Nonlinearity

LONDON MATHEMATICAL SOCIETY EST. 1865

IOP Publishing

PAPER

A detectability criterion and data assimilation for nonlinear differential equations

To cite this article: Jason Frank and Sergiy Zhuk 2018 *Nonlinearity* **31** 5235

IOP Publishing | London Mathematical Society Nonlinearity

Nonlinearity **31** (2018) 5235–5257 https://doi.org/10.1088/1361-6544/aaddcb

A detectability criterion and data assimilation for nonlinear differential equations*

Jason Frank[1] **and Sergiy Zhuk**[2]

[1] Mathematical Institute, Utrecht University, PO Box 80010, 3508 TA Utrecht, Netherlands
[2] IBM Research, IBM Tech. Campus, Dublin D15HN66, Ireland

E-mail: j.e.frank@uu.nl and sergiy.zhuk@ie.ibm.com

Received 13 November 2017, revised 19 August 2018
Accepted for publication 30 August 2018
Published 18 October 2018

Recommended by Professor Bruno Eckhardt

Abstract

In this paper we propose a new sequential data assimilation method for nonlinear ordinary differential equations with compact state space. The method is designed so that the Lyapunov exponents of the corresponding estimation error dynamics are negative, i.e. the estimation error decays exponentially fast. The latter is shown to be the case for generic *regular* flow maps if and only if the observation matrix H satisfies detectability conditions. In particular this implies that the rank of H must be at least as great as the number of nonnegative Lyapunov exponents of the underlying attractor. Numerical experiments illustrate the exponential convergence of the method and the sharpness of the theory for the case of Lorenz '96 and Burgers equations with incomplete and noisy observations.

Keywords: data assimilation, synchronization, filtering, detectability, Lyapunov exponents
Mathematics Subject Classification numbers: 62M20, 37C50, 34D06, 37M25

(Some figures may appear in colour only in the online journal)

* Submitted to the editors.

1. Introduction

Consider a process described by the following ordinary differential equation (ODE):

$$\dot{z} = f(t,z), \qquad z(t) \in \mathcal{D} \subset \mathcal{R}^d, \quad z(0) = z_0. \tag{1}$$

The state $z(t)$ and initial condition z_0 are presumed to be unknown, but information about the state is obtained via a noisy observation process:

$$y(t) = H(t)z(t) + \eta(t), \qquad H : \mathcal{R}^{s\times d}, \quad t \geqslant 0, \tag{2}$$

where $\eta(t)$ is a squared-integrable function modelling the noise. Consider a filter, that is, the accompanying process described by the following ODE:

$$\dot{x} = f(t,x) + L(t,x)(y(t) - H(t)x), \qquad x(0) = x_0. \tag{3}$$

Given f (perfect model), and possibly incomplete observations y ($s < d$), the problem is: *to find conditions on H which guarantee that there exists a gain $L(t,x) \in \mathcal{R}^{d\times s}$ such that:*

$$\|\xi(t)\| \leqslant Ce^{-at}\|\xi(0)\|, \qquad \xi \triangleq z - x, \qquad t > 0, C > 0, a > 0, \tag{4}$$

and *to construct the gain $L(t,x)$.*

In this work we solve the above problem for generic f by combining ideas from modern Lyapunov stability theory [7], optimal control and numerical analysis. For the case of noise-free observations $\eta(t) \equiv 0$, we prove a theoretical convergence result for nonlinear, differentiable vector fields f based on a linearization about the estimator x. Namely, assuming that the Lyapunov exponents (LEs) associated with the equation $\dot{X} = \mathrm{D}f(t,x(t))X$ are *forward regular* (see [7]), we reformulate a classical notion from control theory, *detectability* for the pair of matrix-valued functions $(\mathrm{D}f(t,x(t)), H(t))$, in terms of LEs and prove that for a class of gains L (4) holds for every $\xi : \|\xi(0)\| \leqslant \varepsilon$ if and only if the pair of matrices $(\mathrm{D}f(t,x(t)), H(t))$ is detectable. This criterion represents our key contribution from the theoretical standpoint. We then apply this rather theoretical result to construct a gain L, and design a numerical algorithm for computing the filter (3), and we demonstrate the effectiveness of this algorithm numerically even in the presence of observational noise $\eta \not\equiv 0$. The computation of the gain L relies upon numerically stable procedures for computing LEs of the fundamental matrix differential equation $\dot{X} = \mathrm{D}f(t,x(t))X$ (see [16–18, 20]). Specifically, using a lemma of Perron (see [7]) and QR-decomposition we design the gain L to guarantee negativity of the LEs associated with the equation governing the estimation error ξ in the absence of noise:

$$\dot{\xi} = (\mathrm{D}f(t,x(t)) - L(t,x(t))H(t))\xi + N(t,\xi(t),x(t)), \quad \xi(0) = x_0 - z_0, \quad \eta \equiv 0, \tag{5}$$

where the nonlinear residual is given by $N(t,\xi,x) = f(t,x+\xi) - f(t,x) - \mathrm{D}f(t,x)\xi$ and satisfies a second order Lipschitz condition. Negativity of all LEs guarantees (4) under a mild condition on N and for $\|\xi(0)\| < \varepsilon$. In other words, the trivial solution $\xi(t) \equiv 0$ of (5) attracts an ε-neighborhood of itself.

Instead of solving the ill-conditioned fundamental matrix differential equation directly we compute an orthonormal basis for the non-stable tangent space by solving a matrix differential equation on a manifold of orthogonal matrices [18]. Since the number of the nonnegative Lyapunov exponents is typically much smaller then the dimension of x, one can reduce the associated computational cost significantly. We stress that this reduction preserves the exponential decay (4), and hence the estimation quality is not compromised. The resulting numerical algorithm for computing L and x is our main contribution from the computational standpoint. Note that our analysis applies to the case where the observations are noise-free

($\eta \equiv 0$). Nevertheless, the proposed filter proves to be efficient in the presence of the observational noise as suggested by our numerical experiments with chaotic nonlinear systems.

1.1. Motivation and related work

Problems like (1)–(3) are fundamental in diverse fields including synchronization in complex networks, data assimilation and control engineering. Syncronizing agents interacting in complex network topologies is of key interest in physical, biological, chemical and social systems to name just a few. In the literature on synchronization of chaotic systems [9, 12, 39, 40, 42], the coupled system (1)–(3) is referred to as a driver-receiver process, and the gain L is to be selected so that the receiver x is synchronized with the driver z. In their early papers on synchronization, Pecora and Carroll [39, 40] note that a necessary condition for synchronization (independent of x_0) is that the conditional Lyapunov exponents of (3) be negative. Similarly to our detectability conditions, conditions for stability of the synchronization manifold, based on a master stability function (MSF) [41], also require that all the LEs of an equation similar to (5) be negative: given an L of the form $L = \nu L_0$ for ν ranging over the spectrum of the coupling matrix, the MSF assigns a maximal LE of (5) to each ν. The synchronization manifold is stable provided the MSF maps this spectrum into $(-\infty, 0)$. In this work, we use a similar idea to describe a class of observation matrices H for which (5) has negative LEs: as per definition 3.5 below, our detectability condition requires that the non-stable tangent space of $\dot{X} = \mathrm{D}f(t, x(t))X$ have only trivial intersection with $\ker H^{\intercal}(t)H(t)$ most of the time. The latter is crucial for the design of the gain L, and allows us to achieve (4). Note that our algorithm does not require evaluating the MSF for different, possibly complex ν; instead we need to know the dimension of the unstable tangent subspace. We refer the reader to [2] for a review of recent results on synchronization and the MSF.

In the control literature, the problem of this paper is known as a filtering problem (if η and z_0 are stochastic) or an observer design problem or state estimation problem (if η and z_0 are deterministic). Theoretically, solution of the stochastic filtering problem for Markov diffusions is given by the so-called Kushner–Stratonovich (KS) equation, a stochastic partial differential equation (PDE) which describes evolution of the conditional density of the states of the underlying diffusion process [28]. For linear systems, the KS equation is equivalent to the Kalman–Bucy filter [27]. In contrast, deterministic state estimators assume that errors have bounded energy and belong to a given bounding set. The state estimate is then defined as a minimax center of the reachability set, a set of all states of the physical model which are reachable from the given set of initial conditions and are compatible with observations. Dynamics of the minimax center is described by a minimax filter. The latter may be constructed by using dynamic programming, i.e. the set $V \leqslant 1$, where V is the so-called value function V solving a Hamilton–Jacobi–Bellman (HJB) equation [6], coincides with the reachability set [5]. Statistically, the uncertainty description in the form of a bounding set represents the case of uniformly distributed bounded errors in contrast to stochastic filtering, where all the errors are usually assumed to be in the form of 'white noise'. However, in many cases $\exp\{-V\}$ coincides with the solution of the KS equation: in fact, for linear dynamics, equations of the minimax filter coincide with those of Kalman–Bucy filter [34].

For generic nonlinear models both minimax and stochastic filters are infinite-dimensional, i.e. to get an optimal estimate one needs to solve either the KS or HJB equation. Hence, if the state space of the model (1) is of high dimension then both filters become computationally intractable due to the 'curse of dimensionality'. To compute the filter one usually compromises optimality to gain computationally tractable approximations. One such approximation,

the Luenberger observer is obtained when the gain L in (3) is chosen so that the estimation error (5) converges to 0 exponentially, provided $\eta = 0$. In fact, there is a deep relationship between observers and (optimal) minimax filters: in the linear case the minimax/Kalman filter uniformly converges to the observer if the observational noise/model error 'disappears' as $t \to \infty$; see [4]. Motivated by this relationship, we construct an observer x for (1) as an approximation of the minimax filter. Note that the proposed detectability condition is most related to [4] where a so-called uniform detectability condition[3] was used together with uniform controllability to establish (4). Similar conditions based on relative degrees are used to design so called high-gain observers [15], minimax sliding mode controllers [49] and minimax filters for differential-algebraic equations [46–48]. A somewhat less restrictive condition of global convergence for skew-symmetric bilinear systems, e.g. Fourier–Galerkin discretization of the 2D Navier–Stokes equations, was reported in [44]. The result of this paper is a lot less restrictive[4] than those in the aforementioned papers, as our detectability condition requires that the non-stable tangent space is regularly observed as $t \to \infty$; see remark 3.6. Note that filters of the form (3) in the context of Navier–Stokes equations were studied in [3, 23, 26] but in the infinite-dimensional setting.

Data assimilation (DA) improves the accuracy of forecasts provided by physical models and evaluates their reliability by optimally combining *a priori* knowledge encoded in equations of mathematical physics with *a posteriori* information in the form of sensor data. Mathematically, many DA methods rely upon various approximations of stochastic filters. We refer the reader to [35, 43] for further discussions on mathematics behind data assimilation. In what follows we discuss a popular family of algorithms based on extended Kalman filter (ExKF) which is the most related to the present work.

In discrete time, ExKF is based on the following idea: given an accurate estimate of the state at time instant t, one linearizes the dynamics around that estimate and applies Kalman filtering for the resulting linear system to obtain an estimate for the next time step. This procedure is then repeated. A computational bottleneck associated with ExKF is the requirement to recompute the state error covariance matrix. The ensemble Kalman filter (EnKF) [22] suggests to overcome this issue by generating an ensemble of trajectories and by computing the ensemble variance to approximate the state error covariance matrix. Recently, this approximation scheme has been combined with ideas from Lyapunov stability theory to construct convergent square-root implementations of the EnKF, a so called EKF-AUS filter [14, 38]: the key idea is to sample and propagate the ensemble state error covariance matrix in the unstable/neutral tangent subspace only. Apart from solving the ensemble 'deflation problem', and as noted above, this may provide significant dimension reduction; see [13, 45]. The asymptotics of the error covarince matrix of the EKF-AUS, and of general reduced-rank discrete-time linear Kalman filter was studied in [11, 32], and in the latter paper it was demonstrated that, under uniform observability conditions, the span of the covariance matrix 'converges' towards the unstable-neutral subspace when the dynamics and the observation operator are linear and when the dynamical model is error free, for any, possibly rank-deficient, initial error covariance matrix. As a result, the corresponding reduced-rank discrete-time linear Kalman filter is equivalent to the optimal Kalman filter, asymptotically. [10] extends the former two works in the case of the discrete-time smoothing problem. Very recently, robustness of the LEs to the presence of model errors in the dynamical model were studied in [31] for discrete-time systems. The importance of Lyapunov exponents and Lyapunov vectors for analysis of DA methods has also been stressed in [8, 30, 36].

[3] Uniform detectability: $q^{\intercal}(\mathrm{D}f(t,x) + \Lambda(x)H(t))q \leqslant -\alpha_0\|q\|^2$, for some $\alpha_0 > 0$, Λ and all $q, x \in \mathcal{R}^d$.
[4] $q^{\intercal}(\mathrm{D}f(t,x(t)) + \Lambda(x(t))H(t))q \geqslant 0$ for all t within any *finite* number of compact intervals of $\mathcal{R}^+$.

Finally, we note that computation of LEs is technically challenging [20], as one needs to rely upon some regularity assumptions, e.g. the aforementioned forward regularity [7] or exponential dichotomy [17], to compute LEs for continuos time ODEs. These assumptions are not easy to verify in practice. On the other hand, the regularity for generic discrete time dynamical systems is provided by the so-called Oseledec theorem [37], hence 'most of the time' the discrete flow maps resulting from the time/spatial discretizations should be regular. In particular, our experiments with Lorenz '96 (L96) and Burgers(–Hopf) equations confirmed exponential convergence of the filter in the case of incomplete observations. We stress that Burgers equation is not dissipative in contrast to the *L*96 system, and it is unclear if this system possesses an invariant ergodic measure making the corresponding discrete flow map subject to the conditions of the Oseledec theorem, so our theory may not apply to this test case. Nevertheless, the proposed filter demonstrates convergence.

Notation. The notation throughout the paper is rather standard. $\|\cdot\|$ denotes the Euclidean norm of $\mathcal{R}^d$, $L^2(0,T)$ is the space of square-integrable functions with norm $\|\cdot\|_{L^2(0,T)}$. $\mathrm{D}f$ and $\mathrm{D}^2 f$ denote the Jacobian and Hessian of a smooth map $f : \mathcal{R}^d \to \mathcal{R}^d$.

This paper is organised as follows. Section 2 briefly recalls the notions of Lyapunov exponents, Lyapunov vectors and forward regularity, and reviews how one can compute LEs by means of continuous *QR*-decomposition. Section 3 introduces the notion of detectability for linear non-autonomous systems in terms of LEs, section 4 presents the design of the gain L and the filter x. Section 5 illustrates the application of the filter to Lorenz '96 and Burgers equations. Conclusions are in section 6 and appendix contains the proofs.

2. Review of Lyapunov stability theory

The stability theory of Lyapunov provides conditions for global stability of nonautonomous linear differential equations, for local stability of (the trivial solution of) certain quasi-linear differential equations, and by extension for local stability of orbits of nonlinear systems. We recall first the result for nonautonomous linear ODEs,

$$\dot{\xi} = A(t)\xi, \qquad \xi(t) \in \mathcal{R}^d, \tag{6}$$

where $A(t) \in \mathcal{R}^{d\times d}$ is continuous and bounded, can be formulated in terms of Lyapunov exponents. Following the exposition of Barreira and Pesin [7], we define the function

$$\lambda(\xi_0) = \limsup_{t\to\infty} \frac{1}{t} \log \|\xi(t;\xi_0)\|, \tag{7}$$

which measures the asymptotic rate of exponential growth or decay of the solution of (6) with initial condition $\xi(0) = \xi_0$. Considered over all $\xi_0 \in \mathcal{R}^d$, this function assumes s distinct values $\tilde{\lambda}_1 > \cdots > \tilde{\lambda}_s$, $s \leqslant d$. To the function λ we assign a *filtration* $\mathcal{V} := \{V_i\}$, where $V_i := \{v \in \mathcal{R}^d : \lambda(v) \leqslant \tilde{\lambda}_i\}$. The filtration $\mathcal{V}$ has the following properties: $\emptyset = V_{s+1} \subsetneq V_s \subsetneq \cdots \subsetneq V_1 = \mathcal{R}^d$, and $\lambda(v) = \tilde{\lambda}_i$ provided $v \in V_i \setminus V_{i+1}$. Letting $n_i := \dim(V_i)$, $i \geqslant 1$, $n_{s+1} = 0$, the number $d_i := n_i - n_{i+1}$, $i = 1, \ldots, s$, is the multiplicity of $\tilde{\lambda}_i$. Hence, the function λ assumes $d = d_1 + \cdots + d_s$ values up to multiplicity: the Lyapunov exponents $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$. A basis $\{v_i\}_{i=1}^d$ of $\mathcal{R}^d$ is called a *normal Lyapunov basis* if any V_i can be represented as a linear span of $\{v_{i_1}, \ldots, v_{i_{n_i}}\}$, $n_i := \dim(V_i)$. The normal basis is ordered if V_i is equal to the span of $\{v_1, \ldots, v_{n_i}\}$. Clearly, $\lambda(v_j) = \lambda_j$ for all $n_i - d_i < j \leqslant n_i$, $i = 1, \ldots, s$.

The stability theory of Lyapunov states (see [7, p 9]) that the trivial solution $\xi(t) = 0$ of (6) is exponentially stable if and only if $\lambda_1 < 0$, that is:

$$\forall \varepsilon > 0 \quad \exists C_\varepsilon > 0 : \qquad \|\xi(t;\xi_0)\| \leqslant C_\varepsilon \mathrm{e}^{(\lambda_1+\varepsilon)t}\|\xi_0\|, \qquad \forall t \geqslant 0. \tag{8}$$

Lyapunov stability theory also applies to quasi-linear systems of the form

$$\dot{\xi} = A(t)\xi + N(t,\xi),$$

where $N(t,\xi)$ satisfies

$$N(t,0) = 0, \qquad \|N(t,\xi) - N(t,\eta)\| \leqslant C\|\xi - \eta\|^p, \quad p > 1, \forall \xi, \eta \in \mathcal{R}^d.$$

In this case $\xi(t) \equiv 0$ also defines a solution, and this solution exponentially attracts an open neighborhood of the origin if the matrix $A(t)$ is bounded and regular and the Lyapunov exponents of the associated linear system (6) are negative [7]. By extension, one can establish the exponential stability of a particular solution $x^*(t)$ of a nonlinear differential equation $\dot{x} = f(t,x)$ by considering the dynamics of the error $\xi(t) = x(t) - x^*(t)$, which satisfies

$$\dot{\xi} = A(t)\xi + N(t,\xi,x), \qquad A(t) = \mathrm{D}f(x^*(t)), \quad N(t,\xi,x^*) = f(t,x^*+\xi) - f(t,x^*) - \mathrm{D}f(x^*)\xi.$$

For precise statements, we refer to [7].

2.1. Computation of Lyapunov exponents

The fundamental matrix equation associated to (6) is

$$\dot{X} = A(t)X(t), \quad X(t) \in \mathcal{R}^{d\times d}, \tag{9}$$

whose solution for the initial condition $X(0) = I$ yields the exact flow of (6). If (instead) the columns of $X(0) = (x_1(0), \ldots, x_d(0))$ form an ordered normal Lyapunov basis, one finds that $\lambda_i = \lambda(x_i(0))$, $i = 1, \ldots, d$. The vector $x_i(t)$ is referred to as the ith Lyapunov vector corresponding to the ith Lyapunov exponent.

The fundamental matrix equation (9) is numerically ill-conditioned, and stabilized formulations are used to compute Lyapunov exponents [18, 19, 21]. The algorithm used in this paper relies upon Perron's lemma[7, lemma 1.3.3] which suggests to transform (9) to the upper-triangular form $\dot{R} = BR$ by means of a Lyapunov coordinate transformation: $R = Q^\intercal X$. The Lyapunov exponents of $\dot{R} = BR$ coincide with those of (9). Moreover, under a regularity assumption, they can be computed by 'averaging' the diagonal elements of B. For the convenience of the reader we sketch out this procedure below.

Recall from [29, p 246] that for any $Y \in \mathcal{R}^{d\times k}$ there exists a QR-decomposition, i.e. an orthogonal matrix $\mathbb{Q} \in \mathcal{R}^{d\times d}$ and upper-triangular matrix $\mathbb{R} \in \mathcal{R}^{d\times k}$ such that $Y = \mathbb{Q}\mathbb{R}$. If the columns of Y are linearly-independent, then, by using the modified Gram–Schmidt (mGS) algorithm[5], one can also construct a so called *thin QR-decomposition*: $Y = QR$ where $Q \in \mathcal{R}^{d\times k}$ and $R \in \mathcal{R}^{k\times k}$. We stress that the thin QR-decomposition is unique if one chooses $R_{ii} > 0$. Moreover, the 'full' QR-decomposition is related to the thin one by

$$\mathbb{Q} = \begin{bmatrix} Q & Q_\perp \end{bmatrix}, \quad \mathbb{R} = \begin{bmatrix} R \\ 0 \end{bmatrix}, \quad Q_\perp^\intercal Q_\perp = I, Q^\intercal Q_\perp = 0 \tag{10}$$

[5] If the columns of Y are nearly linearly-dependent, and so the condition number of Y is large, the mGS algorithm may generate $\mathbb{Q}$ which is not quite orthogonal. In this case a more computationally expensive Housholder transformation can be used instead of the relatively cheap mGS [29, p 255] or one may apply mGS twice [24, 33].

where the columns of $Q_\perp \in \mathcal{R}^{d\times(d-k)}$ span the orthogonal complement of the range of Q in $\mathcal{R}^{d\times d}$. If $k = d$ and the columns of Y are linearly-independent, then $\mathbb{Q} = Q$ and $\mathbb{R} = R$.

If $Y \in \mathcal{R}^{d\times k}$, $k \leqslant d$, and $s = \mathrm{rank}(Y) < k$, then a simple modification of the Gram–Schmidt procedure[6] generates an orthogonal matrix $Q \in \mathcal{R}^{d\times s}$ and an upper-triangular matrix $R \in \mathcal{R}^{s\times s}$ such that $Y = QR$. In this case the diagonal of R may contain zeros.

Now, let $X \in \mathcal{R}^{d\times d}$ solve (9), and consider its QR-decomposition $X(t) = \mathbb{Q}(t)\mathbb{R}(t)$. Since $\mathrm{rank}(X(t)) \equiv d$, it follows that this QR-decomposition is unique, and $\mathbb{R}_{ii} > 0$, $i = 1, \ldots, d$. Moreover, the matrices $\mathbb{Q}$ and $\mathbb{R}$, constructed by the mGS algorithm, are continuously differentiable. It is easy to find, by differentiating the equality $X = \mathbb{Q}\mathbb{R}$, that $\mathbb{Q}$ and $\mathbb{R}$ solve[7] the following system of differential equations:

$$\dot{R}(t) = B(t)R(t), \quad R(0) = \mathbb{R}_0, \quad B = Q^\intercal AQ - S, \tag{11}$$

$$\dot{Q}(t) = Q(t)S(t), \quad Q(0) = \mathbb{Q}_0, \quad S = -S^\intercal, S_{ij} = Q_i^\intercal AQ_j, i > j, \tag{12}$$

where $X(0) = \mathbb{Q}_0\mathbb{R}_0$. On the other hand, it is not hard to prove that (11) and (12) has a unique solution[8] which coincides with the matrices $\mathbb{Q}$ and $\mathbb{R}$. Hence, the following statements are equivalent:

QR1 $X(t) = \mathbb{Q}(t)\mathbb{R}(t)$ is the unique QR-decomposition of X, $\mathbb{Q}(t)\mathbb{Q}^\intercal(t) = I$, and $\mathbb{R}$ is upper-triangular such that $\mathbb{R}_{ii} > 0$, provided X solves (9), and $X(0) = \mathbb{Q}_0\mathbb{R}_0$.
QR2 $\mathbb{Q}$ and $\mathbb{R}$ solve (11) and (12).

Slightly reformulating lemma 1.3.5 from [7, p 21] we introduce the following definition.

Definition 2.1 (forward regularity). Let $\mathbb{R}$ denote the unique solution of (11), and let $R_j(t; r_0)$ denote the jth column of $\mathbb{R}(t)$ such that $R_j(0; r_0) = r_0$, $j = 1 \ldots d$. The function $r \mapsto \kappa(r_0) := \limsup_{t\to\infty} \frac{1}{t}\log(\|R_j(t; r_0)\|)$ is called *forward regular* provided

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t B_{ii}(s)\mathrm{d}s = \liminf_{t\to\infty} \frac{1}{t}\int_0^t B_{ii}(s)\mathrm{d}s. \tag{13}$$

We stress that for bounded $A(t)$, i.e. $\sup_{t\geqslant 0}\|A(t)\| < +\infty$, Perron's lemma guarantees that $\mathbb{R}(t) = \mathbb{Q}^\intercal X(t)$ is a Lyapunov transformation which preserves regularity and Lyapunov exponents [1, p 49, theorem 3.3.1], that is: *$\lambda(\xi_0)$, defined by (7), is forward regular, and its range coincides with the range of $\kappa(r_0)$, provided (13) holds true.* Moreover, if $\lambda(\xi_0)$ is forward regular then [7, p 24, theorem 1.3.1], for any ordered normal Lyapunov basis $\{v_i\}$ of $\mathcal{R}^d$ it holds that:

$$\lambda_i = \lambda(v_i) = \lim_{t\to+\infty} \frac{1}{t}\int_0^t B_{ii}(s)\mathrm{d}s = \lim_{t\to+\infty} \frac{1}{t}\int_0^t Q_i^\intercal(s)A(s)Q_i(s)\mathrm{d}s, \ i = 1, \ldots, d. \tag{14}$$

[6] If the columns of $Y^{(k)} := [Y_1, \ldots, Y_k]$ are linearly-independent, and $Y^{(k)} = Q^{(k)}R^{(k)}$ is the thin QR-decomposition of $Y^{(k)}$, and $Y_{k+1} = Q^{(k)}x$, then we set $Q^{(k+1)} := Q^{(k)}$ and $R^{(k+1)} := [R^{(k)}x]$. If Y_{k+2} is again in the range of $Q^{(k+1)}$ we repeat the above steps, otherwise we add a row of zeros to the bottom of $R^{(k+1)}$ and append a column of size $k+1$ on the right, and $Q^{(k+2)}$ is computed from $[Y^{(k)}Y_{k+2}]$ by using the mGS procedure.

[7] Note that, by construction, $\mathbb{R}$ is upper-triangular, and so is $\dot{\mathbb{R}}$. This implies that B is upper-triangular too, hence the expression for S_{ij} in (12).

[8] Indeed, (11) is linear in R, and B is a continuous function of t, hence (11) has a unique solution. The solutions of (12) are sought on $O(d)$—the compact manifold of orthogonal square matrices, and since the r.h.s. of (12) is Lipschitz in Q w.r.t. the Frobenius norm on $O(d)$, it follows by the Grönwall-Bellman inequality that there is a unique orthogonal square matrix satisfying (12). Therefore, the unique solution of (11) and (12) coincides with the matrices $\mathbb{Q}$ and $\mathbb{R}$.

Finally, we stress that the orthogonalization process ensures that the Lyapunov exponents are ordered: $\lambda_1 \geqslant \cdots \geqslant \lambda_d$. This important fact allows us to directly compute a basis for the non-stable tangent space, i.e. the columns $Q_1, \ldots, Q_k$ of $\mathbb{Q}$ corresponding to $\lambda_1 \geqslant \ldots \lambda_k \geqslant 0$, if we know its dimension $k \leqslant d$ *a priori*, rather than solving (11) and (12) for $\mathbb{Q}$ and $\mathbb{R}$ as suggested above. Indeed, let $X(t) \in \mathcal{R}^{d\times k}$, where $k \leqslant d$, and assume that its thin QR-decomposition is given by:

$$X(t) = Q(t)R(t), \qquad Q(t) \in \mathcal{R}^{d\times k}, \quad R(t) \in \mathcal{R}^{k\times k}.$$

By differentiating the above equality we find that the equations

$$\dot{Q} = (I - QQ^{\intercal})AQ + Q\tilde{S}, \tag{15}$$

$$\dot{R} = \tilde{B}R, \tag{16}$$

can be used to compute $Q \in \mathcal{R}^{d\times k}$ and $R \in \mathcal{R}^{k\times k}$, where as before the skew-symmetric $\tilde{S}^{\intercal} = -\tilde{S} \in \mathcal{R}^{k\times k}$ is chosen to ensure that $\tilde{B} = Q^{\intercal}AQ - \tilde{S} \in \mathcal{R}^{k\times k}$ is upper triangular. If $k = d$, the first term on the right in (15) vanishes, and (15) and (16) reduce to (11) and (12). The key computational advantage here is that a substantial dimension reduction is achieved when solving (15) and (16) instead of (11), provided $k \ll d$.

When $X(0) = Q(0)R(0)$ is of dimension $d \times k$, and assuming that $X(0)$ has nontrivial projection onto the first k elements of some ordered normal Lyapunov basis, the span of the columns of $Q(t)$, found from (15), will coincide with the span of the Lyapunov vectors corresponding to the leading k Lyapunov exponents. The latter are gleaned from the diagonal of $\tilde{B}(t)$ via the relation (14).

3. Detectability for linear systems

In this section we recall the notions of *detectability* and *linear observers* for linear autonomous systems, and we reveal the fundamental relationship between detectability, Lyapunov exponents and existence of linear observers (lemma 3.4). This relationship is a well-known fact from the classical control theory, and we reformulate it here in terms of the QR-decomposition discussed in section 2.1 to illustrate using a simple example how to design a linear observer by using Lyapunov vectors and Lyapunov exponents. In particular, we prove (see lemma 3.4 and proposition 3.7) that the linear observer exists if and only if the system is detectable, and that for non-detectable systems it is impossible, in principle, to design a linear observer as there exists an initial condition such that the corresponding estimation error does not converge to 0. This simple yet powerful observer design tool is then used in section 4 to construct an observer for generic nonlinear dynamical systems.

3.1. Autonomous linear systems

Let us recall the definition of a linear observer, a so-called Luenberger observer. Assume that we observe a function $y(t) \in \mathcal{R}^s$ which represents a possibly incomplete output of a linear system:

$$\begin{aligned} &\dot{z}(t) = Az(t), \quad z(0) = v \in \mathcal{R}^d, \\ &y(t) = Hz(t). \end{aligned} \tag{17}$$

Definition 3.1. The following non-homogeneous linear system

$$\dot{x} = Ax + L(y - Hx), \quad x(0) = x_0 \tag{18}$$

is called a *linear observer* with gain L, the Luenberger gain, provided L is chosen so that the estimation error $\xi := z - x$ converges to 0 exponentially for any x_0.

Remark 3.2. Suppose that $y(t) = Hz(t) + \eta(t)$ for some $\eta \in L^2(0,T)$, and define $J(v) := \int_0^T \|y(t) - He^{At}v\|^2 \,\mathrm{d}t$. Clearly, the best least-squares estimate of the trajectory z given $y(t)$ can be computed by minimizing J w.r.t. $v \in \mathcal{R}^d$ and choosing $z(t) = \exp(At)v$ for this minimizer. Define $W(0,T) \triangleq \int_0^T e^{A^\intercal t}H^\intercal He^{At}\mathrm{d}t$ and set $b \triangleq \int_0^T e^{A^\intercal t}H^\intercal y(t)\mathrm{d}t$. We find that $J(v) = \|y\|^2_{L^2(0,T)} - 2v^\intercal b + v^\intercal W(0,T)v$, and thus the set of minimizers of J is given by $\{v : W(0,T)v = b\}$. For instance, the minimizer with smallest 2-norm can be expressed using the (Moore–Penrose) pseudoinverse $W^\dagger$ of W, i.e. $v^\dagger = W^\dagger(0,T)b$. It is not hard to see that $J(v^\dagger) = \|\eta\|^2_{L^2(0,T)} + \widetilde{\eta}^\intercal W^\dagger(0,T)\widetilde{\eta} + 2\widetilde{\eta}^\intercal(I - W^\dagger(0,T)W(0,T))z(0)$, where $\widetilde{\eta} \triangleq \int_0^T e^{A^\intercal t}H^\intercal\eta(t)\mathrm{d}t$. If $\eta = 0$, one can reconstruct z exactly if and only if $W(0,T)$ is of full rank: indeed, $J(v^\dagger) = 0$ in this case. If $\eta = 0$ but $W(0,T)$ is not of full rank, any solution v of $W(0,T)v = b$ can be represented as $v = v^\dagger \oplus v_0$ where $W(0,T)v_0 = 0$. We stress that v_0 cannot be determined from $W(0,T)v = b$. In fact, $He^{At}v_0 = 0$, and so the observed data y does not contain any information about the null-space of $W(0,T)$. In the control literature, this sub-space is sometimes referred to as an unobservable subspace. Finally, if $\eta \neq 0$ one cannot reconstruct z exactly, independent of the rank of W: the best estimate of z will have the mean-squared error $J(v^\dagger)$.

The unobservable subspace can be computed efficiently for the case of linear autonomous systems. Define

$$O^s := \begin{bmatrix} H \\ HA \\ \vdots \\ HA^s \end{bmatrix}, \quad s \text{ the smallest integer such that } \operatorname{rank}(O^s) = \operatorname{rank}(O^{s+p}), \forall p \geqslant 1.$$

In turns out (see proof of lemma 3.4) that $\ker W(0,T) = \ker O^s$. Note that the state space of (17) splits into two invariant sub-spaces: $\mathcal{R}^d = \ker O^s \oplus \ker^\perp O^s$, and that the part of any trajectory (17) in $\ker^\perp O^s$, i.e. the projection of the state vector $z(t)$ onto $\ker^\perp O^s$, can be recovered from the data y. The 'invisible' part, i.e. the projection of $z(t)$ onto $\ker O^s$, cannot be recovered unless the matrices A and H have a special structure:

Definition 3.3. (A,H) is *detectable* if $\lim_{t\to\infty} \|e^{At}v\| = 0$ for any $v \in \ker(O^s)$.

Detectability implies that the projection of $z(t)$ onto $\ker O^s$, the 'invisible' part, decays to zero exponentially fast. The following lemma establishes the connection between detectability, existence of linear observers and Lyapunov exponents.

Lemma 3.4. *(A,H) is detectable if and only if all the Lyapunov exponents corresponding to $\ker(O^s)$ are negative, i.e. $\lambda(v) < 0, \forall v \in \ker(O^s)$. If (A,H) is detectable then there exists a linear observer (18) in the sense of definition 3.1. If, on the contrary, there exist a vector $v^\dagger \in \ker(O^s)$ such that the corresponding Lyapunov exponent is non-negative, $\lambda(v^\dagger) \geqslant 0$, then for any gain matrix L, the estimation error $\xi = z - x$ either grows unbounded, i.e. $\lim_{t\to+\infty} \|e(t)\| = +\infty$, or stays bounded but can be made arbitrarily large.*

3.2. Non-autonomous linear systems

In this subsection we generalize lemma 3.4 to the case of time-dependent matrices A and H. Specifically, a time-variant version of (17) takes the following form:

$$\begin{aligned} \dot{z}(t) &= A(t)z(t), \quad z(0) = v, \\ y(t) &= H(t)z(t). \end{aligned} \tag{19}$$

In what follows we generalize definition 3.3 to the case of time-dependent A and H, and then this definition is applied to design the linear observer for (19). Assume that $t \mapsto H(t) \in \mathcal{R}^{s\times d}$, $s \leqslant d$ and $t \mapsto A(t) \in \mathcal{R}^{d\times d}$ are such that

$$\sup_{t\geqslant 0}\sup_{\|x\|=1}\|A(t)x\| < +\infty, \quad \sup_{t\geqslant 0}\sup_{\|x\|=1}\|H(t)x\| < +\infty. \tag{20}$$

Definition 3.5. Assume that (13) holds true, and let $Q \in \mathcal{R}^{d\times k}$, $k \leqslant d$ solve (15). Let $\widetilde{Q}(t) \in \mathcal{R}^{d\times k}$ and $\widetilde{R}(t) \in \mathcal{R}^{k\times k}$ be the thin QR-decomposition of $H^{\intercal}HQ$, i.e.

$$\widetilde{Q}(t)\widetilde{R}(t) = H^{\intercal}(t)H(t)Q(t).$$

We say that (A, H) is detectable in the direction Q_j if

$$\limsup_{t\to+\infty}\frac{1}{t}\int_0^t \widetilde{R}_{jj}(s)\,\mathrm{d}s > 0.$$

Furthermore we say that (A, H) is detectable if it is detectable in any direction Q_j from the non-stable tangent space:

$$\lambda_j = \lim_{t\to+\infty}\frac{1}{t}\int_0^t B_{jj}(s)\,\mathrm{d}s = \lim_{t\to+\infty}\frac{1}{t}\int_0^t Q_j(s)^{\intercal}A(s)Q_j(s)\,\mathrm{d}s \geqslant 0. \tag{21}$$

Remark 3.6. Note that $\widetilde{R}_{jj} \geqslant 0$ by construction, with equality $\widetilde{R}_{jj}(t) = 0$ holding only if $H^{\intercal}(t)H(t)Q(t)$ is rank deficient[9] The latter is the case if and only if the linear sub-space generated by the columns of Q, $\{Q(t)x, x \in \mathcal{R}^k\}$ has a non-trivial intersection with $\ker H^{\intercal}(t)H(t)$: $H^{\intercal}(t)H(t)Q(t)x = 0$ for some $x \in \mathcal{R}^k$. Recall from section 2.1 that the way we compute Q ensures that the Lyapunov exponents are ordered: $\lambda_1 \geqslant \cdots \geqslant \lambda_{k^\star}$, and the $k^\star$ leading columns of $Q(t)$, $Q_1, \ldots, Q_{k^\star}$ correspond to the $k^\star$ leading Lyapunov exponents as in (21). Now, if $k^\star$ is the number of non-negative Lyapunov exponents of (19), then, according to definition 3.5, $(A(t), H(t))$ is detectable iff the sub-space generated by the first $k^\star$ columns of Q, i.e. the non-stable tangent space, has only trivial intersection with $\ker H^{\intercal}(t)H(t)$ most of the time, i.e. the measure of the set $\{0 \leqslant s \leqslant t : \widetilde{R}_{jj}(s) > 0\}$ grows at least linearly as $t \to +\infty$, outside perhaps a *finite* number of compact intervals of $\mathcal{R}^+$, for $j = 1, \ldots, k^\star$.

Intuitively, detectability requires that the non-stable tangent space, spanned by $Q_1, \ldots, Q_{k^\star}$, is regularly observed as $t \to \infty$. If, however, there is a vector Q_ℓ in the non-stable tangent space that is 'unseen' by the observation operator H most of the time, i.e. the measure of the set $\{0 \leqslant \tau \leqslant t : R_{\ell\ell}(\tau) > 0\}$ is finite or grows at a sub-linear rate as $t \to +\infty$, then $(A(t), H(t))$ is not detectable.

[9] In fact, $\tilde{R}_{ii} = 0$ if and only if the ith column of $H^{\intercal}(t)H(t)Q(t)$ can be represented as a linear combination of the first $i-1$ columns. In this case, the ith column of $\tilde{Q}$ is set to 0 by the mGS process.

Remark 3.6 suggests that *a necessary condition for detectability of* $(A(t), H(t))$ is

$$\min\{\text{rank}(H(t)), k\} \geqslant k^\star, \text{ for almost all } t \in (\mathcal{R}^+ \setminus K), \tag{22}$$

provided K is a union of a finite number of compact intervals of $\mathcal{R}^+$, and $k^\star$ is the number of nonnegative Lyapunov exponents of (19). Indeed, $\ker H^\intercal(t)H(t) = \ker H(t)$ as the range of H is orthogonal to $\ker H^\intercal$. Hence, $\text{rank}H^\intercal HQ = \text{rank}HQ$. By construction, $Q(t)$ has full column rank k. Hence, $\text{rank}H^\intercal(t)H(t)Q(t) = \text{rank}H(t)Q(t) = \min\{\text{rank}(H(t)), k\}$ at any time t such that $\ker(H(t)Q(t)) = \{0\}$. In the latter case, $\text{rank}\widetilde{R}(t) = \min\{\text{rank}(H(t)), k\}$. Thus, the diagonal of $\widetilde{R}(t)$ may have at most $\min\{\text{rank}(H(t)), k\}$ positive elements for any $t > 0$. As noted above, detectability of $(A(t), H(t))$ implies that, for any $j = 1, \ldots, k^\star$, the measure of the set $\{0 \leqslant s \leqslant t : \widetilde{R}_{jj}(s) > 0\}$ grows at least linearly as $t \to +\infty$, outside perhaps a *finite* number of compact intervals of $\mathcal{R}^+$, K. Hence, there exist $T > 0$ such that $\widetilde{R}_{jj}(s) > 0$ for all $j = 1, \ldots, k^\star$, and almost all $s > T$. But this is possible only if (22) holds true. In particular, (22) implies that *the rank of the observation operator* $H(t)$ *must equal or exceed the number of nonnegative Lyapunov exponents for detectability to hold.*

The following proposition relates detectability to the existence of the linear observer.

Proposition 3.7. *Assume that* $\eta = 0$, A *and* H *satisfy (20), and (13) holds true. Let* $Q \in \mathcal{R}^{d\times k}$ *and* $R \in \mathcal{R}^{k\times k}$ *solve (15) and (16), provided* k *is the number of non-negative Lyapunov exponents of (9), and let* $\widetilde{Q}\widetilde{R} = H^\intercal HQ$ *be the thin QR-decomposition of* $H^\intercal HQ(t)$. *Define the gain*

$$L := p\, Q\, \widetilde{Q}^\intercal H^\intercal, \quad p > 0, \quad H^\intercal HQ = \widetilde{Q}\widetilde{R}, \tag{23}$$

and let x *solve the following system:*

$$\dot{x} = A(t)x + L(t)(y(t) - H(t)x(t)), \quad x(t) = 0. \tag{24}$$

Then there exists $p > 0$ *such that the estimation error* $\xi = z - x$ *decays to* 0 *exponentially fast if and only if* (A, H) *is detectable in the sense of definition 3.5.*

Remark 3.8. Note that the gain L is designed so that all the non-negative LEs of $A(t) - L(t)H(t)$ are made negative. In other words, the gain L (asymptotically) stabilizes the error equation $\dot{e} = (A(t) - L(t)H(t))e$, provided the observations are exact. The latter is possible iff $(A(t), H(t))$ is detectable as per definition 3.5.

The proof of proposition 3.7 is given in the appendix.

4. Approximation of the minimax filter for nonlinear ODEs

As noted in the introduction, for linear systems the minimax/Kalman filter uniformly (in time) converges to the linear observer, provided the observational noise/model error 'disappears' as $t \to \infty$; see [4]. In the general case of nonlinear dynamics and only bounded observational noise, observers represent an approximation of the minimax filter. In this section we design an observer x for a nonlinear ODE

$$\dot{z} = f(t, z), \quad z(t) \in \mathcal{D} \subset \mathcal{R}^d, \quad z(0) = z_0, \tag{25}$$

with compact state space $\mathcal{D}$, given incomplete and noisy[10] observations:

$$y(t) = H(t)z(t) + \eta(t), \quad H : \mathcal{R}^{s\times d}, s \leqslant d, \quad \eta \in L^2(0,+\infty), \quad t \geqslant 0. \tag{26}$$

Assuming that (25) has the unique solution for any $z_0 \in \mathcal{D}$ we prove that, locally the estimation error $\xi = z - x$ decays to 0 exponentially iff there exists a gain $L(t,x)$ such that all the LEs of $A - LH$ are negative, provided $A(t) := \mathrm{D}f(x(t))$, and $\eta = 0$. If the function λ defined by (7) is, in addition, forward regular, then the LEs of $A - LH$ are negative iff A, H is detectable, and the gain L can be constructed as per the recipe of proposition 3.7. We study the case of non-trivial noise η numerically in the following section.

The following proposition utilizes the Lyapunov stability theorem [7, p 29, T 1.4.3] to prove that under mild regularity conditions a vicinity of 0 is attracted to $\xi \equiv 0$ by the error dynamics. Define $N(t,\xi,x) := (N_1(t,\xi,x)\dots N_d(t,\xi,x))^\intercal$ where

$$N_i(t,\xi,x) := \xi^\intercal \left(\int_0^1 \int_0^1 s\mathrm{D}^2 f_i(x + \tau s\xi, t)\mathrm{d}s\,\mathrm{d}\tau \right) \xi. \tag{27}$$

(See appendix for a derivation of the N_i.) We require the following assumption on the linearized dynamics of the observer.

Assumption 4.1. *The following conditions hold for the Jacobian* $\mathrm{D}f(t,x(t))$ *evaluated along the observer process:*

A1 $A(t) := \mathrm{D}f(t,x(t))$ *is bounded:* $\sup_{t\geqslant 0}\sup_{\|x\|=1}\|A(t)x\| < +\infty$.
A2 $\max_{x\in K}\|D^2 f_i(x,t)\| \leqslant C_i(K) < +\infty$ *for every compact subset* K *of* $\mathcal{R}^d$.
A3 the Lyapunov exponent $\lambda(\xi)$ *of* $\dot{X} = A(t)X$ *is forward regular.*

Theorem 4.2. *Let* $\eta \equiv 0$ *and assume the conditions of assumption 4.1 hold. Let* x *be the unique solution of the following system:*

$$\dot{x} = f(t,x) + L(t,x)(y(t) - H(t)x(t)), \qquad x(0) = x_0. \tag{28}$$

Then the following statements are equivalent:

S1 There exist $\varepsilon > 0$ *and* $a > 0$ *such that* $\|\xi(t)\| \leqslant C\mathrm{e}^{-at}\|\xi(0)\|$ *for all* $t \geqslant 0$ *where* $\xi = z - x$ *is the unique solution of the error equation:*

$$\dot{\xi} = (A(t) - L(t,x)H(t))\xi + N(t,\xi,x(t)), \quad \xi(0) = \xi_0, \|\xi_0\| < \varepsilon. \tag{29}$$

S2 All Lyapunov exponents of $\dot{X} = (A - LH)X$ *are negative.*

Remark 4.3. In fact, theorem 4.2 establishes robustness of the LEs w.r.t. small perturbation. Indeed, one may interpret $\xi(t)$ as a perturbation about a particular solution x, i.e. $z(t) = x + \xi$, and the dynamics of this perturbation is defined by the 'perturbed' equation (29): here N is considered as a perturbation of the linear unperturbed equation $X = (A - LH)X$. By theorem 4.2 it follows that the exponential decay of solutions of the 'perturbed' equation (29) is equivalent to the exponential decay of the solutions of the unperturbed linear equation provided the LEs are forward-regular. Osedelec's theorem [37] establishes regularity for a wide class of nonlinear systems possessing an ergodic invariant measure. For a nonlinear system with a global ergodic attractor, the LEs are independent of any particular trajectory x.

[10] $t \mapsto \eta(t) \in \mathcal{R}^s$ is a bounded square-integrable function such that $\|\eta(t)\| \leqslant \varepsilon$.

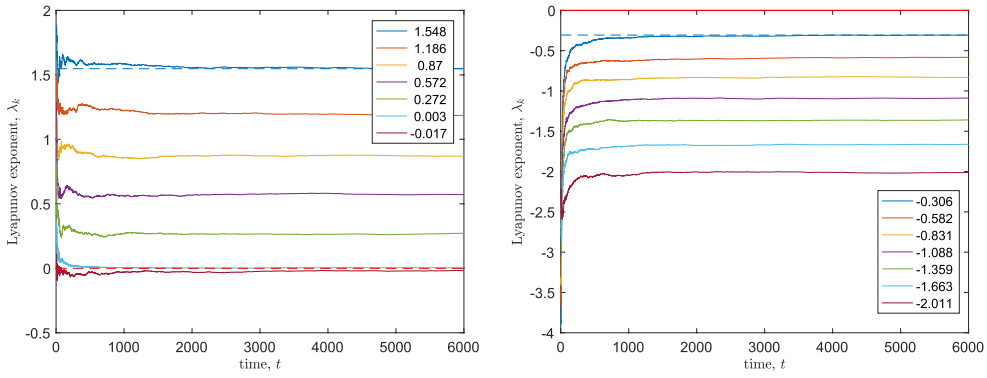

Figure 1. Leading Lyapunov exponents of the Lorenz '96 model (33) (left) and the tangent space dynamics (29) (right).

Corollary 4.4. *Assume that all the conditions of theorem 4.2 hold and let x solve*

$$\dot{x} = f(t,x) + L(t,x)(y(t) - H(t)x(t)), \qquad L = p\,Q\widetilde{Q}^{\mathsf{T}}H^{\mathsf{T}}, p > 0, \quad x(0) = x_0 \tag{30}$$

$$\dot{Q} = (I - QQ^{\mathsf{T}})\mathrm{D}f(t,x)Q + QS, \qquad Q(0) = Q_0 \in \mathcal{R}^{d\times k}, \tag{31}$$

$$S = -S^{\mathsf{T}}, S_{ij} = (Q^{\mathsf{T}}\mathrm{D}f(t,x)Q)_{ij}, i > j, \qquad H^{\mathsf{T}}HQ = \widetilde{Q}\widetilde{R} \tag{32}$$

for appropriate $k \in 1,\dots,d$. Then the estimation error $\xi = z - x$ converges to zero exponentially if and only if (A,H) is detectable in the sense of definition 3.5.

5. Numerical experiments with noisy observations

In this section we apply (30) to the Lorenz '96 (L96) model and Burgers equation, and compare it to the extended Kalman(–Bucy) filter (ExKF). We consider both exact and noisy observations. Abusing the standard control terminology we will refer to (30) as a filter to stress that in the experiments the observations are allowed to be noisy.

5.1. L96

The Lorenz '96 model is a system of ODEs

$$\dot{z}_i = -z_{i-2}z_{i-1} + z_{i-1}z_{i+1} - z_i + \mathcal{F}, \quad z_i(0) = \sin\left(2\pi\frac{i-1}{d}\right), \; i = 1,\dots,d, \tag{33}$$

defined on a periodic lattice with state space dimension $d = 18$, and constant forcing $\mathcal{F} \equiv 8$. We solve (33) using a fourth order explicit Runge–Kutta method (RK4) with the time step $\Delta t = 0.01$ to sample the observations y. The rows of the matrix H are taken to be the first k eigen-vectors of the discrete Laplacian (i.e. $d \times d$ circulant matrix having tridiagonal elements $\begin{bmatrix}1 & -2 & 1\end{bmatrix}$). As per estimate (22), the rank of the observation operator H, k must be greater or equal to the number of nonnegative LEs of $\dot{X} = \mathrm{D}f(x(t))X$ along the trajectory x of (30). We computed RK4 approximations of these LEs by averaging over the interval $t \in [0, 6000]$: on the left panel of figure 1 one can see that 6 leading exponents are nonnegative,

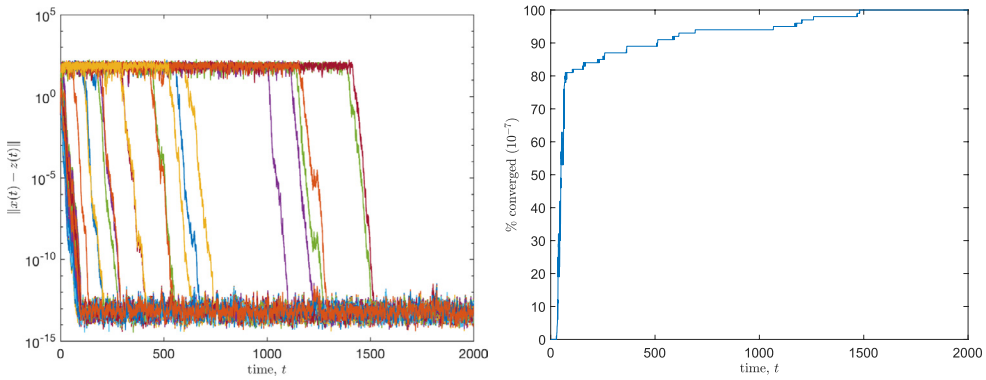

Figure 2. Convergence of filter (30) for the Lorenz '96 model (33) with $k = 7$. Left, the errors $\|\xi(t)\|$ for a 100-member ensemble of perturbed initial conditions. Right, the number of ensemble members converged to tolerance $\|\xi(t)\| < 10^{-7}$ at time t.

but the 7th exponent, $\lambda_7 \approx -0.017$. In addition, for $k = 6$ the approximation of the leading exponent of the error dynamics (29) equals -0.051. The latter two observations suggest that the choice $k = 6$ may lead to very slow convergence or convergence only within a very small neighborhood of the truth $x(t)$, and, for the 'boundary' case $k = 6$, the convergence is very sensitive both to the size of the initial perturbation, and to the accuracy with which one is able to approximate the basis $Q(t)$ of the non-stable tangent space. To alleviate this sensitivity, we set the rank of H to $k = 7$. The latter ensures that indeed, as per condition S2 of theorem 4.2, all the exponents of the error dynamics (29) are negative, and, more importantly, they are well separated from 0; see the right panel of figure 1.

For the case $k = 7$ we generate an ensemble of 100 initial conditions $x_0 = z(0) + 0.01\eta$, $\eta_i \sim \mathcal{N}(0, 1)$, $i = 1, \ldots, d$. For each x_0 we set $p = 10$, and integrate (30) using the same time integrator until $\|x - z\| \leqslant 10^{-14}$. The resulting 2-norm estimation error $\|\xi(t)\| = \|x(t) - z(t)\|$ as a function of time are shown in the left panel of figure 2. We see that all samples ultimately converge at an exponential rate, as predicted by the theory of this paper. However, in some cases there is a long delay before exponential convergence is observed. In the right panel of figure 2 we plot the number of ensemble members that has converged to within tolerance $\|\xi(t)\| < 10^{-7}$ at time t. We see that 80% of the ensemble converges by time $t = 100$, the remaining 20% converges more slowly, with the last ensemble member converging only at time $t = 1500$.

To make the convergence of (30) faster and even less sensitive to the size of the initial perturbation, and to the errors of RK4 approximation of the basis $Q(t)$, we set $k = 8$ and increase the amplitude of the initial perturbation to $x_0 = z(0) + 0.1\eta$, $\eta_i \sim \mathcal{N}(0, 1)$, $i = 1, \ldots, d$ so that $\frac{\|z_0 - x_0\|}{z_0} \approx 0.3$, i.e. the magnitude of the perturbation is up to 30%. Figure 3 demonstrates the convergence. Even with a much larger initial perturbation, all ensemble members converge to machine precision within time $t = 200$, and more than 95% converge to within $\|\xi(t)\| < 10^{-7}$ by time $t = 100$.

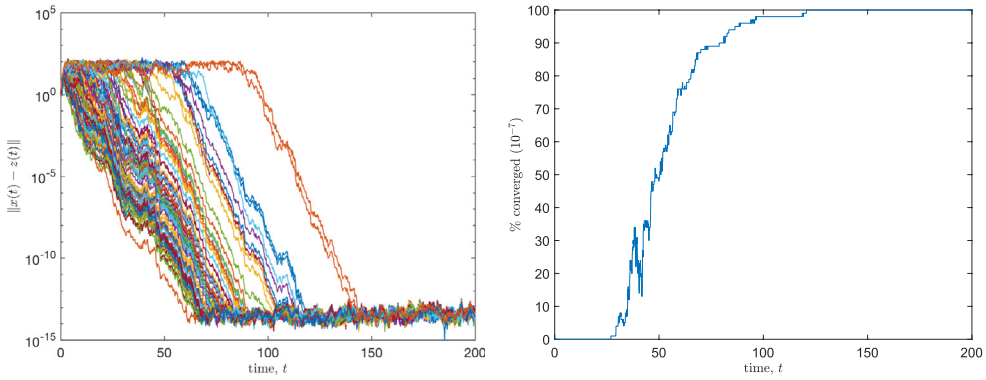

Figure 3. Convergence of filter (30) for the Lorenz '96 model (33) with $k = 8$. Left, the errors $\|\xi(t)\|$ for a 100-member ensemble of perturbed initial conditions. Right, the number of ensemble members converged to tolerance $\|\xi(t)\| < 10^{-7}$ at time t.

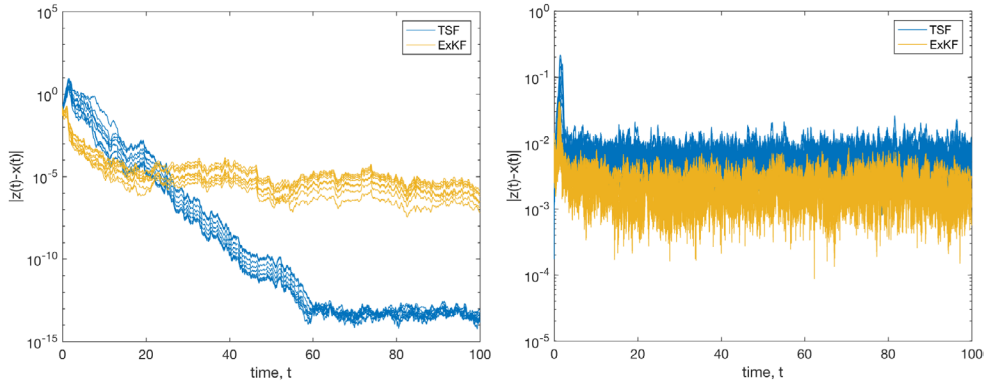

Figure 4. Comparison of the tangent splitting filter (TSF) (30) and the ExKF (34) for the Lorenz '96 model (33) with $k = 8$. Left, the errors $\|\xi(t)\|$ for a 10-member ensemble of perturbed initial conditions. Right, the errors $\|\xi(t)\|$ for a 10-member ensemble with random observational error.

5.1.1. Comparisons with ExKF. Recall from [34] that the Kalman–Bucy filter for linear systems (with no model error), i.e. $f(t,m) = A(t)m$ is equivalent to the minimax filter:

$$\begin{aligned} &\dot{m} = f(t,m) + PH^{\intercal}C(y - Hm), \quad m(0) = x_0, \\ &\dot{P} = \mathrm{D}f(t,m(t))P + P\mathrm{D}f(t,m(t))^{\intercal}(t) - PH^{\intercal}CHP \quad P(0) = P_0, \\ &x_0^{\intercal}P_0x_0 + \int_0^T \eta^{\intercal}(t)C(t)\eta(t)\mathrm{d}t \leqslant 1. \end{aligned} \tag{34}$$

Instead of dealing with stochastic differential equations, for which the Kalman–Bucy filter is formulated, we stay within the (equivalent) deterministic framework. Namely, we take η to be a measurable function satisfying the inequality in (34). We discretize (34) using the explicit fourth order Runge–Kutta method[11], reducing the step size to $\Delta t = 0.001$ to ensure stability on the time interval $T = 100$. We begin with noise-free observations, and use an ensemble of

[11] A more appropriate integrator for (34) is an implicit symplectic integrator as detailed in [25, 46]. However for the sake of comparison we retain the explicit Runge–Kutta method here.

10 initial conditions, $x_0 = z(0) + 0.01\eta$. In this case, to make sure that the error in the initial condition satisfies the inequality in (34), we set $P(0) = \frac{1}{4d\times 10^{-4}} I \approx 139I$ and $C = I$, as $\|\sigma\eta\|^2$ is χ^2-distributed with mean $k\cdot\sigma^2$. Clearly, large $P(0)$ and C represent high trust in the initial condition and observations. As seen in the left panel of figure 4, the estimation error of the ExKF estimates is around 10^{-5} at the end of the interval, whereas the error of (30) decays to machine precision.

5.1.2. Noisy observations. We next simulate a 10-member ensemble for which the initial condition is exact, $x_0 = z_0$, but the observations $y(t)$ are perturbed by random noise at each time step, $y(t) = Hz(t) + 0.01\eta$. Note that the expected value of the norm of the observational noise is given by: $\mathbb{E}\,\sigma\|\eta\| \approx \frac{\sigma\sqrt{2k-1}}{\sqrt{2}}$ for k large. As demonstrated in the right panel of figure 4, the estimation errors $\|x - z\|$ of both the filter (30) and the ExKF (34) are less than the mean of the norm of the observational noise given by $\mathbb{E}\,\sigma\|\eta\| \approx 0.1275$, on the interval $(0, 100)$. The ensemble mean estimation errors level off at around 10^{-2} for both methods, with the averaged error of the filter (30) being approximately 0.01 which is twice that of the averaged error of ExKF, 0.005 (see (34)). Recall from remark 3.2 that this is as good as can be hoped for linear systems in the presence of noisy observations.

We stress that the observational noise introduces an additional nonlinear term $pQ\widetilde{Q}^\intercal H^\intercal\eta$ in the error equation (29); hence, large $p > 0$ not only makes the discretized equation stiff but, importantly amplifies the noise! On the other hand, $\widetilde{Q}^\intercal$ acts as a projection onto the range of $H^\intercal HQ$, and thus $\widetilde{Q}^\intercal H^\intercal\eta$ in fact represents the projection of η onto the range of HQ. Hence, the norm of $Q\widetilde{Q}^\intercal H^\intercal\eta$ is not increasing, and it may even be 0 if η is not in the range of HQ. The experiment shows that the amplification provided by $p = 10$ is minor.

5.2. Burgers equation

As a second example, we discretize the Burgers(–Hopf) equation $u_t = -uu_x$ using the finite difference scheme:

$$\dot{u} = f(u), \qquad \dot{u}_i = -\frac{1}{6\Delta x}\left(u_i(u_{i+1} - u_{i-1}) + (u_{i+1}^2 - u_{i-1}^2),\right) \tag{35}$$

again taken on a periodic lattice ($\Delta x = 2\pi/d$), which has the properties that (i) the quadratic energy $\sum_i u_i^2$ is conserved, implying that every sphere in $\mathcal{R}^d$ is invariant under the motion of the system and $\|u\|$ is constant, and (ii) $\mathrm{tr}(\mathrm{D}f(u)) = 0$, implying that the flow conserves the volume of the phase element. Consequently, this equation is not dissipative in contrast to the *L*96 system, and so the effects of error in the initial condition, and the observational noise are expected to be more pronounced. On the other hand, it is unclear if this system possesses an invariant ergodic measure making it subject to the conditions of the Oseledec theorem, so our theory may not apply to this test case.

5.2.1. Noise-free observations. We set $d = 18$ and take $z_i(0) \sim U(0, 1)$, $i = 1, \ldots, d$. For the filter (30) we again take H to be the eigenvectors of the discrete Laplace operator corresponding to the k leading eigenvalues. The filter (30) was integrated by RK4 with $\Delta t = 0.01$ and $p = 20$ on the interval $t \in [0, 400]$. For $k = 9$ we found that $\dot{X} = \mathrm{D}f(x(t))X$ possesses 10 nonnegative exponents, and tangent dynamics $\dot{\xi} = (A(t) - L(t, x)H^\intercal(t)H(t))\xi$ has 1 nonnegative exponent. Taking $k = 10$ we observed convergence, i.e. $\|x - z\| \leqslant 10^{-14}$, provided $x_0 = z(0) + 0.0005\eta$. This suggests that the basin of attraction of the trivial solution of (29) is rather small. However, for $k = 11$ this basin increases significantly: figure 5 demonstrates

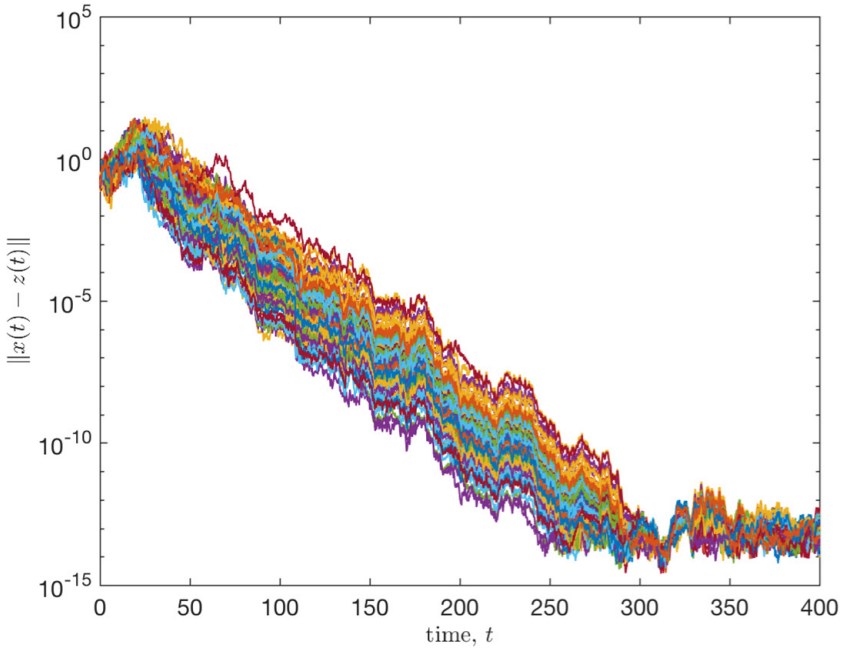

Figure 5. Convergence of filter (30) for the discretized Burgers equation (35) with $k = 11$, showing the errors $\|\xi(t)\|$ for a 100-member ensemble of perturbed initial conditions.

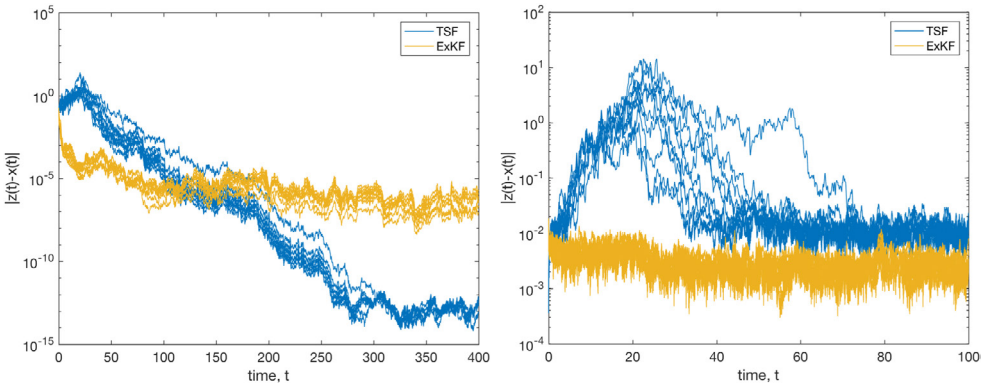

Figure 6. Comparison of the filter (30) and the ExKF (34) for the discretized Burgers equation (35) with $k = 11$. Left, the errors $\|\xi(t)\|$ for a 10-member ensemble of perturbed initial conditions. Right, the errors $\|\xi(t)\|$ for a 10-member ensemble with random observational error.

the exponential decay of the estimation error for an ensemble of 100 initial conditions $x_0 = z(0) + 0.01\eta$.

5.2.2. Comparisons with ExKF. To compare with the ExKF method (34) we again reduce the step size to $\Delta t = 0.001$. We begin with noise-free observations, and use an ensemble of 10 initial conditions, $x_0 = z(0) + 0.01\eta$. In this case, to make sure that the error in the initial

condition satisfies the inequality in (34), we set $P(0) = \frac{1}{4d\times 10^{-4}} I \approx 139I$ and $C = I$, as $\|\sigma\eta\|^2$ is χ^2-distributed with mean $k \cdot \sigma^2$. As seen in the left panel of figure 6, the error of the ExKF estimates is again around 10^{-5} at the end of the interval whereas the error of (30) decays to machine precision.

5.2.3. Noisy observations. We simulate a 10-member ensemble for which the initial condition is exact, $x_0 = z_0$, but the observations $y(t)$ are perturbed by random noise at each time step, $y(t) = Hz(t) + 0.01\eta$. As shown in the right panel of figure 6 the ensemble mean estimation error levels off at around 2×10^{-2}. Here the mean of the norm of the observational noise is 0.033. The convergence of the filter (30) is irregular at the beginning of the interval, probably due to slow convergence of $Q(t)$ to the basis of the nonstable tangent space from its random initial condition. The error in the ExKF method is ultimately smaller than that of the filter (30) by a factor three.

6. Concluding remarks

In this paper we have drawn an explicit connection between the dimension of the nonstable tangent space of a continuous dynamical system, as quantified by the number of nonnegative Lyapunov exponents, and the necessary dimension of the (time-variant) observation operator H of a sequential data assimilation process. We formulated a detectability condition that, when satisfied, provides a necessary and sufficient condition for convergence of the new filter (30) that utilizes an explicit partition of the tangent space into stable and nonstable subspaces. The new filter is comparable to the extended Kalman filter for perturbed initial conditions and noisy observations, and appears to be more robust in the sense that convergence is observed for an order of magnitude larger than the time step.

Acknowledgment

The authors express gratitude to Erik van Vleck, Alberto Carrassi and Colin Grudzien for insightful discussions.

Appendix. Proofs

Proof of lemma 3.4. The first statement follows directly from definition 3.3 and (7). Let us prove the second statement. Recall that $\mathcal{R}^d = \ker O^s \oplus \ker^{\perp} O^s$ and take $w \in \ker^{\perp} O^s$. Definition 3.3 implies that $\int_0^T \|He^{At}w\|\mathrm{d}t > 0$ for any $T > 0$. Hence, $H\mathbb{Q}_i(t) \neq 0$ for any $\mathbb{Q}_i$ from the non-stable tangent subspace of $\dot{X} = AX$; see (21). The latter shows that definition 3.3 implies definition 3.5, which is equivalent to the existence of the observer by proposition 3.7.

Let us prove the last statement. Take $w \in \ker O^s$ and assume that $e^{At}w$ does not decay, and, for some L, the observer exists. Note that $y(t) = He^{At}w = 0$. Hence, by definition 3.1, the solution of (18) converges to 0 for any x_0 as in this case (18) coincides with the error equation $\dot{\xi} = (A - LH)\xi$, and, by definition 3.1, the solution of the latter decays to 0 exponentially fast for any $\xi(0)$. But this contradicts the original assumption ($e^{At}w$ does not decay!). This contradiction proves the last statement of the lemma.

Let us also prove that $\ker(W(0,T)) = \ker(O^s)$. Indeed,

$$W(0,T)v = 0 \Leftrightarrow \int_0^T \|He^{At}v\|^2_{R^p}\mathrm{d}t = 0 \Leftrightarrow \|He^{At}v\|^2_{R^p} \equiv 0,$$

and since $t \mapsto He^{At}v$ is a smooth function, it follows that $He^{At}v = 0$ for any $t \in [0,T]$. By differentiating the latter equality we find that $HA^j e^{At}v = 0$ for any $j \geqslant 0$ and $t \in [0,T]$. Setting $t = 0$ we get that $W(0,T)v = 0$ implies $O^s v = 0$. On the contrary, by definition of s we have that

$$(A^{s+p})^\intercal H^\intercal \in \mathrm{span}\{H^\intercal, A^\intercal H^\intercal, \ldots, (A^\intercal)^s H^\intercal\}, \quad p \geqslant 1.$$

Hence, $O^s v = 0$ implies that $v^\intercal (O^s)^\intercal = 0$ and so $v^\intercal (A^j)^\intercal H^\intercal = 0$ for any $j \geqslant 0$. Therefore, $v^\intercal e^{A^\intercal t} H^\intercal = 0$. ■

Proof of proposition 3.7. Note that $\dot{\xi} = (A(t) - L(t)H(t))\xi$, and A, H are bounded matrix-valued functions by (20). Hence, ξ decays to 0 exponentially fast iff the LEs $\mu_1 \geqslant \cdots \geqslant \mu_d$ of

$$\dot{W} = (A(t) - L(t)H(t))W, \tag{A.1}$$

satisfy $\mu_1 < -\kappa < 0$ (see [7, p 6]). We stress that the forward regularity is not required for the latter statement to hold as per (8). However, it becomes important when we invoke definition 3.5 to prove that ξ decays to 0 exponentially fast. Our proof is based on the following simple observation: if X solves (9), L satisfies (23), and (13) holds, then:

(U) if $X(t) = \mathbb{Q}(t)\mathbb{R}(t)$ and $\mathbb{R}_1 = \mathbb{Q}^\intercal W$ then $\mathbb{R}_1(t)$ is upper-triangular with positive diagonal, i.e. $\mathbb{Q}\mathbb{R}_1$ represents the unique QR-decomposition of W.

Assume for now that (U) holds true. Recall from (14) that, given the unique QR-decomposition of X, i.e. $X = \mathbb{Q}\mathbb{R}$, one can compute the ith Lyapunov exponent of (9), λ_i by evaluating the limit of the quantity $\frac{1}{t}\int_0^t Q_i^\intercal A Q_i \mathrm{d}s$, which depends only on A and the ith column of $\mathbb{Q}$, Q_i. Hence, by (U), the ith Lyapunov exponent of (A.1), μ_i, depends only on $A - LH$ and the same Q_i:

$$\mu_i = \lim_{t\to+\infty} \frac{1}{t}\int_0^t Q_i^\intercal (A - LH) Q_i \mathrm{d}s = \lambda_i(X) - \lim_{t\to+\infty} \frac{1}{t}\int_0^t Q_i^\intercal LH Q_i \mathrm{d}s.$$

Now, by (23) it follows that $LH = p Q \widetilde{Q}^\intercal H^\intercal H$, and hence, by (10), we get:

$$\mathbb{Q}^\intercal LH\mathbb{Q} = \begin{bmatrix} Q^\intercal \\ Q_\perp^\intercal \end{bmatrix} p\, Q\, \widetilde{Q}^\intercal H^\intercal H \begin{bmatrix} Q & Q_\perp \end{bmatrix} = \begin{bmatrix} p\widetilde{R} & p\,\widetilde{Q}^\intercal H^\intercal H Q_\perp \\ 0 & 0 \end{bmatrix}. \tag{A.2}$$

Clearly, $Q_i^\intercal LH Q_i = p\widetilde{R}_{ii}$, $i = 1, \ldots, k$, and so:

$$\mu_i = \lambda_i - p \lim_{t\to+\infty} \frac{1}{t}\int_0^t \widetilde{R}_{ii}\mathrm{d}s.$$

Since $\widetilde{R}_{ii} \geqslant 0$, it follows by the previous equality that $\mu_i < 0$ for $i > k$, and that the k leading LEs $\mu_1 \geqslant \cdots \geqslant \mu_k$ of (A.1) are negative iff

$$\exists p > 0: \quad p \lim_{t\to+\infty} \frac{1}{t}\int_0^t \widetilde{R}_{ii}\mathrm{d}s > \lambda_i > 0, \quad i = 1, \ldots, k.$$

But this is the case if and only if (A, H) is detectable in the sense of definition 3.5. Note that selecting

$$p > \frac{\kappa + \max_j\{\lambda_j\}}{\min_j\{\lim_{t\to+\infty} \frac{1}{t}\int_0^t \widetilde{R}_{jj}(t)\mathrm{d}t\}}$$

we achieve the desired inequality, namely $\mu_1 < -\kappa$. This completes the proposition's proof.

Let us now prove (U). Let X solve (9), $X(0) = \mathbb{Q}_0\,\mathbb{R}_0$, and $\mathbb{Q} \in \mathcal{R}^{d\times d}$, $\mathbb{R} \in \mathcal{R}^{d\times k}$ denote the unique solution of (11) and (12), and $Q \in \mathcal{R}^{d\times k}$, $R \in \mathcal{R}^{k\times k}$ solve (15) and (16) for $k \leqslant d$. Define $\mathbb{R}_1(t) := \mathbb{Q}(t)^\intercal W(t)$ and assume that $W(0) = X(0)$. Then

$$\dot{\mathbb{R}}_1 = \dot{\mathbb{Q}}^\intercal W + \mathbb{Q}^\intercal \dot{W} = (\mathbb{Q}^\intercal A\mathbb{Q} - \mathbb{S} - \mathbb{Q}^\intercal L(t)H\mathbb{Q})\mathbb{R}_1, \quad \mathbb{R}_1(0) = \mathbb{R}_0.$$

Here $\mathbb{S}$ denotes S defined in (12). Recall from (11) that $\mathbb{B} = \mathbb{Q}^\intercal A\mathbb{Q} - \mathbb{S}$ is upper-triangular. Furthermore, by (A.2), $\mathbb{Q}^\intercal L(t)H\mathbb{Q}$ is also upper-triangular. Consequently, $\mathbb{R}_1(t)$ is upper-triangular with positive main diagonal if it is so initially. Thus $\mathbb{R}_1$ and $\mathbb{Q}$ verify QR2, and hence, by QR1, $\mathbb{R}_1(t)\mathbb{Q}(t)$ coincides with the unique QR-decomposition of W. ■

Proof of theorem 4.2. We first prove S2 $\Rightarrow$ S1. Note that $\dot{\xi} = \dot{z} - \dot{x} = f(t,z) - f(t,x) - L(t,x)H(t)\xi(t)$. Recall that $f = (f_1, \ldots, f_d)^\intercal$, and by A1 and A2 of assumption 4.1 f has bounded Jacobian (w.r.t. x) $A(t) = \mathrm{D}f(t,x(t))$ and N_i is defined by (27), the Hessian of each f_i w.r.t. x is bounded on every compact set K, uniformly w.r.t. time $t > 0$. Moreover, by A2 we have that:

$$\|N(t,\xi_1,x(t)) - N(t,\xi_2,x(t))\| \leqslant C(K)\|\xi_1 - \xi_2\|^2, \quad C(K) := \max_i C_i(K) \tag{A.3}$$

is uniformly Hölder continuous with exponent 2 on every compact set $K \subset \mathcal{R}^d$. Define $g(s) := f_i(s\xi + x)$. Then $g(1) - g(0) = \int_0^1 g'(s)\mathrm{d}s$, or, equivalently

$$f_i(t,\xi(t) + x(t)) - f_i(t,x(t)) = \int_0^1 \xi^\intercal(t)\nabla f_i(t, s\xi(t) + x(t))\mathrm{d}s.$$

Now, define $g(\tau; s, t) := \xi^\intercal(t)\nabla f_i(t, \tau s\xi(t) + x(t))$. We get that

$$f_i(t,\xi(t) + x(t)) - f_i(t,x(t)) - \xi^\intercal(t)\nabla f_i(t,x(t)) = \int_0^1 (g(1;s,t) - g(0;s,t))\mathrm{d}s = N_i(t,\xi(t),x(t))$$

where the last equality follows from $g(1;s,t) - g(0;s,t) = \int_0^1 \partial_\tau g(\tau,s,t)\mathrm{d}\tau$. This implies that

$$f(t,z) - f(t,x) = f(t,\xi + x) - f(t,x) = \mathrm{D}f(t,x) + N(t,\xi,x)$$

and so the estimation error ξ solves (29). Now, S1 follows[12] from [7, p 27, theorem 1.4.1] and [7, p 29, theorem 1.4.3].

[12] See (1.4.14) on p 29 of [7].

We next prove S1 $\Rightarrow$ S2. Note that (29) is equivalent to the following integral equation:

$$\xi(t) = X(t)X^{-1}(0)\xi_0 + \int_0^t X(t)X^{-1}(s)N(s,\xi(s),x(s))\mathrm{d}s$$

provided $\dot{X} = (A - LH)X$, $X(0) = X_0$. Take some q and consider a linear equation $\dot{v} = (A - LH)v + q$, $v(0) = \xi_0$ or its integral representation:

$$v(t) = X(t)X^{-1}(0)\xi_0 + \int_0^t X(t)X^{-1}(s)q(s)\mathrm{d}s.$$

Since X is independent of ξ, it follows that $v(t) = \xi(t)$ provided $q(s) = N(s,\xi(s,\xi_0),x(s))$. Hence, S2 is verified if we can show that $X(t)X^{-1}(0)\xi_0$ decays to 0 exponentially fast for any ξ_0: $\|\xi_0\| < \varepsilon$, i.e. all the LEs $\mu_d \leqslant \cdots \leqslant \mu_1$ of $\dot{X} = (A - LH)X$ are negative. Let $X = \mathbb{Q}\mathbb{R}$ be the (unique) full QR decomposition of $X \in \mathcal{R}^{d\times d}$, and set $z := \mathbb{Q}^\intercal v$, $p := \mathbb{Q}^\intercal q$. Then $\dot{z} = B(t)z + p$, $z(0) = \mathbb{Q}^\intercal(0)\xi_0$ and $\|z\| = \|v\|$. Here B is an upper-triangular matrix defined as in (11), and $\mu_i = \lim_{t\to+\infty}\frac{1}{t}\int_0^t B_{ii}(s)\mathrm{d}s$ by (14). Note that $\dot{z}_d = B_{dd}z_d + p_d$, or, equivalently:

$$z_d(t) = \mathrm{e}^{\int_0^t B_{dd}(s)\mathrm{d}s}z_d(0) + \mathrm{e}^{\int_0^t B_{dd}(s)\mathrm{d}s}\int_0^t \mathrm{e}^{-\int_0^s B_{dd}(\tau)\mathrm{d}\tau}p_d(s)\mathrm{d}s.$$

Assume that $\mu_d > 0$. Then $\frac{1}{t}\int_0^t B_{dd}(s)\mathrm{d}s > \mu_d - \delta$ for a small $\delta > 0$ and all $t > t^\star$. Thus

$$-\int_0^s B_{dd}(\tau)\mathrm{d}\tau \leqslant 0 \Rightarrow \gamma := \int_0^{+\infty}\mathrm{e}^{-\int_0^s B_{dd}(\tau)\mathrm{d}\tau}p_d(s)\mathrm{d}s < +\infty$$

as $p(s) = \mathbb{Q}^\intercal(s)N(s,\xi(s,\xi_0),x(s))$, and N satisfies (A.3), and $\|\xi(t)\| \leqslant C\mathrm{e}^{-at}\|\xi_0\|$ by S1 Let

$$\theta(t) := \gamma - \int_0^t \mathrm{e}^{-\int_0^s B_{dd}(\tau)\mathrm{d}\tau}p_d(s)\mathrm{d}s.$$

Then $\|\theta(t)\| \leqslant \varepsilon$ provided $t > t_1^\star$. Without loss of generality we can assume that $z_d(0) < \gamma$. But then, for $t > t_1^\star$ we have that

$$z_d(t) = \mathrm{e}^{\int_0^t B_{dd}(s)\mathrm{d}s}\left(z_d(0) + \gamma - \theta(t)\right) \to +\infty,$$

as $\int_0^t B_{dd}(s)\mathrm{d}s > t(\mu_d - \delta)$ provided $t > t^\star$ and $\mu_d - \delta > 0$ so that $\mathrm{e}^{\int_0^t B_{dd}(s)\mathrm{d}s}$ grows unbounded. This condradicts 2. as $|z_d| \leqslant \|z\| = \|x\| = \|\xi\| \leqslant C\mathrm{e}^{-at}\|\xi_0\|$. Now, if $\mu_d = 0$ then, for any δ such that $0 < \delta < a$ there exists $t_\delta > 0$ such that $-\delta t < \int_0^t B_{ii}(s)\mathrm{d}s < \delta t$, for all $t > t_\delta$. Hence, γ is still bounded as $|p_d| < \|N\| \leqslant \tilde{C}\|\xi_0\|\mathrm{e}^{-at}$. Hence, $z_d > \mathrm{e}^{-at}\tilde{C}_1$ which again contradicts 2. Hence $\mu_d < 0$. Noting that $\dot{z}_{d-1} = B_{d-1d-1}z_{d-1} + B_{d-1d}z_d + p_{d-1}$, and that B_{d-1d} is proportional to the function $s \mapsto Q_{d-1}^\intercal(s)A(s)Q_d(s)$ which is bounded from above as $\|A\| < +\infty$ we can rewrite the equation for z_{d-1} as follows:

$$z_{d-1}(t) = \mathrm{e}^{\int_0^t B_{d-1d-1}(s)\mathrm{d}s}z_{d-1}(0) + \mathrm{e}^{\int_0^t B_{d-1d-1}(s)\mathrm{d}s}\int_0^t \mathrm{e}^{-\int_0^s B_{d-1d-1}(\tau)\mathrm{d}\tau}\tilde{p}_{d-1}(s)\mathrm{d}s,$$

where $\tilde{p}_{d-1} = p_{d-1} + B_{d-1d}z_d$ decays to zero exponentially fast. Then it is easy to demonstrate that $\mu_{d-1} = \lim_{t\to+\infty}\frac{1}{t}\int_0^t B_{d-1d-1}(s)\mathrm{d}s < 0$ by the same argument as was used above

to show that $\mu_d < 0$. Repeating this argument for every $d - i$, $i < d$ we obtain that indeed S1 $\Rightarrow$ S2. This completes the proof. ■

Proof of corollary 4.4. In the proof of proposition 3.7 we demonstrated that the equation $\dot{e} = (A(t) - L(t)H(t))e$ has negative LEs iff (A, H) is detectable, and the gain L is defined by (23). Note that the aforementioned statement about LEs is exactly the statement S2 of theorem 4.2 which is equivalent to S1. This completes the proof. ■

References

[1] Adrianova L Y 1995 *Introduction to Linear Systems of Differential Equations* (*Translations of Mathematical Monographs* vol 146) (Providence, RI: American Mathematical Society) (Transl. from Russian by Peter Zhevandrov)

[2] Arenas A, Diaz-Guilera A, Kurths J, Moreno Y and Zhou C 2008 Synchronization in complex networks *Phys. Rep.* **469** 93–153

[3] Azouani A, Olson E and Titi E S 2014 Continuous data assimilation using general interpolant observables *J. Nonlinear Sci.* **24** 277–304

[4] Baras J S, Bensoussan A and James M R 1988 Dynamic observers as asymptotic limits of recursive filters: special cases *SIAM J. Appl. Math.* **48** 1147–58

[5] Baras J S and Kurzhanski A 1995 Nonlinear filtering: the set-membership and the H_∞ techniques *Proc. 3rd IFAC Symp. Nonlinear Control Sys. Design* (Oxford: Pergamon)

[6] Bardi M and Capuzzo-Dolcetta I 1997 *Optimal Control and Viscosity Solutions of Hamilton–Jacobi Equations* (Basel: Birkhäuser)

[7] Barreira L and Pesin Y B 2002 *Lyapunov Exponents and Smooth Ergodic Theory* vol 23 (Providence, RI: American Mathematical Society)

[8] Bergemann K, Gottwald G and Reich S 2009 Ensemble propagation and continuous matrix factorization algorithms *Q. J. R. Meteorol. Soc.* **135** 1560–72

[9] Boccalettia S, Kurths J, Osipov G, Valladares D and Zhou C 2002 The synchronization of chaotic systems *Phys. Rep.* **366** 1–101

[10] Bocquet M and Carrassi A 2017 Four-dimensional ensemble variational data assimilation and the unstable subspace *Tellus* A **69** 1304504

[11] Bocquet M, Gurumoorthy K S, Apte A, Carrassi A, Grudzien C and Jones C K 2017 Degenerate kalman filter error covariances and their convergence onto the unstable subspace *SIAM/ASA J. Uncertain. Quantification* **5** 304–33

[12] Brown R and Rulkov N F 1997 Synchronization of chaotic systems: transverse stability of trajectories in invariant manifolds *Chaos* **7** 395–413

[13] Carrassi A, Trevisan A, Descamps L, Talagrand O and Uboldi F 2008 Controlling instabilities along a 3DVar analysis cycle by assimilating in the unstable subspace: a comparison with the EnKF *Nonlinear Process. Geophys.* **15** 503–21

[14] Carrassi A, Ghil M, Trevisan A and Uboldi F 2008 Data assimilation as a nonlinear dynamical systems problem: stability and convergence of the prediction-assimilation system *Chaos* **18** 023112

[15] Deza F, Busvelle E and Gauthier J 1992 High gain estimation for nonlinear systems *Syst. Control Lett.* **18** 295–9

[16] Dieci L, Elia C and Van Vleck E S 2010 Exponential dichotomy on the real line: SVD and QR methods *J. Differ. Equ.* **248** 287–308

[17] Dieci L, Elia C and Van Vleck E S 2011 Detecting exponential dichotomy on the real line: SVD and QR algorithms *BIT* **51** 555–79

[18] Dieci L, Jolly M S and Van Vleck E S 2011 Numerical techniques for approximating lyapunov exponents and their implementation *J. Comput. Nonlinear Dyn.* **6** 011003

[19] Dieci L, Russell R D and Van Vleck E S 1997 On the computation of Lyapunov exponents for continuous dynamical systems *SIAM J. Numer. Anal.* **34** 402–23

[20] Dieci L and Van Vleck E S 2005 On the error in computing Lyapunov exponents by QR methods *Numer. Math.* **101** 619–42

[21] Dieci L and Van Vleck E S 2015 Lyapunov exponents: computation *Encyclopedia of Applied and Computational Mathematics* ed B Engquist (Berlin: Springer)
[22] Evensen G 2003 The ensemble kalman filter: theoretical formulation and practical implementation *Ocean Dyn.* **53** 343–67
[23] Foias C, Mondaini C F and Titi E S 2016 A discrete data assimilation scheme for the solutions of the two-dimensional Navier–Stokes equations and their statistics *SIAM J. Appl. Dyn. Syst.* **15** 2109–42
[24] Frank J and Vuik C 1999 Parallel implementation of a multiblock method with approximate subdomain solution *Appl. Numer. Math.* **30** 403–324
[25] Frank J and Zhuk S 2014 Symplectic Möbius integrators for LQ optimal control problems *Proc. of IEEE Conf. on Decision and Control* (Piscataway, NJ: IEEE)
[26] Gesho M, Olson E and Titi E S 2016 A computational study of a data assimilation algorithm for the two-dimensional Navier–Stokes equations *Commun. Comput. Phys.* **19** 1094–110
[27] Ghil M, Cohn S, Tavantzis J, Bube K and Isaacson E 1981 Applications of estimation theory to numerical weather prediction *Dynamic Meteorology: Data Assimilation Methods* (Berlin: Springer) pp 139–224
[28] Gihman I and Skorokhod A 1997 *Introduction to the Theory of Random Processes* (*Dover Books on Mathematics*) (New York: Dover)
[29] Golub G H and Van Loan C F 2012 *Matrix Computations* vol 3 (Baltimore, MA: JHU Press)
[30] González-Tokman C and Hunt B R 2013 Ensemble data assimilation for hyperbolic systems *Physica* D **243** 128–42
[31] Grudzien C, Carrassi A and Bocquet M 2017 Asymptotic forecast uncertainty and the unstable subspace in the presence of additive model error (in preparation)
[32] Gurumoorthy K S, Grudzien C, Apte A, Carrassi A and Jones C K 2017 Rank deficiency of kalman error covariance matrices in linear time-varying system with deterministic evolution *SIAM J. Control Optim.* **55** 741–59
[33] Hoffmann W 1989 Iterative algorithms for Gram–Schmidt orthogonalization *Computing* **41** 335–48
[34] Krener A J 1980 Kalman–Bucy and minimax filtering *IEEE Trans. Autom. Control* **25** 291–2
[35] Law K, Stuart A and Zygalakis K 2015 *Data Assimilation: a Mathematical Introduction* (Berlin: Springer)
[36] de Leeuw B M, Dubinkina S, Frank J, Steyer A, Tu X and van Vleck E S 2017 Projected shadowing-based data assimilation (in preparation)
[37] Oseledec V I 1968 Multiple ergodic theorem. Lyapunov characteristic numbers for dynamical systems *Trudy Mosk. Mat. Obsc.* **19** 197
[38] Palatella L, Carrassi A and Trevisan A 2013 Lyapunov vectors and assimilation in the unstable subspace: theory and applications *J. Phys. A: Math. Theor.* **46** 254020
[39] Pecora L M and Carroll T L 1990 Synchronization in chaotic systems *Phys. Rev. Lett.* **64** 821–4
[40] Pecora L M and Carroll T L 1991 Driving systems with chaotic signals *Phys. Rev.* A **44** 2374–83
[41] Pecora L M and Carroll T L 1998 Master stability functions for synchronized coupled systems *Phys. Rev. Lett.* **80** 2109–12
[42] Pecora L M, Carroll T L, Johnson G A, Mar D J and Heagy J F 1997 Fundamentals of synchronization in chaotic systems, concepts, and applications *Chaos* **7** 520–43
[43] Reich S and Cotter C 2015 *Probabilistic Forecasting and Bayesian Data Assimilation* (Cambridge: Cambridge University Press)
[44] Tchrakian T, Frank J and Zhuk S 2017 Exponentially convergent data assimilation algorithm for Navier–Stokes equations *Proc. American Control Conf.* (Piscataway, NJ: IEEE)
[45] Trevisan A, D'Isidoro M and Talagrand O 2010 Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension *Q. J. R. Meteorol. Soc.* **136** 487–96
[46] Zhuk S, Frank J, Herlin I and Shorten R 2015 Data assimilation for linear parabolic equations: minimax projection method *SIAM J. Sci. Comp.* **37** A1174–96
[47] Zhuk S and Petreczky M 2017 *Solutions of Differential–Algebraic Equations as Outputs of LTI systems: Application to LQ Control Problems* (Automatica) vol 84 pp 166–73
[48] Zhuk S and Petreczky M 2017 Minimax observers for linear differential-algebraic equations *IEEE Trans. Autom. Control* **62** 4101–08
[49] Zhuk S and Polyakov A 2017 Note on minimax sliding mode control design for linear systems *IEEE Trans. Autom. Control* **62** 3395–400