Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

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

Stefano Fazzino

sfazzino@gmail.com

Independent Researcher

Research Article

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

Posted Date: October 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11431473/v1

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

Additional Declarations: No competing interests reported.

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

Stefano Fazzino[1*]

[1*]Independent researcher, Catania, Italy.

Corresponding author(s). E-mail(s): sfazzino@gmail.com;

Abstract

Lyapunov exponents of fractional-order systems are widely computed, but a definition of these exponents for the Caputo system has not been given. A Caputo system has the whole past as its state, is not invariant under time translation, and its linearisation is not a cocycle, so that the multiplicative ergodic theorem 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 holds in the form of Thieullen. We show that the memory of a bounded Caputo trajectory grows in general as a power of time, so that the system ages and the exponents have to be defined on the stationary dynamics. The positive exponents are well defined and independent of any truncation of memory, and the essential exponent is zero. We conjecture, on the basis of the linear case, that all non-positive exponents vanish and that the negative exponents reported by existing codes are contraction rates of the observable. A Benettin algorithm on the augmented system supports these results on the fractional Chen and Rössler systems. It also shows that memory truncation, applied to the Grünwald-Letnikov derivative, leaves exponent and attractor unchanged, whereas, applied to the integral form of the equation, it deforms the attractor, lowers the maximal exponent by a fifth to a half and can suppress chaos, and in finite-mode realisations it introduces spurious negative exponents of the order of the inverse of the memory window.

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

MSC Classification: 37H15 , 34A08 , 26A33 , 37M25 , 37D45

1 Introduction

Lyapunov exponents measure the average rate at which nearby trajectories of a dynamical system separate or approach each other, and a positive exponent is the standard quantitative indicator of chaos [1–3]. Their definition is based on the theorem 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 initial point. Fractional-order systems, in which the time derivative is replaced by a derivative of non-integer order in the sense of Caputo [5], are widely used to model systems with memory, for instance in viscoelasticity, electrochemistry and anomalous diffusion. Chaotic fractional-order versions of the Lorenz, Rössler, Chen and Lü systems have been studied for about two decades [6–8], and their Lyapunov exponents are computed routinely. However, none of the hypotheses of the theorem of Oseledets holds for these systems.

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 that the state of the system at time t is not a point of $\mathbb{R}^n$ but the entire trajectory up to t. This has three consequences. First, there is no flow on $\mathbb{R}^n$, since the value of the solution at a given time does not determine its future. Second, the system is not invariant under time translation: a solution shifted in time is a solution of an equation with a different lower limit of integration, i.e., the system has an age. Third, the linearisation along a trajectory also depends on the past, and the family of linearised evolution operators is not a cocycle. The multiplicative ergodic theorem requires a cocycle over a measure-preserving dynamical system, and neither of them is available here.

Several numerical approaches have been proposed to compute Lyapunov exponents of fractional-order systems. 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, with a memory truncated by the multistep construction. The 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, i.e., the history of the perturbation is reset at each renormalisation. A method based on the memory principle of the Grünwald-Letnikov derivative [13] keeps the history of the perturbation. 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 orthogonalisation the exponents of the resulting augmented system, together with a set of quantities that they call pseudo-exponents. Finally, the short-memory principle [5], which discards the history beyond a fixed window, is used in most long integrations. It is applied in two forms: to the Grünwald-Letnikov derivative, for which the error bound of Podlubny holds, and to the Volterra integral form of the equation, for which Ford and Simpson [15] and Deng [16] showed that a fixed memory severely degrades the accuracy. To our knowledge, the effect of either form on the Lyapunov exponents has not been quantified.

These works do not specify which mathematical object is being computed, and three numerical observations have not been explained. The first is the dependence of

the estimates on the integration horizon, with values that oscillate and no convergence analysis [12]. The second is the absence of the zero exponent that every autonomous flow possesses, observed numerically by Danca, Fečkan, Kuznetsov and Chen [17]. The third is the presence of the pseudo-exponents of [14], which do not correspond to any quantity of the original system.

In this paper we give a definition of the Lyapunov exponents of fractional-order systems and derive these observations from it. Our tool is the diffusive representation of the Caputo operator [18, 19]. Since the power-law kernel is the Laplace transform of a positive measure, the fractional system is equivalent to an autonomous system of ordinary differential equations on a continuum of first-order memory modes, each relaxing at its own rate and 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 is available in the form proved by Ruelle [20] for compact cocycles in Hilbert space and extended by Mañé [21] and Thieullen [22] to asymptotically compact cocycles.

We show that, in general, 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 as $t^{1-\nu}$ (Section 2.3). Because of this ageing, the exponents have to be defined on the stationary dynamics, which we assume to be approached by the Caputo trajectory 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 larger than zero are well defined, form a discrete set and do not depend on any truncation of memory (Section 3). We show that the essential exponent of the cocycle is zero and that the flow has a zero exponent, we prove that in the linear case the observable does not contract exponentially, and we conjecture, on the basis of this result and of the numerical computations, that all non-positive exponents are zero. This would explain why the negative exponents of the integer-order system are not found for $\nu < 1$ (Section 4). We then implement a Benettin algorithm on the augmented system, based on the sum-of-exponentials quadratures developed for fast fractional integration [23–26], validate it against the Mittag-Leffler function (Section 5), and apply it to the fractional Chen and Rössler systems (Section 6). The numerical results support this structure of the spectrum and also show that truncating the memory to a window τ_m has very different effects depending on where the truncation is applied. When it is applied to the derivative, as in the short-memory principle, it does not change the results. When it is 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 fifth to a half for windows of five hundred to a thousand time units, destroys the dynamics for a window of one hundred, and, in finite-mode realisations, introduces negative exponents of the order of $-1/\tau_m$. The difference is due to the integrability of the kernel that is truncated, as we show by comparing the two truncations 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, and Section 4 describes the structure of the spectrum. Section 5 presents the numerical method and its validation, and Section 6

the results on chaotic systems. Section 7 contains the discussion, including a conjecture on the rate of convergence of the finite-time exponents, and Section 8 the conclusions. All proofs are given in the appendices.

2 The diffusive representation as a state space

2.1 The Caputo system and the kernel identity

We consider the autonomous fractional-order system

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

where $x(t) \in \mathbb{R}^n$ is the state, $f : \mathbb{R}^n \to \mathbb{R}^n$ 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 as

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

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

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

where k_ν is the kernel of the Riemann-Liouville fractional integral of order ν, which we denote by I^ν. The kernel decays as a power of τ and is not integrable on $[0,\infty)$, so that the state at time t depends on the whole past with weights that do not vanish.

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

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

Here $\omega > 0$ is a relaxation rate and μ_ν is the density with which the rates contribute to the kernel. Substituting (4) into (3) and exchanging the order of integration, which is allowed by Tonelli's theorem since $\mu_\nu e^{-\omega\tau}$ is positive and $f(x(s))$ is bounded on $[0,t]$, the solution can be written as

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

where $z(\omega,t) \in \mathbb{R}^n$ is the state of the memory mode with rate ω, i.e., of a first-order filter driven by $f(x(t))$ and relaxing at rate ω. Equation (5) is the diffusive

representation of Montseny [18], the infinite-state representation of Trigeassou and Maamri [14], and the basis of the numerical schemes of Yuan and Agrawal [23] and Diethelm [24]. The fractional system is therefore an ordinary differential system on a continuum of modes, and its power-law memory is due to the slow modes: the function $\mu_\nu(\omega)/\omega$, i.e., the weight of a mode multiplied by its relaxation time, is not integrable at $\omega = 0$, and the modes with $\omega \to 0$ retain their content for arbitrarily long times.

Two properties of (5) will be used in the following. First, the initial condition x_0 is not stored in z but appears as an additive constant, and the Caputo problem corresponds to the choice $z(\cdot, 0) = 0$, i.e., to zero initial memory. Second, in the limit $\nu \to 1$ one has $\mu_\nu(\omega) \to 0$ for every $\omega > 0$, while the mass of μ_ν concentrates at $\omega = 0$, where the mode is a pure integrator, and the ordinary system $\dot{x} = f(x)$ is recovered. We use this limit as a check on the results for the observable x, which must reduce to those of the ordinary system as $\nu \to 1$. As discussed in Section 4.3, the same is not expected for the spectrum in the augmented state space, which is expected to be discontinuous at $\nu = 1$.

2.2 The state space

The space $L^2(\mu_\nu \, \mathrm{d}\omega)$ cannot be used for the memory profile $z(\cdot, t)$, because the output map $z \mapsto \langle \mu_\nu, z \rangle := \int_0^\infty \mu_\nu(\omega) z(\omega) \, \mathrm{d}\omega$, which reconstructs x from the modes and feeds it to f, is not bounded on it: by the Cauchy-Schwarz inequality its norm is $(\int_0^\infty \mu_\nu \, \mathrm{d}\omega)^{1/2}$ and $\int_0^\infty \omega^{-\nu} \, \mathrm{d}\omega$ diverges at infinity. As usual in the theory of diffusive systems [18, 19], this is remedied by giving more weight to the high rates. We define

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

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

$$|\langle \mu_\nu, z \rangle| \le \Big(\int_0^\infty \frac{\mu_\nu(\omega)}{1 + \omega} \, \mathrm{d}\omega \Big)^{1/2} \|z\|_{\mathcal{H}} = \|z\|_{\mathcal{H}} \,, \tag{7}$$

since $\int_0^\infty \omega^{-\nu} (1 + \omega)^{-1} \, \mathrm{d}\omega$ is the Beta integral $B(1 - \nu, \nu) = \pi / \sin(\pi\nu)$. The operator of multiplication by $-\omega$ is self-adjoint and non-positive on $\mathcal{H}$, and generates the contraction semigroup $S(t)z = e^{-\omega t} z$. Because of the modes with $\omega \to 0$, its norm is $\|S(t)\| = 1$ for every t.

The state must also include x_0, which is a constant in (5) but has to be perturbed when sensitivity to initial conditions is studied. The state space is therefore the product

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

and the augmented system, which is autonomous, reads

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

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 initial datum $(a, z_0) \in \mathcal{X}$ the system (9) has a unique local mild solution in $\mathcal{X}$,*

$$z(t) = S(t)z_0 + \int_0^t S(t-s)\, f\big(a + \langle \mu_\nu,\, z(s)\rangle\big)\, \mathrm{d}s,$$

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.

A sketch of the proof is given in Appendix A.1.

2.3 The system ages

To define an attractor in $\mathcal{X}$ and an invariant measure on it, the memory $z(\cdot, t)$ should remain bounded. The following proposition shows that a bounded trajectory $x(t)$ has, in general, an unbounded memory.

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

$$z(\omega, t) \simeq \int_0^t f(x(s))\, \mathrm{d}s \simeq \frac{\bar{x} - x_0}{\Gamma(2-\nu)}\, t^{1-\nu} \qquad (\omega t \to 0), \tag{10}$$

and, if $\bar{x} \neq x_0$, there exist positive constants c_1 and c_2, with c_1 depending on $|\bar{x} - x_0|$, such that for t large enough

$$c_1\, t^{(1-\nu)/2} \le \|z(\cdot, t)\|_{\mathcal{H}} \le c_2\, t^{(1-\nu)/2}. \tag{11}$$

In this case the orbit of the Caputo problem in $\mathcal{X}$ is neither bounded nor relatively compact.

The proof is given in Appendix A.2. According to the proposition, a Caputo solution with bounded observable x has a memory that keeps changing and moves towards lower and lower rates, i.e., the system does not forget the initial condition x_0. This is consistent with a known property, the non-existence of exactly periodic solutions of Caputo systems [8]. The norm of the memory grows as a power of time, with an exponent $(1-\nu)/2$ that vanishes for $\nu \to 1$.

For this reason two dynamics have to be distinguished. The first is that of the stationary solutions of the augmented system, defined for all $t \in \mathbb{R}$, with memory

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

i.e., the solutions of the equation with the Liouville-Weyl derivative, which is invariant under time translation. For these solutions the memory can have a finite norm. For instance, if x is a stationary signal whose spectral density P_x is bounded at zero frequency and decays fast enough, a direct computation in the frequency domain shows that the mean value of $\|z(\cdot, t)\|^2_{\mathcal{H}}$ is finite for every $\nu \in (0, 1)$, although the content of the slowest modes, $\lim_{\omega\to 0} z(\omega, t)$, has finite variance only for $\nu > 1/2$. Finite norm

is however much weaker than the existence of a compact invariant set, which is the property needed below. The second dynamics is that of the Caputo orbit, which starts from zero memory. A connection between the two is suggested by a classical identity between the Caputo and Riemann-Liouville derivatives of a function with $x(0) = x_0$ [5],

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

according to which the Caputo problem differs from the problem with memory starting at $t = 0$ by a non-autonomous forcing that decays as $t^{-\nu}$, with a prefactor $1/\Gamma(1-\nu) \simeq 1-\nu$ that vanishes as $\nu \to 1$. The identity concerns the lower limit $t = 0$ and not the stationary solutions, which have lower limit $-\infty$. It nevertheless suggests that the Caputo orbit approaches the stationary dynamics algebraically, as $t^{-\nu}$. Since this is not proven for chaotic systems, we formulate it as a hypothesis.

Hypothesis 3 (invariant set) The augmented system (9) has a compact invariant set $\mathcal{A} \subset \mathcal{X}$, contained in a single fibre $\{a = a^*\}$ and formed by stationary solutions of the form (12) bounded in $\mathcal{X}$, and the observable of the Caputo orbit approaches the observable of a stationary solution on $\mathcal{A}$: there exists a solution $(a^*, z^*(t)) \in \mathcal{A}$ such that $|x(t) - x^*(t)| = O(t^{-\nu})$ for $t \to \infty$, where $x^* = a^* + \langle \mu_\nu,\, z^* \rangle$.

The hypothesis contains two assumptions of different nature. The existence of $\mathcal{A}$ requires that the Liouville-Weyl system admits solutions bounded on the whole real line, i.e., that $f(x(s))$ along the attractor has enough cancellations for the fractional integral from $-\infty$ to converge. This is plausible but not proven. The rate $t^{-\nu}$ is exact for the observable in the linear case (Proposition 9 below). In general it is suggested by (13): if the invariant set is stable in the sense of shadowing, a forcing of amplitude $O(t^{-\nu})$ produces a deviation $O(t^{-\nu})$. The assumption that $\mathcal{A}$ lies in a single fibre means that the stationary dynamics is described by one value of the constant a.

3 Lyapunov exponents

3.1 Linearisation and cocycle property

Let $(a, z(t))$ be a solution of (9), $x(t)$ the corresponding trajectory and $J(t) := Df(x(t))$ the Jacobian matrix of the vector field along it. The linearised equation reads

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

where $(\delta a, \delta z) \in \mathcal{X}$ is a tangent vector and $\delta x \in \mathbb{R}^n$ is the corresponding perturbation of the observable. Let $M(t,s) : \mathcal{X} \to \mathcal{X}$ be the operator that maps $(\delta a, \delta z(s))$ into $(\delta a, \delta z(t))$. Since (9) is autonomous, $M(t,s)$ depends only on the state at time s and on $t-s$, and it satisfies $M(t,r) = M(t,s)M(s,r)$ for all $r \le s \le t$. Hence M is a cocycle over the flow of (9). The fractional variational equation, which is not autonomous, does not define a cocycle over a flow on $\mathbb{R}^n$, and it is the cocycle structure that allows us to define Lyapunov exponents in the usual sense.

The relation between (14) and the fractional variational equation is useful to understand 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 $(\delta x_0, 0)$, i.e., to perturbations of the initial condition with zero tangent memory, observed through $\delta x = \delta a + \langle \mu_\nu, \delta z\rangle$. Codes that integrate the fractional variational equation thus follow an n-dimensional subspace of tangent data in an infinite-dimensional tangent space. Moreover, when they renormalise by QR at time t_k, they replace the tangent vector $(\delta a, \delta z(t_k))$ with a vector of the form $(\delta x(t_k), 0)$, i.e., they erase its memory. This is the "impulsive algorithm with variable lower limit" described in [12]. Such codes therefore do not compute the cocycle (14), but a modified quantity that depends on the renormalisation interval. We return to this point in Section 4.3.

3.2 Structure of the evolution operator

By the variation-of-constants formula, the cocycle from time 0 can be written as

$$\begin{aligned} M(t,0) &= S(t) + K(t), \qquad S(t)(\delta a, \delta z_0) = (\delta a,\ e^{-\omega t}\delta z_0), \\ K(t)(\delta a, \delta z_0) &= \Big(0,\ \int_0^t e^{-\omega(t-s)} J(s)\, \delta x(s)\, \mathrm{d}s\Big), \end{aligned} \tag{15}$$

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

The proof is given in Appendix B.1.

3.3 Multiplicative ergodic theorem and definition

The theorem of Oseledets has a version for cocycles on Hilbert spaces due to Ruelle [20], extended to non-compact cocycles by Mañé [21] and Thieullen [22], and in a general form by Lian and Lu [27]. In this version compactness is replaced by a condition on the measure of non-compactness $\alpha(\cdot)$ of the cocycle. For a bounded operator, α is the infimum of the radii of the finite coverings of the image of the unit ball; it is subadditive, it vanishes on compact operators and it does not exceed the operator norm. Defining the essential exponent

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

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 $\lambda_1 > \lambda_2 > \cdots > \kappa$, finite or accumulating only at κ, each with finite multiplicity, together with the Oseledets filtration and the associated invariant manifolds. Below κ the theorem does not provide any structure.

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

The proof, which follows from Lemma 4, is given in Appendix B.3.

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, since $\mathcal{A}$ is compact. 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 the theorem of Thieullen, restricted to the invariant tangent subspace $\{\delta a = 0\}$ of the 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, since they are defined for the exact system. In particular, whether the system is chaotic, i.e., whether* $\lambda_1 > 0$, *is a property of the system, which numerical methods approximate.*

The theorem follows from Proposition 5 and from the theorem of Thieullen, whose integrability condition is satisfied because $\mathcal{A}$ is compact and J is therefore bounded on it. It provides a definition and an existence result, not an algorithm or a value, and it relies on a hypothesis that is not proven for chaotic systems. Since $\kappa = 0$, the theorem gives no information on the exponents that are not positive; these are discussed in Section 4.

It remains to relate the definition, which is given on $\mathcal{A}$, to what a computation measures, i.e., the Caputo orbit starting 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* $t \to \infty$ *to the exponents* $\lambda_j > 0$ *of* $\mathcal{A}$ *with respect to the physical measure, with algebraic corrections whose form is discussed in Section 7.*

The conjecture is based on Hypothesis 3, suggested by (13): the observable of the Caputo orbit approaches that of $\mathcal{A}$ as $O(t^{-\nu})$, so that the local expansion rate, if it depends regularly on the observable state, differs from its value on $\mathcal{A}$ by $O(t^{-\nu})$. In the linear case the corresponding statement holds for the unstable directions

(Proposition 9). In the chaotic case it can be checked numerically, and we do so in Section 6 by comparing 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

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 can be verified by differentiating (9) with respect to t and using $\dot{x} = \langle \mu_\nu, \partial_t z \rangle$. Its norm is bounded from above and from below also along the Caputo orbit. Indeed, from $\partial_t z = -\omega z + f(x)$ and Proposition 2, for $\omega \ll 1/t$ one has $\omega z = O(t^{-\nu})$ and therefore $\partial_t z \simeq f(x(t))$, which does not depend on ω and is integrable against ρ_ν near the origin, while for large ω one has $\partial_t z = O(\omega^{-1})$. Hence $\|v(t)\|$ remains bounded, although $\|z(\cdot, t)\|_{\mathcal{H}}$ grows, and it does not tend to zero because $x(t)$ does not converge to an equilibrium. The Lyapunov exponent of v is therefore zero, both on $\mathcal{A}$ and along the Caputo orbit.

Danca, Fečkan, Kuznetsov and Chen [17] observed numerically that fractional-order systems have no zero exponent. The two results refer to different definitions. The time shift of a solution is a solution of the augmented system with a different initial memory, and the zero exponent is associated with this direction. It is not a solution of the fractional variational equation with zero memory, which is the only one followed by fractional codes. The absence of the zero exponent is therefore due to the restriction to data of the form $(\delta x_0, 0)$.

4.2 No exponential contraction: the linear case

Proposition 9 *Let $f(x) = \Lambda x$ with Λ diagonalisable, and consider the perturbation δx of the observable due to a perturbation δx_0 of the initial condition along an eigenvector of Λ with eigenvalue η. If $|\arg \eta| > \nu\pi/2$, i.e., in the stable sector of Matignon [28], the perturbation decays algebraically,*

$$\delta x(t) = E_\nu(\eta t^\nu)\, \delta x_0 \simeq -\frac{\delta x_0}{\eta\, \Gamma(1-\nu)}\, t^{-\nu} \qquad (t \to \infty),$$

where $E_\nu(w) = \sum_{k \geq 0} w^k / \Gamma(k\nu + 1)$ is the Mittag-Leffler function, so that its growth rate $t^{-1} \ln(|\delta x(t)| \,/\, |\delta x_0|)$ tends to zero. If $|\arg \eta| < \nu\pi/2$, then $\delta x(t) \simeq \nu^{-1} \exp(\eta^{1/\nu} t)\, \delta x_0$, the exponent is $\operatorname{Re} \eta^{1/\nu} > 0$, and for real η the finite-time exponent is $\eta^{1/\nu} - t^{-1} \ln \nu + o(t^{-1})$. In this case the exponent of the tangent vector in the norm of $\mathcal{X}$ coincides with that of the observable.

The proof is given in Appendix C. The proposition shows that, in a linear Caputo system, the observable does not contract exponentially along any direction, because the memory modes with $\omega \to 0$ release the perturbation at an algebraic rate. This is in contrast with the negative exponents reported in the numerical literature, for instance -2.60 for order 0.995 in [12], which are finite-time values. For a real stable eigenvalue $\eta = -\gamma$, with $\gamma > 0$, the function $E_\nu(-\gamma t^\nu)$ decays approximately as the stretched

exponential $\exp(-\gamma t^\nu/\Gamma(1+\nu))$ up to a crossover time $t_*(\gamma,\nu)$, and algebraically afterwards. For $t \gg t_*$ the finite-time growth rate of the observable is therefore

$$\frac{1}{t}\ln\frac{t^{-\nu}}{\gamma\,\Gamma(1-\nu)} = -\nu\,\frac{\ln t}{t} - \frac{\ln\big(\gamma\,\Gamma(1-\nu)\big)}{t} \qquad (t \gg t_*), \tag{17}$$

i.e., 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. For this reason, at order 0.995 the drift towards zero occurs beyond the horizons usually explored.

4.3 The tangent subspace and the meaning of negative exponents

Two further points concern the computation of the exponents. The first is related to 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 a norm not smaller than $|\delta a|$, and therefore a non-negative exponent. Perturbations with $\delta a \neq 0$ are perturbations of the Caputo initial condition, i.e., displacements to a different fibre $\{a = \text{const}\}$, whereas the perturbations tangent to the invariant set $\mathcal{A}$, on which the exponents are defined, have $\delta a = 0$. For this reason Definition 6 refers to the invariant subspace $\{\delta a = 0\}$, which also contains the direction of the flow. The maximal exponent is the same in the whole space and in this subspace, because the corresponding Oseledets direction is tangent to $\mathcal{A}$ and therefore lies in $\{\delta a = 0\}$.

The second point concerns the non-positive part of the spectrum. Since $\kappa = 0$, the multiplicative ergodic theorem gives no information on the exponents below zero. Moreover, a tangent vector whose observable δx contracts exponentially still has components on the slow memory modes, which decay as $e^{-\omega t}$ with $\omega \to 0$, so that its norm in $\mathcal{H}$ is not expected to contract exponentially. The negative exponents of the integer-order system, such as the value -14.6 of the Lorenz system, are therefore not expected to appear in the augmented system for any $\nu < 1$. They are contraction rates of the observable, and they are expected to appear only at $\nu = 1$, when the memory band disappears, so that the spectrum in $\mathcal{X}$ would be discontinuous at $\nu = 1$. We checked this numerically on the Lorenz system at $\nu = 0.99999$. In the subspace $\{\delta a = 0\}$ we obtain the exponents $(0.95,\ -0.005,\ -0.005)$, instead of the integer-order values $(0.91,\ 0,\ -14.6)$: the negative exponent is replaced by values close to zero. The maximal exponent differs from the integer-order value by about five per cent, a difference that we attribute to the discretisation of the memory band, which we did not refine at this order.

This has a direct consequence for the interpretation of fractional codes. What they measure as negative exponents are contraction rates of the observable on the time scale of the renormalisation interval. If the QR renormalisation resets the history every Δ time units, each segment starts again in the initial, fast phase of the Mittag-Leffler function, and the method measures the short-time contraction rate on the scale Δ. This quantity is reproducible but depends on Δ, and it tends to zero as $\Delta \to \infty$. The Benettin algorithm on the augmented system, instead, returns after the positive

exponents only the essential band. The zero exponent of the flow lies inside the band and can be identified only by its direction. We expect that the pseudo-exponents of [14] correspond to the essential band computed with a small number of 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 that drift slowly to zero or, for truncated memory, are of the order of minus the slowest retained rate.*

The theorem of Thieullen gives no information below $\kappa = 0$, and Proposition 9 covers only the observable in the linear case, so that this statement is formulated as a conjecture. It is supported by the following argument: below κ every tangent vector has a component on the slow modes that cannot contract faster than $e^{-\omega t}$ with $\omega \to 0$, and the compact part K can redistribute the perturbation but cannot cancel this component. The numerical test is reported 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 N_m rates ω_i with weights w_i such that

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

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 [23–26], which reduces the cost of a step from $O(N)$ to $O(N_m)$, with N the number of steps and $N_m = O(\log(T/h) \cdot \log \epsilon_q^{-1})$ for a relative accuracy ϵ_q. Our quadrature of the measure $\mu_\nu(\omega)\,\mathrm{d}\omega$ uses a Gauss-Jacobi rule with eight nodes and weight $\omega^{-\nu}$ on the interval $[0, \omega_{\min}]$, which treats the singularity at the origin exactly, and Gauss-Legendre rules in the variable $\ln \omega$ on geometric intervals, one per decade, from $\omega_{\min}$ to $\omega_{\max}$, with eight nodes per interval. The lower edge $\omega_{\min}$ is not a truncation of memory, since the Jacobi interval extends to $\omega = 0$, but it determines the resolution with which the slow band is represented, and for this reason we vary it. The upper edge is $\omega_{\max} = 10^4/h$. The modes with $\omega > 1/h$ represent the part of the kernel that acts within one step, and the neglected part beyond $\omega_{\max}$ has a weight of order $(\omega_{\max} h)^{-\nu}$. With $\omega_{\max} = 20/h$ and $\nu = 0.5$ this part was twelve per cent of the contribution within one step and changed the maximal exponent by two per cent, which motivates the larger value. With N_m between 88 and 112 the relative error of (18) is below 10^{-8} on $[h, T]$.

To study the effect of a truncation of memory we also use a variant in which the Jacobi interval is removed. In this case no mode slower than $\omega_{\min}$ is retained, so that no memory survives beyond $\tau_m = 1/\omega_{\min}$. This is the truncation implicit in every finite-mode realisation with 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 $\int_0^{\omega_{\min}} \mu_\nu(\omega) e^{-\omega\tau}\,\mathrm{d}\omega$, which for

$\tau \ll \tau_m$ is equal to the constant $\sin(\pi\nu)\,\omega_{\min}^{1-\nu}/(\pi(1-\nu))$. Since the mass of μ_ν below $\omega_{\min}$ decreases only as $\omega_{\min}^{1-\nu}$, the slow modes carry a large fraction of the kernel at all lags: the ratio between the subtracted part and $k_\nu(\tau)$ is $(\omega_{\min}\tau)^{1-\nu}/\Gamma(2-\nu)$, which for $\nu = 0.9$ and $\tau_m = 10^3$ is about one third at lag 10^{-2} and one half at lag one. The truncation therefore modifies the kernel at all lags, and not only beyond τ_m.

5.2 Truncated-memory codes in the original formulation

For comparison with common practice, we also integrate the Chen system in its original formulation with the two truncations found in the literature. In both cases we use the tangent system obtained by exact linearisation of the discrete scheme, with the same window, and we rescale the whole stored history of the tangent vector at each renormalisation, which is correct because the tangent scheme is linear. The first scheme is the Grünwald-Letnikov form, for which the short-memory principle is formulated,

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

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

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

i.e., the explicit fractional Euler scheme with the sum restricted to the last L steps, which is the fixed memory of [15, 16]. Here the weights decay as $j^{\nu-1}$ and are not summable, and the neglected part of the sum is comparable with the retained one. Both schemes are of first order. They are run with $h = 0.002$, the largest step for which the explicit Euler scheme is stable on the Chen system, over $T = 1500$ time units, and with $h = 0.001$ over $T = 1000$; in both cases the first hundred time units are discarded, and the uncertainty is the standard deviation of ten block averages. At these steps the schemes are not yet in the asymptotic regime: halving h changes the full-memory exponent by three per cent, and both values differ by ten to fifteen per cent from the result of the second-order scheme of Table 1. The reason is that the first-order error acts on the kernel as a small shift of the effective order, and the maximal exponent of the Chen system depends strongly and non-monotonically on ν in this range, as shown in Section 6. For this reason we compare the truncations within each scheme, and we report the results for both steps.

5.3 Integration and QR algorithm

The discretised state and tangent equations are

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

with $i = 1, \ldots, N_m$ and δa constant, so that the tangent space has dimension $n(N_m+1)$, or nN_m in the subspace $\{\delta a = 0\}$. Since the modes are linear and diagonal, their exact update over a step is $z_i(t+h) = e^{-\omega_i h} z_i(t) + \int_0^h e^{-\omega_i (h-s)} f\, \mathrm{d}s$. We use the second-order exponential time-differencing Runge-Kutta scheme (ETD-RK2), which approximates the forcing as linear within a step and avoids the stiffness due to $\omega_{\max} \sim 1/h$ without implicit steps. Since the Caputo solution behaves as t^ν near $t = 0$, with a singular derivative 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 and 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 decomposition is the discretised inner product of $\mathcal{X}$, $\|(\delta a, \delta z)\|^2 = |\delta a|^2 + \sum_i w_i (1+\omega_i)\, |\delta z_i|^2$, implemented by rescaling the tangent coordinates. The exponents do not depend on the choice among equivalent norms, but at finite N_m this choice affects the speed of convergence. We compute $k = 5$ exponents. The code consists of a few hundred lines of Python with NumPy.

5.4 Validation on the linear case

The linear system $f(x) = \Lambda x$ with $\Lambda = \mathrm{diag}(1, -2)$ has an exact solution in terms of the Mittag-Leffler function, and its exponents are known from 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. Fig. 1 shows the finite-time exponents at $\nu = 0.7$ with $h = 0.005$ and $T = 60$. The average from $t = 0$ contains a transient due to the random initial tangent vector, which decays as $1/t$, while the average from $T/3$ converges to $\lambda_1 = 0.9999$. For $\nu = 0.5$ and $\nu = 0.995$ the same procedure gives $\lambda_1 = 0.9994$ and 1.0000. The computed trajectory agrees with the Mittag-Leffler function within $5 \cdot 10^{-3}$ on $[0, 6]$ for $\nu = 0.5$ and within $5 \cdot 10^{-4}$ for $\nu = 0.7$; beyond $t \simeq 8$ the power series of E_ν at negative argument loses precision because of cancellations and cannot be used as a reference. Along the stable direction the finite-time exponent is negative and tends to zero, as predicted by (17).

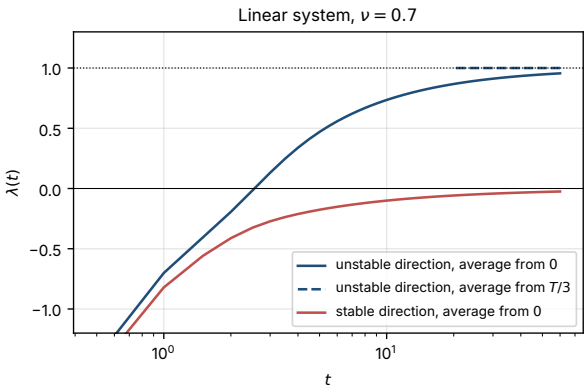

Fig. 1 Validation on the linear system $\Lambda = \mathrm{diag}(1, -2)$ at $\nu = 0.7$. The finite-time exponent along the unstable direction converges to the exact value $\eta^{1/\nu} = 1$ (dotted line) when the transient of the initial tangent vector is discarded. Along the stable direction the finite-time exponent is negative and tends to zero, in agreement with Proposition 9.

6 Results on chaotic systems

6.1 Systems and procedure

We consider the fractional Chen system [7],

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

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],

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

with $a = 0.5$, $b = 0.2$, $c = 10$, at $\nu = 0.9$. In these two equations a, b, c denote the parameters of the systems; this notation is used only here and should not be confused with the component a of the state space. The Lorenz system with the standard parameters converges to an equilibrium for $\nu \leq 0.99$ in our simulations and is chaotic at $\nu = 0.995$. This agrees with the stability condition of Matignon [28]: the eigenvalues $0.094 \pm 10.19\,i$ of the equilibria $C_{\pm}$ have argument 0.497π, so that $C_{\pm}$ are asymptotically stable for $\nu < 0.994$. At $\nu = 0.995$ the memory effects are too small to be informative, and we use the Lorenz system only for the check at $\nu = 0.99999$ reported in Section 4.3. The horizon is $T = 2000$ for the Chen system ($h = 0.002$) and $T = 4000$ for the Rössler system ($h = 0.005$), and the initial conditions are $(-9, -5, 14)$ and $(1, 1, 0.1)$, respectively. We compare three estimates of each exponent: the average from $t = 0$ with zero initial memory (cold start), the average from $T/4$ along the same trajectory (delayed average), and, for the Chen system, the average over a second run started from the state $x(T/4)$ with zero memory (restart). The uncertainty on λ_1 is the standard deviation of block averages over 250 time units. For the exponents with

system, ν	memory	N_m	λ_1	$\langle\lambda_{2..5}(T)\rangle$	q
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 interval extends to $\omega = 0$, with $\omega_{\min}$ setting the resolution of the slow band; "truncated" means that no mode below $\omega_{\min}$ is retained, i.e., a memory window $\tau_m = 1/\omega_{\min}$. The exponent q describes the algebraic decay of the norms of the Benettin vectors from the second one onwards.

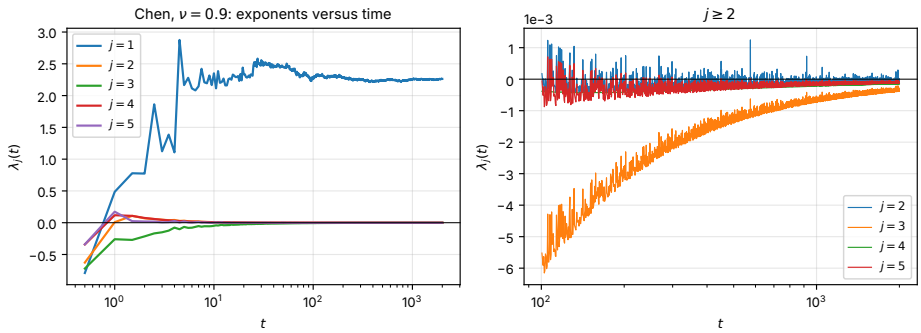

Fig. 2 Fractional Chen system at $\nu = 0.9$, full memory with $\omega_{\min} = 10^{-6}$. Left: the five finite-time exponents along the cold-start run; the maximal one settles at 2.25, while the others tend to zero. Right: the exponents with $j \geq 2$ on an enlarged scale for $t > 100$; at the end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

$j \geq 2$ we also report the exponent q of the algebraic decay of the norms of the Benettin vectors, obtained by fitting $\lambda_j(t)\, t = -q \ln t + \text{const}$ on the last nine tenths of the run.

6.2 Maximal exponent and essential band

The results are summarised in Table 1 and Fig. 2. The maximal exponent does not depend on the resolution of the memory band. In the three full-memory runs of the Chen system at $\nu = 0.9$ it is 2.26 ± 0.02, independently of $\omega_{\min}$, and the cold-start, delayed and restart averages agree within the ergodic fluctuations (for the run with $\omega_{\min} = 5 \cdot 10^{-4}$ they are 2.254, 2.266 and 2.28, respectively). At $\nu = 0.999$ the same procedure gives 2.01, close to the value 2.03 of the integer-order Chen system. A scan in ν shows a strong and non-monotonic dependence on the order, with a periodic window at $\nu = 0.99$ and values 2.13, 2.41 and 2.63 at $\nu = 0.98$, 0.97 and 0.95, followed by 2.25 at $\nu = 0.9$ and 2.41 at $\nu = 0.85$. All the other exponents are compatible with zero, and their small negative value is due to the finite resolution. With the same retained memory, extending the resolved interval from $\omega_{\min} = 5 \cdot 10^{-4}$ to 10^{-6} reduces

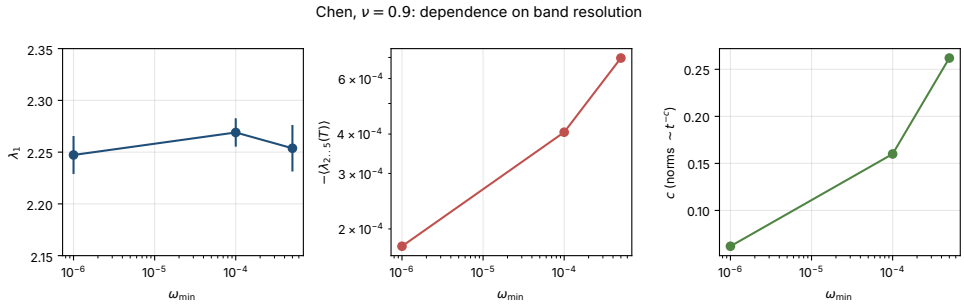

Fig. 3 Chen system at $\nu = 0.9$ with full memory. As the resolution of the slow band increases (smaller $\omega_{\min}$), the maximal exponent does not change, while the non-positive exponents and the algebraic decay of the norms tend to zero.

the residual from $-7 \cdot 10^{-4}$ to $-1.8 \cdot 10^{-4}$ and the decay of the norms from $t^{-0.26}$ to $t^{-0.06}$ (Fig. 3). The decrease is monotonic and slow, a factor of four over three and a half decades of $\omega_{\min}$, as expected from an average over a continuum of modes with weight $\omega^{-\nu}$ rather than from a single slow mode. With the same quadrature, the residual and its decay are of the same order in the three cases considered, which indicates that they are due to the discretisation of the memory continuum rather than to the systems. Finally, no exponential contraction is observed in any case, whereas the corresponding integer-order systems have negative exponents of order -10.

6.3 Effect of memory truncation

The runs with the slow modes removed are shown in Fig. 4. Without modes below $\omega_{\min}$ the non-positive exponents of the Chen system settle at values of the order of $-\omega_{\min}$: $-1.7 \cdot 10^{-3}$ for $\tau_m = 10^3$ and $-1.3 \cdot 10^{-2}$ for $\tau_m = 10^2$, i.e., between 1.3 and 1.7 times $-\omega_{\min}$. The slowest retained mode sets the scale of the apparent contraction rate, so that the negative exponents of finite-mode codes are of the order of the inverse of their slowest time constant. The truncation, however, also affects the maximal exponent. With $\tau_m = 10^3$ it decreases 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 agrees with the estimate given in the second paragraph of Section 5: removing the slow modes changes the kernel at all lags, and the resulting system is less chaotic.

We now consider the truncations used in practice (Table 2 and Fig. 5). In the Grünwald-Letnikov form (19) the maximal exponent at $h = 0.002$ 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 $h = 0.001$ the values are 2.562 ± 0.029, 2.583 ± 0.017 and 2.566 ± 0.028. The attractor has the same extent in all runs. The short-memory principle, applied to the derivative as formulated by Podlubny, therefore leaves exponent and geometry unchanged even with a window of one hundred time units, which for the Chen system corresponds to 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. With $\tau_m = 500$ it decreases to 1.85 ± 0.14 and 2.07 ± 0.17, with an uncertainty seven times larger, and the attractor becomes larger: its extent

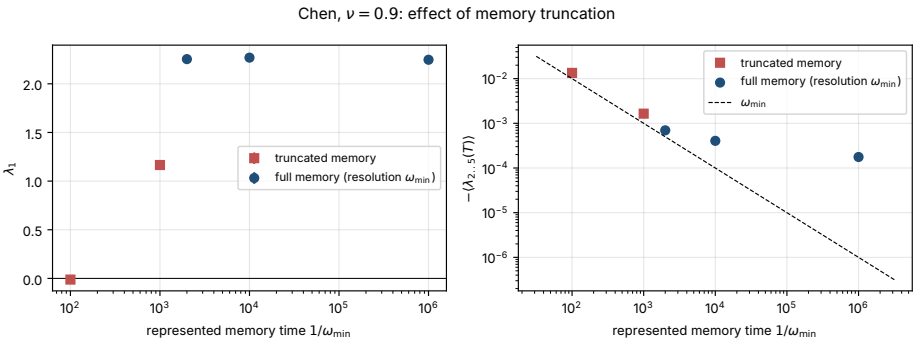

Fig. 4 Chen system at $\nu = 0.9$, augmented system: maximal exponent (left) and non-positive exponents (right) as a function of the represented memory time. With the slow modes removed (squares) the non-positive exponents lie close to the line $-\omega_{\min}$, and the maximal exponent decreases from 2.25 to 1.17 and then becomes negative. With full memory (circles) the maximal exponent does not change and the negative residual decreases with the resolution.

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	equilibrium

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. *In the last row the windows are $\tau_m = 10^3$ and $\tau_m = 10^2$ instead of 500 and 100.

in z changes from $[6, 40]$ to $[1, 63]$ and $[3, 51]$ for the two steps. With $\tau_m = 100$ the trajectory diverges after about five hundred time units at the larger step and eight hundred at the smaller one, and no exponent can be defined.

The three truncations can therefore be grouped into two classes. Truncating the derivative removes the tail of a summable kernel and has no visible effect. Truncating the integral, either by fixing the memory of a Volterra sum or by removing the slow modes of a finite-mode realisation, removes the tail of a non-summable kernel. In this case the system changes: for the shorter windows the trajectory diverges in the first case and converges to an equilibrium in the second. In the finite-mode realisation the change is accompanied by spurious negative exponents of the order of $-1/\tau_m$. As a consequence, values obtained with fixed memory in the integral form, or with finite-mode realisations, can be compared neither with each other nor with ours unless the window is specified, and the threshold of chaos in ν obtained in this way may be displaced. Values obtained with the short-memory principle applied to the Grünwald-Letnikov derivative are not affected.

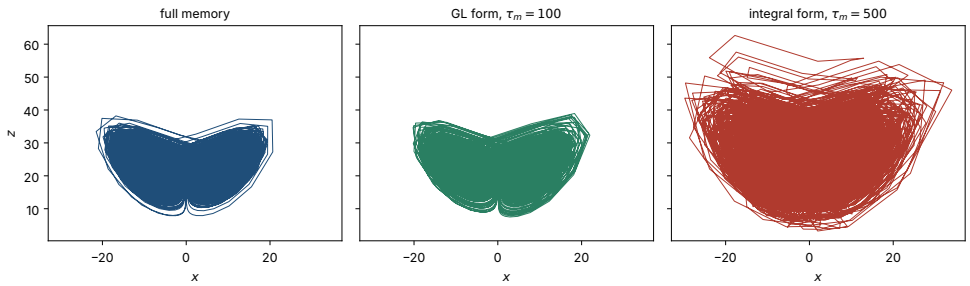

Fig. 5 Chen system at $\nu = 0.9$, projection on the (x, z) plane of the last three hundred time units: full memory (left), Grünwald-Letnikov form with $\tau_m = 100$ (centre), Volterra integral form with $\tau_m = 500$ (right). Truncating the derivative leaves the attractor unchanged, while truncating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

7 Discussion

7.1 Rate of convergence of the finite-time exponents

The dependence of the numerical estimates on the horizon can be interpreted within the present framework. The finite-time exponent along the Caputo orbit is the average over $[0, t]$ of the local expansion rate $\ell(x(s))$, which we assume to depend regularly on the observable state, and its distance from the limit has two sources. The first is the same as in ordinary systems: 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 $O(t^{-1/2})$, as follows from the central limit theorem. The second is specific to the fractional case: by Hypothesis 3 the observable of the Caputo orbit is at distance $O(s^{-\nu})$ from that of $\mathcal{A}$ at time s, so that $\ell(x(s))$ differs by $O(s^{-\nu})$ from its value on $\mathcal{A}$, and the average of this difference over $[0, t]$ is $t^{-1} \int^t s^{-\nu}\, \mathrm{d}s \propto t^{-\nu}/(1-\nu)$. The prefactor deriving 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 $\nu \to 1$, so that the memory term connects continuously to the ordinary term proportional to $1/t$. This argument leads to the following conjecture.

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

$$\lambda_1(t) - \lambda_1 = \frac{b_1}{\Gamma(2-\nu)}\, t^{-\nu} + \frac{b_2}{t} + O_{\mathrm{fluct}}\big(t^{-1/2}\big), \tag{24}$$

where b_1 and b_2 are constants, b_1 depending on the initial condition and on the sensitivity of ℓ to the state; starting from a state on $\mathcal{A}$ the term in $t^{-\nu}$ is absent.

The law (24) is consistent with the limit $\nu \to 1$ and with the linear case of Proposition 9, where the convergence along the unstable direction is exactly $1/t$ without algebraic term. It also explains why it is difficult to verify: 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 $\nu = 0.85$–0.9 at which the standard systems are chaotic, the term is hidden by the ergodic fluctuations, and indeed the differences

between cold-start, delayed and restart averages in our runs are within the error. A test of the conjecture requires chaotic systems with $\nu < 1/2$, which the standard oscillators are not, or ensemble averages over many initial conditions at fixed ν. We also note that for $\nu \leq 1/2$ the slowest modes of a stationary solution have unbounded variance, as remarked in Section 2.3, so that Hypothesis 3 is more delicate in this range.

7.2 Comparison with existing 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 an estimate with memory reset, which should approach ours as the renormalisation interval increases, while their negative exponents are contraction rates of the observable on the scale of that interval. The Grünwald-Letnikov method of [13], which keeps 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 and its negative exponents to drift slowly to zero with the horizon. The augmented-system computations of [14] refer to the same object as ours, computed with a coarse quadrature. We expect their pseudo-exponents to correspond to the essential band of Section 4.3, whose dependence on the resolution is shown in Fig. 3.

7.3 Practical consequences

Models with Caputo derivatives are used in viscoelasticity, in the modelling of batteries and supercapacitors, in biology and in circuits with constant phase elements. In several of these applications one is interested in whether the model is chaotic, how predictable it is and how sensitive it is to the initial conditions. The answer is usually obtained by computing Lyapunov exponents with the available codes, almost always with some truncation of the memory to reduce the cost of long integrations. Our results affect this practice in three ways.

The first concerns the positive exponent, which is reliable only if the memory is complete or is 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 fifth to a quarter, and a window of one hundred units destroys the attractor. The removal of the slow modes in a finite-mode realisation halves the exponent with a window of one thousand units and suppresses chaos with a window of one hundred. Bifurcation diagrams and chaos thresholds in the order ν obtained with such truncations may therefore refer to a different system, 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, instead, leaves the exponent unchanged within one per cent even with a window of one hundred units. We therefore recommend truncating the Grünwald-Letnikov derivative rather than the integral, or using the augmented-system algorithm, whose cost per step grows only logarithmically with the integration horizon, and always reporting the memory window.

The second concerns the negative exponents reported by these codes, which depend on the treatment of the memory. In finite-mode realisations they are of the order of

the inverse of the slowest retained time constant, and in codes with memory reset they are contraction rates on the scale of the renormalisation interval. If they are used to estimate a Lyapunov dimension or a rate of loss of memory of the initial conditions, or to compare two models, the result depends on the window. Section 4.3 shows that they are short-time contraction rates of the observable.

The third concerns ageing. A fractional-order system does not forget its initial condition, and, when the time average of the trajectory differs from the initial condition, the memory of the initial state grows in the slow modes for the whole life of the system, as $t^{(1-\nu)/2}$ in the norm of the memory, while the observable is expected to approach the stationary dynamics as $t^{-\nu}$. For models of materials and devices with memory, this means that the initial transient never vanishes completely and that two samples started from different states remain distinguishable. For the identification of fractional models from experimental data, it means that the data of an aged system differ from those of a stationary one, and that the time elapsed since the start of the dynamics should be reported together with the data.

7.4 Limitations

The existence of the invariant set $\mathcal{A}$ for chaotic fractional systems (Hypothesis 3) is not proven, and the definition relies on it. The zero exponent of the flow is present by construction, but it cannot be distinguished by its value from the essential band for any of the horizons we considered, so that its existence follows from the autonomy of the augmented system and is not confirmed by the computations. The residual negative values of the band decrease only slowly with the resolution, so that a computation reporting them should specify $\omega_{\min}$. Finally, the exponents we report refer to the memory state in the norm of $\mathcal{X}$. The contraction rates of the observable, which are relevant for some applications such as synchronisation, are a different quantity, which the present framework identifies but does not compute.

8 Conclusions

In this paper we have given a definition of the Lyapunov exponents of a fractional-order system that does not depend on any truncation of memory. The definition is obtained by rewriting the system, through the diffusive representation, as an autonomous evolution equation on a Hilbert space of memory modes, and by applying the multiplicative ergodic theorem in the form valid for asymptotically compact cocycles. We have shown that, when its time average differs from the initial condition, the memory of a bounded Caputo trajectory is not bounded, but grows as $t^{(1-\nu)/2}$. The system therefore ages, and the exponents have to be defined on the stationary dynamics, which the Caputo orbit is assumed to approach algebraically in time. The positive exponents are well defined and form a discrete set, the essential exponent is zero, and the flow has a zero exponent, which fractional codes do not detect. In the linear case the observable does not contract exponentially, and we have conjectured that, also in the nonlinear case, all exponents other than the positive ones are zero. The negative exponents of the integer-order system are contraction rates of the observable and were not found in any of our computations with $\nu < 1$.

A Benettin algorithm on the augmented system, validated against the Mittag-Leffler function, supports this structure on the fractional Chen and Rössler systems. The maximal exponent is reproducible within one per cent and does not depend on the resolution of the memory band, while all the other exponents vanish within a residual that decreases with the resolution and is of the same order in all the systems considered. The numerical results also show that truncating the memory has opposite effects depending on where it is applied. Applied to the Grünwald-Letnikov derivative, as prescribed by the short-memory principle, it leaves exponent and attractor unchanged. 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 fifth to a half and can suppress chaos or destroy the attractor; in finite-mode realisations it also introduces spurious negative exponents of the order of $-1/\tau_m$. The difference is due to the integrability of the truncated kernel. Results obtained by truncating the integral therefore describe a different system and can be compared only if the window is specified. Finally, we have conjectured that the finite-time estimates converge with a memory term proportional to $t^{-\nu}$, which is visible only for orders below one half. The proof of this conjecture and of the existence of the invariant set for chaotic fractional systems remain open problems.

Acknowledgements. This work is the fruit of many years of study and hard work with continous support and patience from my family, to which I express my sincere gratitude .

Declarations

- **Funding:** No funding was received for this work.
- **Competing interests:** The author declares no competing interests.
- **Ethics approval and consent to participate:** Not applicable.
- **Consent for publication:** Not applicable.
- **Data availability:** The numerical data generated in this study are available from the author on reasonable request; all results can be reproduced with the method described in Section 5.
- **Materials availability:** Not applicable.
- **Code availability:** The Python code used for the computations is available from the author on reasonable request.
- **Author contribution:** S.F. is the sole author and carried out all parts of the work.

Appendix A Proofs for the state space and the ageing of memory

A.1 Sketch of the proof of Proposition 1

The main difficulty is that a constant vector $f \in \mathbb{R}^n$, regarded as a function of ω, does not belong to $\mathcal{H}$, because $\int_0^\infty \rho_\nu \, \mathrm{d}\omega = \infty$. The forcing therefore lies outside the state space, as in 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

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

so that $\|S(\tau)f\|_{\mathcal{H}} \sim \tau^{\nu/2-1}$ as $\tau \to 0$, which is integrable near $\tau = 0$ for every $\nu > 0$. With this estimate, and since the nonlinearity $z \mapsto f(a + \langle \mu_\nu, z\rangle)$ is locally Lipschitz from $\mathcal{H}$ to $\mathbb{R}^n$ by (7), local existence and uniqueness follow from the standard theory of semilinear equations [29] by a Picard iteration in a space of continuous functions on $[0, T]$ with values in $\mathcal{H}$, weighted to account for the integrable singularity $\tau^{\nu/2-1}$. The equivalence with (3) is obtained by applying the output functional to the mild solution and using the kernel identity (4), as in the derivation of (5).

A.2 Proof of Proposition 2

Let $D := \sup_{t\geq 0} |x(t) - x_0|$ and $F_\infty := \sup_{t\geq 0} |f(x(t))|$, which are finite because x is bounded and f is continuous. We denote by $F(t) := \int_0^t f(x(s))\,\mathrm{d}s$ the integral of the vector field and proceed in four steps.

Step 1: expression of F. By (3), $x - x_0 = I^\nu f$. Applying $I^{1-\nu}$ and using the semigroup property of the Riemann-Liouville integrals, $I^{1-\nu} I^\nu = I^1$, one obtains

$$F(t) = I^{1-\nu}(x - x_0)(t) = \frac{1}{\Gamma(1-\nu)} \int_0^t (t-s)^{-\nu}\big(x(s) - x_0\big)\,\mathrm{d}s.$$

Step 2: growth of F. With the change of variable $s = tu$ the previous expression becomes

$$F(t) = \frac{t^{1-\nu}}{\Gamma(2-\nu)} \int_0^1 w(u)\,\big(x(tu) - x_0\big)\,\mathrm{d}u, \qquad w(u) := (1-\nu)(1-u)^{-\nu},$$

where w is a probability density on $[0, 1]$. If w is replaced by the indicator function of an interval $[0, u_0]$, the integral is u_0 times the average of $x - x_0$ over $[0, u_0 t]$, which tends to $u_0(\bar{x} - x_0)$. By linearity the same holds for every step function, and since the step functions are dense in $L^1(0, 1)$ and $|x - x_0| \leq D$, the integral tends to $\int_0^1 w\,\mathrm{d}u\,(\bar{x} - x_0) = \bar{x} - x_0$. Hence

$$F(t) = \frac{t^{1-\nu}}{\Gamma(2-\nu)}\big(\bar{x} - x_0 + o(1)\big) \qquad (t \to \infty). \tag{A1}$$

Moreover, F is Hölder continuous with exponent $1 - \nu$: splitting the integral that defines $I^{1-\nu}$ at s one verifies that, for $0 \leq s \leq t$,

$$|F(t) - F(s)| \leq C_H\,(t-s)^{1-\nu}, \qquad C_H := \frac{2D}{\Gamma(2-\nu)}. \tag{A2}$$

Step 3: representation of the memory. Integrating by parts in $z(\omega,t) = \int_0^t e^{-\omega(t-s)}\,\mathrm{d}F(s)$ and using $F(0)=0$ gives

$$z(\omega,t) = F(t)\,e^{-\omega t} + R(\omega,t), \qquad R(\omega,t) := \omega\int_0^t e^{-\omega(t-s)}\big(F(t)-F(s)\big)\,\mathrm{d}s. \tag{A3}$$

By (A2), $|R(\omega,t)| \le C_H\,\omega\int_0^t e^{-\omega\tau}\tau^{1-\nu}\,\mathrm{d}\tau$, which yields the two bounds

$$|R(\omega,t)| \le \frac{C_H}{2-\nu}\,\omega\,t^{2-\nu}, \qquad |R(\omega,t)| \le C_H\,\Gamma(2-\nu)\,\omega^{\nu-1}. \tag{A4}$$

In addition, directly from the definition of z, $|z(\omega,t)| \le F_\infty/\omega$.

Step 4: estimates. For $\omega t \le \varepsilon$, (A3), (A1) and the first bound in (A4) give $z(\omega,t) = F(t)\big(1+O(\varepsilon)\big) + O(\varepsilon\,t^{1-\nu})$, which proves (10). If $\bar{x} \ne x_0$, for ε small enough and t large enough one has $|z(\omega,t)| \ge c_0\,t^{1-\nu}$ for $\omega \le \varepsilon/t$, with $c_0 > 0$ independent of t and proportional to $|\bar{x}-x_0|$ for small ε, and therefore

$$\|z(\cdot,t)\|_{\mathcal{H}}^2 \ge c_0^2 t^{2-2\nu}\int_0^{\varepsilon/t}\mu_\nu(\omega)\,\mathrm{d}\omega = c_0^2\,\frac{\sin\pi\nu}{\pi(1-\nu)}\,\varepsilon^{1-\nu}\,t^{1-\nu}.$$

For the upper bound we split the integral that defines the norm into three intervals. For $\omega \le 1/t$, the first bound in (A4) and (A1) give $|z| \le C_1 t^{1-\nu}$, and the contribution is at most $2C_1^2 t^{2-2\nu}\int_0^{1/t}\mu_\nu\,\mathrm{d}\omega \le C_2\,t^{1-\nu}$ for $t \ge 1$. For $1/t < \omega \le 1$, the second bound in (A4) gives $|z| \le |F(t)|\,e^{-\omega t} + C_3\omega^{\nu-1}$, and the contribution is at most a constant times $t^{2-2\nu}\int_0^\infty \omega^{-\nu}e^{-2\omega t}\,\mathrm{d}\omega + \int_{1/t}^1 \omega^{\nu-2}\,\mathrm{d}\omega$, which is $O(t^{1-\nu})$. For $\omega > 1$, the bound $|z| \le F_\infty/\omega$ gives a contribution not larger than $F_\infty^2\int_1^\infty \rho_\nu\,\omega^{-2}\,\mathrm{d}\omega$, which is finite since $\rho_\nu\omega^{-2} \sim \omega^{-1-\nu}$ at infinity. Hence $\|z(\cdot,t)\|_{\mathcal{H}}^2 \le C_4\,t^{1-\nu}$, where C_4 depends on D and F_∞ but not on t, and together with the lower bound this proves (11). Since the norm diverges, the orbit is not bounded in $\mathcal{X}$, and therefore not relatively compact.

Appendix B Proofs for the evolution operator

B.1 Proof of Lemma 4

On the $\mathbb{R}^n$ component $S(t)$ is the identity, and on the $\mathcal{H}$ component it is the multiplication by $e^{-\omega t}$. A multiplication operator by a continuous function is bounded with norm equal to the supremum of the function, here one, and its spectrum is the closure of the range of the function, here $[0,1]$. Since every point of this interval is in the spectrum and not only isolated eigenvalues of finite multiplicity, $S(t)$ is not compact. As for $K(t)$, its image is the set of functions $\omega \mapsto \int_0^t e^{-\omega(t-s)}g(s)\,\mathrm{d}s$, with g ranging over an infinite-dimensional set of functions of s, so that its rank is not finite. The compactness of $K(t)$ is proved in Appendix B.2. The idea is that $K(t)$ factors through the space of continuous functions on $[0,t]$: the first factor is bounded by a Mittag-Leffler estimate on the linearised Volterra equation, and the second is compact by a Kolmogorov-Riesz argument in the weighted space $\mathcal{H}$.

B.2 Compactness of $K(t)$

We prove the second statement of Lemma 4: for every $t > 0$ and every bounded measurable $J : [0,t] \to \mathbb{R}^{n\times n}$ with $\sup_{[0,t]} \|J\| =: L < \infty$, the operator $K(t) : \mathcal{X} \to \mathcal{X}$ 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),

$$\delta z(s) = S(s)\delta z_0 + \int_0^s S(s-r)\, J(r)\, \delta x(r)\, \mathrm{d}r,$$

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 $\langle \mu_\nu,\, S(\tau)v\rangle = v \int_0^\infty \mu_\nu(\omega) e^{-\omega\tau}\, \mathrm{d}\omega = k_\nu(\tau)\, v$, valid for $\tau > 0$, we obtain that $\delta x(s) = \delta a + \langle \mu_\nu,\, \delta z(s)\rangle$ satisfies the linear Volterra equation

$$\delta x(s) = \varphi(s) + \int_0^s k_\nu(s-r)\, J(r)\, \delta x(r)\, \mathrm{d}r, \qquad \varphi(s) := \delta a + \langle \mu_\nu,\, S(s)\delta z_0\rangle\,. \tag{B5}$$

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

$$\|\varphi\|_\infty \le |\delta a| + \sup_s \|S(s)\delta z_0\|_{\mathcal{H}} \le |\delta a| + \|\delta z_0\|_{\mathcal{H}} \le \sqrt{2}\, \|(\delta a, \delta z_0)\|_{\mathcal{X}}\,. \tag{B6}$$

Step 2: the map from the datum to the observable perturbation is bounded

Let $\mathcal{T}$ be the Volterra operator $(\mathcal{T}g)(s) = \int_0^s k_\nu(s-r) J(r) g(r)\, \mathrm{d}r$ on $C([0,t];\mathbb{R}^n)$. Its iterates are controlled by the semigroup property of the Riemann-Liouville kernels, $\int_0^s k_\alpha(s-r) k_\beta(r)\, \mathrm{d}r = k_{\alpha+\beta}(s)$: by induction,

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

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

$$\left\|(I-\mathcal{T})^{-1}\right\| \le \sum_{m\ge 0} \frac{(Lt^\nu)^m}{\Gamma(m\nu+1)} = E_\nu(Lt^\nu), \tag{B7}$$

the Mittag-Leffler function of Proposition 9. Equation (B5) therefore has a unique continuous solution $\delta x = (I-\mathcal{T})^{-1}\varphi$, and by (B6)

$$\|\delta x\|_\infty \le \sqrt{2}\, E_\nu(Lt^\nu)\, \|(\delta a, \delta z_0)\|_{\mathcal{X}}\,. \tag{B8}$$

We denote by $\Phi_t : \mathcal{X} \to C([0,t];\mathbb{R}^n)$ 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 $V_t : C([0,t];\mathbb{R}^n) \to \mathcal{H}$ by $(V_t g)(\omega) = \int_0^t e^{-\omega(t-s)} g(s)\,\mathrm{d}s$. For every g one has the pointwise bound

$$|(V_t g)(\omega)| \le \|g\|_\infty \int_0^t e^{-\omega(t-s)}\,\mathrm{d}s = \|g\|_\infty \frac{1-e^{-\omega t}}{\omega} \le \|g\|_\infty \min\left(t, \omega^{-1}\right), \tag{B9}$$

and $\int_0^\infty \rho_\nu(\omega)\min(t,\omega^{-1})^2\,\mathrm{d}\omega < \infty$, 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 $\mathcal{B}$ be the image of the unit ball of $C([0,t];\mathbb{R}^n)$; we show that $\mathcal{B}$ is relatively compact in $\mathcal{H} = L^2(\rho_\nu\,\mathrm{d}\omega)$ 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 (B9):

$$\sup_{u\in\mathcal{B}} \int_R^\infty \rho_\nu\,|u|^2\,\mathrm{d}\omega \le \int_R^\infty \rho_\nu(\omega)\,\omega^{-2}\,\mathrm{d}\omega \xrightarrow[R\to\infty]{} 0,$$

$$\sup_{u\in\mathcal{B}} \int_0^\varepsilon \rho_\nu\,|u|^2\,\mathrm{d}\omega \le t^2 \int_0^\varepsilon \rho_\nu(\omega)\,\mathrm{d}\omega \xrightarrow[\varepsilon\to 0]{} 0.$$

On a compact interval $[\varepsilon, R]$ the functions of $\mathcal{B}$ are uniformly bounded, by (B9), and uniformly Lipschitz, because differentiation under the integral sign gives

$$|\partial_\omega (V_t g)(\omega)| = \Big| \int_0^t (t-s)\,e^{-\omega(t-s)} g(s)\,\mathrm{d}s \Big| \le \|g\|_\infty \int_0^t (t-s)\,\mathrm{d}s = \frac{t^2}{2}\,\|g\|_\infty\,;$$

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

Step 4: conclusion

By (15), $K(t)(\delta a, \delta z_0) = \big(0,\, V_t \Phi_t(\delta a, \delta z_0)\big)$, that is, $K(t) = \iota \circ V_t \circ \Phi_t$ with $\iota : \mathcal{H} \to \mathcal{X}$ 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 $\|K(t)\| \le \sqrt{2}\,L\,E_\nu(Lt^\nu)\,\big(\int_0^\infty \rho_\nu \min(t,\omega^{-1})^2\,\mathrm{d}\omega\big)^{1/2}$, and that it uses only the boundedness of J, 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 necessary. 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 (B5) and the Mittag-Leffler bound (B7) are what make the continuity, and the bound on it, available without any further regularity.

B.3 Proof of Proposition 5

The measure of non-compactness α is subadditive, vanishes on compact operators and is not larger than the operator norm. Since $K(t)$ is compact by Lemma 4, one has $\alpha(M(t,0)) = \alpha(S(t))$, and $\alpha(S(t)) \le \|S(t)\| = 1$. To show that $\alpha(S(t)) = 1$, let $\omega_0 > 0$ and consider the infinite-dimensional subspace of $\mathcal{X}$ formed by the vectors $(0,z)$ with z supported in $[0,\omega_0]$. The operator $S(t)$ maps this subspace onto itself, and its inverse on the subspace, the multiplication by $e^{\omega t}$ with $\omega \le \omega_0$, has norm at most $e^{\omega_0 t}$. Hence the image of the unit ball of $\mathcal{X}$ contains the ball of radius $e^{-\omega_0 t}$ of this infinite-dimensional subspace, whose measure of non-compactness is $e^{-\omega_0 t}$. Hence $e^{-\omega_0 t} \le \alpha(S(t)) \le 1$ for every $\omega_0 > 0$, i.e., $\alpha(M(t,0)) = 1$ for every t, and $\kappa = 0$.

Appendix C Proof for the linear case

The asymptotic expansions of the Mittag-Leffler function are classical [5]. For $|\arg w| > \nu\pi/2$ one has, for every integer $m \ge 1$,

$$E_\nu(w) = -\sum_{k=1}^{m} \frac{w^{-k}}{\Gamma(1-k\nu)} + O\big(\,|w|^{-m-1}\,\big),$$

while for $|\arg w| < \nu\pi/2$

$$E_\nu(w) = \frac{1}{\nu}\exp\left(w^{1/\nu}\right) - \sum_{k=1}^{m} \frac{w^{-k}}{\Gamma(1-k\nu)} + O\big(\,|w|^{-m-1}\,\big).$$

Along an eigenvector of Λ with eigenvalue η, the perturbation of the initial datum evolves as $\delta x(t) = E_\nu(\eta t^\nu)\,\delta x_0$. In the stable sector the first expansion with $m = 1$ gives the algebraic decay stated in the proposition, so that $t^{-1}\ln(|\delta x(t)|\,/\,|\delta x_0|) \to 0$. In the unstable sector the second expansion gives $\delta x(t) = \nu^{-1}\exp(\eta^{1/\nu} t)\,\delta x_0\,(1+o(1))$, since the algebraic terms are exponentially small with respect to the leading one, and the growth rate is $\operatorname{Re}\eta^{1/\nu}$. For real η the correction is exponentially small, so that $t^{-1}\ln(|\delta x(t)|\,/\,|\delta x_0|) = \eta^{1/\nu} - t^{-1}\ln\nu + o(t^{-1})$, which converges as $1/t$ and not as $t^{-\nu}$.

It remains to show that, in the unstable case, the exponent of the tangent vector $(\delta x_0, \delta z(t))$ in the norm of $\mathcal{X}$ is equal to that of the observable. By (7), $|\delta x(t)| \le |\delta x_0| + \|\delta z(t)\|_{\mathcal{H}} \le \sqrt{2}\,\|(\delta x_0, \delta z(t))\|_{\mathcal{X}}$, so that the norm in $\mathcal{X}$ is bounded from below by both $|\delta x(t)|\,/\sqrt{2}$ and $|\delta x_0|$. Conversely, $\delta z(t) = \int_0^t e^{-\omega(t-s)}\Lambda\,\delta x(s)\,\mathrm{d}s$, and the pointwise bound (B9) of Appendix B.2 gives $\|\delta z(t)\|_{\mathcal{H}} \le \|\Lambda\|\sup_{s\le t}|\delta x(s)|\,\big(\int_0^\infty \rho_\nu \min(t,\omega^{-1})^2\,\mathrm{d}\omega\big)^{1/2}$, where the last factor grows as $t^{(1+\nu)/2}$.

The norm in $\mathcal{X}$ is therefore bounded from above by $|\delta x_0|$ plus $\sup_{s\le t}|\delta x(s)|$ times a factor growing algebraically in t. In the unstable case both bounds grow as $\exp(\mathrm{Re}\,\eta^{1/\nu}\,t)$ up to algebraic factors, so that the exponent in $\mathcal{X}$ coincides with that of the observable.

References

[1] Benettin, G., Galgani, L., Giorgilli, A., Strelcyn, J.-M.: 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] Benettin, G., Galgani, L., Giorgilli, A., Strelcyn, J.-M.: 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] Eckmann, J.-P., Ruelle, D.: Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. **57**, 617–656 (1985)

[4] Oseledets, V.I.: A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. Trans. Moscow Math. Soc. **19**, 197–231 (1968)

[5] Podlubny, I.: Fractional Differential Equations. Academic Press, San Diego (1999)

[6] Li, C., Chen, G.: Chaos and hyperchaos in the fractional-order Rössler equations. Physica A **341**, 55–61 (2004)

[7] Li, C., Peng, G.: Chaos in Chen's system with a fractional order. Chaos Solitons Fractals **22**, 443–450 (2004)

[8] Tavazoei, M.S., Haeri, M.: A proof for non existence of periodic solutions in time invariant fractional order systems. Automatica **45**, 1886–1890 (2009)

[9] Caponetto, R., Fazzino, S.: 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] Danca, M.-F., Kuznetsov, N.: Matlab code for Lyapunov exponents of fractional-order systems. Int. J. Bifurcation Chaos **28**, 1850067 (2018)

[11] Danca, M.-F.: Improved Matlab code for Lyapunov exponents of fractional order systems. arXiv:2604.09334 (2026)

[12] Danca, M.-F., Fečkan, M.: Memory principle of the Matlab code for Lyapunov exponents of fractional order. arXiv:2406.04686 (2024)

[13] Li, H., Shen, Y., Han, Y., Dong, J., Li, J.: Determining Lyapunov exponents of fractional-order systems: a general method based on memory principle. Chaos

Solitons Fractals **168**, 113167 (2023)

[14] Trigeassou, J.-C., Maamri, N.: Modeling and initialization of nonlinear and chaotic fractional order systems based on the infinite state representation. Fractal Fract. **7**, 713 (2023)

[15] Ford, N.J., Simpson, A.C.: The numerical solution of fractional differential equations: speed versus accuracy. Numerical Algorithms **26**, 333–346 (2001)

[16] Deng, W.: Short memory principle and a predictor–corrector approach for fractional differential equations. Journal of Computational and Applied Mathematics **206**, 174–188 (2007)

[17] Danca, M.-F., Fečkan, M., Kuznetsov, N., Chen, G.: Fractional-order PWC systems without zero Lyapunov exponents. Nonlinear Dyn. **92**, 1061–1078 (2018)

[18] Montseny, G.: Diffusive representation of pseudo-differential time-operators. ESAIM Proc. **5**, 159–175 (1998)

[19] Matignon, D., Prieur, C.: Asymptotic stability of linear conservative systems when coupled with diffusive systems. ESAIM: COCV **11**, 487–507 (2005)

[20] Ruelle, D.: Characteristic exponents and invariant manifolds in Hilbert space. Ann. Math. **115**, 243–290 (1982)

[21] Mañé, R.: Lyapunov exponents and stable manifolds for compact transformations. In: Geometric Dynamics. Lecture Notes in Math., vol. 1007, pp. 522–577. Springer, Berlin (1983)

[22] Thieullen, P.: Fibrés dynamiques asymptotiquement compacts, exposants de Lyapounov, entropie, dimension. Ann. Inst. H. Poincaré Anal. Non Linéaire **4**, 49–97 (1987)

[23] Yuan, L., Agrawal, O.P.: A numerical scheme for dynamic systems containing fractional derivatives. J. Vib. Acoust. **124**, 321–324 (2002)

[24] Diethelm, K.: An investigation of some nonclassical methods for the numerical approximation of Caputo-type fractional derivatives. Numer. Algorithms **47**, 361–390 (2008)

[25] Jiang, S., Zhang, J., Zhang, Q., Zhang, Z.: Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations. Commun. Comput. Phys. **21**, 650–678 (2017)

[26] Baffet, D., Hesthaven, J.S.: A kernel compression scheme for fractional differential equations. SIAM J. Numer. Anal. **55**, 496–520 (2017)

[27] Lian, Z., Lu, K.: Lyapunov exponents and invariant manifolds for random

dynamical systems in a Banach space. Mem. Amer. Math. Soc. **206**(967) (2010)

[28] Matignon, D.: Stability results for fractional differential equations with applications to control processing. In: Computational Engineering in Systems Applications, Lille, pp. 963–968 (1996). IMACS

[29] Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Springer, New York (1983)