HAL
open science

Homogeneous Observers for Projected Quadratic Partial Differential Equations

Sergiy Zhuk, Andrey Polyakov

To cite this version:

Sergiy Zhuk, Andrey Polyakov. Homogeneous Observers for Projected Quadratic Partial Differential Equations. IEEE Conference on Decision and Control, Dec 2020, Jesu Island, South Korea. hal-02904474

HAL Id: hal-02904474

https://inria.hal.science/hal-02904474v1

Submitted on 22 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONFIDENTIAL. Limited circulation. For review only.

Homogeneous Observers for Projected Quadratic Partial Differential Equations

Sergiy Zhuk and Andrey Polyakov

***Abstract*— The paper proposes a new homogeneous observer for finite-dimensional projections of quadratic homogeneous hyperbolic PDEs with compact state space. The design relies upon new sufficient conditions for fixed-time convergence of observer's gain, described as a solution of a non-linear homogeneous matrix differential equations, towards an ellipsoid in the space of symmetric non-negative matrices. Convergence of the observer is analyzed, and a numerical convergence test is proposed: numerical experiments confirm the test on ODEs obtained by finite-difference discretization of Burgers-Hopf equation.**

I. Introduction

Nonlinear filtering and observer design are fundamental in diverse fields including synchronization in complex networks, data assimilation and control engineering to name just a few. Theoretically, solution of the stochastic filtering problem for Markov diffusions is given by the so-called Kushner-Stratonovich (KS) equation [1], a stochastic Partial Differential Equation (PDE) which describes evolution of the conditional density of the states of the underlying diffusion process. For linear systems, KS equation is equivalent to the Kalman-Bucy Filter.

In contrast, deterministic state estimators (observers), including the algorithm presented in this paper, assume that errors have bounded energy and belong to a given bounding set. The optimal state estimate (or minimum energy estimate) is then defined as the minimax center of the reachability set, and temporal dynamics of the minimax center is described by a minimax filter [2]. There is a fundamental connection between stochastic filters and observers: namely, an observer can be obtained as asymptotic limit of stochastic filters associated with the "noisy" version of the state equation and observations; when "noise" disappears (e.g. noise variance approaches zero) the mean-squared error between stochastic and deterministic state estimates goes to zero, and, in fact, the conditional measure associated with the stochastic filter converges to a degenerate measure concentrated at the observer [2]. Having this in mind we consider the case of exact measurements and zero model disturbances.

Quadratic ODEs studied below can be obtained as finite-dimensional projections of quadratic homogeneous hyperbolic PDEs including Euler equation in two spatial dimensions [3], [4], or a finite-difference discretization of the Burgers equation in 1D [5], [6]. Such ODEs have a number of conserved quantities, e.g. energy and volume, and so all the solutions evolve on a given bounded manifold (e.g. a sphere of a certain radius) and never approach zero. In other words, the original system has no diffusion or damping, all the Lyapunov exponents equal zero and so the system is not asymptotically stable. State estimation for such systems is a challenging problem especially in the case of high dimension of the state vector. Indeed, in the latter case it is hard to implement the differential algebraic approach conventional for nonlinear systems [7], [8], [9]. Methods based on linearization, e.g. popular Extended Kalman Filters could fail here too if linearized dynamics has positive Lyapunov exponents especially for Burgers equations which have only zero Lyapunov exponents [5]. Our design does not rely upon linearization. Instead we make use of the intrinsic symmetries provided by homogeneity of the state equation in combination with rather mild uniform observability assumptions along observer's trajectory to design a new homogeneous observer which together with the proof of existence and uniqueness of solutions for the system of nonlinear differential equations, which describe dynamics of the observer and its gain, represent our key contribution.

The proposed design relies upon an important feature of the matrix differential equation, describing observer's gain, yet another important contribution, namely the practical fixed-time convergence (see e.g. [10]) of the gain, which means that independently of the initial condition the gain converges to a neighbourhood of zero in a fixed time. We demonstrate that the uniform complete observability [11] along the observer trajectory is sufficient for fixed time convergence. We also stress that standard tools of robust design, e.g. various growth estimates for LTV systems, or invariance of uniformly exponentially convergent LTV-systems to Lipschitz perturbations [12] could turn out to be very conservative for the considered systems (see remarks in Section IV-C).

Finally we stress that our design is minimal in that one cannot make use of Lyapunov subspace reduction (e.g. [5], [13]) and restrict the observer's gain corrections just to the stable Lyapunov subspace due to the following fact: all the Lyapunov exponents of the considered dynamical system equal to zero. Nevertheless, unlike [5] we do not require to observe all "unstable directions" instead requiring the uniform complete observability along the observer's trajectory. Design of finite-time homogeneous observers for different classes of homogeneous systems was discussed in [14], [15].

The research is supported by the project "RECoT" of Inria North-European Associate Team Program.

Sergiy Zhuk is with IBM Research, Dublin, Ireland (sergiy.zhuk@ie.ibm.com).

Andrey Polyakov is with Inria Lille-Nord Europe, France (andrey.polyakov@inria.fr)

Manuscript 1178 submitted to 2020 59th IEEE Conference on Decision and Control (CDC). Received March 31, 2020.

The generalized homogeneous dynamical systems where introduced in [16].

II. Mathematical preliminaries

a) Notation: $\mathbb{R}^n$ – Euclidean space of n-dimensional column vectors with real-valued entries and canonical basis $\{\psi_1 \dots \psi_n\}$; $C(t_0, T, \mathbb{R}^n)$ – the space of continuous $\mathbb{R}^n$-valued functions; $\mathbb{S}^n$ – Hilbert space of symmetric non-negative definite $n \times n$-matrices. $\lambda_{\min}(P)$ and $\lambda_{\max}(P)$ denote minimal and maximal eigen-values of a matrix P. $\mathrm{Tr}(P)$ – trace of P; $I_n \in \mathbb{S}^n$ – the $n \times n$-identity matrix;

b) Trace Inequalities: If A is symmetric and B is skew-symmetric then $\mathrm{Tr}(AB + B^\top A) = 0$.

Lemma 1: If $0 < A \in \mathbb{S}^n$ then

$$\|A\|_\infty = \max_{ij} |a_{ij}| \le \mathrm{Tr}\, A \le \sqrt{n\, \mathrm{Tr}\, A^2}.$$

c) Fixed-time convergence: Recall [10] that $\Omega \subset \mathbb{R}^n$ is said to be *globally uniformly fixed-time attractive* set of the system

$$\dot{\xi}(t) = f(t, \xi(t)), \quad t > t_0 \tag{1}$$

if there exists $T_{\max} > 0$ such that

$$\xi(t) \in \Omega, \quad \forall t \ge t_0 + T_{\max},$$

for any initial value $\xi(t_0) \in \mathbb{R}^n$ and any $t_0 \in \mathbb{R}$.

III. Problem statement

Let $x(t) \in \mathbb{R}^n$ and $y(t) \in \mathbb{R}^m$ denote the state vector and output of the following system:

$$\dot{x}(t) = B(x(t))\, x(t), \quad x(t_0) = x_0\,, \tag{2}$$

$$y(t) = C(t)x(t) \tag{3}$$

provided $x \mapsto B(x) \in \mathbb{R}^{n\times n}$ is a linear matrix-valued mapping such that $B(x) = -B^\top(x)$ and, $C(t) \in \mathbb{R}^{m\times n}$ is a given measurable matrix-valued function such that

$$\lambda_{\max}(C^\top(t)C(t)) \le \bar{c} < +\infty\,.$$

Consider the following system of equations:

$$\dot{z} = B(z)z + PC^\top R(y - Cz), \qquad z(t_0) = 0 \tag{4}$$

$$\dot{P} = B(z)P + PB^\top(z) - PC^\top RCP + Q(y), \quad P(t_0) = 0 \tag{5}$$

where $R = R^\top$ and $Q = Q^\top$ are given continuous matrix-valued functions such that:

$$0 < \underline{r} \le \lambda_{\min}(R(t)) \le \lambda_{\max}(R(t)) \le \bar{r} < +\infty, \tag{6}$$

$$Q(t) = \underline{q}\|y(t)\|^2 I, \quad \underline{q} > 0\,. \tag{7}$$

In the forthcoming sections we report the following results:

- existence and uniqueness of the continuous solution for the system of nonlinear differential eqs. (2), (4) and (5) on $[t_0, T)$ for any real $T > t_0$,
- sufficient conditions of fixed-time convergence for the gain P solving the non-linear matrix differential eq. (5),
- sufficient conditions of asymptotic convergence of $z(t)$, the solution of eq. (4) to $x(t)$ for any $x(t_0) \in \mathbb{R}^n$ verifying natural observability conditions.

A. Motivating examples

As noted in Section I the class of dynamical systems studied here includes finite dimensional projections of some important PDEs including Burgers(-Hopf) equation (in 1D) and Euler equations (in 2D).

Following [5], [6] we consider the ODE obtained by discretizing the Burgers(-Hopf) equation $u_t = -\frac{\partial_\xi u^2}{2}$ on $(0, 1)$ with periodic boundary conditions by using the finite difference scheme:

$$\dot{u}_i = -\frac{n}{6}\left(u_i(u_{i+1} - u_{i-1}) + (u_{i+1}^2 - u_{i-1}^2)\right) \tag{8}$$

taken on a periodic lattice ($i = 1 \dots n$, $u_{-1} = u_n$, $u_{n+1} = u_1$) which has the properties that

- the quadratic energy $\sum_i u_i^2$ is conserved, implying that every sphere in $\mathbb{R}^n$ is invariant under the motion of the system and $\|u\|$ is constant,
- the trace of the Jacobian of the r.h.s. of (8) is zero, implying that the flow conserves the volume of the phase element.

Denoting $x = (u_1 \dots u_n)^\top$ and setting $D = \{D_{ij}\}_{i,j=1}^n$ with $D_{ji} = -D_{ij}$ for $j \le i$, and $D_{ii+1} = 1$, $D_{ij} = 0$ for $j > i + 1$ except for $D_{1n} = -1$, we can rewrite (8) as follows: $\dot{x} = B(x)x$ with

$$B(x) = -\frac{n}{6}(\mathrm{diag}(x)D + D\,\mathrm{diag}(x)).$$

Clearly,

$$B^\top(x) = -\frac{n}{6}(D^\top \mathrm{diag}(x) + \mathrm{diag}(x)D^\top) = -B(x)$$

since $D^\top = -D$.

Euler[1] equations in 2D

$$\begin{aligned} &\partial_t\omega + \vec{u}\cdot\nabla\omega = 0\,, \quad -\Delta\psi = \omega\,, \\ &\vec{u} = \bar{u} + \nabla^\perp\psi\,, \omega(0) = \mathrm{curl}(\vec{u}_0)\,, \end{aligned} \tag{9}$$

with periodic boundary conditions at the boundary of a rectangular domain can be approximated by an ODE which has similar properties. Namely, assuming periodic boundary conditions, and applying Fourier-Galerkin (FG) approximation one can project Euler equation onto a $2N + 1$-dimensional subspace generated by $\{e^{ikx}e^{isy}\}_{|k|,|s|\le\frac{N}{2}}$ and obtain an ODE for the projection coefficients in the form $\dot{x} = B(x)x$ with skew-symmetric linear (in x) matrix B. In this case the

[1]More specifically, vorticity-streamfunction formulation of Euler equation

Manuscript 1178 submitted to 2020 59th IEEE Conference on Decision and Control (CDC). Received March 31, 2020.

components of x will represent projection coefficents of the solution. Analogously, for homogeneous Dirichlet conditions one can take sin-basis. We refer the reader to [3] for further details.

IV. Main results

A. Existence and uniqueness of continuous solutions

Theorem 1: For any $x(t_0) = x_0 \in \mathbb{R}^n$ and any $T \in (t_0, +\infty)$ the system of eqs. (2), (4) and (5) has the unique solution (x, z, P) such that $x, z \in C(t_0, T, \mathbb{R}^n)$ and $P \in C(t_0, T, \mathbb{S}^n)$. Moreover, $P(t) > 0$ for all $t \in (t_1, T)$, where $t_1 \geq t_0$ is an instant of time such that $y(t) = 0$ for $t \leq t_1$ and $y(t_1 + \delta) \neq 0$ for a small $\delta > 0$.

Proof: Since B is skew-symmetric then $\dot{x}^\top(t)x(t) = 0$, i.e. $\|x(t)\| = \|x(t_0)\| = \theta < +\infty$ and $\|y(s)\| \leq \bar{c}\theta < +\infty$ for all $t \geq t_0$. Taking into account that the right-side of the system (4)-(5) is locally Lipschitz continuous w.r.t. state variables and continuous w. r. t. to time variable, we conclude (e.g. by using standard argument based on Picard theorem) that this system has the unique solution (z, P) defined on a time interval $[t_0, T)$ such that $P(t)$ is a symmetric matrix for all $t \geq t_0$.

If $y(t) = 0$ for all $t \leq t_1$ then $Q(y(t)) = 0$, $P(t) = 0$ and $z(t) = 0$ for all $t \leq t_1$. However, $\lim_{\delta \to 0^+} \dot{P}(t_1 + \delta) = \lim_{\delta \to 0^+} Q(y(t_1 + \delta)) > 0$, i.e. $P(t) > 0$ for $t > t_1$.

Let us show that $0 < P(t) = P^\top(t) < aI$, $a < +\infty$ for all $t \in (t_1, T)$, where $T > t_1$ is given by one of the following cases:

1) $T = +\infty$;
2) $T < +\infty$, $\sup_{t \in [t_1, T)} \|z(t)\| < +\infty$ and $\|P(t)\| \to +\infty$ as $t \to T$;
3) $T < +\infty$, $\|z(t)\| \to +\infty$ as $t \to T$ and $\sup\limits_{t \in [t_1, T)} \|P(t)\| < +\infty$;
4) $T < +\infty$, $\|z(t)\| \to +\infty$ and $\|P(t)\| \to +\infty$ as $t \to T$;
5) $T < +\infty$, $\sup\limits_{t \in [t_1, T)} \|z(t)\| < +\infty$, $\sup\limits_{t \in [t_1, T)} \|P(t)\| < +\infty$ and $P(T)$ is not positive definite.

Let us now prove that case 1) is the only possible one, i.e. $T = +\infty$ for any $z_0 \in \mathbb{R}^n$ and $P_0 > 0$. Suppose the contrary.

Case 2). In this case, $\|B(z)\| \leq \|B\| \|z\|_\infty < +\infty$ and we can write (recall that $\mathrm{Tr}(B(z)P + PB^\top(z)) = 0$):

$$\frac{d}{dt} \mathrm{Tr}(P) = -\mathrm{Tr}(PC^\top RCP) + \mathrm{Tr}(Q)$$

and

$$\mathrm{Tr}\, P(t) \leq n \int_{t_0}^{T} \|y(s)\|^2 ds < +\infty, \quad \forall t \in [t_1, T].$$

Hence, using Lemma 1 we derive the contradiction.

Case 3). If

$$V_0 = z^\top P^{-1} z$$

then for $t > t_1$ we have

$$\begin{aligned}
\dot{V}_0 =& -z^\top C^\top RCz + 2z^\top C^\top RCx - z^\top P^{-1} Q P^{-1} z^\top \\
=& -z^\top C^\top RCz + 2z^\top C^\top RCx - (x)^\top C^\top RCx + \\
& x^\top C^\top RCx - z^\top P^{-1} Q P^{-1} z^\top \\
=& -(z-x)^\top C^\top RC(z-x) + x^\top C^\top RCx - z^\top P^{-1} Q P^{-1} z \\
\leq& x^\top C^\top RCx.
\end{aligned}$$

Hence $V_0(T) \leq V_0(T - \varepsilon) + \bar{r}\bar{c}\theta^2\varepsilon$, where $0 < \varepsilon < T - t_1$. By assumption, $\|P(t)\|$ is uniformly bounded on $[T - \varepsilon, T)$. Hence, using $V_0(z(T)) < +\infty$ we conclude $\|z(T)\| < +\infty$.

Case 4). Let V be defined as follows

$$V(z, P) = z^\top P^{-1} z + \mathrm{Tr}(P).$$

We have:

$$\begin{aligned}
\dot{V} =& -z^\top C^\top RCz + 2z^\top C^\top RCx - z^\top P^{-1} Q P^{-1} z^\top + \\
& \mathrm{Tr}(Q) - \mathrm{Tr}(PC^\top RCP) \\
=& -(z-x)^\top C^\top RC(z-x) + x^\top C^\top RCx + \\
& \mathrm{Tr}(Q) - z^\top P^{-1} Q P^{-1} z^\top - \mathrm{Tr}(PC^\top RCP) \\
\leq& x^\top C^\top RCx + \mathrm{Tr}(Q)
\end{aligned} \tag{10}$$

and

$$V(t) \leq V(T - \varepsilon) + \underline{q} \int_{T-\varepsilon}^{T} \|y\|^2 ds + \int_{T-\varepsilon}^{T} x^\top C^\top RCx ds < +\infty$$

for all $t \in [T - \varepsilon, T]$, where $0 < \varepsilon < T - t_1$. Using Lemma 1 we derive the contradiction.

Case 5). By assumption, $P(t) > 0$ on (t_1, T) and $P(T)x = 0$ for some $x \neq 0$. Hence, $W = P^{-1}(t)$ is defined for every $t \in (t_1, T)$, and $\lim_{t \uparrow T} \|W(t)\| = +\infty$. In this case, $\mathrm{Tr}\, W(t) \to \infty$ as $t \to T$ due to Lemma 1. For any $t \in [t_0, T)$ we have:

$$\dot{W} = -W\dot{P}W = -B^\top(z)W - WB(z) - WQW + C^\top RC,$$

$W(t_0) = P^{-1}(t_0)$. Since $\mathrm{Tr}(WB(x) + B^\top(x)W) = 0$ then

$$\frac{d}{dt} \mathrm{Tr}\, W = \mathrm{Tr}(C^\top RC) - \mathrm{Tr}(WQW) \leq \bar{r}\bar{c}n.$$

Hence, $\mathrm{Tr}(W(t))$ cannot tend to ∞ as $t \to T$. We derive the contradiction. ■

The following trivial corollary is derived from Theorem (1).

Corollary 1: Theorem remains true for $P(t_0) > 0$ and $t_1 = t_0$ in this case.

B. Fixed-time convergence for the gain P

Now, let us demonstrate practical fixed-time convergence of the gain P. To this end we introduce

Definition 1: Let z solve (4), and let $\dot{X} = B(z)X$, $X(0) = X_0$, and set $\Phi(t, s) = X(t)X^{-1}(s)$. $\mathcal{N}(t, t - \sigma; R, z)$ is said

to be *a gramian along a trajectory* z *of* (4) if $\mathcal{N}(t,t-\sigma;R,z)=\int_{t-\sigma}^{t}\Phi^\top(s,t)C^\top RC\Phi(s,t)ds$.

Corollary 2: Let z and P solve (4)-(5) on $[t_0,+\infty)$, and let $\mathcal{N}(t,t-\sigma;R,z)$ denote the gramian along the trajectory z. If there exist $\alpha>0$ and $\sigma>0$ such that for any $t>t_0+\sigma$ we have that $\alpha I<\mathcal{N}(t,t-\sigma;I,z)$ then

$$\lambda_{\max}(P(t))\leq\frac{1}{\underline{r}\alpha}+\underline{q}\int_{t-\sigma}^{t}\|y(s)\|^2ds\,,\quad t\geq t_0+\sigma. \quad (11)$$

Proof: Assume that $t<+\infty$ and consider the standard LQR design problem for the system $\dot q=-B^\top(z)q+C^\top u$, $q(t)=h$ with the cost $J(u)=q^\top(t_0)P(t_0)q(t_0)+\int_{t_0}^{t}u^\top R^{-1}u+q^\top Qqds$. It is known that the feed-back $\hat u=RCPq$ minimizes J and $J(\hat u)=h^\top P(t)h$. Hence for any other control u we have that $J(\hat u)=h^\top P(t)h\leq J(u)$. Let us select u as follows: set $u(s)=RC(s)\Phi(s,t)\mathcal{N}^{-1}(t,t-\sigma;R,z)h$ for $s\in[t-\sigma,t]$ and set $u(s)=0$ for $s\in[t_0,t-\sigma)$. Let us show that for this u one gets: $q(s)=0$ for all $s\in[t_0,t-\sigma]$. To this end recall that

$$q(s)=\Phi^\top(t,s)h-\int_{s}^{t}\Phi^\top(\tau,s)C^\top(\tau)u(\tau)d\tau$$

and so $q(t-\sigma)=0$ if[2]

$$\Phi^\top(t,t-\sigma)h=\Phi^\top(t,t-\sigma)\int_{t-\sigma}^{t}\Phi^\top(\tau,t)C^\top(\tau)u(\tau)d\tau$$

Clearly, for the above choice of u we get that $q(s)=0$ for all $s\in[t_0,t-\sigma]$. Hence, for this u we get that

$$\begin{aligned}J(\hat u)=h^\top P(t)h&\leq\int_{t-\sigma}^{t}u^\top R^{-1}u+q^\top Qqds\\&=h^\top\mathcal{N}^{-1}(t,t-\sigma;R,z)h+\int_{t-\sigma}^{t}q^\top Qqds \quad (12)\\&\leq\frac{\|h\|^2}{\underline{r}\alpha}+\int_{t-\sigma}^{t}q^\top Qqds\end{aligned}$$

since $\mathcal{N}(t,t-\sigma;R,z)\geq\underline{r}\mathcal{N}(t,t-\sigma;I,z)\geq\underline{r}\alpha$. To compute $\int_{t-\sigma}^{t}q^\top Qqds$ we first note that

$$q(s)=\Phi^\top(t,s)(I-\mathcal{N}(t,s;R,z)\mathcal{N}^{-1}(t,t-\sigma;R,z))h=$$
$$\Phi^\top(t,s)\mathcal{N}(s,t-\sigma;R,z)\mathcal{N}^{-1}(t,t-\sigma;R,z)h$$

for $s\geq t-\sigma$ and so

$$\int_{t-\sigma}^{t}q^\top Qqds=$$
$$h^\top\mathcal{N}^{-1}(t,t-\sigma;R,z)W(t,t-\sigma)\mathcal{N}^{-1}(t,t-\sigma;R,z)h$$

with

$$W(t,t-\sigma)=$$
$$\int_{t-\sigma}^{t}\mathcal{N}(s,t-\sigma;R,z)\Phi(t,s)Q\Phi^\top(t,s)\mathcal{N}(s,t-\sigma;R,z)ds$$

[2]We used the obvious equality $\Phi(\tau,t-\sigma)=\Phi(\tau,t)\Phi(t,t-\sigma)$

Recalling that $\Phi(t,s)\Phi^\top(t,s)=I$ for skew-symmetric systems we get that:

$$W(t,t-\sigma)\leq\int_{t-\sigma}^{t}\lambda_{\max}(Q)\mathcal{N}^2(s,t-\sigma;R,z)ds.$$

Finally, we note that $\mathcal{N}^2(s,t-\sigma;R,z)\leq\mathcal{N}^2(t,t-\sigma;R,z)$ hence $W(t,t-\sigma)\leq\mathcal{N}^2(t,t-\sigma;R,z)\int_{t-\sigma}^{t}\lambda_{\max}(Q)ds$ and so $\int_{t-\sigma}^{t}z^\top Qzds\leq\|h\|^2\int_{t-\sigma}^{t}\lambda_{\max}(Q)ds$. This and (12) completes the proof. ■

The straightforward consequence of Corollary 2 is practical fixed-time convergence of the Riccati equation.

Corollary 3: Let the conditions of Corollary 2 hold true. Then

$$\Omega=\{P\in\mathbb{S}^n:0<P\leq(\underline{q}\bar c\sigma\|x_0\|^2+(\underline{r}\alpha)^{-1})I\}$$

is globally uniformly fixed-time attractive set of the system (5) with $T_{\max}=\sigma$.

By using the same LQR-based argument as in the proof of Corollary 2 one can demonstrate the practical fixed-time convergence for P^{-1}:

Corollary 4: Let z and P solve (4)-(5) on $[t_0,+\infty)$. If there exists $\tilde\sigma>0$ such that

$$\int_{t-\tilde\sigma}^{t}\|y(s)\|^2ds\geq\alpha_y>0$$

then

$$\lambda_{\max}(P^{-1}(t))\leq\frac{1}{\alpha_y\underline{q}}+\tilde\sigma\bar c\bar r,\quad\forall t\geq t_0+\tilde\sigma. \quad (13)$$

Proof: Given z let us denote

$$\mathcal{N}(t,s,Q)=\int_{s}^{t}\Phi(t,\tau)Q(y(\tau))\Phi^\top(t,\tau)d\tau.$$

Since Φ is orthogonal matrix ($B(z)$ is skew-symmetric!) it follows that $\mathcal{N}(t,s,Q)=\int_{s}^{t}\|y(\tau)\|^2d\tau$. By theorem 1 there is $t^*\in(t_0,T)$ such that $P(t^*)>0$. Consider the following standard LQR control problem

$$\dot q=B(z)q+Q^{\frac12}(y)u,\quad q(t)=h\in\mathbb{R}^n,$$
$$J(u)=q(t^*)^\top P^{-1}(t^*)q(t^*)+\int_{t^*}^{t}u^\top u+q^\top C^\top RCqd\tau.$$

Since $W(t)=P^{-1}(t)$ satisfies

$$\dot W=-B^\top(z)W-WB(z)-WQW+C^\top RC,W(t^*)=P^{-1}(t^*)$$

then $\tilde u=Q^{\frac12}Wq$ minimizes the functional J and $J(\hat u)=h^\top W(t)h$. For any other control u we have $J(\hat u)\leq J(u)$. Let us select u as follows: set $u(s)=Q^{1/2}(s)\Phi^\top(s,t)\mathcal{N}^{-1}(t,t-\tilde\sigma,Q)h$ for $s\in[t-\tilde\sigma,t]$ and $u(s)=0$ for $s\in[t^*,t-\tilde\sigma]$. Let us show that for this u we have $q(s)=0$ for all $s\in[t^*,t-\tilde\sigma]$. Since for $s\in[t-\tilde\sigma,t]$ we have

$$q(s)=\Phi(t,s)h-\int_{s}^{t}\Phi(s,\tau)Q^{1/2}(\tau)u(\tau)d\tau=$$

$$\Phi(t,s)\left(I_n-\int_s^t\Phi(t,\tau)Q(\tau)\Phi^\top(\tau,t)d\tau\mathcal{N}^{-1}(t,t-\tilde{\sigma},Q)\right)h=$$

$$\Phi(t,s)\left(I_n-\mathcal{N}(t,s,Q)\mathcal{N}^{-1}(t,t-\tilde{\sigma},Q)\right)h=$$

$$\Phi(t,s)\left(\mathcal{N}(t,t-\tilde{\sigma},Q)-\mathcal{N}(t,s,Q)\right)\mathcal{N}^{-1}(t,t-\tilde{\sigma},Q)h=$$

$$\Phi(t,s)\mathcal{N}(s,t-\tilde{\sigma},Q)\mathcal{N}^{-1}(t,t-\tilde{\sigma},Q)h$$

then $q(s)=0$ for $s\in[t^*,t-\tilde{\sigma}]$. Hence,

$$\int_{t^*}^t q^\top C^\top RCqds=$$

$$h^\top\mathcal{N}^{-\top}(t,t-\tilde{\sigma},Q)V(t,t-\tilde{\sigma})\mathcal{N}^{-1}(t,t-\tilde{\sigma},Q)h$$

where

$$V(t,t-\tilde{\sigma})=$$

$$\int_{t-\tilde{\sigma}}^t\mathcal{N}(s,t-\tilde{\sigma},Q)\Phi^\top(t,s)C^\top RC\Phi(t,s)\mathcal{N}(s,t-\tilde{\sigma},Q)ds.$$

Since $\Phi^\top(t,s)\Phi(t,s)=I$ for skew-symmetric matrices then

$$V(t,t-\tilde{\sigma})\le\bar{r}\bar{c}\int_{t-\tilde{\sigma}}^t\mathcal{N}^2(s,t-\tilde{\sigma},Q)ds.$$

Taking into account $\mathcal{N}^2(s,t-\tilde{\sigma},Q)\le\mathcal{N}^2(t,t-\tilde{\sigma},Q)$ we derive

$$\int_{t^*}^t q^\top C^\top RCqds\le\tilde{\sigma}\bar{c}\bar{r}\|h\|^2.$$

On the other hand, we have

$$\int_{t^*}^t u^\top uds=$$

$$h^\top\mathcal{N}^{-1}(t,t-\tilde{\sigma},Q)\int_{t-\tilde{\sigma}}^t\Phi(s,t)Q(s)\Phi^\top(s,t)ds\mathcal{N}^{-1}(t,t-\tilde{\sigma},Q)h$$

and

$$\int_{t^*}^t u^\top uds=h^\top\mathcal{N}^{-1}(t,t-\tilde{\sigma},Q)h\le\frac{\|h\|^2}{\alpha_y\underline{q}}.$$

Hence, we derive

$$J(\hat{u})=h^\top W(t)h\le J(u)\le\frac{\|h\|^2}{\alpha_y\underline{q}}+\tilde{\sigma}\bar{c}\bar{r}\|h\|^2,\quad\forall h\in\mathbb{R}^n.$$

■

Remark 1: Note that for the case of observable LTI systems the corresponding Riccati equation (5) has the property of the practical fixed-time convergence as defined above in 3.

C. Asymptotic convergence of the observer z

Let us define $B_1(x)=\left[\begin{array}{ccc}B(\psi_1)x & \dots & B(\psi_n)x\end{array}\right]$ and

$$W(x,t)=P^{\frac{1}{2}}(t)B_1^\top(x)P^{-\frac{1}{2}}(t)+P^{-\frac{1}{2}}(t)B_1(x)P^{\frac{1}{2}}(t),$$

and $Z=P^{\frac{1}{2}}C^\top RCP^{\frac{1}{2}}-\underline{q}\|y\|^2P^{-1}$.

Theorem 2: Let C and R be constant matrices. Let x solve (2), and z, P solve (4)-(5). Define $e=x-z$ and $V(t)=\|P^{-\frac{1}{2}}(t)e(t)\|^2$. If $x(t_0)=x_0$ is so that there exists $\sigma_{\|x_0\|}>0$ such that $\forall t>t_0+\sigma_{\|x_0\|}$:

$$\int_{t-\sigma_{\|x_0\|}}^t\lambda_{\max}(W(x(s),s)-Z(s))ds<0,\qquad(14)$$

then $V(t)\to0$ for every $x(t_0)=\lambda x_0$, $\lambda>0$. If in addition $\lambda_{\max}(P(t))<\bar{p}<+\infty$, e.g. if conditions of corollary 2 hold, then $\|x(t)-z(t)\|\to0$ for every λx_0, $\lambda>0$.

Proof: Upon differentiating e we find: $\dot{e}=B(x)x-B(z)z-PC^\top R(y-Cz)=(B(z)-PC^\top RC)e+B(e)x$. Recalling that $\dot{P}^{-1}=-P^{-1}\dot{P}P^{-1}$ and using (5) we find that

$$\dot{V}(e,t)=2(P^{-1}e,B(e)x)-(RCe,Ce)-q_0\|y(t)\|^2\|P^{-1}e\|^2\qquad(15)$$

Using the definitions of W and Z we get:

$$2(P^{-1}e,B(e)x)=((B_1^\top(x)P^{-1}+P^{-1}B_1(x))e,e)$$

so that $2(P^{-1}e,B(e)x)=(WP^{-\frac{1}{2}}e,P^{-\frac{1}{2}}e)$ and

$$\dot{V}(e,t)=\left((W-Z)\frac{P^{-\frac{1}{2}}e}{V^{\frac{1}{2}}},\frac{P^{-\frac{1}{2}}e}{V^{\frac{1}{2}}}\right)V$$

Since $\|\frac{P^{-\frac{1}{2}}e}{V^{\frac{1}{2}}}\|=1$ for any $e(t)$ and $P(t)$ we get that

$$\left((W-Z)\frac{P^{-\frac{1}{2}}e}{V^{\frac{1}{2}}},\frac{P^{-\frac{1}{2}}e}{V^{\frac{1}{2}}}\right)\le\lambda_{\max}(W-Z)$$

This demonstrates that $\dot{V}\le\lambda_{\max}(W-Z)V$. By Gronwall lemma we get:

$$V(t)\le e^{\int_{t-\sigma_{\|x_0\|}}^t\lambda_{\max}(W(x(s),s)-Z(s))ds}V(t-\sigma_{\|x_0\|})$$

Thus eq. (14) implies $V(t)\to0$ but this convergence is not necessarily monotone. To demonstrate that $V(t)\to0$ for every $x(t_0)=\lambda x_0$, $\lambda>0$ we note that eqs. (2), (4) and (5) taken together form a homogeneous system: $\dot{\xi}=F(\xi)$, $F(\lambda\xi)=\lambda^2F(\xi)$ where $\xi=(x,z,P)$. Hence $\xi(t,\lambda\xi_0)=\lambda\xi(\lambda t,\xi_0)$, where $\xi(\cdot,\bar{\xi})$ is a solution of $\dot{\xi}=F(\xi)$ with the initial condition $\xi(0)=\bar{\xi}$. Since $\xi_0=(x_0,0,0)$ (recall eqs. (2), (4) and (5)) it follows that $\lambda\xi_0=(\lambda x_0,0,0)$ and so $V(t)=\|P^{-\frac{1}{2}}(\lambda t,0)(x(\lambda t,x_0)-z(\lambda t,0))\|^2\to0$ for any $\lambda>0$ since, as we saw above, it holds for x_0.

If $\lambda_{\max}(P(t))<\bar{p}<+\infty$ then $P^{-1}(t)>p_1>0$ for all $t>t_0$ and hence $V(t)\to0$ implies $\|e(t)\|\to0$. In particular, conditions of corollary 2 imply that $P^{-1}(t)>p_1>0$ for all $t>t_0$. ■

Manuscript 1178 submitted to 2020 59th IEEE Conference on Decision and Control (CDC). Received March 31, 2020.

The above result do not provide means to check the convergence of the observer numerically. In what follows we provide a sufficient condition which relies upon splitting the state vector x into observable and unobservable parts, i.e. $x = x^u \oplus x^o$, $x^o = C^+y$, $x^u = x - x^o$, and noting that $\|x^u(t)\|^2 = \theta^u(t)$ provided $\theta^u(t) = \|x(t)\|^2 - \|x^o(t)\|^2 = \|x_0\|^2 - \|C^+y(t)\|^2$, where C^+ denotes the pseudo-inverse matrix to C. It then follows that $x^u(t) = \sum_{j=1}^{N^u} \alpha_j \theta^u(t) v_j$ where $\alpha_j \geq 0$, $\sum_j \alpha_j = 1$, and the vectors v_j are taken as vertices of a polytope $\mathcal{P}$ containing $\{x : \|x\| = 1, Cx = 0\}$, the projection of the unit ball of $\mathbb{R}^n$ onto the null-space of C. If $C = [I_m, 0_{m,n-m}]$ then the most trivial example of such polytop $\mathcal{P}$ is given by convex combinations of vectors $v_j = (0 \dots 0, \tilde{v}_j)^\top$ with $\tilde{v}_j$ being jth vertex of the $n-m$-hypercube $[-1,1]^{n-m}$ so that $x_j^u(t) \in \theta^u(t)[-1,1]$. A non-trivial example of $\mathcal{P}$ is given in Section V. Clearly, the tighter $\mathcal{P}$ approximates the projected unit ball of $\mathbb{R}^n$ onto the unobservable subspace given by the null-space of C the more accurate representation for x^u one gets. In the following corollary we suggest to evaluate an upper bound of $\lambda_{\max}(\cdot)$ from eq. (14) over a polytope $\mathcal{P}$ to derive a practical way of checking eq. (14).

Corollary 5: Let $C(t) \equiv C, R(t) \equiv R$ and $\|x_0\| = \theta$, and set $\theta^u(t) = \theta - \|C^+y(t)\|$. Let $\mathcal{P}$ be a polytope containing the projected ball $\{x : \|x\| = 1, Cx = 0\}$ with vertices v_j, $j = 1, \dots, N_u$. Define

$$\bar{g}(t) = \max_{j \in \{1 \dots N_u\}} \{\lambda_{\max}\left(W(\theta^u(t)v_j) + W(C^+y(t)) - Z(t)\right)\}$$

If $\int_{t-\sigma_{\|x_0\|}}^t \bar{g}(s)ds < 0$ for some $\sigma_{\|x_0\|} > 0$ and all $t > t_0 + \sigma_{\|x_0\|}$ then $V(t) \to 0$ (non-monotonically!) for every $x(t_0) = \lambda x_0$, $\lambda > 0$.

Proof: We first note that, by definition of $\bar{g}(t)$,

$$\alpha_i(W(\theta^u(t)v_i, t) + W(x^o(t), t) - Z(t)) \leq \bar{g}(t) I_n \,.$$

and thus

$$\sum_{i=1}^n \alpha_i \left(W(\theta^u(t)v_i, t) + W(x^o(t), t) - Z(t)\right) \leq \bar{g}(t) I_n \,.$$

On the other hand, recalling that $x = x^u \oplus x^o$, and that $x^u(t) = \sum_{j=1}^{N^u} \alpha_j \theta^u(t) v_j$ where $\alpha_j \geq 0$, $\sum_j \alpha_j = 1$, and the vectors v_j are the vertices of $\mathcal{P}$, we get:

$$W(x,t) - Z(s) = \sum_{i=1}^n \alpha_i \left(W(\theta^u(t)v_i, t) + W(x^o(t), t) - Z(t)\right)$$

Hence $\lambda_{\max}(W(x(s), s) - Z(s)) \leq \bar{g}(s)$. Now, the statement follows from theorem 2. ■

Remark 2: Note that the estimate of Corollary 5 could turn out to be quite conservative but in many cases including the one considered in the following section it closely follows the $\lambda_{\max}(W - Z)$ computed for the exact x.

Remark 3: We note that a more restrictive and thus conservative assumption can be imposed to guarantee convergence of e to 0: indeed, note that the exact equation for e, $\dot{e} = (B(z) - PC^\top RC)e + B(e)x$, $e(0) = x_0$ can be considered as a perturbed version of $\dot{e}_1 = (B(z) - PC^\top RC)e_1$, which is uniformly exponentially stable if corollary 2 holds, i.e. its fundamental solution $\Phi(t,s)$ verifies $\|\Phi(t,s)\| \leq Ke^{-\gamma(t-s)}$. Note that the perturbation, $B(e)x$ is Lipschitz as $B(e)x = B_1(x)e$, and $\|x(t)\| = \|x_0\| = \theta$ so that $\|B_1(x)e\| \leq \|B_1\|\theta\|e\|$. By noting that $e(t) = \Phi(t,s)e(s) + \int_s^t \Phi(t,\tau)B_1(x)e(\tau)d\tau$, and applying the integral form of Grownwall lemma one can easily derive that the Lyapunov spectrum of the exact error dynamics is bounded from above by $\gamma - K\theta\|B_1\|$. By using estimates of $\lambda_{\max}(P)$ and $\lambda_{\min}(P)$ obtained in eqs. (11) and (13) one could express K and γ in terms of $\underline{q}, \bar{r}$ and some unknown values σ, α and then require that $\gamma - K\theta\|B_1\| < 0$ to guarantee uniform exponential stability of the error dynamics but it is not clear how to verify this condition numerically without a priori knowledge of σ and α.

Remark 4: Yet another way to establish stability of the error dynamics is by using Vazhevskii estimates (see [12, p.110]) of the growth of solutions of perturbed LTV systems. Indeed, since $B(z)z = B_1(z)z$ one can use $B_1(z)$ instead of $B(z)$ in eq. (5). In this case, one would need to change the proofs of the existence and uniqueness theorems given in Section IV-A, but it would still be possible to prove the existence results under some extra assumptions. In the latter case, the error dynamics will change to $\dot{e} = (B_1(z) - PC^\top RC)e + B(x)e$, $e(0) = x_0$. Now, Vazhevskii estimate provides

$$\|e(t)\| \leq \|e(t_0)\| \exp\{\int_{t_0}^t q(s)ds\}$$

with $q(s) = \lambda_{\max}((B_1(z) + B(x) - PC^\top RC) + (B_1(z) + B(x) - PC^\top RC)^\top)$. Since $B(x) = -B^\top(x)$ it follows that $q(s)$ is independent of x and e decays to 0 exponentially (but not uniformly!) provided $\int_{t_0}^t q(s)ds < 0$. We stress that this estimate is easy to check numerically since we have all the parameters. However, at least for the system we consider in Section V, we observed that numerically $q(s)$ was always non-negative even though the error e was converging to 0.

V. Experimental validation

We perform a number of numerical experiments for the ODE obtained by discretizing the Burgers(-Hopf) equation $u_t = -\frac{\partial_\xi u^2}{2}$ on $(0,1)$ with periodic boundary conditions by using the finite difference scheme eq. (8). Recall that this ODE conserves energy and multidimensional volume: $\|x(t)\| = \|x_0\|$ and the Lebesgue measure of the set of initial conditions is preserved by ODE's semigroup.

We set $n = 8$ and select C so that $Cx = (x_1, x_2, x_3, x_4, x_5, x_6)^\top$, i.e. 3 out of 8 components of x were measured. To perform the simulations, we first generate 10 different "true" initial conditions by drawing 10 vectors x_0 from a normal distribution, normalizing each so that $\|x_0\| = 1$ and removing the mean (to account for the periodicity). Then we set $z(0) = 0$, $P(0) = 0$, $\underline{q} = \bar{r} =$

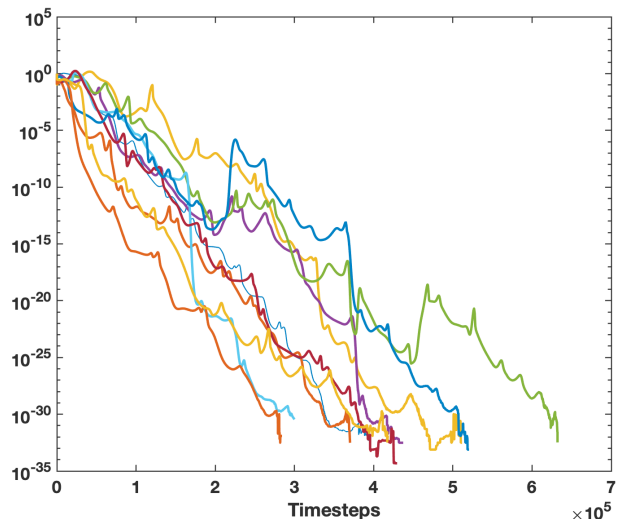

Fig. 1: Convergence of the estimation error $\|x(t) - z(t)\|$ (log-scale) for 10 simulations: for each simulation the intial condition $x_0 \in \mathbb{R}^8$ for the "truth" trajectory was picked randomly and normalized so that $\|x_0\| = 1$; the same $z(0) = 0$, $P(0) = 0$, $\underline{q} = \bar{r} = 10^4$ were used across all simulations.

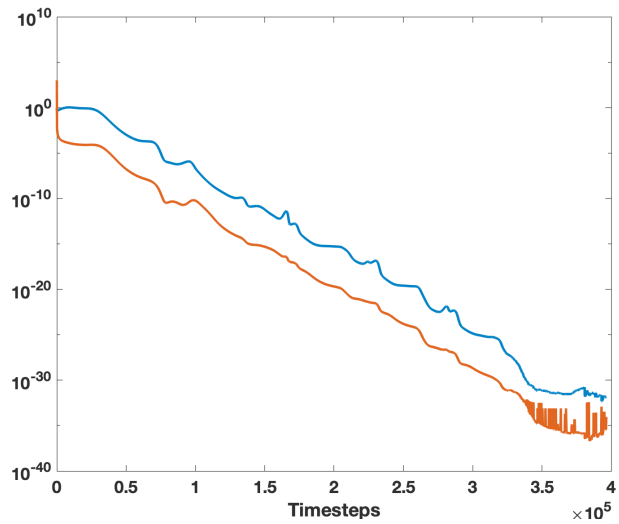

Fig. 2: Convergence of the estimation error (blue) and function V (red).

10^4, and solve the state equation (2), the filter (4), and Ricatti equation (5) simultaneously by RK4 method with $\Delta t = 1.333 \times 10^{-4}$ on a time interval $[0, 80]$. Fig. 1 demonstrates the non-monotone character of the convergence of estimation error, and Fig. 2 compares convergence of the estimation error with dynamics of V. Both figures justify our choice of sufficient conditions of convergense of integral type: each figure suggests that $V(t) - V(t-\sigma) < 0$ and $\|e(t)\|^2 - \|e(t-\sigma)\|^2 < 0$ for some $\sigma > 0$. This intuition is further confirmed in Fig. 3 where we plot $\lambda_{\max}(W - Z)$: its dynamics suggets that $\int_{t-\sigma}^{t} \lambda_{\max}(W - Z)ds < 0$, and indeed the latter holds even for the integral of the worst-case approximation of $\lambda_{\max}(W - Z)$ given by $\bar{g}$ as it is evident from Fig. 3.

To compute $\bar{g}$ we approximated the circle containing the two unobserved components of x by an outer polytope with 40 vertices: we took $N_u = 40$ points of the form $\tilde{v}_k = [\cos(\frac{2\pi}{k}), \sin(\frac{2\pi}{k})]$ on the unit circle and multiplied each by 1.02 to get an outer politope. Then we set $v_k = (0, \ldots, 0, \tilde{v}_k)^\top$, $k = 1 \ldots 40$ and compute

$$\bar{g}(t) = \max_{j \in \{1 \ldots N_u\}} \{\lambda_{\max}\left(W(\theta^u(t)v_j) + W(C^+ y(t)) - Z(t)\right)\}$$

with $\theta^u(t) = (1 - \|C^\top C x(t)\|^2)^{\frac{1}{2}}$. Dynamics of $\bar{g}$ closely follows that of $\lambda_{\max}(W - Z)$, and $\int_{t-\sigma}^{t} \bar{g}(s)ds < 0$ for a quite large $\sigma = 40$, see Fig. 3.

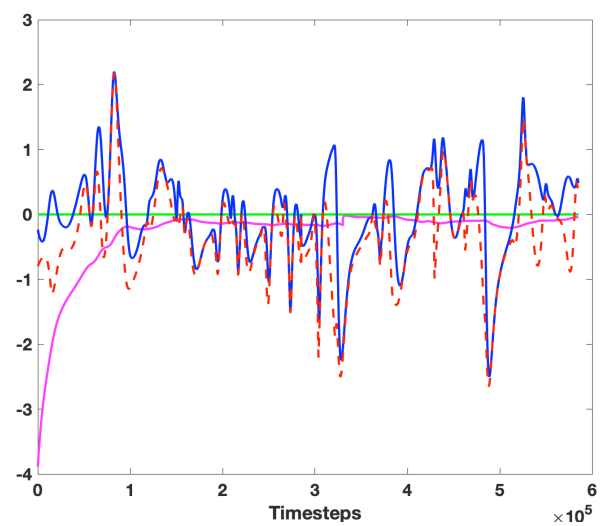

Fig. 3: Comparison of $\lambda_{\max}(W - Z)$ (dashed) with $\bar{g}$ (blue) and the corresponding $\int_{t-\sigma_{\|x_0\|}}^{t} \bar{g}(s)ds$ (magenta) for $\sigma_{\|x_0\|} = 40$ on the interval $[0, 80]$.

VI. Conclusions

The paper reveals an important relation between uniform complete observability, fixed-time convergence for non-linear homogeneous matrix differential equations and homogeneous observer design. This relation, on the other hand, requires further exploration of numerically efficient sufficient conditions for convergence, and how to deal with the case of non-trivial model disturbance.

References

[1] R. Stratonovich, "On the theory of optimal non-linear filtering of random functions," *Theory of Probability and Its Applications*, no. 4, p. 223225.

[2] J. Baras, A. Bensoussan, and M. James, "Dynamic observers as asymptotic limits of recursive filters: Special cases," *SIAM Journal on Applied Mathematics*, vol. 48, no. 5, pp. 1147–1158, 1988.

[3] S. Zhuk, T. T. Tchrakian, and J. Frank, "Exponentially convergent data assimilation algorithm for navier-stokes equations," in *2017 American Control Conference (ACC)*, May 2017, pp. 3249–3256.

[4] S. Zhuk and T. Tchrakian, "Parameter estimation for euler equations with uncertain inputs," in *2015 54th IEEE Conference on Decision and Control (CDC)*, Dec 2015, pp. 572–577.

[5] J. Frank and S. Zhuk, "A detectability criterion and data assimilation for nonlinear differential equations," *Nonlinearity*, vol. 31, no. 11, p. 5235, 2018.

[6] A. Jameson, "Energy estimates for nonlinear conservation laws with applications to solutions of the burgers equation and one-dimensional vicous flow in a shock tube by central difference schemes," in *18th AIAA Computational Fluid Dynamics Conference*, 2007, p. 4620.

[7] M. Fliess, J. Lévine, P. Martin, and P. Rouchon, "Flatness and defect of non-linear systems: introductory theory and examples," *International journal of control*, vol. 61, no. 6, pp. 1327–1361, 1995.

[8] A. Isidori, *Nonlinear Control Systems*. Springer-Verlag, N. Y. Inc., 1995.

[9] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 728–740, October 1977.

[10] A. Polyakov, "Nonlinear feedback design for fixed-time stabilization of linear control systems," *IEEE Trans. Autom. Cont.*, vol. 57(8), pp. 2106–2110, 2012.

[11] R. S. Bucy, "The riccati equation and its bounds," *Journal of computer and system sciences*, vol. 6, no. 4, pp. 343–353, 1972.

[12] L. Y. Adrianova, *Introduction to linear systems of differential equations*. American Mathematical Soc., 1995.

[13] M. Tranninger, R. Seeber, S. Zhuk, M. Steinberger, and M. Horn, "Detectability analysis and observer design for linear time varying systems," *IEEE Control Systems Letters*, vol. 4, no. 2, pp. 331–336, 2019.

[14] V. Andrieu, L. Praly, and A. Astolfi, "Homogeneous Approximation, Recursive Observer Design, and Output Feedback," *SIAM Journal of Control and Optimization*, vol. 47, no. 4, pp. 1814–1850, 2008.

[15] F. Lopez-Ramirez, A. Polyakov, D. Efimov, and W. Perruquetti, "Finite-time and fixed-time observer design: Implicit Lyapunov function approach," *Automatica*, vol. 87, no. 1, pp. 52–60, 2018.

[16] A. Polyakov, *Generalized Homogeneity in Systems and Control*. Springer, 2020.