

Lyapunov exponents of fractional-order systems: a definition from the diffusive representation, the ageing of memory, and what memory truncation measures

Stefano Fazzino

sfazzino@gmail.com

University of Catania <https://orcid.org/0000-0002-3015-4938>

Research Article

Keywords: Lyapunov exponents, fractional-order systems, Caputo derivative, diffusive representation, multiplicative ergodic theorem, short-memory principle

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11383967/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Lyapunov exponents of fractional-order systems: a definition from the diffusive representation, the ageing of memory, and what memory truncation measures

Stefano Fazzino

Independent researcher, Catania, Italy

Abstract

Lyapunov exponents of fractional-order systems are routinely computed, but it has never been stated what they are. A system with a Caputo derivative is not a dynamical system in the usual sense: its state is its whole history, it is not invariant under time translation, and the linearised evolution is not a cocycle, so the multiplicative ergodic theorem that gives meaning to Lyapunov exponents does not apply. We use the diffusive representation of the Caputo operator to rewrite the system as an autonomous semilinear equation on a Hilbert space of memory modes, where the theorem exists in the form of Ruelle, Mañé and Thieullen. Three consequences follow. First, the memory of a bounded Caputo trajectory is not bounded: the slow modes accumulate the integral of the vector field, which grows as a power of time, so the system ages and exponents must be defined on the stationary dynamics, of which the Caputo trajectory is an algebraically decaying perturbation. Second, the positive exponents are well defined and independent of any truncation of memory, whereas the non-positive part of the spectrum is an essential band at zero: a fractional-order system has no exponential contraction, and the negative exponents reported by existing codes are rates of the observable over a memory window, not exponents of the system. Third, the picture is confirmed numerically by a Benettin algorithm on the augmented system, validated against the Mittag-Leffler function at three orders and applied to the fractional Chen and Rössler systems: the maximal exponent is reproducible

Email addresses: sfazzino@gmail.com (Stefano Fazzino),
sfazzino@dieei.unict.it (Stefano Fazzino)

and insensitive to the resolution of the memory band, all other exponents vanish within a residual that decreases with resolution, and memory truncation acts in two opposite ways that the literature calls by the same name: applied to the Grünwald-Letnikov derivative, as in the short-memory principle, it leaves exponent and attractor unchanged, whereas applied to the integral form of the equation, as in fixed-memory Volterra codes and in finite-mode realisations, it deforms the attractor, lowers the maximal exponent by a quarter to a half and can suppress chaos or destroy the attractor altogether, and in finite-mode realisations it adds spurious negative exponents equal to the inverse of the memory window.

Keywords: Lyapunov exponents, fractional-order systems, Caputo derivative, diffusive representation, multiplicative ergodic theorem, short-memory principle

1. Introduction

Lyapunov exponents measure the average rate at which nearby trajectories of a dynamical system separate or approach one another, and the presence of a positive exponent is the standard quantitative signature of chaos [1, 2, 3]. Their definition rests on a result of Oseledets [4]: for a flow on a finite-dimensional space with an invariant probability measure, the linearised evolution along almost every trajectory is a cocycle, and the growth rates of its singular values converge to a finite set of numbers that do not depend on the starting point. Fractional-order systems, in which the time derivative is replaced by a derivative of non-integer order in the sense of Caputo [5], have become a standard modelling tool wherever memory effects matter, from viscoelasticity to electrochemistry to anomalous diffusion, and chaotic fractional-order versions of the classical oscillators of Lorenz, Rössler, Chen and Lü have been studied for two decades [6, 7, 8]. It is therefore natural that their Lyapunov exponents are computed as a matter of routine. What is less often noticed is that none of the hypotheses of Oseledets' theorem holds for them.

The reason is the non-locality of the fractional derivative. The Caputo derivative of order $\nu \in (0, 1)$ is a convolution with a power-law kernel over the whole past, so the state of the system at time t is not a point of $\mathbb{R}^n$ but the entire trajectory up to t . Three consequences follow that make the usual notion of Lyapunov exponent ill-posed. There is no flow on $\mathbb{R}^n$: the

value of the solution at a given time does not determine its future. The system is not invariant under time translation: shifting a solution in time does not give a solution of the same equation but of one with a different lower limit of integration, so the system has an age. And the linearisation along a trajectory inherits the memory: the perturbation at time t depends on the perturbations at all earlier times, and the family of linearised evolution operators is not a cocycle. The multiplicative ergodic theorem needs a cocycle over a measure-preserving dynamical system; here both are missing.

Numerical practice has sidestepped the question in several ways. A semi-analytical method [9] converts the fractional system into a discrete map by a multistep differential transform and applies the QR algorithm of Benettin to the map; the memory is truncated by the multistep construction. The widely used code of Danca and Kuznetsov [10], recently improved [11], integrates the fractional variational equations with a predictor-corrector scheme and renormalises with QR; a subsequent note [12] clarifies that the algorithm is impulsive with variable lower limit, that is, the history of the perturbation is effectively reset at each renormalisation. A method based on the memory principle of the Grünwald-Letnikov derivative [13] is presented as more faithful to the non-locality. Trigeassou and Maamri [14] use the infinite-state representation of the fractional operator, which is the diffusive representation under another name, to simulate the fractional Chen system, and compute by Gram-Schmidt the exponents of the resulting augmented system together with what they call pseudo-exponents, describing the approach as experimental and the infinite-mode model as essentially theoretical. In parallel, the short-memory principle [5], which discards the history beyond a fixed window to keep the cost linear in time, is used almost universally in long integrations. It is applied in two forms: to the Grünwald-Letnikov derivative, for which Podlubny's error bound holds, and, more loosely, to the Volterra integral form of the equation, for which Ford and Simpson [15] and Deng [16] showed that a fixed memory degrades the accuracy severely. The effect of either form on the Lyapunov exponents has, to our knowledge, not been quantified.

None of these works says what object is being computed, and three empirical facts have remained without explanation: the dependence of the estimates on the integration horizon, with tables that oscillate and no convergence analysis [12]; the absence of the zero exponent that every autonomous flow possesses, observed numerically by Danca, Fečkan, Kuznetsov and Chen [17] and declared not proven; and the pseudo-exponents of [14], a band of exponents of the augmented system that corresponds to nothing in the original

one.

In this paper we give a definition and derive the three facts from it. The tool is the diffusive representation of the Caputo operator [18, 19]: the power-law kernel is the Laplace transform of a positive measure, so the fractional system is exactly equivalent to an autonomous system of ordinary differential equations on a continuum of first-order memory modes, each relaxing at its own rate, all driven by the same vector field. On a suitable Hilbert space of memory modes this augmented system is a semilinear evolution equation, its linearisation is a cocycle, and the multiplicative ergodic theorem exists in the form proved by Ruelle [20] for compact cocycles in Hilbert space and extended by Mañé [21] and Thieullen [22] to cocycles that are only asymptotically compact. Our contribution is the following. We show that the memory of a bounded Caputo trajectory is not bounded in the natural Hilbert space, because the slow modes accumulate the time integral of the vector field, which grows like $t^{1-\nu}$ (Section 2.3); this ageing forces one to define the exponents on the stationary dynamics, of which the Caputo trajectory is a non-autonomous perturbation decaying as $t^{-\nu}$. We prove that the linearised cocycle is the sum of a contraction semigroup of norm one and a compact operator, so that the exponents strictly above zero are well defined, discrete and independent of any truncation of memory (Section 3). We show that the non-positive part of the spectrum is an essential band at zero, which contains the zero exponent of the flow and in which no exponential contraction can exist, and we explain why the negative exponents of the integer-order system are not exponents of the fractional one for any $\nu < 1$ (Section 4). We then build a Benettin algorithm on the augmented system, using the sum-of-exponentials quadratures developed for fast fractional integration [23, 24, 25, 26], validate it against the Mittag-Leffler function (Section 5), and apply it to the fractional Chen and Rössler systems (Section 6). The numerical results confirm the structure of the spectrum and add a fact that we believe has practical consequences for the literature: truncating the memory to a window τ_m is harmless when applied to the derivative, as in the short-memory principle proper, but when applied to the integral form of the equation, which is what fixed-memory Volterra codes and finite-mode realisations do, it deforms the attractor, lowers the maximal exponent by a quarter to a half for windows of five hundred to a thousand time units, destroys the dynamics for a window of a hundred, and, in finite-mode realisations, adds negative exponents equal to $-1/\tau_m$. The two truncations differ in the integrability of the kernel they cut, and we show this by comparing them on the same system

(Section 6.3).

The paper is organised as follows. Section 2 introduces the diffusive representation, the state space and the ageing of memory. Section 3 contains the definition of the exponents and the existence theorem. Section 4 describes the structure of the spectrum. Section 5 describes the algorithm and its validation, Section 6 the results on chaotic systems, Section 7 the discussion, including a conjecture on the rate of convergence that we state but could not test, and Section 8 the conclusions.

2. The diffusive representation as a state space

2.1. The Caputo system and the kernel identity

We consider an autonomous system of fractional order

$${}^C D_t^\nu x(t) = f(x(t)), \quad x(0) = x_0, \quad \nu \in (0, 1), \quad (1)$$

where $x(t) \in \mathbb{R}^n$ is the state, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a locally Lipschitz vector field, x_0 is the initial condition, and ${}^C D_t^\nu$ is the Caputo derivative of order ν , defined for a differentiable function x by

$${}^C D_t^\nu x(t) = \frac{1}{\Gamma(1-\nu)} \int_0^t (t-s)^{-\nu} \dot{x}(s) ds, \quad (2)$$

with Γ the Euler gamma function. It is classical [5] that (1) is equivalent to the Volterra integral equation

$$x(t) = x_0 + \int_0^t k_\nu(t-s) f(x(s)) ds, \quad k_\nu(\tau) = \frac{\tau^{\nu-1}}{\Gamma(\nu)}, \quad (3)$$

in which k_ν is the kernel of the Riemann-Liouville fractional integral of order ν . The kernel decays as a power and is not integrable on $[0, \infty)$: the state at time t depends on the whole past with weights that never die out. This single fact is the source of everything that follows.

The kernel is the Laplace transform of a positive measure. Writing $\int_0^\infty \omega^{-\nu} e^{-\omega\tau} d\omega = \Gamma(1-\nu) \tau^{\nu-1}$, valid for $\nu < 1$, and using the reflection formula $\Gamma(\nu)\Gamma(1-\nu) = \pi/\sin(\pi\nu)$, one obtains

$$k_\nu(\tau) = \int_0^\infty \mu_\nu(\omega) e^{-\omega\tau} d\omega, \quad \mu_\nu(\omega) = \frac{\sin(\pi\nu)}{\pi} \omega^{-\nu}. \quad (4)$$

Here $\omega > 0$ plays the role of a relaxation rate and μ_ν is the density with which rates are represented in the kernel. Substituting (4) into (3) and exchanging the order of integration, which is legitimate by Tonelli's theorem since the integrand is positive in ω and $f(x(s))$ is bounded on $[0, t]$, the solution can be written as

$$\begin{aligned} x(t) &= x_0 + \int_0^\infty \mu_\nu(\omega) z(\omega, t) d\omega, \\ \partial_t z(\omega, t) &= -\omega z(\omega, t) + f(x(t)), \quad z(\omega, 0) = 0, \end{aligned} \tag{5}$$

where $z(\omega, t) \in \mathbb{R}^n$ is the state of the memory mode of rate ω : a first-order filter driven by the vector field $f(x(t))$ and relaxing at rate ω . This is the diffusive representation of Montseny [18], the infinite-state representation of Trigeassou [14], and the basis of the numerical schemes of Yuan and Agrawal [23] and Diethelm [24]. The fractional system is thus an ordinary differential system on a continuum of modes, and the power-law memory is the collective effect of the slow ones: μ_ν has infinite mass at infinity and, more to the point, $\mu_\nu(\omega)/\omega$, the weight of a mode times its relaxation time, is not integrable at the origin, and it is the modes with $\omega \rightarrow 0$ that never forget.

Two features of (5) should be kept in mind. The initial condition x_0 is not encoded in z but appears as an additive constant, and the Caputo problem corresponds to the choice $z(\cdot, 0) = 0$, that is, to zero initial memory. And the limit $\nu \rightarrow 1$ is singular in the right way: $\mu_\nu(\omega) \rightarrow 0$ for every $\omega > 0$ while its mass concentrates at $\omega = 0$, where the mode is a pure integrator, and one recovers $\dot{x} = f(x)$. We shall use this limit as a check on every formula: anything specifically fractional must carry a prefactor that vanishes, or that reduces to the ordinary case, as $\nu \rightarrow 1$.

2.2. The state space

The natural space for the memory profile $z(\cdot, t)$ would seem to be $L^2(\mu_\nu d\omega)$, but it is not adequate, for a precise reason: the output map $z \mapsto \langle \mu_\nu, z \rangle := \int_0^\infty \mu_\nu(\omega) z(\omega) d\omega$, which reconstructs x from the modes and feeds it to f , is not a bounded functional on that space, because by the Cauchy-Schwarz inequality its norm is $(\int \mu_\nu d\omega)^{1/2}$ and $\int_0^\infty \omega^{-\nu} d\omega$ diverges at infinity. The remedy, standard in the theory of diffusive systems [18, 19], is to weight the high rates more heavily. We set

$$\mathcal{H} := L^2(\mathbb{R}^+, \mathbb{R}^n; \rho_\nu(\omega) d\omega), \quad \rho_\nu(\omega) := \mu_\nu(\omega) (1 + \omega), \tag{6}$$

with inner product $\langle u, v \rangle_{\mathcal{H}} = \int_0^\infty \rho_\nu u \cdot v \, d\omega$ and norm $\|\cdot\|_{\mathcal{H}}$. On this space the output functional is bounded, and with constant exactly one:

$$|\langle \mu_\nu, z \rangle| \leq \left(\int_0^\infty \frac{\mu_\nu(\omega)}{1 + \omega} \, d\omega \right)^{1/2} \|z\|_{\mathcal{H}} = \|z\|_{\mathcal{H}}, \quad (7)$$

because $\int_0^\infty \omega^{-\nu} (1 + \omega)^{-1} \, d\omega$ is the Beta integral $B(1 - \nu, \nu) = \pi / \sin(\pi\nu)$. The multiplication operator $A := -\omega \cdot$ is self-adjoint and negative on $\mathcal{H}$ and generates the contraction semigroup $S(t)z = e^{-\omega t} z$, whose norm is $\|S(t)\| = 1$ for every t because of the modes with $\omega \rightarrow 0$.

The full state must include x_0 , which in (5) is a constant but must be perturbed when one speaks of sensitivity to initial conditions. The state space is therefore the product

$$\mathcal{X} := \mathbb{R}^n \times \mathcal{H}, \quad \|(a, z)\|_{\mathcal{X}}^2 = |a|^2 + \|z\|_{\mathcal{H}}^2, \quad (8)$$

and the augmented system, which is autonomous, reads

$$\dot{a} = 0, \quad \partial_t z = -\omega z + f(a + \langle \mu_\nu, z \rangle), \quad x = a + \langle \mu_\nu, z \rangle. \quad (9)$$

The Caputo problem (1) is the restriction of (9) to the initial data $(x_0, 0)$.

Proposition 1 (well-posedness). *Let f be locally Lipschitz. For every datum $(a, z_0) \in \mathcal{X}$ the system (9) has a unique local mild solution in $\mathcal{X}$,*

$$z(t) = S(t)z_0 + \int_0^t S(t-s) f(a + \langle \mu_\nu, z(s) \rangle) \, ds,$$

and for $(a, z_0) = (x_0, 0)$ the function $x = x_0 + \langle \mu_\nu, z \rangle$ is the unique continuous solution of (3) on every interval on which it remains bounded.

Sketch. The delicate point is that a constant vector $f \in \mathbb{R}^n$, seen as a function of ω , does not belong to $\mathcal{H}$, because $\int \rho_\nu \, d\omega = \infty$: the forcing lies outside the state space, as happens for parabolic equations with boundary forcing. The mild integral nevertheless converges in the sense of Bochner, because $S(\tau)f \in \mathcal{H}$ for every $\tau > 0$ with

$$\begin{aligned} \|S(\tau)f\|_{\mathcal{H}}^2 &= |f|^2 \int_0^\infty \rho_\nu(\omega) e^{-2\omega\tau} \, d\omega \\ &= |f|^2 \frac{\sin \pi\nu}{\pi} \left[\Gamma(1 - \nu)(2\tau)^{\nu-1} + \Gamma(2 - \nu)(2\tau)^{\nu-2} \right], \end{aligned}$$

so that $\|S(\tau)f\|_{\mathcal{H}} \sim \tau^{\nu/2-1}$ as $\tau \rightarrow 0$, which is integrable near $\tau = 0$ for every $\nu > 0$. With this estimate the nonlinearity, which is locally Lipschitz from $\mathcal{H}$ to $\mathbb{R}^n$ by (7), falls within the standard theory of semilinear equations [27] through a weighted Picard iteration. The equivalence with (3) is the exchange of integrals described after (4). $\square$

2.3. The system ages

One would expect that if the trajectory $x(t)$ is bounded, the memory $z(\cdot, t)$ is bounded too, and that one may speak of an attractor in $\mathcal{X}$ and of an invariant measure on it. This is not so.

Proposition 2 (ageing). *Let x be a bounded solution of (3) on $[0, \infty)$ whose time average $\bar{x} := \lim_{t \rightarrow \infty} t^{-1} \int_0^t x(s) ds$ exists. Then the slow memory modes accumulate a content that grows as a power of time,*

$$z(\omega, t) \simeq \int_0^t f(x(s)) ds \simeq \frac{\bar{x} - x_0}{\Gamma(2 - \nu)} t^{1-\nu} \quad (\omega \ll 1/t), \quad (10)$$

and consequently

$$\|z(\cdot, t)\|_{\mathcal{H}} \asymp |\bar{x} - x_0| t^{(1-\nu)/2} \quad (t \rightarrow \infty), \quad (11)$$

unless $\bar{x} = x_0$. The orbit of the Caputo problem in $\mathcal{X}$ is neither bounded nor relatively compact.

Proof. For $\omega \ll 1/t$ the factor $e^{-\omega(t-s)}$ is close to one on all of $[0, t]$, so $z(\omega, t) \simeq \int_0^t f(x(s)) ds$, which is the Riemann-Liouville integral of order one of f , denoted $I^1 f$. By the semigroup property of Riemann-Liouville integrals and by (3), $I^1 f = I^{1-\nu} I^\nu f = I^{1-\nu}(x - x_0)$, and if $x - x_0$ has mean $\bar{x} - x_0$ then $I^{1-\nu}(x - x_0) \simeq (\bar{x} - x_0) t^{1-\nu} / \Gamma(2 - \nu)$. The estimate (11) follows by integrating $\rho_\nu |z|^2$ over $\omega < 1/t$: $\int_0^{1/t} \omega^{-\nu} d\omega \cdot t^{2-2\nu} \propto t^{\nu-1} t^{2-2\nu} = t^{1-\nu}$. The contribution of the rates $\omega > 1/t$ is bounded uniformly in t , since there $|z(\omega, t)| \leq \sup |f| / \omega$ and $\rho_\nu \omega^{-2} \sim \omega^{-1-\nu}$ is integrable at infinity; the estimate is made explicit in [Appendix A](#). $\square$

The proposition says that a Caputo solution that is bounded in the observable x has a memory that keeps changing forever, migrating towards ever lower rates: the system does not forget x_0 , and it is precisely the memory of x_0 that grows. This is the quantitative form of a fact known in another guise, the impossibility of exactly periodic solutions for Caputo systems [8]:

non-stationarity is not a transient but a power-law ageing, with exponent $(1 - \nu)/2$ that vanishes, as it should, when $\nu \rightarrow 1$.

The correct reading of the situation is to separate two dynamics. The first is that of the stationary solutions of the augmented system, defined for all $t \in \mathbb{R}$, with memory

$$z(\omega, t) = \int_{-\infty}^t e^{-\omega(t-s)} f(x(s)) \, ds, \quad x(t) = a + \int_{-\infty}^t k_\nu(t-s) f(x(s)) \, ds, \quad (12)$$

that is, the solutions of the equation with the Liouville-Weyl derivative, which is invariant under time translation. For these the low-rate content is bounded whenever $f(x(s))$ has zero mean with sufficient cancellation, and it is on these that a compact invariant set makes sense. The second is the Caputo orbit, which starts from zero memory. The relation between the two is a classical identity: the Caputo and Riemann-Liouville derivatives differ by the initial term,

$${}^{RL}D_t^\nu x = {}^CD_t^\nu x + \frac{x_0 t^{-\nu}}{\Gamma(1-\nu)}, \quad (13)$$

so the Caputo problem is the problem with memory from 0 plus a non-autonomous forcing that decays as $t^{-\nu}$ and whose prefactor $1/\Gamma(1-\nu) \simeq (1-\nu)$ vanishes as $\nu \rightarrow 1$. The Caputo orbit is thus a non-autonomous, algebraically decaying perturbation of the stationary dynamics. We shall need the following assumption, which we state as such because for chaotic systems it is not proven.

Hypothesis 3 (invariant set). The augmented system (9) possesses a compact invariant set $\mathcal{A} \subset \mathcal{X}$ made of stationary solutions of the form (12) bounded in $\mathcal{X}$, and the Caputo orbit $x(t)$ follows its observable component in the sense that $\text{dist}(x(t), \pi_x \mathcal{A}) = O(t^{-\nu})$ as $t \rightarrow \infty$, where π_x denotes the projection on the observable $x = a + \langle \mu_\nu, z \rangle$.

The hypothesis has two parts of different nature. The existence of $\mathcal{A}$ requires that the Liouville-Weyl system admit solutions bounded on the whole real line, that is, that the “velocity” $f(x(s))$ along the attractor have cancellations sufficient for the fractional integral from $-\infty$ to converge; this is plausible, not proven. The rate $t^{-\nu}$ of approach is exact in the linear case (Proposition 9 below) and is what (13) suggests in general: if the invariant set is stable in the sense of shadowing, a forcing of amplitude $O(t^{-\nu})$ produces a deviation $O(t^{-\nu})$.

3. Lyapunov exponents

3.1. Linearisation and the cocycle property

Along a solution $(a, z(t))$ of (9), with $x(t)$ the reconstructed trajectory and $J(t) := Df(x(t))$ the Jacobian matrix of the vector field along it, the linearised equation reads

$$\dot{\delta a} = 0, \quad \partial_t \delta z = -\omega \delta z + J(t) \delta x, \quad \delta x = \delta a + \langle \mu_\nu, \delta z \rangle, \quad (14)$$

where $(\delta a, \delta z) \in \mathcal{X}$ is a tangent vector and $\delta x \in \mathbb{R}^n$ the corresponding perturbation of the observable. Let $M(t, s) : \mathcal{X} \rightarrow \mathcal{X}$ be the operator that carries $(\delta a, \delta z(s))$ to $(\delta a, \delta z(t))$. Because (9) is autonomous and M is its linearisation, M satisfies the cocycle property $M(t, r) = M(t, s)M(s, r)$ for all $r \leq s \leq t$. This is exactly what the fractional variational equation fails to satisfy, and it is what makes it legitimate, from here on, to speak of Lyapunov exponents in the usual sense.

The relation with the fractional variational equation deserves to be spelled out, because it explains what existing codes compute. The problem ${}^C D^\nu \delta x = J(t) \delta x$, $\delta x(0) = \delta x_0$, is the restriction of (14) to tangent data of the form $(\delta x_0, 0)$, that is, to perturbations of the initial condition alone with zero tangent memory, read through $\delta x = \delta a + \langle \mu_\nu, \delta z \rangle$. Codes that integrate the fractional variational equation therefore follow an n -dimensional subspace of initial tangent data inside an infinite-dimensional tangent space, and when they renormalise by QR at time t_k they replace the tangent vector $(\delta a, \delta z(t_k))$ by one of the form $(\delta x(t_k), 0)$, erasing its memory. This is the precise description of the “impulsive algorithm with variable lower limit” of [12], and it explains why those codes do not compute the cocycle (14) but a surrogate in which the renormalisation interval becomes a physical parameter; we return to this in Section 4.

3.2. Structure of the evolution operator

By the variation-of-constants formula, the cocycle from time 0 splits as

$$\begin{aligned} M(t, 0) &= S(t) + K(t), & S(t)(\delta a, \delta z_0) &= (\delta a, e^{-\omega t} \delta z_0), \\ K(t)(\delta a, \delta z_0) &= \left(0, \int_0^t e^{-\omega(t-s)} J(s) \delta x(s) ds \right), \end{aligned} \quad (15)$$

where $\delta x(s)$ is the perturbation of the observable along the solution of (14) with the given datum.

Lemma 4 (structure of S and K).

1. $S(t)$ is bounded with $\|S(t)\| = 1$, it is not compact, and its spectrum on the $\mathcal{H}$ component is the interval $[0, 1]$.
2. $K(t)$ is not of finite rank, but it is compact on $\mathcal{X}$ for every $t > 0$ if J is bounded.

Proof. The first point is the structure of a multiplication operator by $e^{-\omega t}$, plus the identity on $\mathbb{R}^n$. For the second, the image of $K(t)$ is the set of functions $\omega \mapsto \int_0^t e^{-\omega(t-s)} g(s) ds$ with g ranging over an infinite-dimensional set of functions of s , so the rank is not finite. Compactness is proved in [Appendix B](#): $K(t)$ factors through the space of bounded continuous functions on $[0, t]$, the first factor being bounded by a Mittag-Leffler estimate on the linearised Volterra equation and the second being compact by a Kolmogorov-Riesz argument in the weighted space $\mathcal{H}$. $\square$

3.3. The multiplicative ergodic theorem and the definition

Oseledets' theorem has a version for cocycles on Hilbert spaces due to Ruelle [20], extended by Mañé [21] and Thieullen [22] to non-compact cocycles and in general form by Lian and Lu [28]. The hypothesis that replaces compactness is a condition on the Kuratowski measure of non-compactness $\alpha(\cdot)$ of the cocycle, which for a bounded operator is the infimum of the radii of finite coverings of the image of the unit ball. Defining the essential exponent

$$\kappa := \lim_{t \rightarrow \infty} \frac{1}{t} \ln \alpha(M(t, 0)), \quad (16)$$

the theorem states that, for an ergodic invariant probability measure ϱ of the flow on a compact invariant set, the Lyapunov exponents strictly larger than κ form a discrete set $\lambda_1 > \lambda_2 > \dots > \kappa$, finite or accumulating only at κ , each with finite multiplicity, together with the Oseledets filtration and the associated invariant manifolds; below κ the theorem guarantees no structure.

Proposition 5 (essential exponent). *For the cocycle (15) one has $\kappa \leq 0$.*

Proof. The Kuratowski measure is subadditive and vanishes on compact operators: $\alpha(M(t, 0)) \leq \alpha(S(t)) + \alpha(K(t)) = \alpha(S(t)) \leq \|S(t)\| = 1$ by Lemma 4. Hence $\kappa \leq 0$. $\square$

Definition 6 (Lyapunov exponents of a fractional-order system). Under Hypothesis 3, let ϱ be an ergodic invariant measure of the flow of (9) on $\mathcal{A}$; at least one exists by the Krylov-Bogoliubov theorem. The Lyapunov exponents of the fractional system (1) with respect to ϱ are the Lyapunov exponents of the cocycle (14) on $\mathcal{A}$, in the sense of Thieullen’s theorem, restricted to the invariant tangent subspace $\{\delta a = 0\}$ of perturbations tangent to $\mathcal{A}$.

The restriction to $\{\delta a = 0\}$ is discussed in Section 4.3; it does not affect the positive exponents.

Theorem 7 (existence of the positive exponents). *Under Hypothesis 3, the strictly positive Lyapunov exponents of the fractional system are well defined, form a discrete set with finite multiplicities, and do not depend on any truncation of the memory, being defined in the exact system. In particular, the statement “the system is chaotic”, meaning $\lambda_1 > 0$, does not depend on the numerical method.*

The theorem is the combination of Proposition 5 with Thieullen’s theorem, and its content should be stated with the corresponding caution: it provides a definition and an existence result, not an algorithm nor a value, and it rests on a hypothesis that for chaotic systems is not proven. The exponents depend on ν in an essential way; it is their definition that does not depend on the truncation.

What remains is to connect the definition, given on $\mathcal{A}$, with what a computation measures, namely the Caputo orbit from $(x_0, 0)$, which by Proposition 2 does not lie on $\mathcal{A}$ and does not converge to it in norm.

Conjecture 8 (the Caputo orbit measures the exponents of $\mathcal{A}$). *The finite-time exponents computed along the Caputo orbit, $\lambda_j(t) = t^{-1} \ln$ of the growth factors of the cocycle (14) with QR renormalisation in $\mathcal{X}$, converge as $t \rightarrow \infty$ to the exponents $\lambda_j > 0$ of $\mathcal{A}$ with respect to the physical measure, with algebraic corrections whose form is given in Section 7.*

The basis of the conjecture is (13): the Caputo orbit is the stationary orbit plus a forcing $O(t^{-\nu})$, and the local expansion rate, a regular function of the observable state, is affected by $O(t^{-\nu})$. In the linear case the conjecture is a theorem (Proposition 9). In the chaotic case it has a direct numerical check, which we perform in Section 6: compare the exponents obtained starting from zero memory with those obtained from a delayed average and from a restart at a later state.

4. Structure of the spectrum

4.1. The zero exponent exists

The augmented system is autonomous. Along a solution that is bounded in the observable and does not converge to an equilibrium, the time derivative $v(t) := (0, \partial_t z(\cdot, t))$ is a tangent vector that solves (14), as one checks by differentiating (9) with respect to t and using $\dot{x} = \langle \mu_\nu, \partial_t z \rangle$. Its norm is bounded above and below also along the Caputo orbit: from $\partial_t z = -\omega z + f(x)$ and Proposition 2, for $\omega \ll 1/t$ one has $\omega z = O(t^{-\nu})$ and hence $\partial_t z \simeq f(x(t))$, a constant in ω that is integrable against ρ_ν near the origin, while for large ω one has $\partial_t z = O(\omega^{-1})$. Ageing affects the state, not its time derivative. The Lyapunov exponent of v is therefore exactly zero, both on $\mathcal{A}$ and along the Caputo orbit.

Danca, Fečkan, Kuznetsov and Chen [17] observed numerically, and declared not proven, that fractional-order systems have no zero exponent. The two statements do not contradict each other: they refer to two different definitions. The time shift of a solution is a solution of the augmented system with a different initial memory, and it is along this direction that the zero lives; it is not a solution of the fractional variational equation with zero memory, which is all that fractional codes follow. The absence of the zero exponent is a property of the restriction to the data $(\delta x_0, 0)$, not of the system.

4.2. No exponential contraction: the linear case is exact

Proposition 9. *Let $f(x) = \Lambda x$ with Λ diagonalisable. For every eigenvalue λ of Λ with $|\arg \lambda| > \nu\pi/2$, the stable sector in the sense of Matignon [29], the perturbation of the initial datum along the eigenvector decays algebraically,*

$$\delta x(t) = E_\nu(\lambda t^\nu) \delta x_0 \simeq -\frac{\delta x_0}{\lambda \Gamma(1-\nu)} t^{-\nu} \quad (t \rightarrow \infty),$$

where $E_\nu(w) = \sum_{k \geq 0} w^k / \Gamma(k\nu + 1)$ is the Mittag-Leffler function, and its Lyapunov exponent is zero. For every real positive eigenvalue, $\delta x(t) \simeq \nu^{-1} \exp(\lambda^{1/\nu} t) \delta x_0$ and the exponent is $\lambda^{1/\nu} > 0$, with finite-time exponent $\lambda(t) = \lambda^{1/\nu} - t^{-1} \ln \nu + o(t^{-1})$.

Proof. The asymptotic expansions of the Mittag-Leffler function are classical [5]: in the stable sector $E_\nu(w) = -\sum_{k \geq 1} w^{-k} / \Gamma(1 - k\nu) + O(|w|^{-K})$, in the unstable sector $E_\nu(w) = \nu^{-1} \exp(w^{1/\nu}) - \sum_{k \geq 1} w^{-k} / \Gamma(1 - k\nu)$. Algebraic decay gives $t^{-1} \ln |\delta x(t)| \rightarrow 0$; exponential growth gives the rate $\lambda^{1/\nu}$, and

the algebraic correction to the logarithm is exponentially small compared with the leading term, so the finite-time exponent converges as $1/t$ and not as $t^{-\nu}$. The exponent in $\mathcal{X}$ coincides with that of the observable because the norm of the tangent vector $(\delta x_0, \delta z(t))$ lies between $|\delta x_0|$ and a constant times $\sup |\delta x|$. $\square$

The proposition says something that the numerical literature contradicts systematically: in a Caputo system exponential contraction is impossible, because the memory modes with $\omega \rightarrow 0$ release the perturbation at an algebraic rate. The negative exponents reported by all codes, for instance -2.60 for order 0.995 in [12], are finite-time exponents, and the proposition predicts their form: since $E_\nu(-at^\nu)$ behaves exponentially up to a crossover time $t_*(a, \nu)$ and algebraically afterwards, the finite-time exponent along a stable direction is

$$\lambda(t) \simeq \frac{1}{t} \ln \frac{t^{-\nu}}{a \Gamma(1-\nu)} = -\nu \frac{\ln t}{t} - \frac{\ln(a \Gamma(1-\nu))}{t} \quad (t \gg t_*), \quad (17)$$

that is, it tends to zero from below as $-\nu \ln t/t$. For ν close to 1 the crossover time grows without bound, because the Mittag-Leffler function tends to the exponential, which is why at order 0.995 negative exponents look stable: the drift is there, beyond the explored horizon.

4.3. The tangent subspace, and what a negative exponent is

Two points become evident as soon as one computes. The first concerns the component δa of a tangent vector, which is conserved by (14): in the norm of $\mathcal{X}$ every vector with $\delta a \neq 0$ has exponent ≥ 0 by construction. Perturbations with $\delta a \neq 0$ are perturbations of the Caputo initial condition, that is, displacements towards a system of a different age; those tangent to the invariant set $\mathcal{A}$, on which the exponents are defined, have $\delta a = 0$. Definition 6 therefore refers to the invariant subspace $\{\delta a = 0\}$, in which the flow direction also lives; the maximal exponent is the same in both subspaces, because the unstable direction crosses both.

The second point is more substantial. In the state space $\mathcal{X}$ the non-positive part of the spectrum is, for every $\nu < 1$, the essential band at 0^- : a tangent vector whose observable δx contracts exponentially still has components on the slow memory modes, which decay as $e^{-\omega t}$ with $\omega \rightarrow 0$, and its $\mathcal{H}$ -norm does not contract exponentially. The negative exponents of the integer-order system, such as the -14.6 of the Lorenz system, are therefore not Lyapunov

exponents of the augmented system for any $\nu < 1$: they are contraction rates of the observable, and they reappear only in the singular limit $\nu \rightarrow 1$, in which the memory band disappears together with its weight $\sin(\pi\nu)$. We verified this directly: the Lorenz system at $\nu = 0.99999$ in the subspace $\{\delta a = 0\}$ gives $(0.95, -0.005, -0.005)$, not $(0.91, 0, -14.6)$. The consequence for reading fractional codes is clear-cut. What they measure as negative exponents are rates of the observable over the scale of the renormalisation interval: if the QR renormalisation resets the history every Δ time units, each segment restarts in the exponential phase of the Mittag-Leffler function, and the method measures the short-time rate on scale Δ , a reproducible and physically meaningful number, but one that depends on Δ and is a Lyapunov exponent of the system only in the limit $\Delta \rightarrow \infty$, where it tends to zero. Fractional codes thus measure memory-reset exponents with Δ as a hidden parameter. The Benettin algorithm on the augmented system, in turn, returns after the positive exponents the essential band and nothing else: the zero of the flow is immersed in it and cannot be separated by value, only by direction. The pseudo-exponents of [14] are this band seen with a few modes.

Conjecture 10. *For a chaotic fractional-order system satisfying Hypothesis 3, all Lyapunov exponents of the cocycle (14) other than the strictly positive ones are zero. The negative exponents measured by numerical methods are finite-time exponents drifting to zero according to a law of the form (17) or, for truncated memory, equal to minus the slowest retained rate.*

Thieullen’s theorem says nothing below $\kappa = 0$, and Proposition 9 covers only the linear case, so this is stated as a conjecture. Its support is structural: below κ every tangent vector has a component on the slow modes that cannot contract faster than $e^{-\omega t}$ with $\omega \rightarrow 0$, and the compact part K can redistribute the perturbation but cannot cancel the low-rate tail. The numerical test is in Section 6.

5. Numerical method and validation

5.1. Discretisation of the memory continuum

The continuum of modes is replaced by a finite set of J rates ω_i with weights w_i such that

$$k_\nu(\tau) \approx \sum_{i=1}^J w_i e^{-\omega_i \tau} \quad \text{for } \tau \in [h, T], \quad (18)$$

where h is the time step and T the integration horizon. This is the sum-of-exponentials approximation of the Caputo kernel developed for fast fractional integration [23, 24, 25, 26], which reduces the cost of a step from $O(N)$ to $O(J)$, N being the number of steps, with $J = O(\log T \cdot \log \varepsilon^{-1})$ for a relative accuracy ε . Our quadrature of the measure $\mu_\nu(\omega) d\omega$ uses a Gauss-Jacobi rule with weight $\omega^{-\nu}$ on the panel $[0, \omega_{\min}]$, which treats the singularity at the origin exactly, and Gauss-Legendre rules in the variable $\ln \omega$ on geometric panels, one per decade, from $\omega_{\min}$ to $\omega_{\max}$, with eight nodes per panel. The lower edge $\omega_{\min}$ is not a truncation of memory, since the Jacobi panel extends to $\omega = 0$, but it fixes the resolution with which the slow band is represented; we shall vary it for exactly this reason. The upper edge is chosen as $\omega_{\max} = 10^4/h$, a choice that deserves a comment: the modes with $\omega > 1/h$ represent the instantaneous part of the kernel within one step, and the missing tail beyond $\omega_{\max}$ weighs as $(\omega_{\max}h)^{-\nu}$; with $\omega_{\max} = 20/h$ at $\nu = 0.5$ this tail was twelve per cent of the instantaneous part and biased the maximal exponent by two per cent. With J between 88 and 112 the relative error of (18) is below 10^{-8} on all of $[h, T]$.

To test the effect of a truncation of memory we also use a variant in which the Jacobi panel is removed: no mode slower than $\omega_{\min}$ is retained, so no memory survives beyond $\tau_m = 1/\omega_{\min}$. This is the truncation implicit in every finite-mode realisation with a smallest rate $\omega_{\min}$, as in [14], and it should not be confused with the short-memory principle of [5], which acts on the derivative. Removing the slow modes subtracts from the kernel the quantity $\int_0^{\omega_{\min}} \mu_\nu(\omega) e^{-\omega\tau} d\omega$, which for $\tau \ll \tau_m$ is the constant $\sin(\pi\nu) \omega_{\min}^{1-\nu} / (\pi(1-\nu))$: because the mass of μ_ν below $\omega_{\min}$ decays only as $\omega_{\min}^{1-\nu}$, the slow modes carry a large share of the kernel at *all* lags, and the ratio of the subtracted part to $k_\nu(\tau)$ is $(\omega_{\min}\tau)^{1-\nu} / \Gamma(2-\nu)$, which at $\nu = 0.9$ and $\tau_m = 10^3$ is already one third at lag 10^{-2} and one half at lag one. The truncated system is therefore not the fractional system with a shorter memory; it is a different system at every time scale.

5.2. Truncated-memory codes in the original formulation

To compare with the classical practice we also integrate the Chen system in its original formulation with the two truncations found in the literature, using in both cases the tangent system obtained by exact linearisation of the discrete scheme, with the same window, and rescaling the whole stored history of the tangent vector at each renormalisation, which is legitimate because the tangent scheme is linear. The first is the Grünwald-Letnikov form, on which

the short-memory principle is stated,

$$h^{-\nu} \sum_{j=0}^L c_j^{(\nu)} (x_{n-j} - x_0) = f(x_{n-1}), \quad c_0^{(\nu)} = 1, \quad c_j^{(\nu)} = \left(1 - \frac{\nu+1}{j}\right) c_{j-1}^{(\nu)}, \quad (19)$$

with $L = \tau_m/h$ terms retained; the weights decay as $j^{-\nu-1}$, are summable, and the neglected tail weighs $O(\tau_m^{-\nu})$, which is Podlubny's bound. The second is the Volterra integral form with product-rectangle weights,

$$x_n = x_0 + \frac{h^\nu}{\Gamma(\nu+1)} \sum_{j=n-L}^{n-1} [(n-j)^\nu - (n-j-1)^\nu] f(x_j), \quad (20)$$

which is the explicit fractional Euler scheme with the sum restricted to the last L steps, the “fixed memory” of [15, 16]; here the weights decay as $j^{\nu-1}$, are not summable, and the neglected part of the sum is comparable with the retained one. Both schemes are of first order and are run with $h = 0.002$, the largest step at which the explicit Euler scheme is stable on the Chen system, over $T = 1500$ time units, and with $h = 0.001$ over $T = 1000$, the first hundred time units being discarded in both cases; the uncertainty is the standard deviation of ten block averages. At these steps the schemes are not yet in the asymptotic regime: halving h moves the full-memory exponent by three per cent, and both values differ by ten to fifteen per cent from the ETD-RK2 result of Table 1. The difference should not surprise: the first-order error acts on the kernel like a small shift of the effective order, and the maximal exponent of the Chen system is a steep and non-monotonic function of ν in this range, as Section 6 shows. Comparisons are therefore made within each scheme, and we report both steps.

5.3. Integration and QR algorithm

The discretised state and tangent equations are

$$\begin{aligned} \dot{z}_i &= -\omega_i z_i + f(x), & x &= x_0 + \sum_{i=1}^J w_i z_i, \\ \dot{\delta z}_i &= -\omega_i \delta z_i + J(x) \delta x, & \delta x &= \delta a + \sum_i w_i \delta z_i, \end{aligned} \quad (21)$$

with δa constant, so that the tangent space has dimension $n(J+1)$, or nJ in the subspace $\{\delta a = 0\}$. Since the modes are linear and diagonal,

their exact update over a step is $z_i(t+h) = e^{-\omega_i h} z_i(t) + \int_0^h e^{-\omega_i(h-s)} f \, ds$; we use the second-order exponential time-differencing Runge-Kutta scheme (ETD-RK2), which treats the forcing as linear within a step and avoids the stiffness of $\omega_{\max} \sim 1/h$ without implicit steps. Because the Caputo solution starts as t^ν , with a singular derivative at $t=0$ when the initial memory is zero, the first time unit is integrated with twenty sub-steps. The Benettin algorithm [1, 2] is then applied in the standard way: a set of k tangent vectors is propagated, re-orthonormalised by QR every $\Delta_{QR} = 0.5$ time units, and the exponents are the time averages of the logarithms of the diagonal elements of the triangular factors. The inner product used for the QR is that of $\mathcal{X}$ discretised, $\|(\delta a, \delta z)\|^2 = |\delta a|^2 + \sum_i w_i(1 + \omega_i) |\delta z_i|^2$, implemented by rescaling the tangent coordinates; exponents do not depend on the norm among equivalent norms, but at finite J the choice affects how fast they settle. We follow $k = 5$ exponents. The code is a few hundred lines of Python with NumPy and is available from the author.

5.4. Validation on the linear case

The linear system $f(x) = \Lambda x$ with $\Lambda = \text{diag}(1, -2)$ has an exact solution in terms of the Mittag-Leffler function and exact exponents by Proposition 9: $\lambda_1 = 1^{1/\nu} = 1$ along the unstable direction, and an exponent that drifts to zero as $-\nu \ln t/t$ along the stable one. Figure 1 shows the finite-time exponents at $\nu = 0.7$ with $h = 0.005$ and $T = 60$. The average from $t = 0$ carries the $1/t$ transient of the random initial tangent vector, of size $\ln(\text{dim})/2T$; the average from $T/3$ converges to $\lambda_1 = 0.9999$. At $\nu = 0.5$ and $\nu = 0.995$ the same protocol gives $\lambda_1 = 0.9994$ and 1.0000 . The trajectory itself agrees with the Mittag-Leffler function to within $5 \cdot 10^{-3}$ on $[0, 6]$ at $\nu = 0.5$ and $5 \cdot 10^{-4}$ at $\nu = 0.7$ (beyond $t \simeq 8$ the power series of E_ν at negative argument loses precision by cancellation and cannot serve as a reference). The stable direction gives an exponent of order -10^{-2} drifting to zero, where a code with memory reset would measure $-2^{1/\nu}$.

6. Results on chaotic systems

6.1. Systems and protocol

We use the fractional Chen system [7],

$${}^C D^\nu x = a(y - x), \quad {}^C D^\nu y = (c - a)x - xz + cy, \quad {}^C D^\nu z = xy - bz, \quad (22)$$

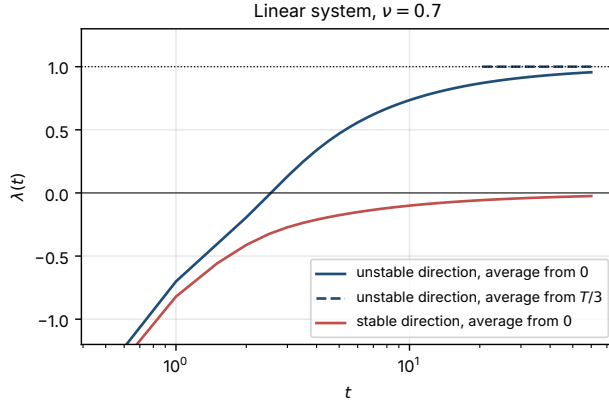


Figure 1: Validation on the linear system $\Lambda = \text{diag}(1, -2)$ at $\nu = 0.7$. The finite-time exponent along the unstable direction converges to the exact value $\lambda^{1/\nu} = 1$ (dotted line) once the transient of the initial tangent vector is discarded; along the stable direction the exponent is of order -10^{-2} and drifts to zero, in accordance with Proposition 9.

with $a = 35$, $b = 3$, $c = 28$, at orders $\nu = 0.9$ and 0.85 , and the fractional Rössler system in the form of Li and Chen [6],

$${}^C D^\nu x = -y - z, \quad {}^C D^\nu y = x + ay, \quad {}^C D^\nu z = b + z(x - c), \quad (23)$$

with $a = 0.5$, $b = 0.2$, $c = 10$, at $\nu = 0.9$. The Lorenz system with canonical parameters converges to an equilibrium for $\nu \leq 0.99$ in our integration and is chaotic at 0.995 ; this is what Matignon’s stability condition [29] predicts, since the eigenvalues $0.094 \pm 10.19i$ of the equilibria $C_\pm$ have argument 0.497π , so that $C_\pm$ become asymptotically stable for $\nu < 0.994$. At 0.995 the memory effects are too small to be instructive; we use the Lorenz system only for the limit check at $\nu = 0.99999$ of Section 4.3. The horizon is $T = 2000$ for Chen ($h = 0.002$) and $T = 4000$ for Rössler ($h = 0.005$); the initial conditions are $(-9, -5, 14)$ and $(1, 1, 0.1)$. Three estimates of each exponent are compared: the “cold” average from $t = 0$ with zero initial memory, the “delayed” average from $T/4$ along the same trajectory, and, for the Chen system, a “restart” in which a second cold run is started from the state $x(T/4)$ with zero memory. The uncertainty on λ_1 is the standard deviation of block averages over 250 time units. For the exponents $j \geq 2$ we also report the exponent c of the algebraic decay of the norms of the Benettin vectors, obtained by fitting $\lambda_j(t) t = -c \ln t + \text{const}$ on the last nine tenths of the run.

system, ν	memory	J	λ_1	$\langle \lambda_{2..5}(T) \rangle$	c
Chen 0.9	truncated, $\tau_m = 10^2$	72	-0.012 ± 0.001 (equilibrium)	$-1.3 \cdot 10^{-2}$	–
Chen 0.9	truncated, $\tau_m = 10^3$	80	1.165 ± 0.005	$-1.7 \cdot 10^{-3}$	0.88
Chen 0.9	full, $\omega_{\min} = 5 \cdot 10^{-4}$	88	2.254 ± 0.023	$-7.0 \cdot 10^{-4}$	0.26
Chen 0.9	full, $\omega_{\min} = 10^{-4}$	96	2.269 ± 0.014	$-4.1 \cdot 10^{-4}$	0.16
Chen 0.9	full, $\omega_{\min} = 10^{-6}$	112	2.247 ± 0.018	$-1.8 \cdot 10^{-4}$	0.06
Chen 0.85	full, $\omega_{\min} = 10^{-6}$	112	2.414 ± 0.008	$-1.0 \cdot 10^{-4}$	0.06
Rössler 0.9	full, $\omega_{\min} = 10^{-6}$	112	0.066 ± 0.003	$-1.2 \cdot 10^{-4}$	0.07

Table 1: Lyapunov exponents in the subspace $\{\delta a = 0\}$. “Full” means that the singular panel extends to $\omega = 0$, $\omega_{\min}$ fixing the resolution of the slow band; “truncated” means no mode below $\omega_{\min}$, that is, a memory window $\tau_m = 1/\omega_{\min}$. The exponent c is the algebraic decay of the norms of the Benettin vectors from the second onwards.

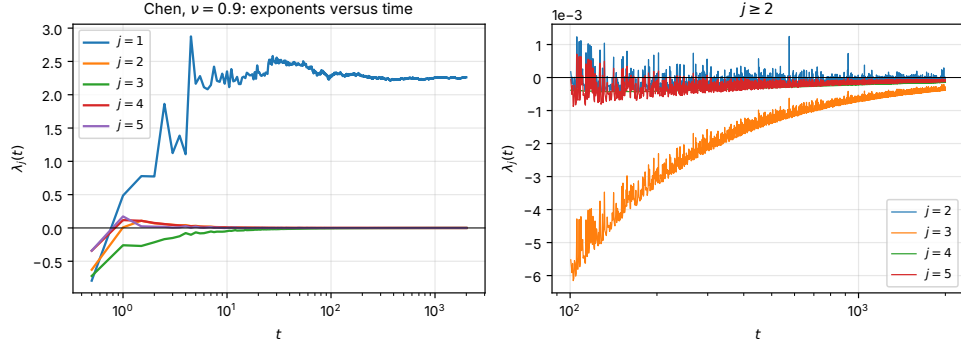


Figure 2: Fractional Chen system at $\nu = 0.9$, full memory with $\omega_{\min} = 10^{-6}$. Left: the five finite-time exponents along the cold run; the maximal one settles at 2.25, the others go to zero. Right: the exponents $j \geq 2$ on an enlarged scale for $t > 100$; they settle within $-4 \cdot 10^{-4}$ of zero by the end of the run and drift slowly towards it.

6.2. Maximal exponent and essential band

Table 1 and Figure 2 summarise the results. Three facts emerge. The maximal exponent is a property of the system: in the three full-memory runs of the Chen system at $\nu = 0.9$ it is 2.26 ± 0.02 independently of $\omega_{\min}$; at $\nu = 0.999$ the same protocol gives 2.01, the value of the integer-order Chen system (2.03), and a scan in ν shows a steep, non-monotonic dependence, with a periodic window at 0.99 and values of 2.13, 2.41, 2.63 at 0.98, 0.97, 0.95 before the 2.25 at 0.9 and the 2.41 at 0.85; and it coincides between cold, delayed and restart averages within the ergodic fluctuations (for the run with $\omega_{\min} = 5 \cdot 10^{-4}$: 2.254, 2.266 and 2.28 respectively). All other exponents are compatible with zero, and their negative residual is a resolution artefact: at

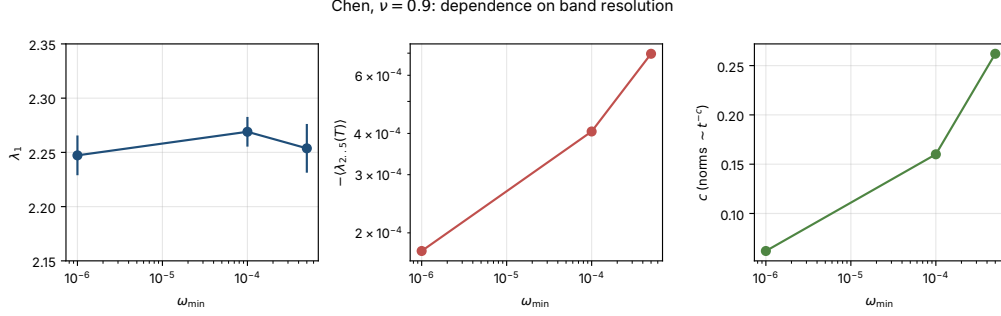


Figure 3: Chen system at $\nu = 0.9$, full memory: as the resolution of the slow band increases (smaller $\omega_{\min}$), the maximal exponent does not change, while the non-positive exponents and the algebraic decay of the norms go to zero.

equal memory retained, extending the resolved panel from $5 \cdot 10^{-4}$ to 10^{-6} brings the residual from $7 \cdot 10^{-4}$ down to $1.8 \cdot 10^{-4}$ and the decay of the norms from $t^{-0.26}$ to $t^{-0.06}$, monotonically and slowly, roughly logarithmically in $\omega_{\min}$, as one expects from an average over a continuum of modes with weight $\omega^{-\nu}$ and not from a single slow mode (Figure 3). The residual and its decay are the same, at equal quadrature, in the three systems and at the two orders: they are properties of the discretisation of the memory continuum, not of the systems. And no exponential contraction is visible in any case, whereas the corresponding integer-order systems have negative exponents of order -10 .

6.3. What memory truncation measures

The control with the slow modes removed gives the sharpest result of the augmented-system computations (Figure 4). Without modes below $\omega_{\min}$ the non-positive exponents of the Chen system settle in the immediate vicinity of $-\omega_{\min}$: $-1.7 \cdot 10^{-3}$ for $\tau_m = 10^3$, $-1.3 \cdot 10^{-2}$ for $\tau_m = 10^2$. This is the predicted signature: the slowest retained mode fixes the apparent contraction rate, and what finite-mode codes report as negative exponents is the inverse of their slowest rate. But the effect does not stop at the non-positive exponents. With $\tau_m = 10^3$ the maximal exponent drops from 2.25 to 1.17, and with $\tau_m = 10^2$ chaos disappears: the trajectory converges to the equilibrium and all exponents are negative. This is what the estimate of Section 5 predicts: removing the slow modes alters the kernel at every lag, and the result is a different, less chaotic system.

The natural question is whether the same happens with the truncations used in practice, and the answer separates them (Table 2, Figure 5). In the

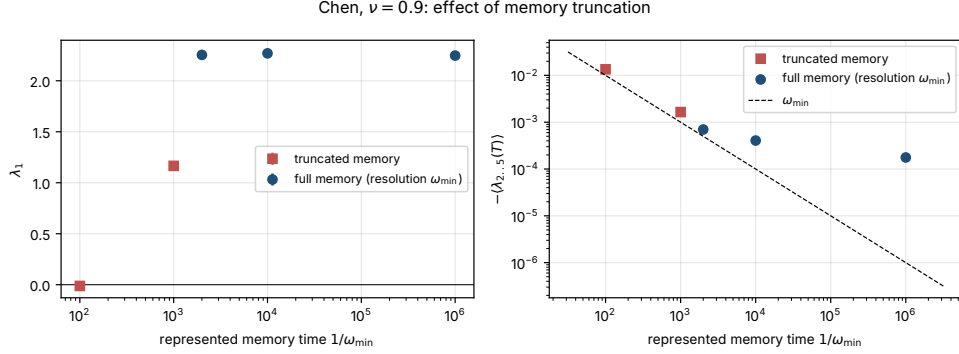


Figure 4: Chen system at $\nu = 0.9$, augmented system: maximal exponent (left) and non-positive exponents (right) versus the represented memory time. With the slow modes removed (squares) the non-positive exponents fall on the line $-\omega_{\min}$ and the maximal exponent drops from 2.25 to 1.17 and then below zero; with full memory (circles) the maximal exponent is invariant and the negative residual decreases with resolution.

Grünwald-Letnikov form (19) the exponent is 2.488 ± 0.018 with full memory, 2.468 ± 0.021 with $\tau_m = 500$ and 2.483 ± 0.024 with $\tau_m = 100$ at $h = 0.002$, and 2.562 ± 0.029 , 2.583 ± 0.017 , 2.566 ± 0.028 at $h = 0.001$; the attractor has the same extent in all runs: the short-memory principle, applied to the derivative as Podlubny states it, leaves exponent and geometry unchanged even with a window of a hundred time units, which for the Chen system is a few hundred oscillations. In the integral form (20) the exponent with full memory is 2.467 ± 0.018 at $h = 0.002$ and 2.564 ± 0.027 at $h = 0.001$, in agreement with the Grünwald-Letnikov values as it must be; with $\tau_m = 500$ it falls to 1.85 ± 0.14 and 2.07 ± 0.17 , with an uncertainty seven times larger, and the attractor swells, its extent in z passing from $[6, 40]$ to $[1, 63]$ and $[3, 51]$; with $\tau_m = 100$ the trajectory escapes to infinity, after about five hundred time units at the larger step and eight hundred at the smaller, and no exponent can be defined. The three truncations thus fall into two families. Truncating the derivative cuts a summable kernel and is harmless; truncating the integral, whether by fixing the memory of a Volterra sum or by removing the slow modes of a realisation, cuts a non-summable kernel and changes the system, in one case towards an equilibrium and in the other towards infinity; in the finite-mode realisation the change carries the spurious rate $-1/\tau_m$ as its signature. This has a direct consequence for the reading of the literature: values obtained with fixed memory in the integral form, or with finite-mode realisations, are comparable neither with one another nor with ours unless the

scheme	h	full memory	$\tau_m = 500$	$\tau_m = 100$
Grünwald-Letnikov form (19)	0.002	2.488 ± 0.018	2.468 ± 0.021	2.483 ± 0.024
	0.001	2.562 ± 0.029	2.583 ± 0.017	2.566 ± 0.028
Volterra integral form (20)	0.002	2.467 ± 0.018	1.85 ± 0.14 (attractor deformed)	unbounded at $t \simeq 500$
	0.001	2.564 ± 0.027	2.07 ± 0.17 (attractor deformed)	unbounded at $t \simeq 800$
augmented system, slow modes removed	0.002	2.254 ± 0.023	1.165 ± 0.005 ($\tau_m = 10^3$)	equilibrium ($\tau_m = 10^2$)

Table 2: Maximal Lyapunov exponent of the Chen system at $\nu = 0.9$ under the three truncations of memory. The first two schemes are the first-order ones of Section 5.2, at two steps ($T = 1500$ and 1000); the third is the ETD-RK2 scheme of Section 5 with the windows of Table 1. The first-order values are not converged in h and comparisons are meaningful within a row.

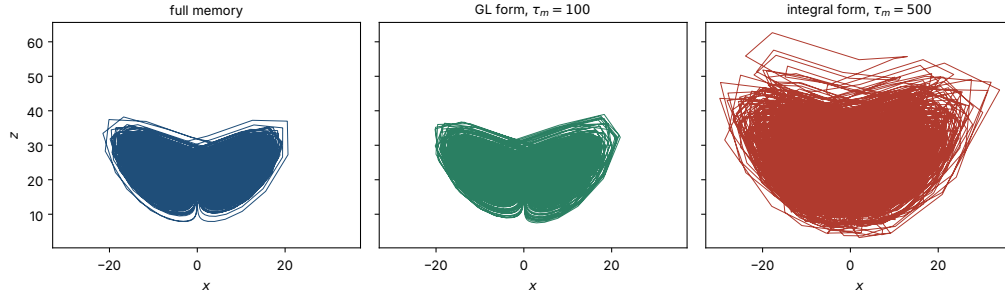


Figure 5: Chen system at $\nu = 0.9$, projection on the (x, z) plane of the last three hundred time units: full memory (left), Grünwald-Letnikov form with $\tau_m = 100$ (centre), Volterra integral form with $\tau_m = 500$ (right). Truncating the derivative leaves the attractor unchanged; truncating the integral deforms it before destroying it.

window is declared, and such truncations can shift the threshold of chaos in ν ; values obtained with the short-memory principle on the Grünwald-Letnikov derivative are not affected.

7. Discussion

7.1. On the rate of convergence of finite-time exponents

The dependence of numerical estimates on the horizon has an explanation within the present framework, which we state as a conjecture because we could not test it. The finite-time exponent along the Caputo orbit is the average over $[0, t]$ of the local expansion rate $\ell(x(s))$, a regular function of the observable state. Its distance from the limit has two sources. The first is ordinary: the trajectory is a finite sample of the measure ϱ , and if the correlations of ℓ decay fast enough the error has a systematic part $O(1/t)$ and

fluctuations $O(t^{-1/2})$ by the central limit theorem. The second is specifically fractional: by (13) the Caputo orbit is at distance $O(s^{-\nu})$ from $\mathcal{A}$ at time s , so $\ell(x(s))$ deviates by $O(s^{-\nu})$ from its value on $\mathcal{A}$, and the average of this deviation over $[0, t]$ is $t^{-1} \int^t s^{-\nu} ds \propto t^{-\nu}/(1-\nu)$. The prefactor inherited from (13) is $1/\Gamma(1-\nu)$, and the product $1/[(1-\nu)\Gamma(1-\nu)] = 1/\Gamma(2-\nu)$ is finite and tends to one as $\nu \rightarrow 1$: the memory term joins continuously the ordinary C/t term.

Conjecture 11 (rate of convergence). *For the QR method on the augmented system, starting from zero memory,*

$$\lambda_1(t) - \lambda_1 = \frac{B}{\Gamma(2-\nu)} t^{-\nu} + \frac{C}{t} + O_{\text{fluct}}(t^{-1/2}), \quad (24)$$

with B proportional to $|x_0 - \bar{x}|$ and to the sensitivity of ℓ to the state; starting from a state on $\mathcal{A}$ the term in $t^{-\nu}$ is absent.

The form (24) passes two checks: the limit $\nu \rightarrow 1$, and the linear case of Proposition 9, where convergence along the unstable direction is exactly $1/t$ with no algebraic term. It also has a consequence that explains why we could not test it: the memory term dominates the fluctuations at long times only if $\nu < 1/2$, because only then does $t^{-\nu}$ decay more slowly than $t^{-1/2}$. At the orders $\nu = 0.85\text{--}0.9$ at which the canonical systems are chaotic, the term is hidden under the ergodic noise, and indeed the differences between cold, delayed and restart averages in our runs are within the error. Testing the conjecture requires chaotic systems at $\nu \leq 1/2$, which the canonical oscillators are not, or an ensemble average over many initial conditions at fixed ν , which we leave for future work.

7.2. Comparison with existing methods

The framework gives a reading of each family of methods. The semi-analytical method of [9] and the predictor-corrector codes of [10, 11] follow the subspace $(\delta x_0, 0)$ and reset the memory at each renormalisation; their maximal exponent is a memory-reset estimate that should converge to ours as the renormalisation interval grows, while their negative exponents are observable rates on the scale of that interval. The Grünwald-Letnikov method of [13], which retains the history of the perturbation through the Grünwald-Letnikov weights, is the closest to the cocycle (14) and we expect its positive exponents to agree with ours; its negative exponents should drift to zero with

the horizon according to (17). The augmented-system computations of [14] are, in our terms, the correct object computed with a coarse quadrature; their pseudo-exponents are the essential band of Section 4.3, and Figure 3 shows how they move with the resolution. A systematic comparison of the four methods on the same system, with declared memory windows, would be the natural next step; the present work provides the reference against which to make it.

7.3. *Practical consequences*

The results have a direct bearing on how fractional-order models are used in practice. Models with Caputo derivatives are by now common in viscoelasticity, in the modelling of batteries and supercapacitors, in biology and in circuits containing constant-phase elements, and in several of these fields one wants to know whether the model is chaotic, how predictable it is and how sensitive to initial conditions; the answer is routinely sought by computing Lyapunov exponents with the available codes, almost always with some truncation of memory to keep the cost of long integrations affordable. Three things follow for this practice.

The first is that the positive exponent is reliable only if the memory is either complete or truncated on the derivative. For the fractional Chen system a fixed memory of five hundred time units in the integral form deforms the attractor and lowers the maximal exponent by a quarter, a window of a hundred destroys it, and the removal of the slow modes in a finite-mode realisation halves the exponent at a thousand units and suppresses chaos at a hundred; bifurcation diagrams and chaos thresholds in the order ν obtained with such truncations may therefore describe a different system from the intended one, and the threshold order beyond which a model becomes chaotic, which in some applications is a design parameter, may be displaced. The short-memory principle applied to the Grünwald-Letnikov derivative, on the other hand, leaves the exponent unchanged within one per cent even at a hundred units. The remedy is cheap: truncate the derivative and not the integral, declare the window, or recompute with the augmented-system algorithm, whose cost is linear in the integration time.

The second is that the negative exponents reported by those codes are not properties of the system but of the memory treatment: in finite-mode realisations they equal the inverse of the slowest retained time, and in memory-reset codes they are contraction rates on the scale of the renormalisation interval. Whoever uses them to estimate a Lyapunov dimension, a rate

of loss of memory of the initial conditions, or to compare two models, is comparing two windows. Section 4.3 says what they actually measure, short-time contraction rates of the observable, and this allows them to be used for what they are.

The third is ageing. A fractional-order system never forgets its initial condition, and the memory of that point keeps growing in the slow modes for the whole life of the system, with the law $t^{(1-\nu)/2}$ in the memory norm and $t^{-\nu}$ in the observable. For whoever models materials or devices with memory this means that the initial transient never fully dies out and that two specimens started from different states remain distinguishable forever; for whoever identifies fractional models from experimental data it means that the data of an aged system are not those of a stationary one, and that the age of the system is a parameter of the measurement.

7.4. Limitations

The existence of the invariant set $\mathcal{A}$ for chaotic fractional systems (Hypothesis 3) is not proven, and the definition rests on it. The zero exponent of the flow, while present by construction, cannot be separated by value from the essential band with any horizon we tried, so its existence is a structural statement about the autonomous system rather than a measurement. The residual negative values of the band decrease only logarithmically with the resolution, so a computation that must report them should declare $\omega_{\min}$. Finally, the exponents we report are those of the memory state in the norm of $\mathcal{X}$; the contraction rates of the observable, which are what matters for some applications such as synchronisation, are a different quantity that the present framework identifies but does not compute.

8. Conclusions

We have given a definition of the Lyapunov exponents of a fractional-order system that does not depend on any truncation of memory, by rewriting the system through the diffusive representation as an autonomous evolution on a Hilbert space of memory modes and applying the multiplicative ergodic theorem in the form valid for asymptotically compact cocycles. On the way we found that the memory of a bounded Caputo trajectory is not bounded, growing as $t^{(1-\nu)/2}$: the system ages, and the exponents must be defined on the stationary dynamics of which the Caputo orbit is an algebraically decaying perturbation. The positive exponents are well defined and discrete;

the rest of the spectrum is an essential band at zero, which contains the zero exponent of the flow that fractional codes do not see and in which exponential contraction is impossible. The negative exponents of the integer-order system are rates of the observable and are not exponents of the fractional one for any order below one.

A Benettin algorithm on the augmented system, validated against the Mittag-Leffler function, confirms this structure on the fractional Chen and Rössler systems: the maximal exponent is reproducible to one per cent and insensitive to the resolution of the memory band, all other exponents vanish within a residual that decreases with resolution and is the same in all systems, and truncating the memory acts in two opposite ways. On the Grünwald-Letnikov derivative, as the short-memory principle prescribes, it leaves exponent and attractor unchanged; on the integral form of the equation, as fixed-memory Volterra codes and finite-mode realisations do, it deforms the attractor, lowers the maximal exponent by a quarter to a half and can suppress chaos or destroy the attractor, and in finite-mode realisations it adds spurious negative exponents equal to $-1/\tau_m$. The difference is the integrability of the kernel that is cut. The last fact concerns the practice of the field directly: results obtained by truncating the integral describe a different system, and they can be compared only if the window is declared. The rate at which finite-time estimates converge is predicted to carry a memory term in $t^{-\nu}$, visible only below order one half; its verification, and the proof of the existence of the invariant set for chaotic fractional systems, are the open questions that this work leaves.

Appendix A. Estimate of the memory norm

For f continuous and bounded by F along the trajectory, $|z(\omega, t)| \leq F \int_0^t e^{-\omega(t-s)} ds \leq F \min(t, \omega^{-1})$. The contribution of the rates $\omega > 1/t$ to the norm is therefore bounded by $F^2 \int_{1/t}^\infty \rho_\nu(\omega) \omega^{-2} d\omega$, which converges at infinity because the integrand decays as $\omega^{-1-\nu}$ and is bounded uniformly in t . The contribution of the rates $\omega < 1/t$ is governed by the effective value of z , which by Proposition 2 is $\simeq (\bar{x} - x_0)t^{1-\nu}/\Gamma(2-\nu)$ and not by the crude bound Ft : it is there that the boundedness of the orbit is decided, and the answer is negative unless $\bar{x} = x_0$. In the particular case of f with zero mean along the trajectory and sufficient cancellation, as for a stationary solution (12), the low-rate content remains $O(1)$ and the norm is bounded: this is the distinction between the Caputo orbit and the invariant set $\mathcal{A}$. The identity (13) is the

classical relation between the Riemann-Liouville and Caputo derivatives of a function with $x(0) = x_0$ [5].

Appendix B. Compactness of $K(t)$

We prove the second statement of Lemma 4: for every $t > 0$ and every bounded measurable $J : [0, t] \rightarrow \mathbb{R}^{n \times n}$ with $\sup_{[0, t]} \|J\| =: L < \infty$, the operator $K(t) : \mathcal{X} \rightarrow \mathcal{X}$ defined in (15) is compact. Throughout, $C([0, t]; \mathbb{R}^n)$ is the space of continuous functions with the supremum norm $\|g\|_\infty$.

Step 1: the observable perturbation solves a Volterra equation. Let $(\delta a, \delta z_0) \in \mathcal{X}$ and let $\delta z(s)$ be the mild solution of (14),

$$\delta z(s) = S(s)\delta z_0 + \int_0^s S(s-r) J(r) \delta x(r) dr,$$

where $S(\tau)$ acts on a constant vector $v \in \mathbb{R}^n$ by $(S(\tau)v)(\omega) = e^{-\omega\tau}v$. Applying the output functional and using the kernel identity (4) in the form $\langle \mu_\nu, S(\tau)v \rangle = v \int_0^\infty \mu_\nu(\omega) e^{-\omega\tau} d\omega = k_\nu(\tau) v$, valid for $\tau > 0$, we obtain that $\delta x(s) = \delta a + \langle \mu_\nu, \delta z(s) \rangle$ satisfies the linear Volterra equation

$$\delta x(s) = \varphi(s) + \int_0^s k_\nu(s-r) J(r) \delta x(r) dr, \quad \varphi(s) := \delta a + \langle \mu_\nu, S(s)\delta z_0 \rangle. \quad (\text{B.1})$$

The forcing φ is continuous on $[0, t]$, because $s \mapsto S(s)\delta z_0$ is continuous in $\mathcal{H}$ (strong continuity of the semigroup) and the output functional is bounded by (7); for the same reasons

$$\|\varphi\|_\infty \leq |\delta a| + \sup_s \|S(s)\delta z_0\|_{\mathcal{H}} \leq |\delta a| + \|\delta z_0\|_{\mathcal{H}} \leq \sqrt{2} \|(\delta a, \delta z_0)\|_{\mathcal{X}}. \quad (\text{B.2})$$

Step 2: the map from the datum to the observable perturbation is bounded. Let $\mathcal{T}$ be the Volterra operator $(\mathcal{T}g)(s) = \int_0^s k_\nu(s-r)J(r)g(r) dr$ on $C([0, t]; \mathbb{R}^n)$. Its iterates are controlled by the semigroup property of the Riemann-Liouville kernels, $\int_0^s k_\alpha(s-r)k_\beta(r) dr = k_{\alpha+\beta}(s)$: by induction,

$$|(\mathcal{T}^m g)(s)| \leq L^m \int_0^s k_{m\nu}(r) dr \|g\|_\infty = L^m \frac{s^{m\nu}}{\Gamma(m\nu + 1)} \|g\|_\infty.$$

Hence the Neumann series $\sum_{m \geq 0} \mathcal{T}^m$ converges in operator norm on $C([0, t]; \mathbb{R}^n)$ and

$$\|(I - \mathcal{T})^{-1}\| \leq \sum_{m \geq 0} \frac{(Lt^\nu)^m}{\Gamma(m\nu + 1)} = E_\nu(Lt^\nu), \quad (\text{B.3})$$

the Mittag-Leffler function of Proposition 9. Equation (B.1) therefore has a unique continuous solution $\delta x = (I - \mathcal{T})^{-1}\varphi$, and by (B.2)

$$\|\delta x\|_\infty \leq \sqrt{2} E_\nu(Lt^\nu) \|(\delta a, \delta z_0)\|_{\mathcal{X}}. \quad (\text{B.4})$$

We denote by $\Phi_t : \mathcal{X} \rightarrow C([0, t]; \mathbb{R}^n)$ the bounded linear map $(\delta a, \delta z_0) \mapsto J(\cdot) \delta x(\cdot)$; its norm is at most $\sqrt{2} L E_\nu(Lt^\nu)$.

Step 3: the Laplace-type map from continuous functions into the memory space is compact. Define $V_t : C([0, t]; \mathbb{R}^n) \rightarrow \mathcal{H}$ by $(V_t g)(\omega) = \int_0^t e^{-\omega(t-s)} g(s) ds$. For every g one has the pointwise bound

$$|(V_t g)(\omega)| \leq \|g\|_\infty \int_0^t e^{-\omega(t-s)} ds = \|g\|_\infty \frac{1 - e^{-\omega t}}{\omega} \leq \|g\|_\infty \min(t, \omega^{-1}), \quad (\text{B.5})$$

and $\int_0^\infty \rho_\nu(\omega) \min(t, \omega^{-1})^2 d\omega < \infty$, because $\rho_\nu \sim \omega^{-\nu}$ at the origin is integrable against t^2 , and $\rho_\nu \omega^{-2} \sim \omega^{-1-\nu}$ is integrable at infinity for $\nu > 0$. So V_t is bounded. To prove compactness, let B be the image of the unit ball of $C([0, t]; \mathbb{R}^n)$; we show that B is relatively compact in $\mathcal{H} = L^2(\rho_\nu d\omega)$ by the Kolmogorov-Riesz criterion for weighted L^2 spaces on the half-line, whose two conditions are uniform smallness of the tails and equicontinuity on compact subintervals. The tails are uniform by (B.5):

$$\begin{aligned} \sup_{u \in B} \int_R^\infty \rho_\nu |u|^2 d\omega &\leq \int_R^\infty \rho_\nu(\omega) \omega^{-2} d\omega \xrightarrow{R \rightarrow \infty} 0, \\ \sup_{u \in B} \int_0^\varepsilon \rho_\nu |u|^2 d\omega &\leq t^2 \int_0^\varepsilon \rho_\nu(\omega) d\omega \xrightarrow{\varepsilon \rightarrow 0} 0. \end{aligned}$$

On a compact interval $[\varepsilon, R]$ the functions of B are uniformly bounded, by (B.5), and uniformly Lipschitz, because differentiation under the integral sign gives

$$|\partial_\omega (V_t g)(\omega)| = \left| \int_0^t (t-s) e^{-\omega(t-s)} g(s) ds \right| \leq \|g\|_\infty \int_0^t (t-s) ds = \frac{t^2}{2} \|g\|_\infty;$$

by the Arzelà-Ascoli theorem B restricted to $[\varepsilon, R]$ is relatively compact in the supremum norm, hence in $L^2(\rho_\nu d\omega; [\varepsilon, R])$ since ρ_ν is bounded there. Combining the uniform tails with the compactness on every $[\varepsilon, R]$, every sequence in B has a subsequence that is Cauchy in $\mathcal{H}$: given $\eta > 0$ choose ε, R so that both tails are below η uniformly on B , extract a subsequence converging in $L^2(\rho_\nu; [\varepsilon, R])$, and observe that its $\mathcal{H}$ -oscillation is eventually below 3η . Thus V_t is compact.

Step 4: conclusion. By (15), $K(t)(\delta a, \delta z_0) = (0, V_t \Phi_t(\delta a, \delta z_0))$, that is, $K(t) = \iota \circ V_t \circ \Phi_t$ with $\iota : \mathcal{H} \rightarrow \mathcal{X}$ the isometric embedding $z \mapsto (0, z)$. The composition of the bounded operator Φ_t with the compact operator V_t and the bounded operator ι is compact. This proves the lemma. We note that the argument gives the explicit bound $\|K(t)\| \leq \sqrt{2} L E_\nu(Lt^\nu) \left(\int_0^\infty \rho_\nu \min(t, \omega^{-1})^2 d\omega \right)^{1/2}$, and that it uses of J only its boundedness, not its continuity.

Remark 12. The factorisation through $C([0, t]; \mathbb{R}^n)$, rather than through the Sobolev space $H^1([0, t]; \mathbb{R}^n)$, is not a matter of convenience. The derivative of the observable perturbation involves the functional $\langle \mu_\nu, \omega \cdot \rangle$, which is not bounded on $\mathcal{H}$; the perturbation δx is continuous but need not be of class H^1 for a general datum in $\mathcal{X}$. The Volterra equation (B.1) and the Mittag-Leffler bound (B.3) are what make the continuity, and the bound on it, available without any further regularity.

References

- [1] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory. *Meccanica*, 15:9–20, 1980.
- [2] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 2: Numerical application. *Meccanica*, 15:21–30, 1980.
- [3] J.-P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.*, 57:617–656, 1985.
- [4] V. I. Oseledets. A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197–231, 1968.
- [5] I. Podlubny. *Fractional Differential Equations*. Academic Press, San Diego, 1999.
- [6] C. Li and G. Chen. Chaos and hyperchaos in the fractional-order Rössler equations. *Physica A*, 341:55–61, 2004.

- [7] C. Li and G. Peng. Chaos in Chen's system with a fractional order. *Chaos Solitons Fractals*, 22:443–450, 2004.
- [8] M. S. Tavazoei and M. Haeri. A proof for non existence of periodic solutions in time invariant fractional order systems. *Automatica*, 45:1886–1890, 2009.
- [9] R. Caponetto and S. Fazzino. A semi-analytical method for the computation of the Lyapunov exponents of fractional-order systems. *Commun. Nonlinear Sci. Numer. Simul.*, 18:22–27, 2013.
- [10] M.-F. Danca and N. Kuznetsov. Matlab code for Lyapunov exponents of fractional-order systems. *Int. J. Bifurcation Chaos*, 28:1850067, 2018.
- [11] M.-F. Danca. Improved Matlab code for Lyapunov exponents of fractional order systems, 2026. arXiv:2604.09334.
- [12] M.-F. Danca and M. Fečkan. Memory principle of the Matlab code for Lyapunov exponents of fractional order, 2024. arXiv:2406.04686.
- [13] H. Li, Y. Shen, Y. Han, J. Dong, and J. Li. Determining Lyapunov exponents of fractional-order systems: a general method based on memory principle. *Chaos Solitons Fractals*, 168:113167, 2023.
- [14] J.-C. Trigeassou and N. Maamri. Modeling and initialization of non-linear and chaotic fractional order systems based on the infinite state representation. *Fractal Fract.*, 7:713, 2023.
- [15] N. J. Ford and A. C. Simpson. The numerical solution of fractional differential equations: speed versus accuracy. *Numerical Algorithms*, 26:333–346, 2001.
- [16] W. Deng. Short memory principle and a predictor–corrector approach for fractional differential equations. *Journal of Computational and Applied Mathematics*, 206:174–188, 2007.
- [17] M.-F. Danca, M. Fečkan, N. Kuznetsov, and G. Chen. Fractional-order PWC systems without zero Lyapunov exponents. *Nonlinear Dyn.*, 92:1061–1078, 2018.
- [18] G. Montseny. Diffusive representation of pseudo-differential time-operators. *ESAIM Proc.*, 5:159–175, 1998.

- [19] D. Matignon and C. Prieur. Asymptotic stability of linear conservative systems when coupled with diffusive systems. *ESAIM: COCV*, 11:487–507, 2005.
- [20] D. Ruelle. Characteristic exponents and invariant manifolds in Hilbert space. *Ann. Math.*, 115:243–290, 1982.
- [21] R. Mañé. Lyapunov exponents and stable manifolds for compact transformations. In *Geometric Dynamics*, volume 1007 of *Lecture Notes in Math.*, pages 522–577. Springer, 1983.
- [22] P. Thieullen. Fibrés dynamiques asymptotiquement compacts, exposants de Lyapounov, entropie, dimension. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 4:49–97, 1987.
- [23] L. Yuan and O. P. Agrawal. A numerical scheme for dynamic systems containing fractional derivatives. *J. Vib. Acoust.*, 124:321–324, 2002.
- [24] K. Diethelm. An investigation of some nonclassical methods for the numerical approximation of Caputo-type fractional derivatives. *Numer. Algorithms*, 47:361–390, 2008.
- [25] S. Jiang, J. Zhang, Q. Zhang, and Z. Zhang. Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations. *Commun. Comput. Phys.*, 21:650–678, 2017.
- [26] D. Baffet and J. S. Hesthaven. A kernel compression scheme for fractional differential equations. *SIAM J. Numer. Anal.*, 55:496–520, 2017.
- [27] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, New York, 1983.
- [28] Z. Lian and K. Lu. Lyapunov exponents and invariant manifolds for random dynamical systems in a Banach space. *Mem. Amer. Math. Soc.*, 206(967), 2010.
- [29] D. Matignon. Stability results for fractional differential equations with applications to control processing. In *Computational Engineering in Systems Applications*, pages 963–968, Lille, 1996. IMACS.