Proceedings of Machine Learning Research vol 144:1–13, 2021

Control of Unknown (Linear) Systems with Receding Horizon Learning

Christian Ebenbauer CE@IST.UNI-STUTTGART.DE
University of Stuttgart, Germany

Fabian Pfitz PFITZ@IST.UNI-STUTTGART.DE
University of Stuttgart, Germany

Shuyou Yu SHUYOU@JLU.EDU.CN
Jilin University, China

Abstract

A receding horizon learning scheme is proposed to transfer the state of a discrete-time dynamical control system to zero without the need of a system model. Global state convergence to zero is proved for the class of stabilizable and detectable linear time-invariant systems, assuming that only input and output data is available and an upper bound of the state dimension is known. The proposed scheme consists of a receding horizon control scheme and a proximity-based estimation scheme to estimate and control the closed-loop trajectory.

Keywords: Unknown Systems, Adaptive Control, Receding Horizon Control and Estimation, Lifting Techniques, Parsimonious Signal Models

1. Introduction

Currently, a lot of research effort is centered around the interplay between control, learning, and optimization. This is driven by extensive research initiatives in artificial intelligence, by the steadily increasing online computing power, and by the wish to build autonomous and intelligent systems in all sorts of application domains. From these developments, a renewed interest in the control of systems where no system model is known, or where the model involves large uncertainties has emerged under the banner of learning-based or data-based control. Traditionally, this is a subject of adaptive control. In this vein, we address a classical problem from adaptive control, namely the stabilization of completely unknown linear time-invariant discrete-time control systems. We aim for a solution that utilizes online optimization and (past and future) receding horizon data, and that provides convergence guarantees. To this end, we propose a scheme that involves estimation, prediction, and feedback control for unknown systems, which we have subsumed under the term learning in the title of this work.

The literature on the adaptive stabilization of unknown systems is huge. Many different solution approaches exist in the adaptive control literature ranging from model-free approaches to model-based approaches Benosman (2016); Tao (2014); Goodwin and Sin (2009); Matni et al. (2019); Recht (2019). The control of unknown systems has also been studied in the area of optimal control and receding horizon control for quite some time, see e.g. Feldbaum (1960); Mosca (1994); Peterka (1984); Bertsekas and Tsitsiklis (1996). Work that is related to our work is for example Mosca (1994) (Chapter 3) and Bemporad et al. (1994), Bittanti et al. (1990), Terzi et al.

(2019) (see the extended work Ebenbauer et al. (2020) for more details). A common approach when controlling completely unknown systems is based on a combination of a control scheme (such as pole-placement) and an online estimation or identification scheme (such as recursive least-squares). Hereby, models are estimated and updated in real-time based on the measured input-output data and these models are utilized in the control scheme. A major challenge when using this so-called certainty equivalence approach to stabilize unknown (and unstructured) systems is the *loss of stabilizability problem*, i.e. how to ensure in a computationally efficient way that the estimated models are, for example, stabilizable so that adaptive pole-placement can be applied. See for example Bittanti and Campi (2006); Hespanha et al. (2003); Morse (1992); Mania et al. (2019); Prandini and Campi (1998) and references therein on this topic. More recent related research on (partially) unknown systems and receding horizon control are for example discussed in Nguyen et al. (2020); Adetola et al. (2009); Tabuada and Fraile (2020); Lucia and Karg (2018); Korda and Mezic (2017); Münzing (2017); Limon et al. (2017); Beckenbach et al. (2018); Berberich et al. (2020); Hewing et al. (2020); Schwenkel et al. (2020); Papadimitrious et al. (2020); Mayne (2014); Coulson et al. (2019), to mention only a few out of the rapidly growing literature.

The proposed approach in this paper is based on the classical certainty equivalence implementation. However, in contrast to the existing literature, we provide a fully online optimization-based solution with provable convergence of the closed-loop for completely unknown linear systems. Several results in the above mentioned literature do not address closed-loop stability or assume that the system is linear and persistently excited in order to identify (directly or indirectly a class of) system models as advocated in Persis and Tesi (2019). Persistent excitation assumptions simplify the addressed problem but may excite undesirable dynamics and are not directly applicable to nonlinear systems (lifting techniques). Hence, in this paper we do not assume that the system is persistently excited nor that some collected finite data is rich enough to robustly stabilize the class of systems consistent with the collected data. Further, we do not assume that the state can be measured nor that the system is stable or controllable. Under these minimal assumptions, satisfaction of state or input constraints is not feasible and we do therefore not consider constraints as it is usually done in the receding horizon (predictive control) literature. Nevertheless, the control of unknown linear systems is an important benchmark problem, and, to the best of our knowledge, a receding horizon approach that provably ensures state convergence under these assumptions has not been reported in the literature. In particular, the contributions of this work are as follows. We propose an online optimization scheme which builds on a receding horizon control scheme and an estimation scheme. The receding horizon control scheme is based on a novel model-independent terminal state weighting in the sense that the objective function and the terminal cost can be chosen independently of a (not necessarily controllable) system model.

The estimation scheme is based on a modified proximal minimization algorithm that guarantees convergence of the estimated quantities, and does not require that the closed loop system is persistently excited. A characteristic feature of the approach is that the estimated quantities do not correspond to a (or to the "true") system model but rather to a signal model (time series or signal predictor) of the closed loop trajectory. The overall computational online effort of the proposed scheme is rather low and requires essentially the solution of (least-squares) regression problems. Finally, the proposed scheme is also applicable to nonlinear systems as demonstrated by simulations. The work Ebenbauer et al. (2020) includes all proofs, more detailed discussions and simulation results.

2. Problem Statement

Consider the discrete-time linear time-invariant system

$$\begin{aligned} z(k+1) &= Fz(k) + Gv(k) \\ y(k) &= Hz(k) \end{aligned} \tag{1}$$

with state $z(k) \in \mathbb{R}^n$, input $v(k) \in \mathbb{R}^q$ and output $y(k) \in \mathbb{R}^p$ at time instant $k \in \mathbb{N}$.

Assumption 1 *We assume that (F, G) is stabilizable and (F, H) detectable. Furthermore, we assume F, G, H are unknown, that an upper bound $m \geq n$ of the state dimension is known, and that only past input and output data is available.*

The goal is to define an efficient algorithmic scheme, which guarantees for any initial state $z(0)$ that the system state $z(k)$ of (1) converges to zero as time index k goes to infinity. Notice that Assumption 1 is essential the minimal requirement needed to solve this problem (see Mårtensson (1985); Mårtensson and Polderman (1993); Helmke et al. (1991)). We develop our scheme in three steps. In a first step, in Section 3, we develop a stabilizing, model-independent receding horizon control scheme based on asymptotically accurate predictor maps for the closed loop trajectory. In a second step, in Section 4, we develop a proximity-based estimation scheme to obtain the asymptotically accurate predictor maps in terms of a so-called signal model for predicting the closed-loop trajectory. Section 3 and 4 are independent of each other and also the contributions therein. In a third step, in Section 5, the control scheme and the estimation scheme are combined in a proper way to solve the stated problem. All proofs and simulations can be found in Ebenbauer et al. (2020).

3. A Model-independent Receding Horizon Control Scheme

3.1. Problem Setup

Consider the system

$$x(k+1) = Ax(k) + Bu(k) \tag{2}$$

with $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}^q$ (n, q do not necessarily coincide with n, q of Section 2). Further, consider the following optimization problem

$$\begin{aligned} V_1(x, p_1) = \min \quad & \sum_{i=0}^{N-1} x_i^\top Q x_i + u_i^\top R u_i + \frac{\Gamma(x)}{\epsilon} x_N^\top Q_N x_N \\ \text{s.t.} \quad & x_i = P_i(k, x_0, u_0, ..., u_{i-1}),\ i = 1...N, x_0 = x \end{aligned} \tag{3}$$

with $p_1^\top = [k, \epsilon]$, $N, k \in \mathbb{N}$, $\epsilon > 0$, $\Gamma : \mathbb{R}^n \to \mathbb{R}$ nonnegative. The decision variables are $u_i \in \mathbb{R}^q$, $i = 0...N-1$ and $x_i \in \mathbb{R}^n$, $i = 0...N$ and we refer to x and p_1 as parameters. The map $P_i : \mathbb{N} \times \mathbb{R}^n \times \mathbb{R}^{iq} \to \mathbb{R}^n$ is an i-th step-ahead state (or signal) predictor. We denote the value of the objective function for some u_i, $i = 1...N-1$ with $V_0(x, p_0, u_0, ..., u_{N-1})$ (since the variables x_i are determined by u_i's and x) and an optimal solution is denoted by $u_i(x, p_1)$, $i = 0...N-1$, $x_i(x, p_1)$, $i = 1...N$. If in (3) we choose x to be the state of (2) at time instant k, i.e. $x = x(k)$ and if we choose $p_1 = p_1(k)$ at time instant k for a given sequence $\{p_1(k)\}_{k\in\mathbb{N}}$, then we refer to a mapping

$$x(k) \mapsto u_0(x(k), p_1(k)) \tag{4}$$

as the *receding horizon control policy* defined by (3) and we call (3) and (2) *closed loop*, if in (2) $u(k) = u_0(x(k), p_1(k))$. We impose the following assumptions.

Assumption 2 *(A, B) in (2) is stabilizable and the state can be measured.*

Assumption 3 *The prediction horizon satisfies $N \geq n = \dim(x(k))$ and $Q > 0, R > 0, Q_N > 0$, $\Gamma : \mathbb{R}^n \to \mathbb{R}$ positive definite.*

Assumption 4 *(a) For any $k \in \mathbb{N}$, we assume that the state predictor maps P_i, $i = 1...N$ have the following linear structure*

$$P_i(k, x_0, u_0, ..., u_i) = A_i(k)x_0 + \sum_{l=0}^{i-1} B_{i-1-l}(k)u_l \tag{5}$$

where $\{A_i(k)\}_{k\in\mathbb{N}}$, $i = 0...N$ with $A_0(k) = I$, and $\{B_i(k)\}_{k\in\mathbb{N}}$, $i = 0...N-1$, $A_i(k) \in \mathbb{R}^{n\times n}$, $B_i(k) \in \mathbb{R}^{n\times q}$, are convergent matrix sequences, i.e.

$$\lim_{k\to\infty} A_i(k) = \hat{A}_i, \ \lim_{k\to\infty} B_i(k) = \hat{B}_i. \tag{6}$$

(b) Moreover, for any $x(0) \in \mathbb{R}^n$ the state predictor maps P_i along the trajectory $x(k)$, $k \in \mathbb{N}$, of the closed loop (3) and (2) with $s_l := u_l(x(k), p_1(k))$ (or of the closed loop (11) and (2) with $s_l := \nu_l(x(k), p_3(k))$) satisfy for $0 \leq i + j \leq N + 1$

$$\begin{aligned} P_j(k, P_i(k, x(k), s_0, ..., s_{i-1}), s_i, .., s_{i+j-1}) &= P_i(k, P_j(k, x(k), s_0, ..., s_{j-1}), s_j, .., s_{i+j-1}) \\ &= P_{i+j}(k, x(k), s_0, ..., s_{i+j-1}). \end{aligned} \tag{7}$$

Further, we assume that state predictor maps predict asymptotically accurate with respect to system (2) in the sense that we have for any $k \in \mathbb{N}$, $i = 0...N$

$$A_i(k)x(k) + \sum_{l=0}^{i-1} B_{i-1-l}(k)s_l = A^i x(k) + \sum_{l=0}^{i} A^{i-1-l} B s_l + e_i(k) \tag{8}$$

where the following error bounds hold for any $i = 0...N$: $\|e_i(k)\|^2 \leq \omega_1(k) + \omega_2(k)\|x(k)\|^2 + \omega_3(k)\sum_{l=0}^{i} \|s_l\|^2$ with $\lim_{k\to\infty} \omega_j(k) = 0$, $j = 1, 2, 3$.

(c) Finally, for any trajectory $x(k)$, $k \in \mathbb{N}$, of the closed loop (3) and (2), there exists functions $\mu_0(k), ..., \mu_{N-1}(k)$ such that

$$\lim_{k\to\infty} \|A_N(k)x(k) + \sum_{l=0}^{N-1} B_{N-1-l}(k)\mu_l(k)\| = 0. \tag{9}$$

Remark 1 *Assumption 4 postulates predictor maps for the closed loop trajectory (input and state sequence) generated by (3) and (2) (or (11) and (2)). In particular, equation (5) and (6) in Assumption 4(a) ensure that we have linear time-varying and converging state predictor maps. Equation (7) ensures a state property in the sense that the predictor maps commute like flow maps of (time-invariant) dynamical state-space models do. Equation (8) ensures that the predictor maps are able to accurately predict $x_1(x(k), p_1(k)), ..., x_N(x(k), p_1(k))$ in (3)* along the closed loop trajectory.

Notice that $e(k)$ *converges to zero, if, for example, the state and input stays bounded. Finally, equation* (9) *in Assumption 4(c) represents a stabilizability condition of the predictor maps along the closed loop trajectory. Notice that if the state predictor maps are determined (learned) online, then Assumption 4 does not imply that the knowledge of such state predictors implies a model (system) identification in the sense that neither the equation* $(\hat{A}_i, \hat{B}_{l-i}) \approx (A^i, A^{l-i}B)$ *must hold nor that for* every *initial data or* every *input sequence the predictions are (asymptotically) accurate.*

The main goal of the next subsection is to show that the state of the closed loop (2), (3) converges to zero under the stated assumptions. A characteristic property of the proposed scheme is that the objective (and potentially constraints) can be chosen independently from the system model (predictor maps) in the sense that no terminal cost or terminal constraint needs to be computed online based on some model information or data. This is a desirable property when controlling unknown systems.

3.2. Results

We define the following auxiliary problems

$$\begin{aligned} V_2(x, p_2) = \min \ & \xi_N^\top Q_N \xi_N \\ \text{s.t.} \ & \xi_{i+1} = A_{i+1}(k)\xi_0 + \sum_{l=0}^{i} B_{i-l}(k)\nu_l, \xi_0 = x, \ i = 0 \ldots N-1 \end{aligned} \tag{10}$$

$$\begin{aligned} V_3(x, p_3) = \min \ & \sum_{i=0}^{N-1} \xi_i^\top Q \xi_i + \nu_i^\top R \nu_i \\ \text{s.t.} \ & \xi_{i+1} = A_{i+1}(k)\xi_0 + \sum_{l=0}^{i} B_{i-l}(k)\nu_l, \xi_0 = x, \ i = 0 \ldots N-1, \xi_N = r \end{aligned} \tag{11}$$

with $p_2 = k$ and $p_3^\top = [k, r^\top]$. A corresponding notation as for (3) is used in (10) and (11).

Lemma 1 *Suppose Assumption 3 holds true. a) Let* $\{x(k)\}_{k\in\mathbb{N}}$ *and* $\{r(k)\}_{k\in\mathbb{N}}$ *be sequences such that* $\lim_{k\to\infty} r(k) = 0$ *and such that problem* (11) *is feasible for every time instant* $k \in \mathbb{N}$ *with* $x = x(k), r = r(k)$*. Then the value* $V_3(x(k), p_3(k))$ *of problem* (11) *is given by* $V_3(x(k), p_3(k)) = x(k)^\top S_3(k) x(k) + x(k)^\top S_4(k) r(k) + r(k)^\top S_5(k) r(k) \geq 0$ *for some matrices* $S_3(k), S_4(k), S_5(k)$ *and the unique solution of* (11) *is linearly parameterized in* $x(k), r(k)$ *in the sense of* $\xi_i(x(k), p_3(k)) = K_{1,i}(k)x(k) + K_{2,i}(k)r(k)$ *and* $\nu_i(x(k), p_3(k)) = K_{3,i}(k)x(k) + K_{4,i}(k)r(k)$*, for all* $i = 0, \ldots, N$*. b) Further, the value function* $V_2(x(k), p_2(k))$ *of problem* (10) *is quadratic and positive semidefinite in* x*, i.e.* $V_2(x(k), p_2(k)) = x(k)^\top S_6(k) x(k) \geq 0$ *for some matrix* $S_6(k)$*, and there exists a solution* $\{\xi_i(x(k), p_2(k))\}_{i=0}^N, \{\nu_i(x(k), p_2(k))\}_{i=0}^{N-1}$ *of* (11) *which is linearly parameterized in* $x(k)$ *in the sense of a).*

Assumption 5 *a) Let* $\{r(k)\}_{k\in\mathbb{N}}$ *be a sequence which converges to zero. We assume that the solutions of* (11) *along the closed loop* (11), (2) *are uniformly bounded, i.e.* $\xi_i(x(k), p_3(k)) = K_{1,i}(k)x(k) + K_{2,i}(k)r(k)$, $\nu_i(x(k), p_3(k)) = K_{3,i}(k)x(k) + K_{4,i}(k)r(k)$ *are bounded in the* *sense that there exists a bound $M > 0$ such that for all $i = 0, ..., N$, $j = 1...4$, $k \in \mathbb{N}$ it holds $\|K_{j,i}(k)\| \leq M$. b) We assume that the solutions of (3) along the closed loop (3), (2) are uniformly bounded, i.e. $x_i(x(k), p_1(k)) = K_{5,i}(k)x(k)$, $u_i(x(k), p_1(k)) = K_{6,i}(k)x(k)$ are bounded in the sense that there exists a bound $M > 0$ such that for all $i = 0, ..., N$, $j = 5, 6$, $k \in \mathbb{N}$ it holds $\|K_{j,i}(k)\| \leq M$.*

The main result of this subsection is Theorem 1, which builds on the following two lemmas.

Lemma 2 *Consider the closed loop (11), (2) and suppose problem (11) is feasible for every time instant $k \in \mathbb{N}$. Let Assumption 3 and 4(a)(b) hold true and let $\{p_3(k)\}_{k\in\mathbb{N}}$ be a sequence such that $\{r(k)\}_{k\in\mathbb{N}}$ converges to zero. Then for any initial state $x(0)$, the state $x(k)$ and the input $u(k)$ of the closed loop converge to zero, i.e. $\lim_{k\to\infty} x(k) = 0$, $\lim_{k\to\infty} u(k) = 0$, assuming that Assumption 5 a) holds true.*

Lemma 3 *Consider (3) and suppose Assumption 4(a) holds true. Further, suppose $\Gamma : \mathbb{R}^n \to \mathbb{R}$ is a function such that for all $x \in \mathbb{R}^n$, all $k \in \mathbb{N}$ and all $\epsilon > 0$ it holds that $\Gamma(x) \geq c(\sum_{i=0}^{N-1} \|\xi_i(x, p_2)\|^2 + \|\nu_i(x, p_2)\|^2)$ for some $c > 0$, where $\{\nu_i(x, p_2)\}_{i=0}^{N-1}$, $\{\xi_i(x, p_2)\}_{i=0}^{N-1}$ is some solution of (10) and $p_2 = k$. Then there exists a $\rho > 0$ such that solution $x_N(x, p_1)$, $p_1^\top = [k, \epsilon]$, of (3) satisfies for all $x \in \mathbb{R}^n$, all $k \in \mathbb{N}$ and all $\epsilon > 0$*

$$x_N(x, p_1)^\top Q_N x_N(x, p_1) \leq V_2(x, p_2) + \epsilon\rho. \tag{12}$$

Theorem 1 *Consider the closed loop (3) and (2), where $\Gamma(x) = \alpha x^\top x$ for some $\alpha > 0$ and suppose Assumption 2, 3 and 4(a)-(c) hold true. Let further $\{p_1(k)\}_{k\in\mathbb{N}}$ be a sequence such that $\{\epsilon(k)\}_{k\in\mathbb{N}}$, $\epsilon(k) > 0$, converges to zero. Then for any initial state $x(0)$, the state $x(k)$ of the closed loop converges to zero as k goes to infinity, assuming that Assumption 5 b) holds true.*

Assumption 5 is our main technical assumption that we impose in the proposed scheme. For example, if the predictor maps P_N are controllable in the sense that $\operatorname{rank}[B_0(k), ..., B_{N-1}(k)] = n$ for all $k \in \mathbb{N}$ and also the limiting predictor has the same property, i.e. $\operatorname{rank}[\hat{B}_0, ..., \hat{B}_{N-1}] = n$, then Assumption 5 holds true (as discussed in the proof of Lemma 1). However, we admit that this assumption is in general difficult to verify.

4. A Proximity-based Estimation Scheme

4.1. Problem Setup

Consider an output sequence (or some observed signal) and an input sequence

$$\{y(k)\}_{k\in\mathbb{N}},\ \{v(k)\}_{k\in\mathbb{N}} \tag{13}$$

with $y(k) \in \mathbb{R}^{\bar{p}_y}$, $v(k) \in \mathbb{R}^{\bar{q}_v}$. Let

$$x(k) = \phi_{\mathrm{y}}(y(k), ..., y(k - \bar{N}_y + 1)) \in \mathbb{R}^{\bar{n}},\ u(k) = \phi_{\mathrm{v}}(v(k), ..., v(k - \bar{N}_v + 1)) \in \mathbb{R}^{\bar{q}} \tag{14}$$

and $\phi_{\mathrm{y}} : \mathbb{R}^{\bar{p}_y \bar{N}_y} \to \mathbb{R}^{\bar{n}}$, $\phi_{\mathrm{v}} : \mathbb{R}^{\bar{q}_v \bar{N}_v} \to \mathbb{R}^{\bar{q}}$ be some given basis (lifting) functions, e.g. $\phi_{\mathrm{y}}(y_1(k), y_2(k), y_1(k-1), y_2(k-1)) = [y_1(k), y_2(k), y_1(k-1), y_2(k-1), y_1(k)y_2(k), y_1(k)^2, y_2(k)^2]^T$, $\bar{p} = 2, \bar{n} = 6, \bar{N}_y = 2$. Notice that in principle one could also consider cross-terms between input and output data, like $u(k)y(k)^3$, but such terms are not considered here for the sake of simplicity. Consider, further, at time instant k the optimization problem

$$\begin{aligned} \theta^*(k) =& \arg\min \;\; c(e,k) + D(\theta, \tilde{\theta}(k-1)) \\ & \text{s.t.} \;\; s(k) - R(k)\theta = e \\ \tilde{\theta}(k) =& (1-\lambda_k)\theta^*(k) + \lambda_k \tilde{\theta}(k-1) \end{aligned} \tag{15}$$

with $\lambda_k \in [0, \lambda_{max})$, $\lambda_{max} \in (0,1)$, where $\bar{N} \in \mathbb{N}$, $c : \mathbb{R}^{\bar{n}\bar{N}} \times \mathbb{N} \to \mathbb{R}$ and $D(x,y) = g(x) - g(y) - (x-y)^\top \nabla_y g(y)$ defines the Bregman distance induced by a function $g : \mathbb{R}^{\bar{n}\bar{N}} \to \mathbb{R}$. The vector $s(k)$ is defined as

$$s(k) = \left[x(k)^\top \ldots x(k-\bar{N}+1)^\top\right]^\top \tag{16}$$

and the matrix $R(k)$ is defined as

$$R(k) = \begin{bmatrix} x(k-1)^T \otimes I & u(k-1)^T \otimes I \\ \vdots & \vdots \\ x(k-\bar{N})^T \otimes I & u(k-\bar{N})^T \otimes I \end{bmatrix}. \tag{17}$$

Decision variables are the parameter vector $\theta \in \mathbb{R}^{\bar{n}(\bar{n}+\bar{q})}$ and e. We refer to θ when using $\arg\min$ since the (slack) variables e can be eliminated and have been introduced just for the notational convenience. Also we define in the following $y(k) = 0, v(k) = 0, v_k = 0$ etc. whenever $k < 0$. We impose now the following assumptions.

Assumption 6 *The objective function c in (15) is continuously differentiable and strictly convex in the first argument and it satisfies for all k and $e \neq 0$: $c(e,k) > c(0,k)$. Further, the function g, which defines the Bregman distance D, is continuously differentiable and strictly convex.*

Assumption 7 *For the given sequences in (13) and given $x(k) = \phi_{\mathrm{y}}(y(k), ..., y(k-\bar{N}_y+1))$ and $u(k) = \phi_{\mathrm{v}}(v(k), ..., v(k-\bar{N}_v+1))$ in (14), there exist matrices A, B and $x_0 \in \mathbb{R}^{\bar{n}}$ that satisfy*

$$x(k+1) = Ax(k) + Bu(k), \quad x(0) = x_0. \tag{18}$$

Remark 2 *a) Notice that $s(k) = R(k)\theta$ with $\theta^\top = [\mathrm{vec}(A)^\top, \; \mathrm{vec}(B)^\top]$, where $\mathrm{vec}(A)$ corresponds to the (column-wise) vectorization of a matrix A, is the linear system of equations $x(j) = Ax(j-1) + Bu(j-1)$, $j = k...k-\bar{N}+1$. b) If $c(e,k) = \|e\|^2$, $g(x) = \|x\|^2$, then (17) reduces to a least squares parameter estimation problem, where a closed form solution to it is known. The motivation for a general convex cost is its flexibility in tuning the estimator. Similarly as in a recently proposed state estimation scheme based on proximal minimization Gharbi et al. (2020), specifying different c, D allows to take into account various aspects like outliers in the data, sparsity in the parameters or cost-biased objectives Bittanti and Campi (2006).*

Remark 3 *Assumption 7 imposes that the given (lifted) signal $\{x(k)\}_{k\in\mathbb{N}}$ can be reproduced by some linear time-invariant system that is driven by the given (lifted) input sequence $\{u(k)\}_{k\in\mathbb{N}}$. Notice that reproducing a given signal by (18) does not imply that the signal $x(k)$ itself originates from (18) nor by a linear time-invariant system at all. For example, a given* (single) *trajectory of a* *nonlinear system or even all trajectories of a large class of nonlinear systems can be reproduced by or embedded into high dimensional linear (not necessarily controllable) systems using for example Carleman or Koopman lifting techniques. Hence, (18) represents a signal model of the actual closed-loop trajectory, rather than a system model of all possible trajectories of the plant. Similarly to adaptive filter theory, our signal model is therefore a parsimonious modeling approach in the sense that it aims to predict nothing more than the closed-loop trajectory.*

The main goal of the next subsection is to show that the (parameter) estimates $\tilde{\theta}(k)$ $(\theta^*(k))$ obtained from (15) converge and that the estimates can be used to define ith step-ahead signal predictor maps $v_i = P_i(k, v_0, u_0, ..., u_{i-1})$ which have the properties as described in Assumption 4 for the signal model (18) and for the given data (14). The convex combination in (15) is introduced to deal with the loss of stabilizability problem (see next subsection).

4.2. Results

Lemma 4 *Let $f : \mathbb{R}^n \times \mathbb{N} \to \mathbb{R}$ be convex and continuous differentiable in the first argument. Suppose the set of minimizers $\mathcal{X}_k = \{x_k^* \in \mathbb{R}^n : \; f(x,k) \geq f(x_k^*, k) := 0 \; \forall x \in \mathbb{R}^n\}$ of f at any time instant k is nonempty and also their intersections*

$$\mathcal{X} = \bigcap_{k=0}^{\infty} \mathcal{X}_k \neq \emptyset, \tag{19}$$

i.e. there exists a common (time-invariant) minimizer $x^* \in \mathcal{X}$ which minimizes f for any k with a common minimum value zero. Let further $g : \mathbb{R}^n \to \mathbb{R}$ be strictly convex, continuous differentiable, $\lambda_{max} \in [0, 1)$, and let D denotes the Bregman distance induced by g. Assume that D is convex in the second argument, then the proximal minimization iterations $x_k^*, \tilde{x}_k$ given by

$$\begin{aligned} x_{k+1}^* &= \arg\min_x f(x,k) + D(x, \tilde{x}_k) \\ \tilde{x}_{k+1} &= (1 - \lambda_{k+1}) x_{k+1}^* + \lambda_{k+1} \tilde{x}_k \end{aligned} \tag{20}$$

with $\lambda_{k+1} \in [0, \lambda_{max}]$, converge to a point in $\mathcal{X}$, i.e. $\lim_{k\to\infty} x_k^* = \lim_{k\to\infty} \tilde{x}_k = \tilde{x}^0 \in \mathcal{X}$.

Notice that the Bregman distance is in general not convex in the second argument, but there are important cases, such as $g(x) = x^T Q x$, $Q > 0$, where this holds Bauschke and Borwein (2001). Notice further that in a classical proximal minimization scheme $\tilde{x}_{k+1} = x_{k+1}^*$ $(\lambda_{k+1} = 0)$. Here, we pick $\tilde{x}_{k+1} = (1 - \lambda_{k+1}) x_{k+1}^* + \lambda_{k+1} \tilde{x}_k$ instead of x_{k+1}^* as the next iterate, because with appropriately chosen λ_{k+1}'s, the loss of stabilizability problem in the estimation scheme (15) can be avoided and it can be guaranteed that every estimated signal model is controllable, if the initial model is controllable, as shown next.

Lemma 5 *Consider $A_c, A_u \in \mathbb{R}^{n\times n}$, $B_c, B_u \in \mathbb{R}^{n\times q}$, $q \leq n$, and let (A_c, B_c) be controllable and (A_u, B_u) be not controllable. Then for any $\lambda_{max} \in (0, 1)$, there exists a $\lambda \in (0, \lambda_{max})$ such that $(A(\lambda), B(\lambda))$ with $A(\lambda) = (1 - \lambda) A_u + \lambda A_c$, $B(\lambda) = (1 - \lambda) B_u + \lambda B_c)$ is controllable. In particular, take some λ_j's with $0 < \lambda_1 < ... < \lambda_{2n^2+1} < \lambda_{max}$, then there exists an $i \in \{1, ..., 2n^2 + 1\}$ such that $(A(\lambda_i), B(\lambda_i))$ is controllable.*

The next theorem is the main result of this subsection.

Theorem 2 *Consider the sequences* (13) *with some given basis functions* (14) *and consider the optimization problem* (15)*. Suppose Assumption 6 and 7 hold true and assume that D is convex in the second argument. Then the following statements hold true.*

(i) The solution sequence $\{\tilde{\theta}(k)\}_{k\in\mathbb{N}}$ converges, i.e. $\lim_{k\to\infty}\tilde{\theta}(k)=\theta^*$.

(ii) If one defines $\tilde{\theta}(k)^\top=[\mathrm{vec}(A(k))^\top,\ \mathrm{vec}(B(k))^\top]$ and predictor maps (5) *according to*

$$A_i(k)=A(k)^i,\ \ B_i(k)=A(k)^iB(k), \tag{21}$$

$i=0...\bar{N}$, then the predictor fulfills the properties (6)*,* (7) *and* (8) *in Assumption 4(a)(b) with respect to the signal model* (18) *and the sequences* (14)*.*

(iii) In addition, if the initialization $(A(0),B(0))$, $\tilde{\theta}(0)^\top=[\mathrm{vec}(A(0))^\top,\ \mathrm{vec}(B(0))^\top]$, of (15) *is controllable, then there exists a sequence $\{\lambda_k\}_{k\in\mathbb{N}}$, $\lambda_{max}\in(0,1)$, (constructed for example according to Lemma 5) such that for any $k\in\mathbb{N}$ the pair $(A(k),B(k))$ is controllable and hence also Assumption 4(c) is fullfilled.*

5. The Overall Scheme

In Section 3, Theorem 1, we have established a control scheme which drives the state of the linear system (2) to zero assuming that the system is stabilizable and that state measurements as well as asymptotically accurate predictor maps are available. In Section 4, Theorem 2, we have established an estimation scheme which delivers asymptotically accurate predictor maps for any lifted signals (14) assuming that these signals can be embedded in a linear signal model of the form (18). Utilizing the estimation scheme (15) in the control scheme (3) means now that the signal model (18) replaces the system model (2) and the predictor maps in (21) are used to define the predictor (5). However, it needs to be clarified how the output and input sequence $\{y(k)\}_{k\in\mathbb{N}}$, $\{v(k)\}_{k\in\mathbb{N}}$ of (1) under Assumption 1 can be related to a signal model of the form (18) such that (A,B) is stabilizable and such that the (observable part of the) state $x(k)$ is available. This issue is addressed in the next lemma.

Lemma 6 *Consider an arbitrary output and input sequence $\{y(k)\}_{k\in\mathbb{N}}$, $\{v(k)\}_{k\in\mathbb{N}}$ of system* (1) *and suppose Assumption 1 holds true. Let*

$$\begin{aligned} x(k) &= \phi_{\mathrm{y}}(y(k),...,y(k-m+1)) = \left[y(k)^\top,...,y(k-m+1)^\top\right]^\top, \\ u(k) &= \phi_{\mathrm{v}}(v(k),...,v(k-m+1)) = \left[v(k)^\top,...,v(k-m+1)^\top\right]^\top \end{aligned} \tag{22}$$

with $m\geq n$. Then the sequences $\{u(k)\}_{k\in\mathbb{N}}$, $\{x(k)\}_{k\in\mathbb{N}}$ satisfy Assumption 7 with a stabilizable pair of matrices (A,B). In addition, if the sequences $\{u(k)\}_{k\in\mathbb{N}}$, $\{x(k)\}_{k\in\mathbb{N}}$ converge to zero when k goes to infinity, then so do the sequences $\{v(k)\}_{k\in\mathbb{N}}$, $\{y(k)\}_{k\in\mathbb{N}}$.

Remark 4 *Notice that the lifted input vector $u(k)$ in* (22) *contain past values of the actual input vector $v(k)$. In order to obtain a state space model with input $v(k)$, one can just add state variables to the signal model. In more detail, define an integrator chain dynamics of the form $\zeta_1(k+1)=\zeta_2(k),...,\zeta_{m-2}(k+1)=\zeta_m(k),\zeta_{m-1}(k+1)=v(k)$, hence $\zeta_1(k)$ corresponds to $v(k-m)$ etc. This augmentation does not effect the stabilizability property, since the states of the integrator chain converge to zero, if $v(k)$ converges to zero. This state augmentation in the signal model leads to matrices with at least the size $A\in\mathbb{R}^{mp+(m-1)q\times mp+(m-1)q}$, $B\in\mathbb{R}^{mp+(m-1)q\times q}$ and needs to be taken into account when implementing the receding horizon scheme.*

We are now ready to close the loop. By Lemma 6, we known that the output and input sequences $\{y(k)\}_{k\in\mathbb{N}}$, $\{v(k)\}_{k\in\mathbb{N}}$ of system (1) satisfy Assumption 7 and Assumption 2 w.r.t. the signal model (18). Assumption 3 and Assumption 6 can be satisfied by setting up the optimization problem accordingly. By Theorem 2, Assumption 4 holds. Hence all assumptions are satisfied and thus the receding horizon scheme guarantees, by Theorem 1 with $\Gamma(x) = \alpha x^T x, \alpha > 0$, a sequence $\epsilon(k) \to 0$, and Assumption 5 b), together with Lemma 6 that the state and the input of (1) converges to zero. These arguments lead to the next theorem.

Theorem 3 *Consider the closed loop system consisting of the system* (1)*, the receding horizon control scheme* (3) *and the proximity-based estimation scheme* (15)*. Define $x(k)$, $u(k)$ $(\phi_{\mathrm{y}}, \phi_{\mathrm{v}})$ according to equation* (22) *and set up the predictor scheme* (15) *according to Assumption 6. Further, set up the receding horizon scheme* (3) *according to Assumption 3 with $n = m$, $\alpha > 0, \Gamma(x) = \alpha x^T x$ and a sequence $\{\epsilon(k)\}_{k\in\mathbb{N}}$, $\epsilon(k) > 0$ that converges to zero. Then, under Assumption 1 and Assumption 5 b), for any initial state $z(0)$, the state $z(k)$ of the closed loop and the input $v(k)$ of the closed loop converges to zero as k goes to infinity.*

Notice that if we postulate that the limiting signal model $\hat{A}, \hat{B}$ is controllable/stabilizable, then we can drop Assumption 5 b) in Theorem 3. This could be enforced by a persistent excitation of the system.

6. Conclusion and Outlook

Motivation of this research was to develop a basic online optimization-based approach that guarantees convergence for a prototypical problem from adaptive control and that may serve as a basis for other online optimization-based (model-free) learning schemes. To this end, a receding horizon learning scheme consisting of a receding horizon control scheme and an proximity-based estimation scheme was proposed. Since the proposed approach relies on predictor maps, it can be considered as an indirect adaptive optimal control method and thus stands in contrast to direct adaptive optimal control methods such as reinforcement learning. From a conceptual point of view, the main ideas of this work are a time-varying terminal state weighting in a model-independent receding horizon control scheme (Section 3), a proximal estimation scheme with guaranteed convergence and the use of signal models and lifting techniques for predicting the closed loop trajectory (Section 4) as well as a proper combination of the control and estimation scheme to achieve guaranteed zero state convergence for completely unknown linear system under suitable assumptions (Section 5). The proposed overall scheme can be extended into several directions by addressing for example constraints, robustness, a priori knowledge about the system model, the peaking phenomena, a Bayesian view, real-time issues or dual (kernel) formulation of the regression problems Scherer and Holicki (2018); Feller and Ebenbauer (2017); Anderson and Dehghani (2007); Mayne et al. (2005), see Ebenbauer et al. (2020) for a detailed discussion on potential extensions. One open technical question is to investigate how Assumption 5 b) can be avoided without invoking a persistency of excitation condition and without significantly increasing the computational complexity. This question is related to the issue whether or not the limiting estimates correctly identify the excited controllable modes of the unknown system and how fast $\epsilon(k)$ converges to zero relative to the uncontrollable modes of the unknown system.

Acknowledgments

The work is supported partially by the National Nature Science Foundation of China (No. 61520 106008).

References

V. Adetola, D. DeHaan, and M. Guay. Adaptive model predictive control for constrained nonlinear systems. *Systems and Control Letters*, 58(5):320 – 326, 2009.

B.D.O. Anderson and A. Dehghani. Historical, generic and current challenges of adaptive control. *IFAC Proceedings Volumes*, 40(13):1 – 12, 2007. ISSN 1474-6670. 9th IFAC Workshop on Adaptation and Learning in Control and Signal Processing.

H.H. Bauschke and J.M. Borwein. Joint and separate convexity of the Bregman distance. *Studies in Computational Mathematics*, 8:23 – 36, 2001.

L. Beckenbach, P. Osinenko, and S. Streif. Addressing infinite-horizon optimization in MPC via Q-learning. *IFAC-PapersOnLine*, 51(20):60 – 65, 2018. 6th IFAC Conference on Nonlinear Model Predictive Control NMPC 2018.

A. Bemporad, L. Chisci, and E. Mosca. On the stabilizing property of SIORHC. *Automatica*, 30 (12):2013 – 2015, 1994. ISSN 0005-1098.

M. Benosman. *Learning-Based Adaptive Control: An Extremum Seeking Approach – Theory and Applications*. Elsevier Science, 2016. ISBN 9780128031513.

J. Berberich, J. Koehler, M.A. Mueller, and F. Allgoewer. Data-driven model predictive control with stability and robustness guarantees. *arXiv*, 2020.

D.P. Bertsekas and J.N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1st edition, 1996.

S. Bittanti and M.C. Campi. Adaptive control of linear time invariant systems: The "bet on the best" principle. *Communications in Information and Systems*, 6(4):299–320, 2006.

S. Bittanti, P. Bolzern, and M. Campi. Recursive least-squares identification algorithms with incomplete excitation: convergence analysis and application to adaptive control. *IEEE Transactions on Automatic Control*, 35(12):1371–1373, 1990.

J Coulson, J. Lygeros, and F. Dörfler. Data-enabled predictive control: In the shallows of the deepc. In *2019 18th European Control Conference (ECC)*, pages 307–312. IEEE, 2019.

C. Ebenbauer, F. Pfitz, and S. Yu. Control of unknown (linear) systems with receding horizon learning. 2020. arXiv (https://arxiv.org/abs/2010.05891).

A. A. Feldbaum. Dual control theory. i. *Avtomat. i Telemekh.*, 21:1240–1249, 1960.

C. Feller and C. Ebenbauer. Relaxed logarithmic barrier function based model predictive control of linear systems. *IEEE Transactions on Automatic Control*, 62:1223–1238, March 2017.

M. Gharbi, B. Gharesifard, and C. Ebenbauer. Anytime proximity moving horizon estimation: Stability and regret. 2020. arXiv.

G.C. Goodwin and K.S. Sin. *Adaptive Filtering Prediction and Control*. Dover Publications, Inc., USA, 2009.

U. Helmke, D. Prätzel-Wolters, and S. Schmid. Necessary and sufficient conditions for adaptive stabilization. In *New Trends in Systems Theory. Progress in Systems and Control Theory, Vol 7.*, pages 348–354. Birkhäuser, Boston, MA., 1991. doi: https://doi.org/10.1007/978-1-4612-0439-8_43.

J.P. Hespanha, D. Liberzon, and A.S. Morse. Overcoming the limitations of adaptive control by means of logic-based switching. *Systems & Control Letters*, 49:49–65, 2003.

L. Hewing, K.P. Wabersich, and M. Menner M.N. Zeilinger. Learning-based model predictive control: Toward safe learning in control. *Annual Review of Control, Robotics, and Autonomous Systems*, 3(1):269–296, 2020.

M. Korda and I. Mezic. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *arXiv*, 2017.

D. Limon, J. Calliess, and J.M. Maciejowski. Learning-based nonlinear model predictive control. *IFAC-PapersOnLine*, 50(1):7769 – 7776, 2017. ISSN 2405-8963. 20th IFAC World Congress.

S. Lucia and B. Karg. A deep learning-based approach to robust nonlinear model predictive control. *IFAC-PapersOnLine*, 51:511–516, 2018.

H. Mania, S. Tu, and B. Recht. Certainty equivalence is efficient for linear quadratic control. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32 (NIPS)*, pages 10154–10164. Curran Associates, Inc., 2019.

N. Matni, A. Proutiere, and A. Rantzerand S. Tu. From self-tuning regulators to reinforcement learning and back again. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 3724–3740. IEEE, 2019.

D.Q. Mayne. Model predictive control: Recent developments and future promise. *Automatica*, 50 (12):2967 – 2986, 2014.

D.Q. Mayne, M.M. Seron, and S.V. Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219 – 224, 2005.

A. S. Morse. Towards a unified theory of parameter adaptive control - Part II: Certainty equivalence and implicit tuning. *IEEE Transactions on Automatic Control*, 37(1):15–29, 1992.

E. Mosca. *Optimal, Predictive, and Adaptive Control*. Prentice-Hall, Inc., 1994. ISBN 0138476098.

B. Mårtensson. The order of any stabilizing regulator is sufficient a priori information for adaptive stabilization. *Systems and Control Letters*, 6(2):87 – 91, 1985. ISSN 0167-6911.

B. Mårtensson and J.W. Polderman. Correction and simplification to "the order of a stabilizing regulator is sufficient a priori information for adaptive stabilization". *Systems and Control Letters*, 20(6):465 – 470, 1993. ISSN 0167-6911.

C. Münzing. *A novel functional exosystem observer with application to relaxed-barrier MPC*. University of Stuttgart (Master Thesis, Advisor: C. Ebenbauer, C. Feller), 2017.

T. W. Nguyen, S.A. Ul Islam, A.L. Bruce, A. Goel, D.S. Bernstein, and I.V. Kolmanovsky. Output-feedback RLS-based model predictive control*. In *2020 American Control Conference (ACC)*, pages 2395–2400, 2020.

D. Papadimitrious, U. Rosolia, and F. Borrelli. Control of unknown nonlinear systems with linear time-varying MPC. *arXiv*, 2020.

C. De Persis and P. Tesi. Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Transactions on Automatic Control*, 65(3):909–924, 2019.

V. Peterka. Predictor-based self-tuning control. *Automatica*, 20(1):39 – 50, 1984. ISSN 0005-1098.

M. Prandini and M.C. Campi. A new recursive identification algorithm for singularity free adaptive control. *Systems and Control Letters*, 34(4):177 – 183, 1998.

B. Recht. A tour of reinforcement learning: The view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, 2:253–279, 2019.

C.W. Scherer and T. Holicki. An IQC theorem for relations: Towards stability analysis of data-integrated systems. *IFAC-PapersOnLine*, 51(25):390 – 395, 2018. 9th IFAC Symposium on Robust Control Design ROCOND 2018.

L. Schwenkel, M. Gharbi, S. Trimpe, and C. Ebenbauer. Online learning with stability guarantees: A memory-based real-time model predictive controller. *Automatica*, 122(12), 2020. (See also arXiv).

P. Tabuada and L. Fraile. Data-driven stabilization of SISO feedback linearizable systems. *arXiv*, 2020.

G. Tao. Multivariable adaptive control: A survey. *Automatica*, 50(11):2737 – 2764, 2014.

E. Terzi, M. Farina, L. Fagiano, and R. Scattolini. Learning-based predictive control for linear systems: A unitary approach. *Automatica*, 108(11):108473, 2019.