2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

On State Reconstruction for Linear Reduced-Order Output Feedback Optimal Control

Roland Schurig*, Eric Lenz*, Andreas Himmel* and Rolf Findeisen*

***Abstract*— The solution of optimal control problems is fundamental to numerous concepts, such as model predictive control. Despite recent advancements, solving these problems remains a challenge, particularly with complex systems involving multiple states. Employing reduced order models is a strategy to simplify these problems, but accurately estimating system states from output measurements continues to be difficult. We address the challenge of reconstructing the state of a linear dynamical system using measured input and output data, intending to use this reconstructed state in an optimal control framework, reducing the problem size. Our objective is to minimize the discrepancy between the optimal solution derived from the true system state and that from the reconstructed state. We approach the reconstruction problem through the lens of observability within a finite horizon, which allows us to confine our search to a subspace of the original state space containing only observable components. This confinement effectively yields a reduced-order system representation. We delineate conditions under which the reconstruction problem can be solved and demonstrate the practicality of our approach with a case study.**

I. INTRODUCTION

Solving optimal control problems is at the core of many methodologies, such as model predictive control (MPC) [1], [2]. One key challenge in optimal control is the trade-off between prediction accuracy and computational complexity of the used mathematical model. Since the optimization is typically performed online in real time, the model must not be too complex.

In this work, we are concerned with high-dimensional linear systems. Despite their linear nature, they are often challenging because the computational complexity associated with solving optimization problems appearing in MPC scales in general cubically with the state dimension [3], [4]. While one could aim to use data-driven control approaches [5], [6], one would still need to consider input/output data with the length of the state dimension to fix the initial state of the predicted trajectory. Furthermore, harvesting the insight contained in an available model might be challenging.

One possibility to alleviate this problem is to *reduce* the mathematical model in dimension, and then use it as a prediction model; see, e.g., [7], [8], [9], [10]. The reduction step constrains the system dynamics to a low-dimensional subspace of the original state space.

A popular method for model reduction is *balanced truncation* [11], [12]. The core idea is to focus on directions of the state space which are equally well to control and to observe. The controllability and observability measures are, however, based on the *infinite-time* behavior of the system [12].

We argue that this is not the best criterion for many methods using optimal control problems such as MPC, as often they are used for transitional processes where the goal should be achieved in finite, not infinite time. Consequently, we are interested in a subspace that captures the system behavior over this finite horizon. For example in MPC this might be sufficient because the prediction happens repeatedly in closed-loop with updated information on the system state.

Another problem in using "off-the-shelf" reduction methods like balanced truncation in our context is that balanced truncation is based on the minimal realization of the transfer function associated with the dynamical system. This means that only system trajectories with zero initial state are considered. In many control applications, this is justifiable, because the system solutions that are not modeled by the minimal realization are not influenced by the control anyway. Furthermore, they decay exponentially fast, so they are neglectable for the long-term system behavior. For example in MPC, we want to find – given the current state of the system – the optimal input sequence over the finite prediction horizon. The unmodeled solutions might very well play a role in the system's *short-term behavior*, so not knowing them might lead to an uneducated choice of the optimal input sequence over the horizon.

The contribution of this work is the characterization of a low-dimensional subspace to which the system description is restricted, based on the finite-time behavior of the system. We focus on the output feedback problem, i.e., full-state measurement is unavailable. Thus, we need to reconstruct the state with a suitably tailored observer. For this, we adapt the scheme from [8]. There, the observer is not based on its asymptotic properties, but the *most recent* input/output data is used to infer the best possible current state. We put special focus on the problem of using the available input/output information to infer a state that contains the most important information *with respect to the system behavior on the finite prediction horizon*. All unnecessary information for this horizon is neglected.

A closed-loop analysis of the optimization-based controller, such as MPC, in conjunction with our proposed reconstruction method, is out of the scope of this paper and will be the focus of future work.

We structure the paper as follows. In Section II, we introduce the considered system class, while Section III

* TU Darmstadt, Control and Cyber-Physical Systems Laboratory, {roland.schurig,elenz,andreas.himmel,rolf.findeisen}@iat.tu-darmstadt.de
The authors acknowledge support within the frame of the LOEWE initiative emergenCITY.

outlines the considered problem. In Section IV, we present our solution strategy for this problem. Section V illustrates the approach for an example system and we give a conclusion in Section VI.

Notation: We denote the field of real numbers by $\mathbb{R}$ and the group of integers by $\mathbb{Z}$. For a map $f\colon X \to Y$ and a subset $M \subseteq X$, the restriction of f to M is $f|_M$.

An *inner product space* over $\mathbb{R}$ is the quadruple $(V, +, \cdot, g_V)$ where $(V, +, \cdot)$ is a $\mathbb{R}$-vector space and $g_V : V \times V \to \mathbb{R}$ is an inner product. We denote the evaluation of g_V at v, v' by $g_V(v, v') = \langle v, v' \rangle_V$. The inner product implies the norm $\|v\|_V \coloneqq \sqrt{\langle v, v \rangle_V}$.

Let V and W be finite dimensional inner product spaces and consider a linear map $A\colon V \to W$. The subspace $\operatorname{ran} A \coloneqq \{Av \mid v \in V\} \subseteq W$ is called its *image* and the subspace $\ker A \coloneqq \{v \mid Av = 0\} \subseteq V$ its *kernel*. Its *adjoint* $A^\dagger\colon W \to V$ is defined by $\langle Av, w \rangle_W = \langle v, A^\dagger w \rangle_V$ for all $v \in V$ and $w \in W$. If A is surjective, its *pseudoinverse* is the map $A^\sharp \coloneqq A^\dagger (AA^\dagger)^{-1}\colon W \to V$, and $A^\sharp w$ is the unique solution of $Av = w$ of smallest possible norm.

Let U_1, U_2 be subspaces of V. We say that V is the *direct sum of* U_1 *and* U_2, denoted $V = U_1 \oplus U_2$, if $V = U_1 + U_2$ and $U_1 \cap U_2 = \{0\}$ holds. There is one and only one linear map $P\colon V \to V$ with $P^2 = P$, $\operatorname{ran} P = U_1$ and $\ker P = U_2$, which we call the *projection on* U_1 *along* U_2. [13, Sec. 6.6]. A vector $v \in V$ is contained in U_1 if and only if $v = Pv$ holds. We can uniquely express any vector $v \in V$ as a sum of vectors in U_1 and U_2, i.e., $v = Pv + (\mathrm{id}_V - P)v$.

The *orthogonal complement* of a subspace $U \subseteq V$ is $U^\perp \coloneqq \{v \in V \mid \forall u \in U \colon \langle v, u \rangle_V = 0\}$. Then, $V = U \oplus U^\perp$.

Additionally, $V \times W$ can be made into an inner product space by letting $(v, w) + \alpha(v', w') \coloneqq (v + \alpha v', w + \alpha w')$ and $\langle (v, w), (v', w') \rangle_{V \times W} \coloneqq \langle v, v' \rangle_V + \langle w, w' \rangle_w$ for $v, v' \in V$, $w, w' \in W$ and $\alpha \in \mathbb{R}$.

Let $I \subset \mathbb{Z}$ be an interval. This allows the construction of a new inner product space $V^I \coloneqq \{f\colon I \to V\}$ where $(f + \alpha g)(t) \coloneqq f(t) + \alpha g(t)$ and

$$\langle f, g \rangle_{V^I} \coloneqq \sum_{t \in I} \langle f(t), g(t) \rangle_V$$

for $f, g \in V^I$, $\alpha \in \mathbb{R}$ and $t \in I$.

If I is an empty interval, then the set V^I only contains one element, which we denote by $\diamond$.

Consider $t_0, t_1, t_2 \in \mathbb{Z}$ with $t_0 \leq t_1 \leq t_2$, and let $f_1 \in V^{[t_0, t_1)}$ and $f_2 \in V^{[t_1, t_2)}$. Then, their *concatenation* – denoted by $f_1 f_2$ – is the map $f \in V^{[t_0, t_2)}$ defined by

$$f(t) \coloneqq \begin{cases} f_1(t) & \text{if } t \in [t_0, t_1), \\ f_2(t) & \text{if } t \in [t_1, t_2). \end{cases}$$

II. Preliminaries

We restrict our analysis to *discrete-time* systems. We thoroughly introduce the system class that we consider, which is mainly motivated by [14], [15], [16]. They are characterized by a discrete underlying *time set* $\mathcal{T}$, which is defined as a subgroup of $(\mathbb{R}, +)$. We use $\mathcal{T} = \mathbb{Z}$ as the time set throughout this work. Since $\mathbb{Z}$ is always implicitly understood to be the time set, we will not explicitly mention it. All intervals are restricted to $\mathbb{Z}$, e.g., $[a, b)$ is to be read as $\{t \in \mathbb{Z} \mid a \leq t < b\}$.

We consider linear systems with outputs defined as follows:

Definition 1. A *linear system with outputs* is the collection $\mathcal{S} = (X, U, Y, \phi, h)$ consisting of:

- An inner product space X with dimension $\dim X = n < \infty$ called the *state space* of $\mathcal{S}$;
- An inner product space U with dimension $\dim U = m < \infty$ called the *control-value space* of $\mathcal{S}$;
- An inner product space Y with dimension $\dim Y = q < \infty$ called the *measurement-value space* of $\mathcal{S}$;
- A map $\phi\colon \mathcal{D} \to X$, with

$$\mathcal{D} = \{(t_1, t_0, x, \omega) \mid t_0, t_1 \in \mathbb{Z},\, t_0 \leq t_1,\, x \in X,\, \omega \in U^{[t_0, t_1)}\},$$

  the *transition map* of $\mathcal{S}$, and;
- A map $h\colon \mathbb{Z} \times X \to Y$, the *measurement map* of $\mathcal{S}$,

such that the following properties hold:

- For each triple $t_0, t_1, t_2 \in \mathbb{Z}$ with $t_0 < t_1 < t_2$, if $\omega_0 \in U^{[t_0, t_1)}$ and $\omega_1 \in U^{[t_1, t_2)}$, and if $x_0 \in X$ is a state so that

$$\phi(t_1, t_0, x_0, \omega_0) = x_1 \quad \text{and} \quad \phi(t_2, t_1, x_1, \omega_1) = x_2,$$

  then $\phi(t_2, t_0, x_0, \omega) = x_2$ for $\omega = \omega_0 \omega_1$. *(semigroup)*
- For each $t \in \mathbb{Z}$ and each $x \in X$, we have $\phi(t, t, x, \diamond) = x$ for the empty sequence $\diamond \in U^{[t,t)}$. *(identity)*
- For each pair $t_0, t_1 \in \mathbb{Z}$ with $t_0 < t_1$, the two maps

$$X \times U^{[t_0, t_1)} \ni (x, \omega) \mapsto \phi(t_1, t_0, x, \omega) \in X$$

  and $X \ni x \mapsto h(t_1, x) \in Y$ are linear. *(linearity)*

We call elements of X *states*, elements of U *control values*, elements of Y *measurement values* and maps $\omega \in U^{[t_0, t_1)}$ *controls*. The transition map ϕ has an intuitive interpretation. It defines – with given control $\omega \in U^{[t_0, t_1)}$ – the state $\phi(t_1, t_0, x_0, \omega) \in X$ of the system at time t_1 if the system was in state $x_0 \in X$ at time t_0. We say that $\xi \in X^{[t_0, t_1]}$ is a *path* or *solution* of $\mathcal{S}$ on $[t_0, t_1]$ if there exists a control $\omega \in U^{[t_0, t_1)}$ such that, for each pair $t, s \in [t_0, t_1]$ with $t \geq s$, the statement $\xi(t) = \phi(t, s, \xi(s), \omega|_{[s,t)})$ is true. For a specific control ω, we denote the set of all such solutions by $\mathcal{L}_{t_0, t_1}(\omega)$.

While we state all our results for systems as in Definition 1, we focus in Section V on time-invariant systems.

Definition 2. A linear system with outputs $\mathcal{S} = (X, U, Y, \phi, h)$ is *time-invariant* if for each $\omega \in U^{[t_0, t_1)}$, each $x \in X$ and each $\tau \in \mathbb{Z}$ the equalities

$$\phi(t_1, t_0, x, \omega) = \phi(t_1 + \tau, t_0 + \tau, x, \omega^\tau)$$
$$h(t_0, x) = h(t_0 + \tau, x)$$

hold, where $\omega^\tau \in U^{[t_0 + \tau, t_1 + \tau)}$, $\omega^\tau(t) \coloneqq \omega(t - \tau)$.

III. Problem Formulation

We consider an arbitrary linear system with outputs $\mathcal{S} = (X, U, Y, \phi, h)$. For clearer readability, we will refer to $\mathcal{S}$ just as "the system" $\mathcal{S}$. We look at the following scenario.

Scenario 1. Let $t_0, t_1 \in \mathbb{Z}$ with $t_0 < t_1$. Suppose that $\mathcal{S}$ is in the state x_s a time t_0, and that the control $\omega_0 \in U^{[t_0,t_1)}$ is applied to the system. The resulting solution of $\mathcal{S}$ on $[t_0, t_1]$ is $\xi_0 \in \mathcal{L}_{t_0,t_1}(\omega_0)$, with initial state $\xi_0(t_0) = x_\mathrm{s}$ and terminal state $x_\mathrm{f} = \xi_0(t_1) = \phi(t_1, t_0, x_\mathrm{s}, \omega_0)$. The associated measurement is $\lambda_0 \in Y^{[t_0,t_1]}$, given by $\lambda_0(t) = h(t, \xi_0(t))$.

Additionally, let $t_2 \in \mathbb{Z}$ with $t_1 < t_2$. We consider the system on the interval $[t_1, t_2]$. In an optimal control setting in finite time, we are interested in selecting a control $\omega \in U^{[t_1,t_2)}$ such that – given the initial condition (t_1, x_f) – the system solution behaves "optimally" on $[t_1, t_2]$:

$$\begin{aligned} \mathscr{P}_{t_1,t_2}(x_\mathrm{f})\colon \quad \underset{\omega \in U^{[t_1,t_2)}}{\text{minimize}} \quad & \langle \lambda, \tilde{Q}\lambda \rangle_{Y^{(t_1,t_2]}} + \langle \omega, \tilde{R}\omega \rangle_{U^{[t_1,t_2)}} \\ \text{subject to} \quad & \lambda(t) = h(t, \xi(t)), \\ & \xi(t) = \phi(t, t_1, x_\mathrm{f}, \omega|_{[t_1,t)}), \end{aligned} \tag{1}$$

where $\tilde{Q} \succeq 0$ and $\tilde{R} \succ 0$, with $\tilde{Q} \neq 0$.

The optimal control problem (1) requires knowledge of the system state x_f at the current time t_1. In our setting, however, the state information itself is not directly available, but only control and measurement values.

For this reason, the optimal control problem is accompanied by a *reconstruction problem*, in which we seek to leverage the available measurement and control information over the past interval $[t_0, t_1]$ to estimate or reconstruct a state $x_1 \in X$ at time t_1, which is a "good" approximation for x_f. Then, x_1 is used as an estimate for x_f, i.e., the optimal control problem $\mathscr{P}_{t_1,t_2}(x_1)$ is solved instead of $\mathscr{P}_{t_1,t_2}(x_\mathrm{f})$.

This approach is motivated by [8]. In particular, we do not rely on the structure of a Luenberger observer to reconstruct a state, as done in, e.g., [7]. The motivation for this is that we do not want to design an observer that *asymptotically* converges to the true system state, but we want to use the available information (ω_0, λ_0) on $[t_0, t_1]$ to infer a state at time t_1 which brings us in the best position to solve (1) without knowing the true system state exactly. This leads to the following problem formulation.

Problem 1. Consider Scenario 1. Use the available information (ω_0, λ_0) to infer a path $\xi \in \mathcal{L}_{t_0,t_1}(\omega_0)$ such that, for each $t \in [t_0, t_1]$, the equality $\lambda_0(t) = h(t, \xi(t))$ holds.

The intuitive explanation for this problem formulation is that we want to find a solution for $\mathcal{S}$ on $[t_0, t_1]$ that explains the available information. Then, the state $x_1 = \xi(t_1)$ can be used as an estimate for the true system state x_f at time t_1. By doing so, we have exploited all information available to us on the interval $[t_0, t_1]$.

The conclusion that we can draw if Problem 1 admits a solution is that $x_\mathrm{f} = x_1$ *could* hold, but – as we will see – it does not have to. We discuss the consequences in the later sections; specifically the question of how possible degrees of freedom can be used in solving Problem 1.

The fact that $x_\mathrm{f} \neq x_1$ might be true stems from our assumption that the dimension of the state space X is "very large". We can now be more precise. In particular, we assume that $n = \dim X$ and $q = \dim Y$ are such that it is not economical from a computational perspective for Scenario 1 to consider intervals that lead to measurement spaces with dimensions larger than n. By this, we mean that $\dim Y^{[t_0,t_1]} < n$ and $\dim Y^{(t_1,t_2]} < n$ certainly holds. This will imply that (ω_0, λ_0) does not infer an unique solution ξ.

IV. Observability-based Reconstruction

First, for notational convenience, we introduce – for $t, s \in \mathbb{Z}$ with $t \geq s$ – the map

$$\varphi(t,s)\colon X \to X,\ x_0 \mapsto \varphi(t,s)x_0 \coloneqq \phi(t, s, x_0, 0)$$

which is linear by linearity of $\mathcal{S}$. With this map, the solutions $\xi \in \mathcal{L}_{t_0,t_1}(0)$ with $t, s \in [t_0, t_1]$ satisfy $\xi(t) = \varphi(t,s)\xi(s)$.

Assumption 1. The map $\varphi(t+1, t)$ is bijective for each $t \in \mathbb{Z}$.

Remark 1. Note that this also implies that $\varphi(s,t)$ is bijective for all $s > t+1$, since $\varphi(t+2,t) = \varphi(t+2,t+1)\varphi(t+1,t)$ is bijective by composition, and so on inductively.

Since we want to study $\mathcal{S}$ both on $[t_0, t_1]$ an $[t_1, t_2]$, we first state some required intermediate results on an arbitrary interval of the form $[\sigma, \tau]$ with $-\infty < \sigma < \tau < \infty$.

Theorem 1. *Let $\omega \in U^{[\sigma,\tau)}$ be arbitrary. For each pair $(t_0, x_0) \in [\sigma, \tau] \times X$, there exists a unique solution $\xi \in \mathcal{L}_{\sigma,\tau}(\omega)$ with $\xi(t_0) = x_0$.*

Proof. Note that, under Assumption 1, the map $f_0\colon x \mapsto \phi(t_0, \sigma, x, \omega|_{[\sigma,t_0)})$ is bijective, with inverse given by $f_0^{-1}\colon \hat{x} \mapsto \varphi^{-1}(t_0,\sigma)(\hat{x} - \phi(t_0, \sigma, 0, \omega|_{[\sigma,t_0)}))$. This implies that a solution $\xi \in \mathcal{L}_{\sigma,\tau}(\omega)$ with $\xi(t_0) = \phi(t_0, \sigma, \xi(\sigma), \omega|_{[\sigma,t_0)}) = x_0$ always exists, and that it satisfies $\xi(\sigma) = f_0^{-1}(x_0)$. We can construct such a solution by defining $\xi(t) = \phi(t, \sigma, f_0^{-1}(x_0), \omega|_{[\sigma,t)})$ for each $t \in [\sigma, \tau]$. We prove uniqueness by contradiction. For this, assume that there exists $\psi \in \mathcal{L}_{\sigma,\tau}(\omega)$ with $\psi(t_0) = x_0$ and $\psi \neq \xi$. Using the same argument as above, $\psi(\sigma) = f_0^{-1}(x_0) = \xi(\sigma)$ must hold. Then, for each $t \in [\sigma, \tau]$, we have

$$\psi(t) = \phi(t, \sigma, \psi(\sigma), \omega|_{[\sigma,t)}) = \phi(t, \sigma, \xi(\sigma), \omega|_{[\sigma,t)}) = \xi(t)$$

since $\xi, \psi \in \mathcal{L}_{\sigma,\tau}(\omega)$, which contradicts the assumption. □

We further examine linear combinations of possible solutions of $\mathcal{S}$ on $[\sigma, \tau]$.

Lemma 1. *Consider arbitrary $\omega_1, \omega_2 \in U^{[\sigma,\tau)}$, and suppose that ξ_1 and ξ_2 are solutions of $\mathcal{S}$ on $[\sigma, \tau]$ with $\xi_1 \in \mathcal{L}_{\sigma,\tau}(\omega_1)$ and $\xi_2 \in \mathcal{L}_{\sigma,\tau}(\omega_2)$. Then, for arbitrary $\alpha, \beta \in \mathbb{R}$, the following holds:*

$$\xi = \alpha\xi_1 + \beta\xi_2 \in \mathcal{L}_{\sigma,\tau}(\alpha\omega_1 + \beta\omega_2).$$

Proof. Let $t, s \in [\sigma, \tau]$ with $t \geq s$. We have

$$\begin{aligned}\xi(t) &= \alpha\phi(t, s, \xi_1(s), \omega_1|_{[s,t)}) + \beta\phi(t, s, \xi_2(s), \omega_2|_{[s,t)}) \\ &= \phi(t, s, \alpha\xi_1(s) + \beta\xi_2(s), \alpha\omega_1 + \beta\omega_2|_{[s,t)}) \\ &= \phi(t, s, \xi(s), \alpha\omega_1 + \beta\omega_2|_{[s,t)}),\end{aligned}$$

where we have used linearity of $\mathcal{S}$. □

We can use this result to decompose solutions of $\mathcal{S}$.

Theorem 2. *Let $\omega \in U^{[\sigma,\tau)}$ and consider an arbitrary $\tilde{\xi} \in \mathcal{L}_{\sigma,\tau}(\omega)$. Then, $\mathcal{L}_{\sigma,\tau}(\omega) = \{\xi + \tilde{\xi} \mid \xi \in \mathcal{L}_{\sigma,\tau}(0)\}$.*

Proof. On the one hand, let $\psi \in \mathcal{L}_{\sigma,\tau}(\omega)$ be arbitrary. Then, Lemma 1 implies that $\xi = \psi - \tilde{\xi} \in \mathcal{L}_{\sigma,\tau}(\omega - \omega) = \mathcal{L}_{\sigma,\tau}(0)$, i.e., we have $\psi = \xi + \tilde{\xi}$ with ξ in $\mathcal{L}_{\sigma,\tau}(0)$. On the other hand, let $\xi \in \mathcal{L}_{\sigma,\tau}(0)$ be arbitrary. Relying again on Lemma 1, we have that $\psi = \xi + \tilde{\xi} \in \mathcal{L}_{\sigma,\tau}(0 + \omega) = \mathcal{L}_{\sigma,\tau}(\omega)$ holds. □

Apparently, the specific solution set $\mathcal{L}_{\sigma,\tau}(0)$ plays a prominent role in describing an arbitrary solution set $\mathcal{L}_{\sigma,\tau}(\omega)$. We proceed by further studying the structure of this set.

Corollary 1. *$\mathcal{L}_{\sigma,\tau}(0)$ is a subspace of $X^{[\sigma,\tau]}$.*

Proof. This follows directly from Lemma 1. □

As a next step, we establish that the state space X and the space $\mathcal{L}_{\sigma,\tau}(0)$ of autonomous solutions are closely related.

Lemma 2. *Let s be an arbitrary element of $[\sigma, \tau]$. Then, the map*

$$\Phi^s_{\sigma,\tau} \colon \mathcal{L}_{\sigma,\tau}(0) \to X,\ \xi \mapsto \Phi^s_{\sigma,\tau}\xi := \xi(s)$$

is a vector space isomorphism.

Proof. For linearity, note that $\Phi^s_{\sigma,\tau}(\alpha\xi_1 + \beta\xi_2) = \alpha\xi_1(s) + \beta\xi_2(s) = \alpha\Phi^s_{\sigma,\tau}\xi_1 + \beta\Phi^s_{\sigma,\tau}\xi_2$ holds. Now consider $\xi \in \mathcal{L}_{\sigma,\tau}(0)$ and assume that $\xi \in \ker \Phi^s_{\sigma,\tau}$ i.e., $\xi(s) = 0$ holds. This implies $\xi(\sigma) = 0$ under Assumption 1 and linearity of $\varphi^{-1}(s, \sigma)$, since $\xi(\sigma) = \varphi^{-1}(s, \sigma)\xi(s)$. But then, for each $t \in [\sigma, \tau]$, $\xi(t) = \varphi(t, \sigma)\xi(\sigma) = 0$ by linearity of $\varphi(t, \sigma)$, i.e., $\ker \Phi^s_{\sigma,\tau} = \{0\}$. Therefore, $\Phi^s_{\sigma,\tau}$ is injective. Additionally, Theorem 1 states that for each $x_0 \in X$, there exists a solution $\xi \in \mathcal{L}_{\sigma,\tau}(0)$ with $\xi(s) = x_0$, so that $\Phi^s_{\sigma,\tau}$ is also surjective. □

Let us denote the inverse of $\Phi^s_{\sigma,\tau}$ by $L^s_{\sigma,\tau}$, i.e., $\Phi^s_{\sigma,\tau} L^s_{\sigma,\tau} = \mathrm{id}_X$. This allows the identification $X \ni x_0 \cong \xi \in \mathcal{L}_{\sigma,\tau}(0)$ via $\xi = L^s_{\sigma,\tau} x_0$. Then, ξ is the unique element of $\mathcal{L}_{\sigma,\tau}(0)$ that satisfies $\xi(s) = x_0$.

A. Observability

We first turn our attention to the optimal control problem (1). Hence, we want to study, for $\omega \in U^{[t_1,t_2)}$ and $x_0 \in X$ arbitrary, the influence of the initial condition (t_1, x_0) on the system solution ξ on $[t_1, t_2]$. For this, let $\tilde{\xi}$ denote the path in $\mathcal{L}_{t_1,t_2}(\omega)$ defined by $\tilde{\xi}(t_1) = 0$. Using Theorem 2, we may decompose $\xi \in \mathcal{L}_{t_1,t_2}(\omega)$, which is defined by $\xi(t_1) = x_0$, as $\xi = \hat{\xi} + \tilde{\xi}$, where $\hat{\xi} \in \mathcal{L}_{t_1,t_2}(0)$ is fixed by $\hat{\xi}(t_1) = x_0$. Next, we map ξ to the measurement space $Y^{(t_1,t_2]}$ via

$$\lambda(t) = h(t, \xi(t)) = h(t, \hat{\xi}(t)) + h(t, \tilde{\xi}(t)) = \hat{\lambda}(t) + \tilde{\lambda}(t),$$

with the "obvious" definition of $\hat{\lambda}$ and $\tilde{\lambda}$, where we have used linearity of $h(t, \cdot)$. We shall denote the linear map that sends ω to $\tilde{\lambda}$ by G_{t_1,t_2}. Furthermore, using Lemma 2, we identify $\hat{\xi} \cong x_0$ via the isomorphism $L^{t_1}_{t_1,t_2}$. Then, we denote by F_{t_1,t_2} the linear map that sends x_0 to $\hat{\lambda}$.

Thus, we obtain $\lambda = F_{t_1,t_2} x_0 + G_{t_1,t_2}\omega \in Y^{(t_1,t_2]}$, which motivates the following notions:

Definition 3. The state $x_0 \in X$ is called *indistinguishable from* 0 *on* $[t_1, t_2]$ if it is an element of $\mathcal{W}_{t_1,t_2} := \ker F_{t_1,t_2}$. The states $x_0 \in \mathcal{W}^{\perp}_{t_1,t_2}$ are called *observable on* $[t_0, t_1]$.

Note that observable is not the negation of indistinguishable from 0. Since $X = \mathcal{W}^{\perp}_{t_1,t_2} \oplus \mathcal{W}_{t_1,t_2}$, any state $x_\mathrm{f} \in X$ can be decomposed as $x_\mathrm{f} = \hat{v}_\mathrm{f} + \hat{w}_\mathrm{f}$, having a component $\hat{v}_\mathrm{f} \in \mathcal{W}^{\perp}_{t_1,t_2}$ that is observable on $[t_1, t_2]$ and a component $\hat{w}_\mathrm{f} \in \mathcal{W}_{t_1,t_2}$ which is indistinguishable from 0 on $[t_1, t_2]$. The states in $\mathcal{V}_{t_1,t_2}$ do not have non-trivial parts that are indistinguishable from 0 on $[t_1, t_2]$.

In Scenario 1, if x_f is indistinguishable from 0 on $[t_1, t_2]$, then it has the same influence on all possible measurements on $[t_1, t_2]$ as the zero state. This implies that, in order to solve Problem 1, we only need to know the parts of x_f that are observable on $[t_1, t_2]$ if we aim to solve $\mathscr{P}_{t_1,t_2}(x_\mathrm{f})$. To be more precise, the optimal solutions to $\mathscr{P}_{t_1,t_2}(x_\mathrm{f})$ and $\mathscr{P}_{t_1,t_2}(\hat{v}_\mathrm{f})$ coincide. One can see this by explicitly computing the optimal solution $\omega^\star \in U^{[t_1,t_2)}$ to $\mathscr{P}_{t_1,t_2}(x_\mathrm{f})$:

$$\omega^\star = -(\tilde{R} + G^\dagger_{t_1,t_2}\tilde{Q}G_{t_1,t_2})^{-1} G^\dagger_{t_1,t_2}\tilde{Q}F_{t_1,t_2}x_\mathrm{f}. \quad (2)$$

B. Reconstruction

In this section, we focus on Problem 1. In particular, considering the insights from Section IV-A, we are interested in the *observable part* of $x_1 = \xi(t_1)$, since we use x_1 as our guess for x_f when we solve the optimal control problem (1). For this reason, we first leverage the results of Theorem 2 and Lemma 2 to state Problem 1 in terms x_1.

Let us use Theorem 2 to decompose $\xi = \bar{\xi} + \tilde{\xi}$, with $\bar{\xi} \in \mathcal{L}_{t_0,t_1}(0)$ and $\tilde{\xi} \in \mathcal{L}_{t_0,t_1}(\omega_0)$. Specifically, we choose the fixed solution $\tilde{\xi}$ in $\mathcal{L}_{t_0,t_1}(\omega_0)$ to be defined by $\tilde{\xi}(t_1) = 0$, c.f. Theorem 1. This implies $\xi(t_1) = \bar{\xi}(t_1) + 0$, so that the terminal state of ξ and $\bar{\xi}$ coincide.

Next, we rely on Lemma 2 to identify $\bar{\xi}$ with its *terminal* state $\bar{\xi}(t_1) = x_1$, i.e., we set $\bar{\xi} = L^{t_1}_{t_0,t_1} x_1$.

Since the reconstruction problem is stated in the space of measurement values, we map ξ to $\lambda \in Y^{[t_0,t_1]}$ by

$$\lambda(t) = h(t, \xi(t)) = h(t, (L^{t_1}_{t_0,t_1} x_1)(t)) + h(t, \tilde{\xi}(t)),$$

where we have used linearity of $h(t, \cdot)$. Thus, our task boils down to finding $x_1 \in X$ such that $\lambda = \lambda_0$ holds. To state that even more concisely in terms of x_1, let $\bar{\lambda}$ be the element of $Y^{[t_0,t_1]}$ given by $\lambda_0 - h(\cdot, \tilde{\xi}(\cdot))$ and introduce the map

$$K_{t_0,t_1} \colon X \to Y^{[t_0,t_1]},\ x_1 \mapsto K_{t_0,t_1} x_1$$

with $(K_{t_0,t_1} x_1)(t) := h(t, (L^{t_1}_{t_0,t_1} x_1)(t))$, which is linear by linearity of $\mathcal{S}$. This leads to the following reformulation.

Problem 2. Consider Scenario 1 and $\bar{\lambda} \in Y^{[t_0,t_1]}$ defined above. Find a state $x_1 \in X$ such that $\bar{\lambda} = K_{t_0,t_1} x_1$ holds.

We introduce an additional assumption to ensure the solvability of Problem 1.

Assumption 2. The map K_{t_0,t_1} is surjective.

Under this assumption, we can *always* find a terminal state x_1 that solves the reconstruction problem. Note that this implies $x_1 = x_\mathrm{f}$ only if $\mathcal{U}_{t_0,t_1} := \ker K_{t_0,t_1} = \{0\}$, i.e., if $\dim Y^{[t_0,t_1]} = \dim X$. Otherwise, if x_1 satisfies $K_{t_0,t_1} x_1 = \bar{\lambda}$, then *all* states $x_1 + w \in X$ with $w \in \mathcal{U}_{t_0,t_1}$ satisfy the equality as well. Thus, in this case, there are infinitely many possible terminal states, living in an affine subspace $\mathcal{A}(\omega_0, \lambda_0) \subset X$, which all fit the given control and measurement, and one of them is the true state x_f.

C. Proposed Solution to the Reconstruction Problem

Now, we analyze solutions to Problem 2, combining the results from the previous sections. For this, we assume throughout the section that $\mathcal{V}$ is a subspace of the state space X such that X is the direct sum of $\mathcal{V}$ and $\mathcal{U}_{t_0,t_1}$. Let P denote the projection on $\mathcal{V}$ along $\mathcal{U}_{t_0,t_1}$.

Theorem 3. *For every $\lambda \in Y^{[t_0,t_1]}$, there exists a unique state $x_1 \in \mathcal{V}$ that satisfies $\lambda = K_{t_0,t_1} x_1$.*

Proof. Fix an arbitrary $\lambda \in Y^{[t_0,t_1]}$. Under Assumption 2, there exists a state $x \in X$ with $\lambda = K_{t_0,t_1} x$. We can decompose x as $x = x_1 + u_1$, with $x_1 = Px \in \mathcal{V}$ and $u_1 = (\mathrm{id}_X - P)x \in \mathcal{U}_{t_0,t_1}$. Since $K_{t_0,t_1} u_1 = 0$ by definition of $\mathcal{U}_{t_0,t_1}$, linearity of K_{t_0,t_1} implies $\lambda = K_{t_0,t_1} x_1$.

We show uniqueness by contradiction. Assume that $x_2 \in \mathcal{V}$ is a state such that $\lambda = K_{t_0,t_1} x_2$ and $x_2 \neq x_1$ holds. Then, by linearity of K_{t_0,t_1}, we have $0 = \lambda - \lambda = K_{t_0,t_1}(x_1 - x_2)$, which implies the inclusion $x_1 - x_2 \in \mathcal{U}_{t_0,t_1}$. However, since $\mathcal{V}$ is a linear space, we have $x_1 - x_2 \in \mathcal{V}$, i.e., the difference $x_1 - x_2$ lies in $\mathcal{V} \cap \mathcal{U}_{t_0,t_1} = \{0\}$, where we have used $X = \mathcal{U}_{t_0,t_1} \oplus \mathcal{V}$. This implies $x_1 - x_2 = 0$, which is equivalent to $x_1 = x_2$, and thus contradicts the assumption. □

Let us discuss the consequences of Theorem 3 on Problem 1. For this, we denote with $S_\mathcal{V} \colon Y^{[t_0,t_1]} \to \mathcal{V}$ the inverse of the restriction $K_{t_0,t_1}|_\mathcal{V}$, i.e., $K_{t_0,t_1} S_\mathcal{V} = \mathrm{id}_{Y^{[t_0,t_1]}}$ holds. By setting $x_1 = S_\mathcal{V} \bar{\lambda}$, we perfectly reconstruct the component $\bar{v}_\mathrm{f} = P x_\mathrm{f} \in \mathcal{V}$ of the true state x_f in the subspace $\mathcal{V}$, i.e., $x_1 = \bar{v}_\mathrm{f}$ holds. This is illustrated for a single-input, single-output example system with a two-dimensional state space in Fig. 1, where we have used $t_0 = t_1 = 0$ and $t_2 = 1$. Since we are only interested in the *observable* part of a state, we project $\bar{v}_\mathrm{f}$ on $\mathcal{W}_{t_1,t_2}^\perp$ along $\mathcal{W}_{t_1,t_2}$. This is our final estimate for the influence of x_f on (1).

We have the free choice in selecting $\mathcal{V}$, under the condition that X is the direct sum of $\mathcal{V}$ and $\mathcal{U}_{t_0,t_1}$. We can state the following about states in $\mathcal{V}$.

Corollary 2. *Consider Problem 2 and set $x_1 = S_\mathcal{V} \bar{\lambda}$. If the inclusion $x_\mathrm{f} \in \mathcal{V}$ holds, then the optimal solutions to $\mathscr{P}_{t_1,t_2}(x_1)$ and $\mathscr{P}_{t_1,t_2}(x_\mathrm{f})$ coincide.*

Proof. The assumption implies $x_1 = S_\mathcal{V} \bar{\lambda} = P x_\mathrm{f} = x_\mathrm{f}$. □

Thus, we are able to perfectly solve $\mathscr{P}_{t_1,t_2}(x_\mathrm{f})$ by only relying on the available information (ω_0, λ_0) if the state x_f happens to lie in a subspace of dimension $r := \dim \mathcal{V} = \dim X - \dim \mathcal{U}_{t_0,t_1} = \dim Y^{[t_0,t_1]} = (t_1 - t_0 + 1)q$.[1]

One particular and possible choice for $\mathcal{V}$ would be to work with the orthogonal complement of $\mathcal{U}_{t_0,t_1}$, i.e., $\mathcal{V} = \mathcal{U}_{t_0,t_1}^\perp$. In this case, the map $S_\mathcal{V}$ coincides with the pseudo-inverse $K_{t_0,t_1}^\sharp$. This would be the "standard" way to solve Problem 2.

In the light of Section IV-A, however, one would also be able, by a suitable choice of $\mathcal{V}$, to *exclude* directions in the reconstruction that are indistinguishable from 0 on $[t_1, t_2]$. If, for instance, the intervals are chosen such that $\dim \mathcal{W}_{t_1,t_2} = \dim X - r$, and if $X = \mathcal{U}_{t_0,t_1} \oplus \mathcal{W}_{t_1,t_2}^\perp$ holds, another choice is to use $\mathcal{V} = \mathcal{W}_{t_1,t_2}^\perp$, which ensures that the dominant directions for the optimal solution (2) – the states that are observable on $[t_1, t_2]$ – are perfectly reconstructed. We show in Section V with an example that the choice of $\mathcal{V}$ does influence the quality of the reconstructed solution.

Remark 2. These results justify calling our method a "reduced-order" method. Since we can restrict the system description to the subspace $\mathcal{W}_{t_1,t_2}^\perp$ for the optimal control problem, and to the subspace $\mathcal{V}$ for the reconstruction problem, we effectively work with a system description that has the dimension of the considered measurement spaces.

V. Numeral Example

We consider the sixth-order linear time-invariant system with outputs $\mathcal{S} = (\mathbb{R}^6, \mathbb{R}, \mathbb{R}, \phi, h)$ that was presented in [10], and which was also used in [7]. Notably, the system has a zero outside the unit circle and is non-minimum phase. For Scenario 1, we choose $t_0 = 0$, $t_1 = 3$ and $t_2 = 7$.

As the initial state $x_\mathrm{s} \in \mathbb{R}^3$ at time $t = 0$, we use the one given in [10]. For this initial state, we create $N_\mathrm{test} = 10^4$ control sequences $\omega_0^i \in \mathbb{R}^3$ with components drawn from the uniform distribution on the open interval $(-1, 1)$. The initial state and the controls imply the measurements $\lambda_0^i \in \mathbb{R}^4$, while the terminal states at time $t = 3$ are $x_\mathrm{f}^i \in \mathbb{R}^6$.

We compare the influence of the choices $\mathcal{V}_\mathrm{a} = \mathcal{U}_{0,3}^\perp$ and $\mathcal{V}_\mathrm{b} = \mathcal{W}_{3,7}^\perp$ on the optimal solution (2). More precisely, we compute $x_{1,\mathrm{a}}^i = S_{\mathcal{V}_\mathrm{a}} \bar{\lambda}^i$ and $x_{1,\mathrm{a}}^i = S_{\mathcal{V}_\mathrm{b}} \bar{\lambda}^i$. Then, for $j \in \{\mathrm{a}, \mathrm{b}\}$, we compute $\tilde{\omega}_j^i$ as the optimal solution to $\mathscr{P}_{t_1,t_2}(x_{1,j}^i)$; see (2). We denote the true optimal solution to $\mathscr{P}_{t_1,t_2}(x_\mathrm{f}^i)$ by ω^i.

As a measure of how well the inferred terminal states capture the optimal solution on $[t_1, t_2]$, we use

$$\varepsilon_j^i := \frac{\|\tilde{\omega}_j^i - \omega^i\|_{\mathbb{R}^4}}{\|\omega^i\|_{\mathbb{R}^4}}, \qquad \varepsilon_j := \frac{1}{N_\mathrm{test}} \sum_{i=1}^{N_\mathrm{test}} \varepsilon_j^i$$

for $j \in \{\mathrm{a}, \mathrm{b}\}$. Hence, the smaller ε_j^i, the better $x_{1,j}^i$ is suited to approximate x_f^i. Finally, we use ε_j as a heuristic approximation of how well method j is suited to predict the measurement on $[t_1, t_2]$ from input/output information on the interval $[t_0, t_1]$ for the initial condition (t_0, x_s).

[1] Under Assumption 2, the dimension of $\mathcal{U}_{t_0,t_1}$ can be inferred straightforwardly from the rank-nullity theorem.

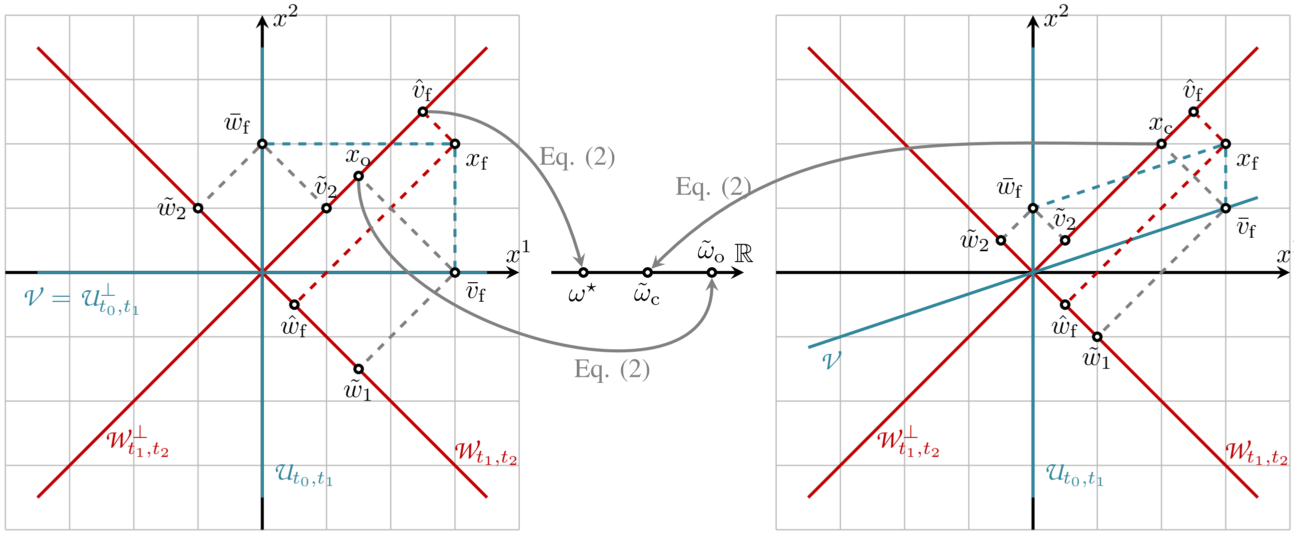

Fig. 1: Illustration of the reconstruction problem. On the left, we use $\mathcal{V} = \mathcal{U}^{\perp}_{t_0,t_1}$, while on the right we use a different $\mathcal{V}$. In both cases, we reconstruct $\bar{v}_f = S_{\mathcal{V}}\bar{\lambda} = Px_f \in \mathcal{V}$, and consider its orthogonal projection on $\mathcal{W}^{\perp}_{t_1,t_2}$ to solve (1). We denote the solution to (1) with initial state x_o and x_c by $\tilde{\omega}_o$ and $\tilde{\omega}_c$, respectively. For the chosen state x_f, the reconstruction in terms of optimal solutions works better for the latter choice. This is because – in this case – the component $\bar{w}_f \in \mathcal{U}_{t_0,t_1}$ of x_f that we are *unable* to reconstruct has a smaller component $\tilde{v}_2$ in $\mathcal{W}^{\perp}_{t_1,t_2}$.

For our simulation, we obtain $\varepsilon_a = 33.55\%$ and $\varepsilon_b = 53.96\%$, so for this academic example, we gain a significant advantage by prioritizing the observable subspace in the reconstruction problem.

VI. Conclusion

This study addresses the challenge of state reconstruction for high-dimensional linear dynamical systems using input/output measurements. The goal is to utilize the reconstructed state within an optimal control framework, leveraging a reduced-order model to mitigate complexity issues.

We presented a novel approach for state reconstruction by exploiting control and measurement information, underpinned by a detailed analysis and revision of system observability over a finite horizon. By constraining the system to an observable subspace of lower dimensionality, we achieved a reduced-order description that encompasses all pertinent information for effective reconstruction. The application to an example system demonstrated the benefits of focusing on this specific subspace, enhancing both reconstruction accuracy and optimal control problem integration.

A critical future direction is to "*close the loop*," i.e., to consider and solve the reconstruction and optimal control problems within a receding horizon framework. This will involve an in-depth study of the dynamics and performance of the resultant closed-loop system.

Moreover, the exploration of the remaining degrees of freedom in selecting the subspace $\mathcal{V}$ – especially when the reconstructed state is ambiguously defined within the observable subspace – promises to refine the reconstruction process further.

References

[1] R. Findeisen and F. Allgöwer, "An Introduction to Nonlinear Model Predictive Control," in *21st Benelux Meet. Syst. Cont.*, pp. 119–141, 2002.
[2] L. Grüne and J. Pannek, *Nonlinear Model Predictive Control : Theory and Algorithms.* Comm. Cont. Eng., Springer, 2. ed., 2017.
[3] Y. Wang and S. Boyd, "Fast model predictive control using online optimization," *IEEE Trans. Cont. Syst. Tech.*, vol. 18, no. 2, pp. 267–278, 2010.
[4] M. Kögel and R. Findeisen, "On efficient predictive control of linear systems subject to quadratic constraints using condensed, structure-exploiting interior point methods," in *2013 Europ. Cont. Conf. (ECC)*, pp. 27–34, 2013.
[5] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the deepc," in *2019 Europ. Cont. Conf. (ECC)*, pp. 307–312, 2019.
[6] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Trans. Autom. Control*, vol. 66, no. 4, pp. 1702–1717, 2020.
[7] J. Lorenzetti, A. McClellan, C. Farhat, and M. Pavone, "Linear reduced-order model predictive control," *IEEE Trans. Autom. Control*, vol. 67, no. 11, pp. 5980–5995, 2022.
[8] M. Kögel and R. Findeisen, "Robust output feedback model predictive control using reduced order models," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 1008–1014, 2015.
[9] P. Sopasakis, D. Bernardini, and A. Bemporad, "Constrained model predictive control based on reduced-order models," in *52nd IEEE Conf. Dec. Cont.*, pp. 7071–7076, 2013.
[10] S. Hovland, C. Lovaas, J. T. Gravdahl, and G. C. Goodwin, "Stability of model predictive control based on reduced-order models," in *47th IEEE Conf. Dec. Cont.*, pp. 4067–4072, 2008.
[11] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Autom. Control*, vol. 26, no. 1, pp. 17–32, 1981.
[12] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. SIAM, 2005.
[13] K. Hoffman and R. Kunze, *Linear Algebra*. Prentice-Hall, 2 ed., 1971.
[14] E. D. Sontag, *Mathematical control theory: deterministic finite dimensional systems*. Springer-Verlag, 2 ed., 1998.
[15] H. W. Knobloch and H. Kwakernaak, *Lineare Kontrolltheorie*. Springer-Verlag, 1985. (in German).
[16] H. W. Knobloch and F. Kappel, *Gewöhnliche Differentialgleichungen*. B.G. Teubner, 1974. (in German).