safeguards against vanishing sensitivity at the distinguished state $x = 0$. Together, they prevent numerical singularities and physical degeneracies, ensuring that the sensitivity mechanism remains operative across admissible states.
- **Analytical significance of (AS2):** Beyond its natural physical interpretation, sensitivity ought to vary continuously with the underlying state, the Lipschitz condition furnishes the regularity necessary for subsequent estimates involving compositions of Λ with continuous functions. This property will be essential when examining the continuity properties of functionals defined through adaptive memory weights.
- **Relation to the kernel hierarchy:** The adaptive sensitivity function Λ operates in conjunction with the kernel classes $\mathcal{K}_{\text{reg}}$ and $\mathcal{K}_{\text{gen}}$ introduced in Section 2. While those classes characterize the temporal weighting patterns through κ , the present construction introduces a complementary mechanism that modulates intensity according to state values. The composite expression $\Lambda(s, f(s))\kappa(t-s)$ thereby encodes the objective, time-driven weighting of past events and the system's capacity to assess their significance based on content. This dual structure enables the description of nuanced memory phenomena such as heightened retention of salient episodes or suppression of routine fluctuations.
- **Balance between generality and analytical utility:** Each condition has been formulated to be as weak as possible while still providing the essential properties required for the subsequent development of the theory. Stronger requirements (e.g., differentiability with respect to x , joint continuity, or monotonicity) have been deliberately avoided to preserve maximal generality, thereby enabling $\mathcal{A}(I)$ to accommodate a broad spectrum of potential applications.

The axiomatic framework established here thus achieves a synthesis of mathematical rigor and physical plausibility, laying a foundation for the construction of adaptive memory structures.

Remark 3.2 (Modular interplay between kernel classes and sensitivity functions). A distinctive feature of the present framework lies in the modular separation between the temporal weighting mechanism, encoded by $\kappa \in \mathcal{K}_{\text{reg}} \cup \mathcal{K}_{\text{gen}}$, and the state-dependent modulation, encoded by $\Lambda \in \mathcal{A}(I)$. This conceptual division offers considerable flexibility: One may select any kernel from the hierarchy established in Section 2 and combine it with any sensitivity function satisfying Definition 3.1, tailoring

the resulting composite memory structure to specific modeling requirements while preserving the benefits of a unified theoretical treatment.

3.2. A constructive example: sensitivity with historical deviation accumulation

The axiomatic framework presented in Definition 3.1 delineates a general class of functions $\mathcal{A}(I)$ characterized by boundedness, Lipschitz regularity in the state variable, measurability in time, and positivity at the zero state. Within this abstract setting, a question arises: Does $\mathcal{A}(I)$ contain functions whose sensitivity mechanism reflects more intricate dependence on the historical evolution of the state, beyond instantaneous values? In this subsection, we address this question by exhibiting a concrete construction in which the sensitivity at a given state incorporates a weighted accumulation of past deviations from a reference trajectory. This construction serves to illustrate the scope of the axiomatic framework and to provide a tangible object for subsequent analysis.

Example 3.1 (Adaptive sensitivity based on historical deviations). Let $I = [0, T]$ and let $r \in C(I)$ be a given continuous function, interpreted as a reference trajectory representing the nominal or expected evolution of the system. Choose parameters $\alpha_0 > 0$ governing the persistence of historical influence, $\gamma_0 > 0$ controlling the sensitivity to instantaneous deviations, and $\beta_0 \geq 0$ modulating the strength of the historical feedback.

For any given trajectory $f \in C(I)$, define the function $\Lambda_f : I \rightarrow [0, \infty)$ by

$$\Lambda_f(s) := \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \frac{\tanh(\gamma_0 |f(s) - r(s)|)}{1 + \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |f(\tau) - r(\tau)|) d\tau}, \quad (3.1)$$

where the constants satisfy $0 < \lambda_{\min} < \lambda_{\max} < \infty$, and $\tanh(z) = (e^z - e^{-z})/(e^z + e^{-z})$ denotes the hyperbolic tangent function.

The object Λ_f defined above has two distinct interpretations depending on the parameter β_0 : For $\beta_0 = 0$, it reduces to a function $\Lambda(s, x) \in \mathcal{A}(I)$; for $\beta_0 > 0$, it is trajectory-dependent and can be regarded as an operator $f \mapsto \Lambda_f$.

Interpretation. For a fixed trajectory f , the denominator integral in (3.1) accumulates the actual weighted deviations of f from the reference r over the interval $[0, s]$. This cumulative historical deviation then modulates the sensitivity at time s through the denominator. The construction thus encodes a genuine historical feedback mechanism: If f has persistently deviated from r in the past, the sensitivity $\Lambda_f(s)$ is attenuated, even if the current deviation $|f(s) - r(s)|$ is large.

When $\beta_0 = 0$, the construction reduces to a purely instantaneous sensitivity that can be represented as a function $\Lambda \in \mathcal{A}(I)$ given by

$$\Lambda(s, x) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \tanh(\gamma_0 |x - r(s)|),$$

so that for any trajectory f , we have $\Lambda_f(s) = \Lambda(s, f(s))$. Here $x \in \mathbb{R}$ denotes a generic state value, while $f(s)$ is the specific value of f at time s . A detailed verification that this Λ indeed belongs to $\mathcal{A}(I)$ is provided in Corollary 3.1 below. For $\beta_0 > 0$, the sensitivity acquires genuine historical dependence and is most naturally viewed as an operator mapping each trajectory f to the function Λ_f defined above.

The construction incorporates several interrelated features:

- **Bounded quantification of instantaneous deviations.** The numerator $\tanh(\gamma_0|f(s) - r(s)|)$ maps the absolute deviation $|f(s) - r(s)| \in [0, \infty)$ to the bounded interval $[0, 1)$. The saturation property of the hyperbolic tangent tempers the influence of extreme deviations while preserving ordinal information; parameter γ_0 controls the slope of this function near the origin, i.e., the sensitivity to small deviations; it thus governs the gain of the instantaneous response.
- **Exponentially weighted accumulation of historical deviations.** Define the weighted historical deviation at time s by

$$\mathcal{D}_f(s) := \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0|f(\tau) - r(\tau)|) d\tau.$$

This quantity aggregates past deviations with an exponential weight $e^{-\alpha_0(s-\tau)}$ that diminishes as the historical time τ recedes from the present s . This weighting pattern reflects the principle that more recent episodes exert stronger influence on the current sensitivity; parameter α_0 determines the temporal extent of this memory: Larger α_0 corresponds to shorter persistence (emphasizing recent deviations), while smaller α_0 permits more persistent historical influence.

- **Adaptive modulation through a feedback mechanism.** The denominator introduces a self-regulating effect: If the trajectory f has historically exhibited sustained deviations from the reference trajectory over an extended period, the accumulated integral becomes large, thereby attenuating the growth of $\Lambda_f(s)$. This negative feedback may be viewed as a mathematical analog of adaptation phenomena observed in certain biological and engineered systems, where sustained anomalous stimuli lead to diminished responsiveness, thereby preventing excessive sensitivity to recurrent deviations [11, 41, 42].

A particularly illustrative scenario arises when the trajectory f closely follows the reference r . In this case, $|f(s) - r(s)|$ and $|f(\tau) - r(\tau)|$ are small for all relevant times. For small arguments, the hyperbolic tangent satisfies $\tanh(z) = z + o(z)$ as $z \rightarrow 0$, so that $\tanh(\gamma_0|f(s) - r(s)|)$ is well approximated by $\gamma_0|f(s) - r(s)|$ up to higher-order corrections; an analogous statement holds for the integrand in the denominator. Consequently, the numerator and the integral term in the denominator are small, and $\Lambda_f(s)$ remains close to its minimum value $\lambda_{\min}$, reflecting a low sensitivity to near-reference behavior. This limiting behavior underscores the consistency of the construction: trajectories that adhere to the nominal evolution receive little modulation, while significant deviations are required to elevate the sensitivity.

Parameter β_0 controls the strength of this feedback: When $\beta_0 = 0$, the construction reduces to the purely instantaneous case described above, while positive β_0 introduces historical dependence.

- **Analytical tractability.** The hyperbolic tangent function is Lipschitz continuous on $[0, \infty)$ with Lipschitz constant 1, strictly increasing, and satisfies $0 \leq \tanh(z) < 1$ for all $z \geq 0$. The composition $\tanh(\gamma_0|\cdot|)$ inherits these properties. The exponential kernel $e^{-\alpha_0(s-\tau)}$ guarantees that the integral term is well-defined and preserves the regularity of the integrand. These properties facilitate the verification that, for each fixed f , the function Λ_f exhibits characteristics analogous to conditions (AS1)–(AS4) in Definition 3.1.

This construction extends the notion of memory sensitivity from a static or instantaneously determined mapping to one that incorporates information about the historical behavior of the system trajectory. The resulting sensitivity function captures a dynamic interplay between present deviation and past experience, suggesting some possible modeling directions for systems where responsiveness adapts based on accumulated history.

Theorem 3.1 (Basic properties of Λ_f). *Under the assumptions of Example 3.1—namely, $r \in C(I)$, $\alpha_0 > 0$, $\gamma_0 > 0$, $\beta_0 \geq 0$, and $0 < \lambda_{\min} < \lambda_{\max} < \infty$ —for any fixed trajectory $f \in C(I)$, the function $\Lambda_f : I \rightarrow [0, \infty)$ defined by (3.1) satisfies the following properties:*

(P1) **Uniform boundedness:** $\lambda_{\min} \leq \Lambda_f(s) \leq \lambda_{\max}$ for all $s \in I$.

(P2) **Lipschitz-type estimate with respect to the supremum norm:** For any two trajectories $f, g \in C(I)$,

$$\|\Lambda_f - \Lambda_g\|_{\infty} \leq L_{\Lambda} \|f - g\|_{\infty}, \quad (3.2)$$

where $\|h\|_{\infty} := \sup_{t \in I} |h(t)|$ and

$$L_{\Lambda} = (\lambda_{\max} - \lambda_{\min})\gamma_0 \left(1 + \frac{\beta_0(1 - e^{-\alpha_0 T})}{\alpha_0} \right).$$

(P3) **Continuity (hence measurability) in time:** For each fixed trajectory $f \in C(I)$, the function $s \mapsto \Lambda_f(s)$ is continuous on I , and consequently Lebesgue measurable.

(P4) **Positivity at zero:** For the zero trajectory $f \equiv 0$, one has $\Lambda_0(s) \geq \lambda_{\min} > 0$ for all $s \in I$.

These properties are the precise analogues of conditions (AS1)–(AS4) in Definition 3.1, now formulated for the operator-induced function Λ_f .

Proof. Fix an arbitrary $f \in C(I)$.

Proof of (P1): uniform boundedness. For any $s \in I$, the properties of the hyperbolic tangent yield

$$0 \leq \tanh(\gamma_0 |f(s) - r(s)|) < 1, \quad 0 \leq \tanh(\gamma_0 |f(\tau) - r(\tau)|) < 1 \quad \text{for all } \tau \in [0, s].$$

Since $\beta_0 \geq 0$ and $e^{-\alpha_0(s-\tau)} \geq 0$, the denominator satisfies

$$1 \leq 1 + \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |f(\tau) - r(\tau)|) d\tau. \quad (3.3)$$

Consequently,

$$0 \leq \frac{\tanh(\gamma_0 |f(s) - r(s)|)}{1 + \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |f(\tau) - r(\tau)|) d\tau} < 1. \quad (3.4)$$

Substituting into the definition of Λ_f and using $0 < \lambda_{\min} < \lambda_{\max}$ gives

$$\lambda_{\min} \leq \Lambda_f(s) \leq \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \cdot 1 = \lambda_{\max},$$

which establishes (P1).

Proof of (P2): Lipschitz-type estimate with respect to the supremum norm. We now establish the Lipschitz-type estimate for Λ_f . For any two trajectories $f, g \in C(I)$ and any $s \in I$, write

$$\Lambda_f(s) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min})G_f(s), \quad \Lambda_g(s) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min})G_g(s),$$

where

$$G_f(s) := \frac{\tanh(\gamma_0 |f(s) - r(s)|)}{1 + \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |f(\tau) - r(\tau)|) d\tau},$$

and $G_g(s)$ is defined analogously with g in place of f . To establish this estimate, we aim to control $|G_f(s) - G_g(s)|$ by $\|f - g\|_{\infty}$.

To this end, we first establish a fundamental Lipschitz estimate for the hyperbolic tangent function.

Lemma 3.1 (Lipschitz continuity of $\tanh(\gamma_0|\cdot - p|)$). *Let $p \in \mathbb{R}$ be fixed and let $\gamma_0 > 0$. Then the function $\psi_{p,\gamma_0}(x) := \tanh(\gamma_0|x - p|)$ is γ_0 -Lipschitz continuous with respect to x ; that is, for all $x, y \in \mathbb{R}$,*

$$|\tanh(\gamma_0|x - p|) - \tanh(\gamma_0|y - p|)| \leq \gamma_0 |x - y|. \quad (3.5)$$

Proof. We first compute the derivative of the hyperbolic tangent:

$$\frac{d}{dz} \tanh(z) = \operatorname{sech}^2(z) = \frac{1}{\cosh^2(z)}.$$

For any real z , by the properties of the exponential function and the arithmetic-geometric mean inequality,

$$\cosh(z) = \frac{e^z + e^{-z}}{2} \geq \sqrt{e^z \cdot e^{-z}} = 1,$$

with equality if and only if $e^z = e^{-z}$, i.e., $z = 0$. Hence, $\cosh^2(z) \geq 1$, and consequently

$$0 < \operatorname{sech}^2(z) = \frac{1}{\cosh^2(z)} \leq 1 \quad (\forall z \in \mathbb{R}). \quad (3.6)$$

Now consider $\psi_{p,\gamma_0}(x) = \tanh(\gamma_0|x - p|)$. Write it in piecewise form:

$$\psi_{p,\gamma_0}(x) = \begin{cases} \tanh(\gamma_0(p - x)), & x \leq p, \\ \tanh(\gamma_0(x - p)), & x \geq p. \end{cases}$$

Set

$$\psi_-(x) = \tanh(\gamma_0(p - x)) \quad (x \leq p), \quad \psi_+(x) = \tanh(\gamma_0(x - p)) \quad (x \geq p),$$

so that $\psi_-(p) = \psi_+(p) = \tanh(0) = 0$. Since $\tanh$ is continuous and the linear functions $t \mapsto \gamma_0(p - t)$ and $t \mapsto \gamma_0(t - p)$ are continuous, ψ_- is continuous on $(-\infty, p]$ and ψ_+ is continuous on $[p, \infty)$. At $x = p$,

$$\lim_{x \rightarrow p^-} \psi_{p,\gamma_0}(x) = \lim_{x \rightarrow p^-} \psi_-(x) = \psi_-(p) = 0, \quad \lim_{x \rightarrow p^+} \psi_{p,\gamma_0}(x) = \lim_{x \rightarrow p^+} \psi_+(x) = \psi_+(p) = 0,$$

and $\psi_{p,\gamma_0}(p) = \tanh(0) = 0$, so ψ_{p,γ_0} is continuous at $x = p$.

For $x < p$, by the chain rule,

$$\psi'_-(x) = \operatorname{sech}^2(\gamma_0(p - x)) \cdot \gamma_0 \cdot (-1),$$

hence

$$|\psi'_-(x)| = \gamma_0 \operatorname{sech}^2(\gamma_0(p - x)) \leq \gamma_0 \quad (\forall x < p),$$

where the last inequality uses (3.6).

For $x > p$, similarly,

$$\psi'_+(x) = \operatorname{sech}^2(\gamma_0(x - p)) \cdot \gamma_0,$$

and thus

$$|\psi'_+(x)| = \gamma_0 \operatorname{sech}^2(\gamma_0(x - p)) \leq \gamma_0 \quad (\forall x > p).$$

At $x = p$, the right derivative is

$$\begin{aligned}\psi'_+(p^+) &= \lim_{h \rightarrow 0^+} \frac{\psi_{p,\gamma_0}(p+h) - \psi_{p,\gamma_0}(p)}{h} \\ &= \lim_{h \rightarrow 0^+} \frac{\tanh(\gamma_0 h)}{h} = \gamma_0 \lim_{h \rightarrow 0^+} \frac{\tanh(\gamma_0 h)}{\gamma_0 h} = \gamma_0,\end{aligned}\quad (3.7)$$

where we used

$$\lim_{z \rightarrow 0} \frac{\tanh z}{z} = 1.$$

The left derivative is

$$\psi'_-(p^-) = \lim_{h \rightarrow 0^+} \frac{\psi_{p,\gamma_0}(p-h) - \psi_{p,\gamma_0}(p)}{-h} = \lim_{h \rightarrow 0^+} \frac{\tanh(\gamma_0 h)}{-h} = -\gamma_0 \lim_{h \rightarrow 0^+} \frac{\tanh(\gamma_0 h)}{\gamma_0 h} = -\gamma_0.$$

Thus, ψ_{p,γ_0} is not differentiable at $x = p$, but the absolute values of the left and right derivatives are γ_0 .

We now prove the Lipschitz inequality. Take arbitrary $x, y \in \mathbb{R}$ and consider two cases.

Case 1: x and y lie on the same side of p (i.e., $x \leq y \leq p$ or $p \leq x \leq y$). If $x = y$, the inequality

$$|\psi_{p,\gamma_0}(x) - \psi_{p,\gamma_0}(y)| \leq \gamma_0 |x - y|$$

holds trivially. Assume $x < y$. Then ψ_{p,γ_0} is differentiable on the interval $[x, y]$ (using ψ_- if $x, y \leq p$, or ψ_+ if $x, y \geq p$). By the Mean Value Theorem, there exists $\xi \in (x, y)$ such that

$$\psi_{p,\gamma_0}(y) - \psi_{p,\gamma_0}(x) = \psi'_{p,\gamma_0}(\xi)(y - x),$$

where $\psi'_{p,\gamma_0}(\xi) = \psi'_-(\xi)$ or $\psi'_+(\xi)$. Consequently,

$$|\psi_{p,\gamma_0}(y) - \psi_{p,\gamma_0}(x)| = |\psi'_{p,\gamma_0}(\xi)| \cdot |y - x| \leq \gamma_0 |y - x|.$$

If $x > y$, exchanging x and y yields the same estimate. Hence, the inequality holds for all x, y on the same side of p :

$$|\psi_{p,\gamma_0}(x) - \psi_{p,\gamma_0}(y)| \leq \gamma_0 |x - y|. \quad (3.8)$$

Case 2: x and y lie on opposite sides of p . Then the interval $[x, y]$ contains the point p where ψ_{p,γ_0} is not differentiable. We consider two subcases.

Subcase 2.1: $x \leq p \leq y$. Split the difference into two parts:

$$|\psi_{p,\gamma_0}(y) - \psi_{p,\gamma_0}(x)| \leq |\psi_{p,\gamma_0}(y) - \psi_{p,\gamma_0}(p)| + |\psi_{p,\gamma_0}(p) - \psi_{p,\gamma_0}(x)|.$$

For $|\psi_{p,\gamma_0}(y) - \psi_{p,\gamma_0}(p)|$: Since $p \leq y$, on the interval $[p, y]$ the function ψ_{p,γ_0} coincides with ψ_+ . By the Mean Value Theorem, there exists $\xi_1 \in (p, y)$, such that

$$\psi_+(y) - \psi_+(p) = \psi'_+(\xi_1)(y - p).$$

Since $|\psi'_+(\xi_1)| \leq \gamma_0$, and $\psi_+(y) = \psi_{p,\gamma_0}(y)$, $\psi_+(p) = \psi_{p,\gamma_0}(p)$, we obtain

$$|\psi_{p,\gamma_0}(y) - \psi_{p,\gamma_0}(p)| = |\psi_+(y) - \psi_+(p)| \leq \gamma_0 |y - p|.$$

For $|\psi_{p,\gamma_0}(p) - \psi_{p,\gamma_0}(x)|$: Since $x \leq p$, on the interval $[x, p]$ the function ψ_{p,γ_0} coincides with ψ_- . By the Mean Value Theorem, there exists $\xi_2 \in (x, p)$, such that

$$\psi_-(p) - \psi_-(x) = \psi'_-(\xi_2)(p - x).$$

Again $|\psi'_-(\xi_2)| \leq \gamma_0$, and $\psi_-(p) = \psi_{p,\gamma_0}(p)$, $\psi_-(x) = \psi_{p,\gamma_0}(x)$, hence

$$|\psi_{p,\gamma_0}(p) - \psi_{p,\gamma_0}(x)| = |\psi_-(p) - \psi_-(x)| \leq \gamma_0|p - x|.$$

Combining these estimates yields

$$|\psi_{p,\gamma_0}(y) - \psi_{p,\gamma_0}(x)| \leq \gamma_0(|y - p| + |p - x|) = \gamma_0|y - x|. \quad (3.9)$$

Subcase 2.2: $y \leq p \leq x$. Similarly, split the difference:

$$|\psi_{p,\gamma_0}(x) - \psi_{p,\gamma_0}(y)| \leq |\psi_{p,\gamma_0}(x) - \psi_{p,\gamma_0}(p)| + |\psi_{p,\gamma_0}(p) - \psi_{p,\gamma_0}(y)|.$$

For $|\psi_{p,\gamma_0}(x) - \psi_{p,\gamma_0}(p)|$: Since $p \leq x$, on $[p, x]$ we use ψ_+ , obtaining

$$|\psi_{p,\gamma_0}(x) - \psi_{p,\gamma_0}(p)| \leq \gamma_0|x - p|.$$

For $|\psi_{p,\gamma_0}(p) - \psi_{p,\gamma_0}(y)|$: Since $y \leq p$, on $[y, p]$ we use ψ_- , obtaining

$$|\psi_{p,\gamma_0}(p) - \psi_{p,\gamma_0}(y)| \leq \gamma_0|p - y|.$$

Hence,

$$|\psi_{p,\gamma_0}(x) - \psi_{p,\gamma_0}(y)| \leq \gamma_0(|x - p| + |p - y|) = \gamma_0|x - y|. \quad (3.10)$$

From Cases 1 and 2, we conclude that for all $x, y \in \mathbb{R}$,

$$|\psi_{p,\gamma_0}(x) - \psi_{p,\gamma_0}(y)| \leq \gamma_0|x - y|. \quad (3.11)$$

Returning to the original notation, this is precisely

$$|\tanh(\gamma_0|x - p|) - \tanh(\gamma_0|y - p|)| \leq \gamma_0|x - y| \quad (\forall x, y \in \mathbb{R}),$$

so ψ_{p,γ_0} is γ_0 -Lipschitz continuous. $\square$

Now we proceed to estimate $G_f(s)$. For clarity, introduce the temporary notation

$$\begin{aligned} A_f(s) &:= \tanh(\gamma_0|f(s) - r(s)|), \\ B_f(s) &:= 1 + \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0|f(\tau) - r(\tau)|) d\tau, \end{aligned}$$

so that $G_f(s) = A_f(s)/B_f(s)$. Analogous quantities $A_g(s)$ and $B_g(s)$ are defined with g in place of f .

From (P1), we already know that $0 \leq A_f(s) < 1$ and $0 \leq A_g(s) < 1$. Moreover, as established in (3.3), the positivity of the parameters ($\beta_0 \geq 0$, $e^{-\alpha_0(s-\tau)} \geq 0$) together with $\tanh(\cdot) \geq 0$ implies $B_f(s) \geq 1$ and $B_g(s) \geq 1$; in particular, $B_f(s)$ and $B_g(s)$ are strictly positive.

Consider the difference $|G_f(s) - G_g(s)|$. By a common denominator,

$$\begin{aligned} |G_f(s) - G_g(s)| &= \left| \frac{A_f(s)}{B_f(s)} - \frac{A_g(s)}{B_g(s)} \right| \\ &= \left| \frac{A_f(s) B_g(s) - A_g(s) B_f(s)}{B_f(s) B_g(s)} \right| \\ &= \frac{|A_f(s) B_g(s) - A_g(s) B_f(s)|}{B_f(s) B_g(s)}. \end{aligned} \quad (3.12)$$

Since $B_f(s)$ and $B_g(s)$ are positive, the denominator $B_f(s)B_g(s)$ is positive and therefore equals its absolute value.

To estimate the numerator, decompose it algebraically:

$$\begin{aligned} A_f(s)B_g(s) - A_g(s)B_f(s) &= A_f(s)B_g(s) - A_f(s)B_f(s) + A_f(s)B_f(s) - A_g(s)B_f(s) \\ &= A_f(s) [B_g(s) - B_f(s)] + B_f(s) [A_f(s) - A_g(s)]. \end{aligned} \quad (3.13)$$

Substituting (3.13) into (3.12) and applying the triangle inequality yields

$$|G_f(s) - G_g(s)| \leq \frac{A_f(s) |B_g(s) - B_f(s)|}{B_f(s) B_g(s)} + \frac{B_f(s) |A_f(s) - A_g(s)|}{B_f(s) B_g(s)}. \quad (3.14)$$

We now estimate the two terms on the right-hand side separately.

Estimate of the first term. Since $A_f(s) \in [0, 1]$ by (P1), we have $A_f(s) \leq 1$. Together with $B_f(s), B_g(s) \geq 1$ from (3.3), we obtain

$$\frac{A_f(s) |B_g(s) - B_f(s)|}{B_f(s) B_g(s)} \leq |B_g(s) - B_f(s)|.$$

Now estimate $|B_g(s) - B_f(s)|$. From the definition,

$$\begin{aligned} |B_g(s) - B_f(s)| &= \left| \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |g(\tau) - r(\tau)|) d\tau - \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |f(\tau) - r(\tau)|) d\tau \right| \\ &= \beta_0 \left| \int_0^s e^{-\alpha_0(s-\tau)} [\tanh(\gamma_0 |g(\tau) - r(\tau)|) - \tanh(\gamma_0 |f(\tau) - r(\tau)|)] d\tau \right|. \end{aligned} \quad (3.15)$$

Applying the triangle inequality for integrals and using $e^{-\alpha_0(s-\tau)} \geq 0$, we obtain

$$|B_g(s) - B_f(s)| \leq \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} |\tanh(\gamma_0 |g(\tau) - r(\tau)|) - \tanh(\gamma_0 |f(\tau) - r(\tau)|)| d\tau.$$

For the integrand difference, apply Lemma 3.1 with $p = r(\tau)$. This yields

$$|\tanh(\gamma_0 |g(\tau) - r(\tau)|) - \tanh(\gamma_0 |f(\tau) - r(\tau)|)| \leq \gamma_0 |g(\tau) - f(\tau)| = \gamma_0 |f(\tau) - g(\tau)|.$$

Substituting this estimate,

$$|B_g(s) - B_f(s)| \leq \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \gamma_0 |f(\tau) - g(\tau)| d\tau$$

$$\begin{aligned}
&= \beta_0 \gamma_0 \int_0^s e^{-\alpha_0(s-\tau)} |f(\tau) - g(\tau)| d\tau \\
&\leq \beta_0 \gamma_0 \|f - g\|_\infty \int_0^s e^{-\alpha_0(s-\tau)} d\tau,
\end{aligned} \tag{3.16}$$

where $\|f - g\|_\infty := \sup_{t \in I} |f(t) - g(t)|$. Computing the remaining integral,

$$\int_0^s e^{-\alpha_0(s-\tau)} d\tau = \left[\frac{e^{-\alpha_0(s-\tau)}}{\alpha_0} \right]_{\tau=0}^{\tau=s} = \frac{1 - e^{-\alpha_0 s}}{\alpha_0}.$$

Thus, we obtain the estimate for the first term:

$$\frac{A_f(s) |B_g(s) - B_f(s)|}{B_f(s) B_g(s)} \leq \beta_0 \gamma_0 \|f - g\|_\infty \cdot \frac{1 - e^{-\alpha_0 s}}{\alpha_0}. \tag{3.17}$$

Estimate of the second term. From $B_f(s) \geq 1$ and $B_g(s) \geq 1$ established in (3.3), we have

$$\frac{B_f(s)}{B_f(s) B_g(s)} = \frac{1}{B_g(s)} \leq 1,$$

and therefore

$$\frac{B_f(s) |A_f(s) - A_g(s)|}{B_f(s) B_g(s)} \leq |A_f(s) - A_g(s)|.$$

For $|A_f(s) - A_g(s)|$, apply Lemma 3.1 with $p = r(s)$. This yields

$$|A_f(s) - A_g(s)| = \left| \tanh(\gamma_0 |f(s) - r(s)|) - \tanh(\gamma_0 |g(s) - r(s)|) \right| \leq \gamma_0 |f(s) - g(s)| \leq \gamma_0 \|f - g\|_\infty.$$

Hence, the second term is bounded by

$$\frac{B_f(s) |A_f(s) - A_g(s)|}{B_f(s) B_g(s)} \leq \gamma_0 \|f - g\|_\infty. \tag{3.18}$$

Combining the two estimates. Substituting (3.17) and (3.18) into (3.14) yields

$$|G_f(s) - G_g(s)| \leq \beta_0 \gamma_0 \|f - g\|_\infty \cdot \frac{1 - e^{-\alpha_0 s}}{\alpha_0} + \gamma_0 \|f - g\|_\infty = \gamma_0 \left(1 + \frac{\beta_0(1 - e^{-\alpha_0 s})}{\alpha_0} \right) \|f - g\|_\infty. \tag{3.19}$$

Define the function

$$L_G(s) := \gamma_0 \left(1 + \frac{\beta_0(1 - e^{-\alpha_0 s})}{\alpha_0} \right).$$

Then for every $s \in I$ and any two trajectories $f, g \in C(I)$,

$$|G_f(s) - G_g(s)| \leq L_G(s) \|f - g\|_\infty.$$

We now examine the behavior of $L_G(s)$ on the interval $I = [0, T]$. Since $\alpha_0 > 0$, $\beta_0 \geq 0$, and $\gamma_0 > 0$, the function $s \mapsto 1 - e^{-\alpha_0 s}$ is nonnegative and strictly increasing on $[0, \infty)$. Consequently, $L_G(s)$ is strictly increasing on $[0, \infty)$ (and constant only when $\beta_0 = 0$). In particular,

- $L_G(0) = \gamma_0$;

- $L_G(s)$ increases with s ;

-

$$\lim_{s \rightarrow \infty} L_G(s) = \gamma_0 \left(1 + \frac{\beta_0}{\alpha_0} \right).$$

On the finite interval $I = [0, T]$, monotonicity implies that the maximum of L_G is attained at the right endpoint:

$$\max_{s \in [0, T]} L_G(s) = L_G(T) = \gamma_0 \left(1 + \frac{\beta_0 (1 - e^{-\alpha_0 T})}{\alpha_0} \right). \quad (3.20)$$

Set

$$L_G := \max_{s \in [0, T]} L_G(s) = \gamma_0 \left(1 + \frac{\beta_0 (1 - e^{-\alpha_0 T})}{\alpha_0} \right) < \infty.$$

Then for all $s \in I$ and any $f, g \in C(I)$,

$$|G_f(s) - G_g(s)| \leq L_G \|f - g\|_\infty.$$

Finally, recalling that $\Lambda_f(s) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min})G_f(s)$ and similarly for $\Lambda_g(s)$, we obtain for any $s \in I$,

$$|\Lambda_f(s) - \Lambda_g(s)| = (\lambda_{\max} - \lambda_{\min}) |G_f(s) - G_g(s)| \leq (\lambda_{\max} - \lambda_{\min}) L_G \|f - g\|_\infty.$$

Taking the supremum over $s \in I$ on the left-hand side yields the desired estimate in the supremum norm:

$$\|\Lambda_f - \Lambda_g\|_\infty \leq L_\Lambda \|f - g\|_\infty, \quad (3.21)$$

where

$$L_\Lambda := (\lambda_{\max} - \lambda_{\min}) L_G = (\lambda_{\max} - \lambda_{\min}) \gamma_0 \left(1 + \frac{\beta_0 (1 - e^{-\alpha_0 T})}{\alpha_0} \right).$$

This completes the verification of (P2).

Proof of (P3): continuity (hence measurability) in time. Fix an arbitrary trajectory $f \in C(I)$. To establish the continuity of $s \mapsto \Lambda_f(s)$ on I , we introduce the auxiliary functions

$$\begin{aligned} \phi(s) &:= \tanh(\gamma_0 |f(s) - r(s)|), \\ \Psi(s) &:= \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |f(\tau) - r(\tau)|) d\tau, \end{aligned}$$

so that

$$\Lambda_f(s) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \frac{\phi(s)}{1 + \beta_0 \Psi(s)}.$$

First, ϕ is continuous on I . Indeed, f and r are continuous by hypothesis, the map $z \mapsto |z|$ is continuous, and the hyperbolic tangent is continuous; therefore, the composition

$$\phi(s) = \tanh(\gamma_0 |f(s) - r(s)|)$$

is continuous.

We now prove that Ψ is continuous on I . Consider the function

$$F(\tau, s) := e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |f(\tau) - r(\tau)|)$$

defined on the compact triangular region

$$\Delta := \{(\tau, s) \in I \times I \mid 0 \leq \tau \leq s \leq T\}.$$

The exponential factor $e^{-\alpha_0(s-\tau)}$ is continuous on Δ , and $\tanh(\gamma_0|f(\tau) - r(\tau)|)$ depends only on τ and is continuous. Hence, F is continuous on the compact set Δ , and therefore uniformly continuous on Δ . Moreover, from $0 \leq \tanh(\cdot) < 1$ and $0 < e^{-\alpha_0(s-\tau)} \leq 1$, we obtain the uniform bound

$$|F(\tau, s)| \leq 1 \quad \forall (\tau, s) \in \Delta. \quad (3.22)$$

Take an arbitrary $s_0 \in I$ and let $s \in I$. To estimate $|\Psi(s) - \Psi(s_0)|$, we consider two cases.

Case 1: $s \geq s_0$. Then

$$\begin{aligned} |\Psi(s) - \Psi(s_0)| &= \left| \int_0^s F(\tau, s) d\tau - \int_0^{s_0} F(\tau, s_0) d\tau \right| \\ &\leq \int_0^{s_0} |F(\tau, s) - F(\tau, s_0)| d\tau + \left| \int_{s_0}^s F(\tau, s) d\tau \right|. \end{aligned} \quad (3.23)$$

Case 2: $s \leq s_0$. Then

$$\begin{aligned} |\Psi(s) - \Psi(s_0)| &= \left| \int_0^s F(\tau, s) d\tau - \int_0^{s_0} F(\tau, s_0) d\tau \right| \\ &\leq \int_0^s |F(\tau, s) - F(\tau, s_0)| d\tau + \left| \int_s^{s_0} F(\tau, s_0) d\tau \right|. \end{aligned} \quad (3.24)$$

Given any $\varepsilon > 0$, by the uniform continuity of F on Δ there exists $\delta_1 > 0$, such that whenever $|s - s_0| < \delta_1$ and $(\tau, s), (\tau, s_0) \in \Delta$,

$$|F(\tau, s) - F(\tau, s_0)| < \frac{\varepsilon}{2T}. \quad (3.25)$$

For such s , the first term in either (3.23) or (3.24) (with the integration limit taken as $s_{\min} := \min\{s, s_0\}$) satisfies

$$\int_0^{s_{\min}} |F(\tau, s) - F(\tau, s_0)| d\tau \leq \frac{\varepsilon}{2T} \cdot T = \frac{\varepsilon}{2}.$$

For the second term, we use (3.22). If $|s - s_0| < \varepsilon/2$, then regardless of which case we are in,

$$\left| \int_{s_{\min}}^{s_{\max}} F(\tau, \max\{s, s_0\}) d\tau \right| \leq 1 \cdot |s_{\max} - s_{\min}| = |s - s_0| < \frac{\varepsilon}{2},$$

where $s_{\max} := \max\{s, s_0\}$ and the integrand $F(\tau, \max\{s, s_0\})$ is understood as $F(\tau, s)$ when $s \geq s_0$, and as $F(\tau, s_0)$ when $s \leq s_0$.

Choosing $\delta := \min\{\delta_1, \varepsilon/2\}$, we obtain $|\Psi(s) - \Psi(s_0)| < \varepsilon$ whenever $|s - s_0| < \delta$. Hence, Ψ is continuous at s_0 ; since s_0 was arbitrary, $\Psi \in C(I)$.

Because $\beta_0 \geq 0$ and $\Psi(s) \geq 0$, the denominator $D(s) := 1 + \beta_0\Psi(s)$ satisfies $D(s) \geq 1$ and is continuous. Consequently,

$$\Lambda_f(s) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \frac{\phi(s)}{D(s)}$$

is a combination of continuous functions by addition, multiplication and division by a non-vanishing continuous denominator, and therefore, Λ_f is continuous on I . Every continuous function is Lebesgue measurable, so (P3) is established.

Proof of (P4): positivity at zero. Consider the zero trajectory $f \equiv 0$. Then $f(s) = 0$ and $f(\tau) = 0$ for all $s, \tau \in I$, and

$$\Lambda_0(s) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \frac{\tanh(\gamma_0 |r(s)|)}{1 + \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |r(\tau)|) d\tau}.$$

We analyze the right-hand side step by step.

- (1) For any $z \geq 0$, $\tanh(z) \in [0, 1]$; in particular, $\tanh(z) \geq 0$.
- (2) Hence, $\tanh(\gamma_0 |r(s)|) \geq 0$ and $\tanh(\gamma_0 |r(\tau)|) \geq 0$ for all s, τ .
- (3) The exponential $e^{-\alpha_0(s-\tau)}$ is always positive, and $\beta_0 \geq 0$; therefore, the integrand and the whole integral are non-negative. Consequently,

$$1 + \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |r(\tau)|) d\tau \geq 1. \quad (3.26)$$

- (4) From the above, we obtain

$$0 \leq \frac{\tanh(\gamma_0 |r(s)|)}{1 + \beta_0 \int_0^s e^{-\alpha_0(s-\tau)} \tanh(\gamma_0 |r(\tau)|) d\tau} < 1.$$

Since $0 < \lambda_{\min} < \lambda_{\max}$, we have $\lambda_{\max} - \lambda_{\min} > 0$, and therefore the fractional term multiplied by $(\lambda_{\max} - \lambda_{\min})$ is non-negative. Thus,

$$\Lambda_0(s) \geq \lambda_{\min} \quad \forall s \in I. \quad (3.27)$$

The constant $\lambda_{\min} > 0$ serves as the required positive lower bound, which is independent of s . This completes the verification of (P4). $\square$

Corollary 3.1 (The purely instantaneous case $\beta_0 = 0$). *Under the assumptions of Theorem 3.1, suppose further that $\beta_0 = 0$. Then the construction in Example 3.1 reduces to a function $\Lambda \in \mathcal{A}(I)$ given explicitly by*

$$\Lambda(s, x) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \tanh(\gamma_0 |x - r(s)|), \quad (s, x) \in I \times \mathbb{R}. \quad (3.28)$$

Consequently, for every trajectory $f \in C(I)$, the associated function Λ_f defined in (3.1) satisfies

$$\Lambda_f(s) = \Lambda(s, f(s)) \quad \text{for all } s \in I.$$

We now verify that Λ indeed belongs to the class $\mathcal{A}(I)$ by checking conditions (AS1)–(AS4) of Definition 3.1.

Verification of (AS1): uniform boundedness. For any $(s, x) \in I \times \mathbb{R}$, the elementary bound $0 \leq \tanh(z) < 1$ for $z \geq 0$ yields

$$0 \leq \tanh(\gamma_0 |x - r(s)|) < 1.$$

Multiplying by $(\lambda_{\max} - \lambda_{\min}) > 0$ and adding $\lambda_{\min}$ gives

$$\lambda_{\min} \leq \Lambda(s, x) \leq \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) = \lambda_{\max}, \quad (3.29)$$

which establishes (AS1) with the same constants $\lambda_{\min}, \lambda_{\max}$.

Verification of (AS2): Lipschitz continuity in the state variable. Fix $s \in I$ and set $p = r(s)$. Applying Lemma 3.1 with this p , we have for any $x, y \in \mathbb{R}$,

$$|\tanh(\gamma_0|x - p|) - \tanh(\gamma_0|y - p|)| \leq \gamma_0|x - y|.$$

Consequently,

$$\begin{aligned} |\Lambda(s, x) - \Lambda(s, y)| &= (\lambda_{\max} - \lambda_{\min}) |\tanh(\gamma_0|x - r(s)|) - \tanh(\gamma_0|y - r(s)|)| \\ &\leq (\lambda_{\max} - \lambda_{\min}) \gamma_0 |x - y|. \end{aligned} \quad (3.30)$$

Thus, $\Lambda(s, \cdot)$ is Lipschitz continuous with constant $L_\Lambda = (\lambda_{\max} - \lambda_{\min})\gamma_0$, satisfying (AS2).

Verification of (AS3): measurability in time. For each fixed $x \in \mathbb{R}$, consider the map $s \mapsto \Lambda(s, x)$. Since r is continuous by hypothesis, the composition $s \mapsto \gamma_0|x - r(s)|$ is continuous; the hyperbolic tangent function is continuous, hence $s \mapsto \tanh(\gamma_0|x - r(s)|)$ is continuous. Adding the constant $\lambda_{\min}$ preserves continuity, so $s \mapsto \Lambda(s, x)$ is continuous on I . Every continuous function is Lebesgue measurable, therefore (AS3) holds.

Verification of (AS4): positivity at zero. Setting $x = 0$ in (3.28) gives

$$\Lambda(s, 0) = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \tanh(\gamma_0|r(s)|).$$

Since $\tanh(\gamma_0|r(s)|) \geq 0$ for all $s \in I$, we obtain

$$\Lambda(s, 0) \geq \lambda_{\min} > 0 \quad \forall s \in I. \quad (3.31)$$

Taking $\varrho := \lambda_{\min}$ (which is precisely the constant appearing in (AS1)), condition (AS4) is satisfied.

Having verified all four conditions (AS1)–(AS4), we conclude that Λ belongs to the class $\mathcal{A}(I)$, thereby confirming that the purely instantaneous case $\beta_0 = 0$ is fully consistent with the framework of Definition 3.1. $\square$

Remark 3.3 (Summary and perspective). Theorem 3.1 establishes that for every $\beta_0 \geq 0$, the function Λ_f generated by the historical deviation construction enjoys four fundamental properties. Properties (P1), (P3) and (P4) are direct analogues of the conditions (AS1), (AS3) and (AS4) that define the class $\mathcal{A}(I)$; they guarantee that each Λ_f is uniformly bounded, continuous in time, and strictly positive at the zero trajectory.

Property (P2), while playing a role analogous to (AS2) in the overall theory, takes a form adapted to the operator perspective: Instead of asserting pointwise Lipschitz dependence on the instantaneous state value $f(s)$, it provides a global estimate in the supremum norm, namely $\|\Lambda_f - \Lambda_g\|_\infty \leq L_\Lambda \|f - g\|_\infty$. This formulation expresses the fact that the map $f \mapsto \Lambda_f$ is Lipschitz continuous from $C(I)$ into itself—a natural and powerful statement when one views the construction as a nonlinear operator acting on entire trajectories. The constant L_Λ captures the combined influence of parameters γ_0, β_0 , and α_0 , and reduces to $(\lambda_{\max} - \lambda_{\min})\gamma_0$ in the purely instantaneous case $\beta_0 = 0$, consistently with Corollary 3.1.

Corollary 3.1 examines the special case $\beta_0 = 0$, where the historical feedback vanishes. In this situation, the operator perspective reverts to the classical function viewpoint: There exists a fixed $\Lambda \in \mathcal{A}(I)$ such that $\Lambda_f(s) = \Lambda(s, f(s))$ for every trajectory f . This observation not only confirms that the construction reproduces the expected instantaneous sensitivity when no historical modulation is present, but also illustrates the flexibility of the framework; the same formula (3.1) interpolates continuously between a standard $\mathcal{A}(I)$ -function ($\beta_0 = 0$) and a genuinely history-dependent operator ($\beta_0 > 0$).

Together, Theorem 3.1 and Corollary 3.1 demonstrate that the historical deviation example is both mathematically rich and perfectly compatible with the framework established in Definition 3.1. The operator viewpoint, with its Lipschitz estimate in the supremum norm, offers a natural language for describing memory mechanisms that depend on whole trajectories, while the special case $\beta_0 = 0$ seamlessly recovers the simpler instantaneous picture. This duality underscores the versatility of the proposed framework and its capacity to accommodate a wide spectrum of adaptive memory phenomena.

This framework thus provides a unified mathematical foundation for capturing nonlinear adaptive memory phenomena, including habituation, state-dependent weighting, and selective retention, that are central to understanding complex systems in fields such as neuroscience, adaptive control, and machine learning.

4. Adaptive memory sets: functional construction and fundamental properties

We finally arrive at the third layer of the framework: The adaptive memory-dependent functional and the associated function set, which integrate the kernel hierarchy and the adaptive sensitivity mechanism developed in the previous two sections. This layer provides a unified quantitative measure that combines instantaneous magnitude with adaptively weighted historical accumulation, thereby completing the three-layer structure.

The kernel hierarchy developed in Section 2 serves a twofold purpose. The regular class $\mathcal{K}_{\text{reg}}$ provides the analytical framework for the main results in Sections 3 and 4. The generalized class $\mathcal{K}_{\text{gen}}$ is introduced as a broader theoretical interface, enabling sign-changing or discontinuous kernels that may arise in applications. The present analysis is developed primarily within the regular class, while the generalized class offers a potential starting point for further extensions under more general conditions.

The preceding sections have established two foundational pillars: A hierarchical classification of memory kernels κ (Section 2) that quantifies the temporal weighting of past states, and a framework for adaptive sensitivity functions Λ (Section 3) that modulates this weight based on the state's value. In this section, we synthesize these components to construct a novel collection of functions tailored for systems exhibiting adaptive memory, a form of nonlinear behavior where the influence of past events depends not only on when they occurred but also on what their values were. This functional will serve as the fundamental analytical tool within this framework, thereby transforming the abstract theoretical setting into concrete functional-analytic machinery.

4.1. Construction and well-posedness of fundamental memory functions

To quantify the cumulative influence of a function's past under the adaptive memory paradigm, we first introduce two auxiliary functions. Their definitions rely on the synergy between the kernel κ and the sensitivity function Λ , and their mathematical legitimacy must be rigorously established through a systematic verification before they can be employed in the definition of the adaptive memory-dependent functional.

Definition 4.1 (Adaptive weighted cumulative function and instantaneous-memory hybrid function). Let $\kappa \in \mathcal{K}_{\text{reg}}$ be a regular admissible kernel (Definition 2.2) and let $\Lambda \in \mathcal{A}(I)$ be an adaptive sensitivity function (Definition 3.1). For any continuous trajectory $f \in C(I)$, where $C(I)$ denotes the space of continuous functions on I and is used throughout in this standard sense, we define the following:

(i) **Adaptive weighted cumulative function** $\mathcal{J}_f : I \rightarrow [0, \infty)$ by

$$\mathcal{J}_f(t) := \int_0^t \Lambda(s, f(s)) \kappa(t-s) |f(s)| \, ds, \quad t \in I. \quad (4.1)$$

This function quantifies the total weighted influence exerted on the present instant t by all past states $f(s)$ for $0 \leq s \leq t$, accumulated through the adaptive memory weight $\Lambda(s, f(s))\kappa(t-s)$.

(ii) **Instantaneous-memory hybrid function** $\mathcal{M}_f : I \rightarrow [0, \infty)$ by

$$\mathcal{M}_f(t) := |f(t)| + \mathcal{J}_f(t), \quad t \in I. \quad (4.2)$$

This function simultaneously captures the system's instantaneous response $|f(t)|$ at time t and the weighted historical accumulation $\mathcal{J}_f(t)$, thereby constituting the fundamental building block for the adaptive memory-dependent functional we seek to establish.

Remark 4.1 (Conceptual rationale behind the terminology). The nomenclature in Definition 4.1 is carefully chosen to reflect both the mathematical structure and the underlying physical intuition. The term adaptive weighted cumulative emphasizes that $\mathcal{J}_f$ is an integral (cumulative) measure, where the integrand is weighted by a factor that is simultaneously time-dependent (through κ) and state-dependent (through Λ). The designation instantaneous-memory hybrid for $\mathcal{M}_f$ accurately conveys its nature as a composite of a purely local quantity $|f(t)|$ and a global, history-dependent quantity $\mathcal{J}_f(t)$. This conceptual separation, distinguishing the accumulation mechanism and the combined instantaneous-historical measure, will prove invaluable in the subsequent analysis of the adaptive memory-dependent functional (see Definition 4.2) and related functional-analytic structures.

To illustrate the relevance of these constructions while respecting the standing regularity assumptions, one may interpret a trajectory $f(t)$ as, for instance, the firing rate of a neuron over time (or, more generally, a suitable continuous approximation thereof). The instantaneous term $|f(t)|$ captures the current level of neural activity, while the cumulative term $\mathcal{J}_f(t)$ aggregates past firing activity weighted by the elapsed time (via κ) and the firing rates (via Λ). The hybrid function $\mathcal{M}_f(t)$ therefore quantifies the combined influence of present and past activity, providing a mathematical description of neuronal adaptation, a nonlinear phenomenon where a neuron's response diminishes under sustained stimulation (habituation) or amplifies when exposed to novel or salient intermittent events (sensitization). This interpretation extends naturally to other domains: In ecological population dynamics, $f(t)$ may represent species abundance, with $\mathcal{M}_f(t)$ capturing the interplay between current

population size and the historical constraints imposed by past population levels; in engineering signal processing, $f(t)$ may represent sensor readings, with $\mathcal{M}_f(t)$ encoding the balance between instantaneous measurements and historical trends. Such examples underscore how the abstract function $\mathcal{M}_f$ serves as a unified mathematical object for modeling state-dependent, history-sensitive adaptive behavior across scientific disciplines.

Before these functions can be utilized as building blocks for the adaptive memory-dependent functional, it is imperative to verify that they are mathematically well-defined for the class of functions for which they are intended. The following lemma establishes their essential properties through a systematic argument.

Lemma 4.1 (Well-posedness of the fundamental memory functions). *Let $f \in C(I)$, $\kappa \in \mathcal{K}_{\text{reg}}$, and $\Lambda \in \mathcal{A}(I)$. Then the functions introduced in Definition 4.1 satisfy the following:*

- (i) **Absolute convergence and finiteness:** *For every $t \in I$, the integral defining $\mathcal{J}_f(t)$ converges absolutely as a Lebesgue integral. Consequently, $\mathcal{J}_f(t)$ is a finite, non-negative real number for all $t \in I$, and $\mathcal{M}_f(t)$ is likewise finite and non-negative.*
- (ii) **Measurability:** *The mappings $t \mapsto \mathcal{J}_f(t)$ and $t \mapsto \mathcal{M}_f(t)$ are Lebesgue measurable on I .*
- (iii) **Uniform boundedness:** *There exists a finite constant $C > 0$, independent of the specific trajectory f and the temporal argument t , such that*

$$\mathcal{M}_f(t) \leq C \|f\|_{\infty}, \quad \forall t \in I. \quad (4.3)$$

More explicitly, one may take $C := 1 + \Lambda_{\infty} \kappa_{\infty} T$, where

$$\Lambda_{\infty} := \sup_{(s,x) \in I \times \mathbb{R}} (\Lambda(s, x)) \leq \lambda_{\max} < \infty, \quad (4.4)$$

$$\kappa_{\infty} := \sup_{\tau \in I} \kappa(\tau) \leq M_{\kappa} < \infty, \quad (4.5)$$

with $\lambda_{\max}$ and M_{κ} being the uniform bounds from conditions (AS1) and (R1), respectively. As established in the proof above, one also obtains the auxiliary estimate $\mathcal{J}_f(t) \leq \Lambda_{\infty} \kappa_{\infty} T \|f\|_{\infty}$ for all $t \in I$.

Proof. We establish each property through a systematic and self-contained argument.

Proof of (i): absolute convergence and finiteness. Since $I = [0, T]$ is compact and f is continuous, f is bounded on I . Denote

$$M := \|f\|_{\infty} = \sup_{s \in I} |f(s)| < \infty. \quad (4.6)$$

From the uniform boundedness condition (AS1) of Definition 3.1, there exists a constant $\lambda_{\max} > 0$ such that

$$|\Lambda(s, f(s))| \leq \lambda_{\max}, \quad \forall s \in I. \quad (4.7)$$

Consider the dominating function for the integrand in (4.1). For any $t \in I$ and $s \in [0, t]$, we have the pointwise estimate

$$|\Lambda(s, f(s)) \kappa(t-s) f(s)| \leq \lambda_{\max} M \cdot \kappa(t-s). \quad (4.8)$$

The kernel κ belongs to $\mathcal{K}_{\text{reg}}$, hence by condition (R2) we have $\kappa \in L^1(I)$ and specifically

$$\int_0^T \kappa(\tau) \, d\tau = 1.$$

Consequently, for each fixed t , the function $s \mapsto \kappa(t - s)$ is Lebesgue integrable on $[0, t]$. By the comparison test for Lebesgue integrals, the integral

$$\int_0^t |\Lambda(s, f(s))\kappa(t - s)|f(s)| \, ds$$

converges, implying that $\mathcal{J}_f(t)$ is absolutely convergent and therefore finite. Non-negativity follows directly from $\kappa(\tau) \geq 0$ (condition (R1)) and $\Lambda(s, x) \geq 0$ (condition (AS1)). The finiteness and non-negativity of $\mathcal{M}_f(t)$ are then immediate from its definition (4.2).

Proof of (ii): measurability. We establish the measurability of $\mathcal{J}_f$ through a product-measurability argument, which provides a rigorous foundation for applying Fubini-type theorems. Define the auxiliary function $\mathcal{G}: I \times I \rightarrow [0, \infty)$ by

$$\mathcal{G}(t, s) := \Lambda(s, f(s))\kappa(t - s)|f(s)|\mathbf{1}_{[0, t]}(s), \quad (4.9)$$

where $\mathbf{1}_{[0, t]}$ denotes the indicator function of the interval $[0, t]$.

Let $\mathcal{L}(I)$ denote the σ -algebra of Lebesgue measurable sets on I , and let $\mathcal{L}(I) \otimes \mathcal{L}(I)$ be the corresponding product σ -algebra. We proceed by verifying the product measurability of each factor constituting $\mathcal{G}$.

(a) **Measurability of $(t, s) \mapsto \Lambda(s, f(s))$:** Consider the mapping $\Phi: I \rightarrow I \times \mathbb{R}$ defined by $\Phi(s) = (s, f(s))$. Since f is continuous, Φ is continuous and therefore Borel measurable. The function

$$\Lambda: I \times \mathbb{R} \rightarrow [0, \infty)$$

is a Carathéodory function: By condition (AS2) in Definition 3.1, $x \mapsto \Lambda(s, x)$ is Lipschitz continuous (hence continuous) for each fixed s ; by condition (AS3) in Definition 3.1, $s \mapsto \Lambda(s, x)$ is Lebesgue measurable for each fixed x . A Carathéodory function is jointly measurable with respect to the product σ -algebra $\mathcal{L}(I) \otimes \mathcal{B}(\mathbb{R})$, where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ -algebra on $\mathbb{R}$. Consequently, the composition

$$H(s) := \Lambda(s, f(s)) = \Lambda \circ \Phi(s)$$

is Lebesgue measurable as the composition of a measurable function Λ with a Borel measurable function Φ . The extension $\tilde{H}(t, s) := H(s)$ is then product measurable, since for any $a \in \mathbb{R}$,

$$\{(t, s) \mid \tilde{H}(t, s) \leq a\} = I \times \{s \mid H(s) \leq a\} \in \mathcal{L}(I) \otimes \mathcal{L}(I).$$

(b) **Measurability of $(t, s) \mapsto \kappa(t - s)$:** Condition (R3) in Definition 2.2 guarantees that κ is Lipschitz continuous on I , hence continuous. The map $(t, s) \mapsto t - s$ is continuous, and the composition of continuous functions preserves continuity; therefore $(t, s) \mapsto \kappa(t - s)$ is continuous on $I \times I$, and consequently Borel measurable (hence $\mathcal{L}(I) \otimes \mathcal{L}(I)$ -measurable).

(c) **Measurability of $(t, s) \mapsto |f(s)|$:** The function f is continuous, so $|f|$ is continuous. As in part (a), the extension $\tilde{F}(t, s) := |f(s)|$ is product measurable.

(d) **Measurability of $(t, s) \mapsto \mathbf{1}_{[0, t]}(s)$:** Define the set

$$E := \{(t, s) \in I \times I : 0 \leq s \leq t\}.$$

The function $\phi(t, s) = t - s$ is continuous, hence $E = \phi^{-1}([0, \infty))$ is a closed set and therefore Borel measurable. Its characteristic function $\chi_E(t, s) = \mathbf{1}_{[0, t]}(s)$ is consequently Borel measurable and belongs to $\mathcal{L}(I) \otimes \mathcal{L}(I)$.

All four factors are non-negative and $\mathcal{L}(I) \otimes \mathcal{L}(I)$ -measurable; hence, their product $\mathcal{G}$ is also $\mathcal{L}(I) \otimes \mathcal{L}(I)$ -measurable and non-negative.

With $\mathcal{G}$ being non-negative and product-measurable, we can legitimately invoke the Fubini-Tonelli theorem for non-negative measurable functions. This classical result guarantees that:

- For almost every (in fact, every) $t \in I$, the section $s \mapsto \mathcal{G}(t, s)$ is $\mathcal{L}(I)$ -measurable;
- The function defined by integrating over the second variable,

$$t \mapsto \int_I \mathcal{G}(t, s) \, ds = \int_0^t \Lambda(s, f(s)) \kappa(t-s) |f(s)| \, ds = \mathcal{J}_f(t),$$

is $\mathcal{L}(I)$ -measurable on I (taking values in $[0, \infty)$).

Finally, the continuity of f implies that $t \mapsto |f(t)|$ is continuous and therefore $\mathcal{L}(I)$ -measurable. The sum

$$\mathcal{M}_f(t) = |f(t)| + \mathcal{J}_f(t)$$

of two $\mathcal{L}(I)$ -measurable functions remains $\mathcal{L}(I)$ -measurable, completing the proof of part (ii).

Proof of (iii): uniform boundedness. From the uniform boundedness condition (R1) for regular kernels, we have $\kappa(\tau) \leq M_\kappa$ for all $\tau \in I$, where M_κ is the constant appearing in Definition 2.2. Define the supremum norm of κ on I as

$$\kappa_\infty := \sup_{\tau \in I} \kappa(\tau) \leq M_\kappa < \infty. \quad (4.10)$$

Similarly, condition (AS1) for adaptive sensitivity functions (Definition 3.1) yields $\Lambda(s, x) \leq \lambda_{\max}$ for all $(s, x) \in I \times \mathbb{R}$. Set

$$\Lambda_\infty := \sup_{(s, x) \in I \times \mathbb{R}} \Lambda(s, x) \leq \lambda_{\max} < \infty. \quad (4.11)$$

For any $t \in I$, using the non-negativity of κ and the estimate $\kappa(t-s) \leq \kappa_\infty$, we obtain

$$\int_0^t \kappa(t-s) \, ds \leq \int_0^t \kappa_\infty \, ds = \kappa_\infty t \leq \kappa_\infty T.$$

Combining these estimates, we bound the cumulative function $\mathcal{J}_f$ as follows:

$$\mathcal{J}_f(t) \leq \int_0^t \Lambda_\infty \kappa(t-s) |f(s)| \, ds \leq \Lambda_\infty \|f\|_\infty \int_0^t \kappa(t-s) \, ds \leq \Lambda_\infty \|f\|_\infty \kappa_\infty T.$$

Consequently, for the hybrid function $\mathcal{M}_f$, we have

$$\begin{aligned} \mathcal{M}_f(t) &= |f(t)| + \mathcal{J}_f(t) \\ &\leq \|f\|_\infty + \Lambda_\infty \kappa_\infty T \|f\|_\infty \\ &= (1 + \Lambda_\infty \kappa_\infty T) \|f\|_\infty. \end{aligned} \quad (4.12)$$

Taking $C := 1 + \Lambda_\infty \kappa_\infty T$, which is finite by virtue of (4.10) and (4.11), establishes the desired uniform estimate (4.3). The bound for $\mathcal{J}_f$ follows directly from the chain of inequalities above. $\square$

Remark 4.2 (Significance of the well-posedness results). Lemma 4.1 provides a foundation for the subsequent theory. It shows that the auxiliary objects $\mathcal{J}_f$ and $\mathcal{M}_f$, which aim to capture the interplay between instantaneous magnitude, temporal weighting, and state-dependent sensitivity, are

not merely formal constructions but objects with clear mathematical meaning. Absolute convergence ensures that no hidden cancellations or singularities affect the definition; measurability renders these functions amenable to integration and further functional-analytic operations, potentially enabling their incorporation into Lebesgue spaces; uniform boundedness offers an initial indication that $\mathcal{M}_f$ relates to the trajectory f in a manner proportional to its supremum norm, with the proportionality constant C depending only on the intrinsic parameters of the kernel and the sensitivity function. Collectively, these properties establish $\mathcal{M}_f$ as a mathematically viable object for quantifying adaptive memory. In particular, the state-dependent nature of Λ within $\mathcal{J}_f$ introduces a nonlinear coupling between the trajectory f and its own history, a feature that enables the resulting framework to capture adaptive phenomena such as habituation and selective retention.

4.2. Definition and basic properties of the adaptive memory-dependent functional

With the well-posedness of the fundamental functions $\mathcal{J}_f$ and $\mathcal{M}_f$ rigorously established in Lemma 4.1, we are now in a position to introduce the central object of our construction, the adaptive memory-dependent functional. Defined through a supremum operation, this functional provides a quantitative measure that integrates instantaneous magnitude and adaptively weighted historical accumulation. Its basic properties, which are essential for the subsequent functional-analytic developments, are systematically examined below.

Definition 4.2 (Adaptive memory-dependent functional). Let $\kappa \in \mathcal{K}_{\text{reg}}$ be a regular admissible kernel (Definition 2.2) and let $\Lambda \in \mathcal{A}(I)$ be an adaptive sensitivity function (Definition 3.1). For any continuous trajectory $f \in C(I)$, the adaptive memory-dependent functional is defined by

$$S_{\kappa, \Lambda}(f) := \sup_{t \in I} \mathcal{M}_f(t) \in [0, \infty), \quad (4.13)$$

where $\mathcal{M}_f$ is the instantaneous-memory hybrid function introduced in Definition 4.1. The uniform boundedness established in Lemma 4.1 (iii) guarantees that this supremum is a finite nonnegative real number.

Lemma 4.2 (Basic properties of the adaptive memory-dependent functional). Let $f \in C(I)$. The functional $S_{\kappa, \Lambda}(f)$ defined in (4.13) satisfies the following fundamental properties:

- (i) **Existence and well-posedness:** The quantity $S_{\kappa, \Lambda}(f)$ exists as a real number and is uniquely determined.
- (ii) **Controlled boundedness via intrinsic parameters:** Recalling the uniform bounds

$$\Lambda_{\infty} := \sup_{(s, x) \in I \times \mathbb{R}} \Lambda(s, x) \quad \text{and} \quad \kappa_{\infty} := \sup_{\tau \in I} \kappa(\tau)$$

introduced in Lemma 4.1 (see (4.4) and (4.5)), we have the estimate

$$S_{\kappa, \Lambda}(f) \leq (1 + \Lambda_{\infty} \kappa_{\infty} T) \|f\|_{\infty}. \quad (4.14)$$

Conditions (AS1) and (R1) ensure $\Lambda_{\infty} \leq \lambda_{\max} < \infty$ and $\kappa_{\infty} \leq M_{\kappa} < \infty$, so the controlling constant is finite.

- (iii) **Positive definiteness:** $S_{\kappa, \Lambda}(f) \geq 0$ for every $f \in C(I)$. Moreover,

$$S_{\kappa, \Lambda}(f) = 0 \quad \text{if and only if} \quad f \equiv 0 \text{ on } I.$$

(iv) **Comparison with the classical supremum norm:** The adaptive memory-dependent functional dominates the classical supremum norm, i.e.,

$$\|f\|_\infty \leq S_{\kappa,\Lambda}(f) \quad \forall f \in C(I). \quad (4.15)$$

Proof. We establish each property through a systematic argument that leverages the results of Lemma 4.1 while remaining self-contained.

Proof of (i): existence and well-posedness. From Lemma 4.1 (i), the function $\mathcal{M}_f$ is well-defined pointwise and takes values in $[0, \infty)$. Specifically, for each $t \in I$, the quantity $\mathcal{M}_f(t)$ is a finite nonnegative real number. Hence the set

$$S_f := \{\mathcal{M}_f(t) : t \in I\} \subset [0, \infty)$$

is nonempty. Lemma 4.1 (iii) provides the uniform bound

$$\mathcal{M}_f(t) \leq (1 + \Lambda_\infty \kappa_\infty T) \|f\|_\infty \quad \text{for all } t \in I, \quad (4.16)$$

which implies that S_f is bounded above. By the completeness axiom of the real numbers, every nonempty subset of $\mathbb{R}$ that is bounded above possesses a unique supremum. Consequently, the quantity $\sup S_f$ exists as a real number and is uniquely determined. Defining $S_{\kappa,\Lambda}(f) := \sup S_f$ therefore yields a well-defined real number. Its non-negativity follows immediately from the fact that every element of S_f is nonnegative, so the supremum of a set of nonnegative numbers is itself nonnegative.

Proof of (ii): controlled boundedness. Recall that $S_{\kappa,\Lambda}(f)$ is by definition the least upper bound of the set $S_f = \{\mathcal{M}_f(t) : t \in I\}$. A fundamental property of the supremum is that it cannot exceed any upper bound of the set. From Lemma 4.1 (iii), we have the explicit upper bound

$$\mathcal{M}_f(t) \leq (1 + \Lambda_\infty \kappa_\infty T) \|f\|_\infty \quad \text{for every } t \in I.$$

Thus the quantity $(1 + \Lambda_\infty \kappa_\infty T) \|f\|_\infty$ serves as an upper bound for the set S_f . Applying the aforementioned property of the supremum yields

$$S_{\kappa,\Lambda}(f) = \sup_{t \in I} \mathcal{M}_f(t) \leq (1 + \Lambda_\infty \kappa_\infty T) \|f\|_\infty, \quad (4.17)$$

which is precisely the estimate (4.14).

Proof of (iii): positive definiteness. We establish the two directions of the equivalence separately.

Non-negativity: For any $f \in C(I)$ and any $t \in I$, Lemma 4.1 (i) guarantees that $\mathcal{M}_f(t) \geq 0$. The set $S_f = \{\mathcal{M}_f(t) : t \in I\}$ therefore consists entirely of nonnegative numbers. The supremum of a set of nonnegative numbers is itself nonnegative, yielding $S_{\kappa,\Lambda}(f) = \sup S_f \geq 0$.

Forward implication ($f \equiv 0 \Rightarrow S_{\kappa,\Lambda}(f) = 0$): Assume that f is identically zero on I , i.e., $f(t) = 0$ for all $t \in I$. Then for any $t \in I$, we have $|f(t)| = 0$. Moreover, from the Definition (4.1) of $\mathcal{J}_f$, the integrand

$$\Lambda(s, 0) \kappa(t - s) |0|$$

vanishes for every $s \in [0, t]$, so $\mathcal{J}_f(t) = 0$. Consequently, $\mathcal{M}_f(t) = |f(t)| + \mathcal{J}_f(t) = 0$ for all $t \in I$. Thus, $S_f = \{0\}$, and its supremum is $\sup\{0\} = 0$. Hence, $S_{\kappa,\Lambda}(f) = 0$.

Reverse implication ($S_{\kappa,\Lambda}(f) = 0 \Rightarrow f \equiv 0$): Suppose that $S_{\kappa,\Lambda}(f) = 0$. Since $S_{\kappa,\Lambda}(f)$ is defined as the supremum of the set $S_f = \{\mathcal{M}_f(t) : t \in I\}$, and all elements of S_f are nonnegative, the condition

$\sup S_f = 0$ forces every element of S_f to be zero. Indeed, if there existed some $t_0 \in I$ with $\mathcal{M}_f(t_0) > 0$, then the supremum would be at least $\mathcal{M}_f(t_0) > 0$, contradicting $S_{\kappa,\Lambda}(f) = 0$. Therefore, $\mathcal{M}_f(t) = 0$ for every $t \in I$.

Now fix an arbitrary $t \in I$. From the definition $\mathcal{M}_f(t) = |f(t)| + \mathcal{J}_f(t)$ and the non-negativity of both terms, we have

$$0 = \mathcal{M}_f(t) = |f(t)| + \mathcal{J}_f(t) \geq |f(t)| \geq 0.$$

The chain of inequalities forces $|f(t)| = 0$, which implies $f(t) = 0$. Since $t \in I$ was chosen arbitrarily, we conclude that $f(t) = 0$ for all $t \in I$, i.e., $f \equiv 0$ on I . This completes the proof of the reverse implication, and together with the forward implication establishes the equivalence

$$S_{\kappa,\Lambda}(f) = 0 \quad \text{if and only if} \quad f \equiv 0 \text{ on } I. \quad (4.18)$$

Proof of (iv): comparison with the classical supremum norm. We first establish the pointwise inequality that underlies the comparison. For any $t \in I$, the definition of $\mathcal{M}_f$ from Definition 4.1 gives

$$\mathcal{M}_f(t) = |f(t)| + \mathcal{J}_f(t).$$

Since $\mathcal{J}_f(t)$ is nonnegative (as established in Lemma 4.1 (i)), we obtain the simple but crucial estimate

$$\mathcal{M}_f(t) \geq |f(t)| \quad \text{for every } t \in I. \quad (4.19)$$

Taking the supremum over $t \in I$ on both sides of (4.19) preserves the inequality, yielding

$$\sup_{t \in I} \mathcal{M}_f(t) \geq \sup_{t \in I} |f(t)|.$$

By Definition 4.2, the left-hand side is $S_{\kappa,\Lambda}(f)$ and the right-hand side is the classical supremum norm $\|f\|_\infty$. Hence,

$$S_{\kappa,\Lambda}(f) \geq \|f\|_\infty \quad \text{for all } f \in C(I),$$

which is precisely the inequality (4.15). $\square$

The inequality established in Lemma 4.2 (iv) is sharp in the sense that for non-zero functions it can be strengthened to a strict inequality, provided the supremum of $|f|$ is attained at a point where the subsequent estimate applies. The following theorem makes this precise.

Theorem 4.1 (Strict comparison when the maximizer lies in $(0, T]$). *Let $f \in C(I)$ be a non-zero function such that the maximum of $|f|$ is attained at some point $t^* \in (0, T]$, i.e., there exists $t^* \in (0, T]$ with*

$$|f(t^*)| = \|f\|_\infty = \max_{t \in I} |f(t)|. \quad (4.20)$$

Then the inequality established in Lemma 4.2 (iv) is strict, namely

$$S_{\kappa,\Lambda}(f) > \|f\|_\infty. \quad (4.21)$$

Proof. Since f is non-zero, we have $\|f\|_\infty > 0$, and consequently $|f(t^*)| > 0$, which implies $f(t^*) \neq 0$.

By the continuity of f at t^* , for

$$\varepsilon = \frac{|f(t^*)|}{2} > 0,$$

there exists $\delta_0 > 0$ such that for all $s \in I$ with $|s - t^*| < \delta_0$, we have

$$|f(s) - f(t^*)| < \frac{|f(t^*)|}{2}.$$

Since $t^* > 0$ by hypothesis, we may choose δ sufficiently small such that $0 < \delta < \min\{\delta_0, t^*\}$. Then, for any $s \in (t^* - \delta, t^*)$ (note that $(t^* - \delta, t^*) \subset (t^* - \delta_0, t^* + \delta_0)$), the reverse triangle inequality yields

$$|f(s)| \geq |f(t^*)| - |f(s) - f(t^*)| > |f(t^*)| - \frac{|f(t^*)|}{2} = \frac{|f(t^*)|}{2} > 0. \quad (4.22)$$

Moreover, the interval $[t^* - \delta/2, t^*]$ is contained in $(t^* - \delta, t^*)$ and has positive length $\delta/2 > 0$. Consequently, the estimate $|f(s)| \geq |f(t^*)|/2$ holds for all $s \in [t^* - \delta/2, t^*]$.

We now estimate the integrand defining $\mathcal{J}_f(t^*)$ on the interval $[t^* - \delta/2, t^*]$. From condition (AS1) of Definition 3.1, we have the uniform lower bound $\Lambda(s, f(s)) \geq \lambda_{\min} > 0$ for all $s \in I$. For the kernel κ , condition (R1) of Definition 2.2 provides the global lower bound $\kappa(\tau) \geq m_\kappa > 0$ for all $\tau \in I$. Since for any $s \in [t^* - \delta/2, t^*]$ we have $t^* - s \in [0, t^*] \subset I$, it follows that $\kappa(t^* - s) \geq m_\kappa > 0$.

Combining these estimates, we obtain for all s in the interval $[t^* - \delta/2, t^*]$ the pointwise lower bound

$$\Lambda(s, f(s))\kappa(t^* - s)|f(s)| \geq \lambda_{\min} m_\kappa \frac{|f(t^*)|}{2} > 0.$$

Now we estimate $\mathcal{J}_f(t^*)$ from below. Restricting the integration domain to $[t^* - \delta/2, t^*]$ and using the pointwise lower bound, we obtain

$$\mathcal{J}_f(t^*) \geq \int_{t^* - \delta/2}^{t^*} \lambda_{\min} m_\kappa \frac{|f(t^*)|}{2} ds = \lambda_{\min} m_\kappa \frac{|f(t^*)|}{2} \cdot \frac{\delta}{2} > 0.$$

With this strict positivity established, we can now compare $\mathcal{M}_f(t^*)$ with $\|f\|_\infty$. Using (4.20) and the fact that $\mathcal{J}_f(t^*) > 0$, we obtain

$$\mathcal{M}_f(t^*) = |f(t^*)| + \mathcal{J}_f(t^*) > |f(t^*)| = \|f\|_\infty.$$

Finally, recalling from Definition 4.2 that $S_{\kappa, \Lambda}(f)$ is the supremum of $\mathcal{M}_f(t)$ over $t \in I$, we have

$$S_{\kappa, \Lambda}(f) = \sup_{t \in I} \mathcal{M}_f(t) \geq \mathcal{M}_f(t^*) > \|f\|_\infty,$$

which establishes the desired strict inequality (4.21). $\square$

Remark 4.3 (Theoretical significance of the adaptive memory-dependent functional). The functional $S_{\kappa, \Lambda}(f)$ introduced in Definition 4.2 provides a quantitative framework for analyzing functions in the context of adaptive memory. Its construction incorporates two complementary mechanisms: Kernel κ encodes the objective variation of memory influence with elapsed time, while the adaptive sensitivity function Λ modulates this influence according to the state values $f(s)$ encountered along the trajectory. The composite weight $\Lambda(s, f(s))\kappa(t - s)$ thus captures a nuanced form of memory, wherein the weight depends on how much time has passed since the event and on the amplitude of the state at the time of the event.

The use of the supremum operation $\sup_{t \in I}$ in the definition ensures that the functional reflects the maximal combined impact that may occur over the time interval. The inequality

$$\|f\|_{\infty} \leq S_{\kappa, \Lambda}(f)$$

(Lemma 4.2 (iv)) shows that the adaptive memory-dependent functional is no weaker than the classical supremum norm; this fundamental comparison guarantees that the new functional does not underestimate the classical magnitude of a function, reflecting its basic compatibility, while laying the groundwork for further analysis of the conditions under which equality holds in this inequality and when it becomes strict. The basic properties established in Lemma 4.2, existence, controlled boundedness, positive definiteness, and further confirm that $S_{\kappa, \Lambda}(\cdot)$ is a well-defined functional on $C(I)$. Moreover, Theorem 4.1 demonstrates that when the maximum of $|f|$ is attained in $(0, T]$, the adaptive memory-dependent functional is strictly greater than the classical supremum norm, reflecting the additional strictness contributed by the memory component.

The construction proceeds in a layered and sequential manner: First, the auxiliary functions $\mathcal{J}_f$ and $\mathcal{M}_f$ are introduced and their well-posedness established (Lemma 4.1); second, the functional is defined via a supremum operation (Definition 4.2); finally, its core properties are systematically verified (Lemma 4.2) and a refined comparison result is obtained (Theorem 4.1).

Returning to the concrete contexts discussed in Remark 4.1, the functional $S_{\kappa, \Lambda}(f)$ quantifies, for instance, the maximal combined impact of present and past neural activity in a single scalar value, thereby capturing nonlinear adaptive phenomena such as the reduced responsiveness after sustained stimulation (habituation) or the heightened sensitivity following a salient event. This unifying perspective illustrates how the abstract mathematical construction translates into a practical tool for analyzing state-dependent, history-sensitive adaptive behavior across applications.

4.3. Construction of the adaptive memory-dependent sensitivity set

Having established the adaptive memory-dependent functional $S_{\kappa, \Lambda}(f)$ and its fundamental properties in the preceding subsection (see Definition 4.2 and Lemma 4.2), we now introduce the collection of functions that will serve as the primary setting for subsequent analysis. This collection is defined as the set of all functions on I for which $S_{\kappa, \Lambda}(f)$ is finite.

Remark 4.4 (Extension of the functional to broader function classes). The adaptive memory-dependent functional $S_{\kappa, \Lambda}(f)$ was introduced in Definition 4.2 for continuous functions $f \in C(I)$. For a broader class of functions, the same expression

$$S_{\kappa, \Lambda}(f) := \sup_{t \in I} \left(|f(t)| + \int_0^t \Lambda(s, f(s)) \kappa(t-s) |f(s)| ds \right)$$

remains meaningful provided the integral exists as a Lebesgue integral. In particular, if f is bounded and measurable, the composition $\Lambda(s, f(s))$ is measurable by the Carathéodory property of Λ , and the integral is well-defined. In the following, we consider functions for which $S_{\kappa, \Lambda}(f)$ is well-defined and finite.

Definition 4.3 (Adaptive memory-dependent sensitivity set). Let $\kappa \in \mathcal{K}_{\text{reg}}$ be a regular admissible kernel (Definition 2.2) and let $\Lambda \in \mathcal{A}(I)$ be an adaptive sensitivity function (Definition 3.1). The

adaptive memory-dependent sensitivity set is defined as

$$\mathcal{M}_{\kappa,\Lambda}(I) := \{f: I \rightarrow \mathbb{R} \mid S_{\kappa,\Lambda}(f) < \infty\}, \quad (4.23)$$

where $S_{\kappa,\Lambda}(f)$ is understood in the sense of Remark 4.4.

Remark 4.5 (On the set-theoretic characterization of $\mathcal{M}_{\kappa,\Lambda}(I)$). Definition 4.3 does not impose any a priori regularity conditions on functions; membership is solely governed by the finiteness of $S_{\kappa,\Lambda}(f)$. As will be shown in Theorem 4.2, every continuous function on I has finite $S_{\kappa,\Lambda}(f)$ value, so $C(I) \subset \mathcal{M}_{\kappa,\Lambda}(I)$. Moreover, Proposition 4.2 demonstrates that the set also contains certain discontinuous functions, such as the indicator function of a subinterval. Consequently, the set $\mathcal{M}_{\kappa,\Lambda}(I)$ is strictly larger than $C(I)$.

Theorem 4.2 (Embedding of $C(I)$ into the adaptive memory-dependent sensitivity set $\mathcal{M}_{\kappa,\Lambda}(I)$ and basic estimates). *Let $\kappa \in \mathcal{K}_{\text{reg}}$ and $\Lambda \in \mathcal{A}(I)$. Then the adaptive memory-dependent sensitivity set $\mathcal{M}_{\kappa,\Lambda}(I)$ satisfies the following properties:*

- (1) **Set-theoretic inclusion:** $C(I) \subset \mathcal{M}_{\kappa,\Lambda}(I)$; in particular, $\mathcal{M}_{\kappa,\Lambda}(I)$ is nonempty.
- (2) **Comparison with the classical supremum norm and controlled boundedness:** For every $f \in C(I)$,

$$\|f\|_{\infty} \leq S_{\kappa,\Lambda}(f) \leq (1 + \Lambda_{\infty}\kappa_{\infty}T)\|f\|_{\infty}, \quad (4.24)$$

where

$$\Lambda_{\infty} := \sup_{(s,x) \in I \times \mathbb{R}} \Lambda(s,x) \quad \text{and} \quad \kappa_{\infty} := \sup_{\tau \in I} \kappa(\tau)$$

are the uniform bounds introduced in Lemma 4.1 (see (4.4) and (4.5)).

Proof. We establish each claim using the results of Lemma 4.2 and the uniform bounds Λ_{∞} , κ_{∞} established therein.

Proof of (i): set-theoretic inclusion $C(I) \subset \mathcal{M}_{\kappa,\Lambda}(I)$. Let $f \in C(I)$ be arbitrary. Since $I = [0, T]$ is compact, f is bounded; consequently, its supremum norm $\|f\|_{\infty} = \sup_{t \in I} |f(t)|$ is finite. By Lemma 4.2 (ii), we have the estimate

$$S_{\kappa,\Lambda}(f) \leq (1 + \Lambda_{\infty}\kappa_{\infty}T)\|f\|_{\infty}. \quad (4.25)$$

The right-hand side is finite because:

- $\Lambda_{\infty} \leq \lambda_{\max} < \infty$ by condition (AS1) of Definition 3.1;
- $\kappa_{\infty} \leq M_{\kappa} < \infty$ by condition (R1) of Definition 2.2;
- $T < \infty$ is the fixed time horizon;
- $\|f\|_{\infty} < \infty$ as noted above.

Thus, $S_{\kappa,\Lambda}(f) < \infty$, and since f is continuous, $S_{\kappa,\Lambda}(f)$ is well-defined. Hence, $f \in \mathcal{M}_{\kappa,\Lambda}(I)$. Since f was arbitrary, we conclude $C(I) \subset \mathcal{M}_{\kappa,\Lambda}(I)$.

Nonemptiness follows immediately: The zero function $\mathbf{0}(t) \equiv 0$ belongs to $C(I)$, hence also to $\mathcal{M}_{\kappa,\Lambda}(I)$. Therefore $\mathcal{M}_{\kappa,\Lambda}(I) \neq \emptyset$.

Proof of (ii): comparison with the classical supremum norm and controlled boundedness. The left-hand inequality $\|f\|_{\infty} \leq S_{\kappa,\Lambda}(f)$ is precisely Lemma 4.2 (iv), which holds for every $f \in C(I)$. The right-hand inequality is exactly (4.25) derived above. Combining these yields the two-sided estimate (4.24). $\square$

Proposition 4.1 (Controlled estimate and continuity of the inclusion mapping into $\mathcal{M}_{\kappa,\Lambda}(I)$). Let $\kappa \in \mathcal{K}_{\text{reg}}$ and $\Lambda \in \mathcal{A}(I)$. The identity mapping

$$\iota : C(I) \rightarrow \mathcal{M}_{\kappa,\Lambda}(I), \quad \iota(f) = f,$$

satisfies the following properties, where $C(I)$ is equipped with the supremum norm $\|\cdot\|_\infty$ and $\mathcal{M}_{\kappa,\Lambda}(I)$ is equipped with the functional $S_{\kappa,\Lambda}$:

(i) **Linearity**: ι is a linear mapping.

(ii) **Controlled estimate**: There exists a constant $C = 1 + \Lambda_\infty \kappa_\infty T$ such that

$$S_{\kappa,\Lambda}(\iota(f)) \leq C\|f\|_\infty \quad \text{for all } f \in C(I). \quad (4.26)$$

(iii) **Continuity**: ι is uniformly continuous (in fact, Lipschitz) when $C(I)$ is equipped with the supremum norm, since for any $f, g \in C(I)$,

$$S_{\kappa,\Lambda}(\iota(f) - \iota(g)) \leq C\|f - g\|_\infty. \quad (4.27)$$

Proof. We establish each property in turn.

Proof of (i): linearity. For any $f, g \in C(I)$ and any scalars $\alpha, \beta \in \mathbb{R}$,

$$\iota(\alpha f + \beta g) = \alpha f + \beta g = \alpha \iota(f) + \beta \iota(g),$$

since ι is defined as the identity mapping. Hence ι is linear. Note that this linearity follows directly from the definition of ι as the identity map and does not require $\mathcal{M}_{\kappa,\Lambda}(I)$ to be equipped with a vector-space structure.

Proof of (ii): controlled estimate. Let $C := 1 + \Lambda_\infty \kappa_\infty T$. For any $f \in C(I)$, Theorem 4.2 (ii) provides the estimate

$$S_{\kappa,\Lambda}(f) \leq C\|f\|_\infty.$$

Since $S_{\kappa,\Lambda}(\iota(f)) = S_{\kappa,\Lambda}(f)$ by definition of ι , the inequality (4.26) follows immediately.

Proof of (iii): continuity. We first note that ι is linear by part (i). Consequently, for any $f, g \in C(I)$,

$$\iota(f) - \iota(g) = \iota(f - g).$$

Now apply the controlled estimate (4.26) to the function $f - g \in C(I)$. This yields

$$S_{\kappa,\Lambda}(\iota(f - g)) \leq C\|f - g\|_\infty.$$

Combining the two equalities, we obtain

$$S_{\kappa,\Lambda}(\iota(f) - \iota(g)) = S_{\kappa,\Lambda}(\iota(f - g)) \leq C\|f - g\|_\infty,$$

which is precisely (4.27). This inequality shows that ι is Lipschitz continuous with Lipschitz constant C with respect to the distance induced by the supremum norm on $C(I)$. In particular, for any $\varepsilon > 0$, choosing $\delta = \varepsilon/C$ gives

$$S_{\kappa,\Lambda}(\iota(f) - \iota(g)) < \varepsilon \quad \text{whenever} \quad \|f - g\|_\infty < \delta,$$

which is the definition of uniform continuity. Hence, ι is uniformly continuous. $\square$

Proposition 4.2 (A bounded discontinuous function belongs to $\mathcal{M}_{\kappa,\Lambda}(I)$). Let $\kappa \in \mathcal{K}_{\text{reg}}$ and $\Lambda \in \mathcal{A}(I)$ be arbitrary. For any fixed $t_\star \in (0, T)$, define

$$f_{t_\star}(t) := \mathbf{1}_{[0, t_\star]}(t) = \begin{cases} 1, & 0 \leq t \leq t_\star, \\ 0, & t_\star < t \leq T. \end{cases} \quad (4.28)$$

Then $f_{t_\star}$ is discontinuous at $t = t_\star$ (hence $f_{t_\star} \notin C(I)$), yet $f_{t_\star} \in \mathcal{M}_{\kappa,\Lambda}(I)$; i.e., $S_{\kappa,\Lambda}(f_{t_\star})$ is well-defined and finite.

Proof. We first verify that $S_{\kappa,\Lambda}(f_{t_\star})$ is well-defined. The function $f_{t_\star}$ defined in (4.28) is piecewise constant, hence Borel measurable. Since Λ is a Carathéodory function (measurable in s and continuous in x), the composition $\Lambda(s, f_{t_\star}(s))$ is Lebesgue measurable. Moreover, $f_{t_\star}$ is bounded by 1, and κ is bounded by $\kappa_\infty < \infty$. Consequently, the integral defining $\mathcal{J}_{f_{t_\star}}(t)$ exists as a Lebesgue integral for each $t \in I$, and $S_{\kappa,\Lambda}(f_{t_\star})$ is well-defined.

Now, by the definitions of $\mathcal{M}_f$ and $S_{\kappa,\Lambda}$ (see Definitions 4.1 and 4.2, together with Remark 4.4 for the extension to broader function classes), for any $t \in I$, we have

$$\mathcal{M}_{f_{t_\star}}(t) = |f_{t_\star}(t)| + \mathcal{J}_{f_{t_\star}}(t),$$

where

$$\mathcal{J}_{f_{t_\star}}(t) = \int_0^t \Lambda(s, f_{t_\star}(s)) \kappa(t-s) |f_{t_\star}(s)| ds.$$

Observe that $|f_{t_\star}(s)| \leq 1$ for all $s \in I$. By condition (AS1) of Definition 3.1, we have

$$\Lambda(s, f_{t_\star}(s)) \leq \Lambda_\infty < \infty$$

for all $s \in I$, where $\Lambda_\infty := \sup_{(s,x) \in I \times \mathbb{R}} \Lambda(s, x)$. By condition (R1) of Definition 2.2, we have

$$\kappa(t-s) \leq \kappa_\infty < \infty$$

for all $t, s \in I$ with $s \leq t$, where $\kappa_\infty := \sup_{\tau \in I} \kappa(\tau)$. Consequently,

$$\mathcal{J}_{f_{t_\star}}(t) \leq \Lambda_\infty \kappa_\infty \int_0^t 1 ds = \Lambda_\infty \kappa_\infty t \leq \Lambda_\infty \kappa_\infty T. \quad (4.29)$$

Since $|f_{t_\star}(t)| \leq 1$ as well, we obtain from (4.29) that for every $t \in I$,

$$\mathcal{M}_{f_{t_\star}}(t) = |f_{t_\star}(t)| + \mathcal{J}_{f_{t_\star}}(t) \leq 1 + \Lambda_\infty \kappa_\infty T.$$

Therefore,

$$S_{\kappa,\Lambda}(f_{t_\star}) = \sup_{t \in I} \mathcal{M}_{f_{t_\star}}(t) \leq 1 + \Lambda_\infty \kappa_\infty T < \infty, \quad (4.30)$$

which proves $f_{t_\star} \in \mathcal{M}_{\kappa,\Lambda}(I)$ by Definition 4.3. $\square$

Remark 4.6 (Mathematical significance of the adaptive memory-dependent sensitivity set and its structure). Theorem 4.2 and Propositions 4.1 and 4.2 collectively elucidate the nature of $\mathcal{M}_{\kappa,\Lambda}(I)$. Several aspects of this structure merit elaboration.

Connection with continuous functions. For continuous functions, Lemma 4.2 (iv) yields the inequality $\|f\|_\infty \leq S_{\kappa,\Lambda}(f)$, showing that the adaptive memory-dependent functional dominates the classical supremum norm. Theorem 4.2 (ii) refines this observation into a two-sided estimate:

$$\|f\|_\infty \leq S_{\kappa,\Lambda}(f) \leq (1 + \Lambda_\infty \kappa_\infty T) \|f\|_\infty \quad \forall f \in C(I).$$

Thus, every continuous function belongs to $\mathcal{M}_{\kappa,\Lambda}(I)$, i.e., $C(I) \subset \mathcal{M}_{\kappa,\Lambda}(I)$, and on this subspace the functional, $S_{\kappa,\Lambda}$, is equivalent to the classical supremum norm. The inclusion mapping $\iota : C(I) \rightarrow \mathcal{M}_{\kappa,\Lambda}(I)$ (where $C(I)$ is equipped with $\|\cdot\|_\infty$ and $\mathcal{M}_{\kappa,\Lambda}(I)$ is equipped with $S_{\kappa,\Lambda}$) is linear, bounded, and continuous; these properties are recorded in Proposition 4.1.

Enlargement beyond continuity. The set $\mathcal{M}_{\kappa,\Lambda}(I)$ is strictly larger than $C(I)$. Proposition 4.2 provides an explicit illustration: The indicator function of a subinterval,

$$f_{t_\star}(t) = \mathbf{1}_{[0,t_\star]}(t),$$

is discontinuous at $t = t_\star$ (hence $f_{t_\star} \notin C(I)$), yet satisfies $S_{\kappa,\Lambda}(f_{t_\star}) < \infty$, and therefore belongs to $\mathcal{M}_{\kappa,\Lambda}(I)$. This example shows that a function can belong to $\mathcal{M}_{\kappa,\Lambda}(I)$ without being continuous. This flexibility is particularly relevant for modeling nonlinear phenomena where signals exhibit jump discontinuities, such as abrupt environmental changes or on-off switching in biological and engineered systems.

Quantitative comparison. The functional $S_{\kappa,\Lambda}(f)$ merges the instantaneous amplitude $|f(t)|$ with the adaptively weighted historical accumulation $\mathcal{J}_f(t)$, thereby capturing the entire evolution of f . Theorem 4.1 reveals a refined quantitative property: When the maximum of $|f|$ is attained in $(0, T]$, the adaptive memory-dependent functional strictly exceeds the classical supremum norm,

$$S_{\kappa,\Lambda}(f) > \|f\|_\infty,$$

which reflects the influence of historical accumulation on the value of $S_{\kappa,\Lambda}(f)$.

Methodological perspective. The construction of $\mathcal{M}_{\kappa,\Lambda}(I)$ follows a layered approach: Starting from the kernel κ and the sensitivity function Λ , we first introduced the auxiliary functions $\mathcal{J}_f$ and $\mathcal{M}_f$ and established their well-posedness (Lemma 4.1); we then defined the functional $S_{\kappa,\Lambda}$ via a supremum operation (Definition 4.2) and verified its basic properties (Lemma 4.2); finally, we introduced the set $\mathcal{M}_{\kappa,\Lambda}(I)$ and examined its relationship with $C(I)$ (Theorem 4.2, Propositions 4.1 and 4.2). This structured development offers a possible framework for investigating functions and functionals within the context of adaptive memory.

Nonlinear significance. The set $\mathcal{M}_{\kappa,\Lambda}(I)$ is defined through $S_{\kappa,\Lambda}(f)$, whose construction fundamentally relies on the state-dependent sensitivity factor $\Lambda(s, f(s))$. This factor introduces a nonlinear coupling between the trajectory f and its own history: The weight assigned to a past event depends not only on when it occurred (via κ) but also on the value of f at that time (via Λ). As illustrated in Remark 4.1, this nonlinearity enables the description of adaptive phenomena such as habituation (diminished sensitivity under sustained stimulation) and selective retention (enhanced memory of salient events). Thus, $\mathcal{M}_{\kappa,\Lambda}(I)$ can be viewed as a mathematical construct designed for analyzing systems where memory is time-dependent and state-dependent.

5. Conclusions

This work has been devoted to a functional-analytic framework for nonlinear adaptive memory, in which the influence of past events is governed by both elapsed time and the state values along a trajectory. This dual dependence is realized through a three-layer organization: A hierarchical classification of kernels, a class of state-dependent sensitivity functions, and a memory-dependent functional with an associated function set. The main contributions of each layer can be summarized as follows.

First, we have introduced a hierarchical classification of memory kernels, consisting of mathematically admissible kernels (Definition 2.1), regular admissible kernels (Definition 2.2), and generalized admissible kernels (Definition 2.3). This classification was illustrated by three prototypical examples—the exponential kernel (Example 2.1), the power-law kernel (Example 2.2), and the finite-memory kernel (Example 2.3)—which together illustrate short-memory, long-memory, and finite-window behaviors. The associated integral estimates (Lemma 2.1 and Proposition 2.1) provide the analytical tools required for the subsequent construction of the adaptive sensitivity framework in Section 3.

Second, we have introduced the class of adaptive sensitivity functions (Definition 3.1), which modulate the memory weight according to the state value $f(s)$ along the trajectory. A concrete construction based on historical deviation accumulation (Example 3.1) was analyzed in detail, and its fundamental properties—uniform boundedness, a Lipschitz-type estimate with respect to the supremum norm, continuity in time, and positivity at the zero trajectory—were established in Theorem 3.1. The special case of purely instantaneous sensitivity was obtained as a special case in Corollary 3.1, showing that the construction interpolates continuously between instantaneous response and genuinely history-dependent sensitivity.

Third, we have constructed the adaptive memory-dependent functional $S_{\kappa,\Lambda}(f)$ (Definition 4.2) and the associated function set $\mathcal{M}_{\kappa,\Lambda}(I)$ (Definition 4.3). The well-posedness of the underlying auxiliary functions $\mathcal{J}_f$ and $\mathcal{M}_f$ was established in Lemma 4.1, and the fundamental properties of the functional—existence, controlled boundedness, positive definiteness, and comparison with the classical supremum norm—were verified in Lemma 4.2. A strict inequality was proved in Theorem 4.1 when the maximum of $|f|$ is attained in $(0, T]$. The embedding $C(I) \subset \mathcal{M}_{\kappa,\Lambda}(I)$ and the two-sided estimate

$$\|f\|_\infty \leq S_{\kappa,\Lambda}(f) \leq (1 + \Lambda_\infty \kappa_\infty T) \|f\|_\infty \quad (5.1)$$

were established in Theorem 4.2, showing that $S_{\kappa,\Lambda}$ and the classical supremum norm control each other on $C(I)$. The inclusion mapping was further shown to be linear, to satisfy a controlled estimate, and to be Lipschitz continuous (Proposition 4.1), and the set $\mathcal{M}_{\kappa,\Lambda}(I)$ was shown to contain bounded discontinuous functions such as the indicator function of a subinterval (Proposition 4.2), thereby confirming that $\mathcal{M}_{\kappa,\Lambda}(I)$ is strictly larger than $C(I)$.

Several directions remain open for future investigation. On the theoretical side, extending the main results from the regular kernel class $\mathcal{K}_{\text{reg}}$ to the generalized kernel class $\mathcal{K}_{\text{gen}}$ would broaden the applicability of the framework to sign-changing and discontinuous memory mechanisms, including inhibitory neural feedback, abrupt system resets, and adaptive gain control. On the applied side, incorporating the proposed functional into concrete models from fields such as neuroscience, adaptive control, and machine learning would help to assess its descriptive power and to explore its potential for

practical implementation. We hope that the present work offers a constructive step toward a broader mathematical understanding of nonlinear adaptive memory phenomena.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (a Chinese national program for basic research in higher education institutions) under Grant No. 2682026CX035. The author would like to express sincere gratitude to the Editors, the Assistant Editor, and the anonymous reviewers for their careful reading of the manuscript and for their insightful, constructive comments, which have greatly improved the clarity and rigor of this work. The author is also grateful for the consistently positive and encouraging feedback received from all reviewers during the review process, and for the efficient handling and thoughtful guidance provided by the editorial team throughout the submission.

Use of Generative-AI tools declaration

The author declares that no Generative-AI (Artificial Intelligence) tools were used in the creation of this article.

Conflict of interest

The author declares that there is no conflict of interest.

References

1. C. Cortázar, F. Quirós, N. Wolanski, A heat equation with memory: large-time behavior, *J. Funct. Anal.*, **281** (2021), 109174. <https://doi.org/10.1016/j.jfa.2021.109174>
2. C. Cortázar, F. Quirós, N. Wolanski, Asymptotic profiles for inhomogeneous heat equations with memory, *Math. Ann.*, **389** (2024), 3705–3746. <https://doi.org/10.1007/s00208-023-02707-6>
3. J. Wang, Q. Ma, W. Zhou, Attractor of the nonclassical diffusion equation with memory on time-dependent space, *AIMS Math.*, **8** (2023), 14820–14841. <https://doi.org/10.3934/math.2023757>
4. N. Mori, Dissipative structure and global existence in critical space for Timoshenko system of memory type, *J. Differ. Equ.*, **265** (2018), 1627–1653. <https://doi.org/10.1016/j.jde.2018.04.014>
5. Y. Qin, X. Pan, Global existence, asymptotic behavior and uniform attractors for a non-autonomous Timoshenko system of thermoelasticity of type III with a time-varying delay, *J. Math. Anal. Appl.*, **484** (2020), 123672. <https://doi.org/10.1016/j.jmaa.2019.123672>
6. B. Baeumer, M. Geissert, M. Kovács, Existence, uniqueness and regularity for a class of semilinear stochastic Volterra equations with multiplicative noise, *J. Differ. Equ.*, **258** (2015), 535–554. <https://doi.org/10.1016/j.jde.2014.09.020>
7. B. S. H. Kashkaria, M. I. Syam, Evolutionary computational intelligence in solving a class of nonlinear Volterra-Fredholm integro-differential equations, *J. Comput. Appl. Math.*, **311** (2017), 314–323. <https://doi.org/10.1016/j.cam.2016.07.027>

8. Z. L. Lu, A posteriori error estimates of fully discrete finite-element schemes for nonlinear parabolic integro-differential optimal control problems, *Adv. Differ. Equ.*, **2014** (2014), 15. <https://doi.org/10.1186/1687-1847-2014-15>
9. W. Feng, S. X. Yang, H. X. Wu, Robust stability analysis of neutral-type hybrid bidirectional associative memory neural networks with time-varying delays, *Abstr. Appl. Anal.*, **2014** (2014), 560861. <https://doi.org/10.1155/2014/560861>
10. S. L. Manu, S. Shikaa, T. Richard, E. P. Dovi, Mathematical model for prediction of tuberculosis in Nigeria using hybrid fractional differential equations and artificial neural network methods, *Franklin Open*, **11** (2025), 100248. <https://doi.org/10.1016/j.fraope.2025.100248>
11. M. D. Ruiz-Medina, Spectral analysis of multifractional LRD functional time series, *Fract. Calc. Appl. Anal.*, **25** (2022), 1426–1458. <https://doi.org/10.1007/s13540-022-00053-z>
12. N. Wulkow, P. Koltai, C. Schütte, Memory-based reduced modelling and data-based estimation of opinion spreading, *J. Nonlinear Sci.*, **31** (2021), 19. <https://doi.org/10.1007/s00332-020-09673-2>
13. D. Alonso-Gutiérrez, M. A. Hernández Cifre, J. Yepes Nicolás, Further inequalities for the (generalized) Wills functional, *Commun. Contemp. Math.*, **23** (2021), 2050011. <https://doi.org/10.1142/S021919972050011X>
14. G. Alberti, G. Crippa, A. L. Mazzucato, Exponential self-similar mixing by incompressible flows, *J. Amer. Math. Soc.*, **32** (2019), 445–490. <https://doi.org/10.1090/jams/913>
15. J. F. Brock, N. M. Dunfield, Norms on the cohomology of hyperbolic 3-manifolds, *Invent. Math.*, **210** (2017), 531–558. <https://doi.org/10.1007/s00222-017-0735-3>
16. K. D. Schmidt, A general Jordan decomposition, *Arch. Math.*, **38** (1982), 556–564. <https://doi.org/10.1007/BF01304831>
17. A. S. Leonov, On the total variation for functions of several variables and a multidimensional analog of Helly’s selection principle, *Math. Notes*, **63** (1996), 61–71. <https://doi.org/10.1007/BF02316144>
18. V. V. Chistyakov, Y. V. Tretyachenko, Maps of several variables of finite total variation. II. E. Helly-type pointwise selection principles, *J. Math. Anal. Appl.*, **369** (2010), 82–93. <https://doi.org/10.1016/j.jmaa.2010.02.042>
19. V. Ene, On the decomposition theorems of Lebesgue and Jordan, *Real Anal. Exchange*, **23** (1998), 313–324.
20. S. Kantorovitz, A Jordan decomposition for operators in Banach space, *Trans. Am. Math. Soc.*, **120** (1965), 526–550. <https://doi.org/10.1090/S0002-9947-1965-0203472-4>
21. S. Saminger-Platz, B. De Baets, H. De Meyer, A generalization of the Mulholland inequality for continuous Archimedean t-norms, *J. Math. Anal. Appl.*, **345** (2008), 607–614. <https://doi.org/10.1016/j.jmaa.2008.03.045>
22. B. Davvaz, V. Leoreanu-Fotea, Applications of interval valued fuzzy n-ary polygroups with respect to t-norms (t-conorms), *Comput. Math. Appl.*, **57** (2009), 1413–1424. <https://doi.org/10.1016/j.camwa.2009.01.015>
23. V. Peiris, V. Roshchina, N. Sukhorukova, Artificial neural networks with uniform norm-based loss functions, *Adv. Comput. Math.*, **50** (2024), 31. <https://doi.org/10.1007/s10444-024-10124-9>

24. N. E. Yudin, Adaptive Gauss–Newton method for solving systems of nonlinear equations, *Dokl. Math.*, **104** (2021), 293–296. <https://doi.org/10.1134/S1064562421050161>
25. D. Kitkuan, P. Kumam, V. Berinde, A. Padcharoen, Adaptive algorithm for solving the SCFPP of demicontractive operators without a priori knowledge of operator norms, *Ann. Univ. Ovidius Math. Ser.*, **27** (2019), 153–175. <https://doi.org/10.2478/auom-2019-0039>
26. P. Sunthrayuth, K. Muangchoo, P. Chalamjiak, W. Nithiarayaphaks, Inertial self-adaptive algorithm with two different inertial factors for solving split feasibility problems in Banach spaces, *J. Anal.*, **33** (2025), 1815–1847. <https://doi.org/10.1007/s41478-025-00896-8>
27. D. Tian, L. Shi, R. Chen, Iterative algorithm for solving the multiple-sets split equality problem with split self-adaptive step size in Hilbert spaces, *J. Inequal. Appl.*, **2016** (2016), 34. <https://doi.org/10.1186/s13660-016-0982-7>
28. P. Jailoka, S. Suantai, On split fixed point problems for multi-valued mappings and designing a self-adaptive method, *Results Math.*, **76** (2021), 133. <https://doi.org/10.1007/s00025-021-01441-2>
29. N. Karmitsa, K. Joki, A. Airola, T. Pahikkala, Limited memory bundle DC algorithm for sparse pairwise kernel learning, *J. Glob. Optim.* **92** (2025), 55–85. <https://doi.org/10.1007/s10898-025-01481-w>
30. T. Linß, G. Radojev, Maximum-norm a posteriori error bounds for an extrapolated upwind scheme applied to a singularly perturbed convection-diffusion problem, *Mediterr. J. Math.*, **21** (2024), 161. <https://doi.org/10.1007/s00009-024-02698-x>
31. A. Esser, A. Mukherjee, S. Sarkar, Memory-efficient attacks on small LWE keys, *J. Cryptol.*, **37** (2024), 36. <https://doi.org/10.1007/s00145-024-09516-3>
32. S. Torregrosa, V. Champaney, A. Ammar, V. Herbert, F. Chinesta, Physics-based active learning for design space exploration and surrogate construction for multiparametric optimization, *Commun. Appl. Math. Comput.*, **6** (2024), 1899–1923. <https://doi.org/10.1007/s42967-023-00329-y>
33. M. E. Belouafi, M. Beggas, N. E. H. Nesba, Uniform convergence of multigrid methods for elliptic quasi-variational inequalities and its implementation, *Commun. Math. Appl.*, **14** (2023), 633. <https://doi.org/10.26713/cma.v14i2.2039>
34. M. Hintermüller, M. Hinze, R. H. W. Hoppe, Weak-duality based adaptive finite element methods for PDE-constrained optimization with pointwise gradient state-constraints, *J. Comput. Math.*, **30** (2012), 101–123. <https://doi.org/10.4208/jcm.1109-m3522>
35. R. Blanquero, E. Carrizosa, N. Gómez-Vargas, On contextual inverse multiobjective problems, *Eur. J. Oper. Res.*, **33** (2025), 192–202. <https://doi.org/10.1016/j.ejor.2025.12.007>
36. T. C. F. Cheng, C. K. Ing, S. H. Yu, Inverse moment bounds for sample autocovariance matrices based on detrended time series and their applications, *Linear Algebra Appl.*, **473** (2015), 180–201. <https://doi.org/10.1016/j.laa.2014.05.017>
37. G. Chen, G. Fang, Probabilistic adaptive width of a multivariate Sobolev space equipped with a Gaussian measure, *Constr. Approx.*, **24** (2006), 245–262.
38. F. Girosi, Some extensions of radial basis functions and their applications in artificial intelligence, *Comput. Math. Appl.*, **24** (1992), 61–80. [https://doi.org/10.1016/0898-1221\(92\)90172-E](https://doi.org/10.1016/0898-1221(92)90172-E)

39. A. R. Nurutdinov, Bio-inspired neural network architecture of embodied intelligence, *Lobachevskii J. Math.*, **45** (2024), 5156–5171. <https://doi.org/10.1134/S1995080224606027>
40. M. E. Cornejo, J. Medina, F. J. Ocaña, Theories, models and bases of attribute implications in multi-adjoint concept lattices with hedges, *Comput. Appl. Math.*, **45** (2026), 30. <https://doi.org/10.1007/s40314-025-03391-9>
41. I. Tsuda, T. Namiki, Some comments on the relationship between the rate function in the large deviation principle and the Kullback–Leibler divergence: toward the interpretation of neural estimation of mutual information, *Jpn. J. Ind. Appl. Math.*, **43** (2026), 7. <https://doi.org/10.1007/s13160-025-00756-9>
42. S. Allana, R. Dara, X. Lin, P. Xiong, Towards integration of privacy enhancing technologies in explainable artificial intelligence, *Knowl.-Based Syst.*, **335** (2025), 115235. <https://doi.org/10.1016/j.knosys.2025.115235>
43. D. Palitta, V. Simoncini, Computationally enhanced projection methods for symmetric Sylvester and Lyapunov matrix equations, *J. Comput. Appl. Math.*, **330** (2018), 648–659. <https://doi.org/10.1016/j.cam.2017.08.011>
44. M. Shamrai, Analysis of perturbations of singular values in concatenated matrices, *Ukrainian Math. J.*, **77** (2025), 1136–1149. <https://doi.org/10.1007/s11253-025-02512-1>
45. J. H. Jiang, B. Miao, A study of anomalous stochastic processes via generalizing fractional calculus, *Chaos*, **35** (2025), 023156. <https://doi.org/10.1063/5.0244009>
46. B. Dong, D. H. Lv, K. Y. Xi, J. H. Li, D. H. Yu, Adaptive frequency evolution decomposition combined with improved fluctuation-based dispersion entropy for muscle fatigue characterization, *IEEE Trans. Instrum. Meas.*, **75** (2026), 1–12. <https://doi.org/10.1109/TIM.2026.3655906>
47. J. H. Jiang, Towards a mathematical theory of adaptive memory: from time-varying to responsive fractional Brownian motion, *arXiv preprint*, 2025. <https://doi.org/10.48550/arXiv.2512.10057>
48. J. H. Jiang, A two-parameter memory-weighted velocity operator for time and state variables: foundations and fundamental properties, *arXiv preprint*, 2026. <https://doi.org/10.48550/arXiv.2601.05122>
49. V. Fischer, S. Mikkelsen, Semiclassical functional calculus on nilpotent Lie groups and their compact nilmanifolds, *Anal. Math. Phys.*, **15** (2025), 60. <https://doi.org/10.1007/s13324-025-01051-z>
50. F. M. Baêta, Asymptotic weighted approximation of convex functions, *Adv. Appl. Math.*, **176** (2026), 103057. <https://doi.org/10.1016/j.aam.2026.103057>



AIMS Press

©2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)