

Data-Enabled Predictive Control: In the Shallows of the DeePC

Jeremy Coulson

John Lygeros

Florian Dörfler

Abstract—We consider the problem of optimal trajectory tracking for unknown systems. A novel data-enabled predictive control (DeePC) algorithm is presented that computes optimal and safe control policies using **real-time feedback** driving the unknown system along a desired trajectory while satisfying system constraints. Using a finite number of data samples from the unknown system, our proposed algorithm uses a **behavioural systems theory approach** to learn a **non-parametric system model** used to predict future trajectories. The DeePC algorithm is shown to be equivalent to the classical and widely adopted Model Predictive Control (MPC) algorithm in the case of deterministic linear time-invariant systems. In the case of nonlinear stochastic systems, we propose regularizations to the DeePC algorithm. Simulations are provided to illustrate performance and compare the algorithm with other methods.

I. INTRODUCTION

As systems are becoming more complex and data is becoming more readily available, scientists and practitioners are beginning to bypass classical model-based techniques in favour of data-driven methods [1]. Data-driven methods are suitable for applications where first-principle models are not conceivable, when models are too complex for control design, and when thorough modelling and parameter identification is too costly.

A challenging problem in systems control is **optimal trajectory tracking**, where a control policy is computed based on output feedback that drives a dynamical system along a desired output trajectory while minimizing a stage cost and respecting safety constraints. One of the most celebrated and widely used control techniques for trajectory tracking is receding horizon Model Predictive Control (MPC), precisely because it allows one to include safety considerations during control design [2]. The key ingredient for MPC is an accurate parametric state space model of the system, but obtaining such a model is often the most time-consuming and expensive part of control design [3].

In the context of unknown black-box systems, there is no approach which solves the optimal trajectory tracking problem subject to constraints and partial (output) observations. However, some more benign variations of the optimal trajectory tracking problem have been approached using data-driven and learning based methods. We single out some references that are representative for the different approaches.

We begin with reinforcement learning approaches in which control actions are chosen to maximize a reward [4]. These

approaches usually require a large number of data samples to perform well, and are often sensitive to hyperparameters leading to non-reproducible and highly variable outcomes [4].

Other approaches propose performing sequential system identification (ID) and control. System ID can be used to produce an approximate model and provide finite sample guarantees quantifying model uncertainty, allowing for robust control design [5]. In this spirit, an end-to-end ID and control pipeline is given in [6] and arrives at a data-driven control solution with guarantees on the sample efficiency, stability, performance, and robustness. The system identification step in these approaches disregards one of the main advantages of a data-driven approach: independence from an underlying parametric representation. Additionally, they only consider regulation, rely on having full state information, and do not enforce constraint satisfaction.

Beyond reinforcement learning and sequential ID and control, there are many other safe learning approaches [7]. However, they rely on a-priori stabilizing controllers and safe regions, and thus apply only to a small class of problems.

MPC based on Dynamic Matrix Control has been historically used as a **data-driven control technique**, in which zero-initial condition step responses are used to predict future trajectories [8]. Although this technique has many limitations [9], it motivates the use of a non-parametric predictive control model. Other non-parametric predictive models have been proposed in [10], [11]. These methods do not solve the problem of optimal trajectory tracking with constraints, but serve as building blocks for our approach.

Here we present a **Data-enabled Predictive Control (DeePC)** algorithm. Unlike classical MPC and the learning-based control techniques outlined above, the DeePC algorithm does not rely on a parametric system representation. Instead, similar to [10], we approach the problem from a behavioural systems theory perspective [12]. Rather than attempting to learn a parametric system model, we aim at learning the **system's "behaviour"** (see Section IV for the precise definition). Our novel predictive control strategy computes optimal controls for unknown systems **using real-time output feedback via a receding horizon implementation**, allowing one to incorporate input/output constraints to ensure safety. We formally show the equivalence of the DeePC algorithm to the classical MPC algorithm in the special case of deterministic linear time-invariant (LTI) systems. Our approach is much simpler to implement than the learning-based techniques above, as well as model-based approaches which require system identification and state observer design. In contrast to most MPC formulations (where full state

All authors are with the Department of Information Technology and Electrical Engineering at ETH Zürich, Switzerland {jcoulson, lygeros, dorfler}@control.ee.ethz.ch. This research was supported by ETH Zürich funds and by the ERC under the project OCAL, grant number 797845.

measurement is required), the DeePC algorithm only requires output measurements. Additionally, since our approach does not rely on a parametric system model, we are hopeful that it can also be applied beyond deterministic LTI systems. To apply the algorithm to such systems, we propose some preliminary though insightful regularizations, and reason as to why these regularizations improve performance. All of the results are validated with a nonlinear and stochastic aerial robotics case study, in which the DeePC algorithm is shown to outperform sequential ID with MPC.

The remainder of the paper is organized as follows: in Section II, we formally define the problem. In Section III we review the basic notion of MPC. In Section IV we introduce behavioural system theory. Section V contains the DeePC algorithm. In Section VI we simulate the DeePC algorithm and compare it with sequential ID and MPC using a quadcopter simulation. We conclude the paper in Section VII. The proofs of the results can be found in [13].

II. PROBLEM STATEMENT

Consider the discrete-time system given by

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t), \end{cases} \quad (1)$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, $D \in \mathbb{R}^{p \times m}$, and $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$ are respectively the state, control input, and output of the system at time $t \in \mathbb{Z}_{\geq 0}$. Given a desired reference trajectory $r = (r_0, r_1, \dots) \in (\mathbb{R}^p)^{\mathbb{Z}_{\geq 0}}$, input constraint set $\mathcal{U} \subseteq \mathbb{R}^m$, output constraint set $\mathcal{Y} \subseteq \mathbb{R}^p$, we wish to apply control inputs such that the system output tracks the reference trajectory r while satisfying constraints and optimizing a cost function. Tracking of the trivial trajectory $r = (0, 0, \dots)$ is simply regulation.

In the case when the model for the system is *known*, i.e., matrices A , B , C and D are known, the problem can be approached using MPC (see Section III). This paper focuses on the above trajectory tracking problem in the case when the model for system (1) is *unknown*, but input/output data samples are available.

III. MPC: A BRIEF OVERVIEW

We outline the well-known receding-horizon model predictive control and estimation algorithm for trajectory tracking when the model of system (1) is known (see, e.g., [14]). Consider the following optimization problem:

$$\begin{aligned} & \underset{u, x, y}{\text{minimize}} && \sum_{k=0}^{N-1} \left(\|y_k - r_{t+k}\|_Q^2 + \|u_k\|_R^2 \right) \\ & \text{subject to} && x_{k+1} = Ax_k + Bu_k, \forall k \in \{0, \dots, N-1\}, \\ & && y_k = Cx_k + Du_k, \forall k \in \{0, \dots, N-1\}, \\ & && x_0 = \hat{x}(t), \\ & && u_k \in \mathcal{U}, \forall k \in \{0, \dots, N-1\}, \\ & && y_k \in \mathcal{Y}, \forall k \in \{0, \dots, N-1\}, \end{aligned} \quad (2)$$

where $N \in \mathbb{Z}_{>0}$ is the time horizon, $u = (u_0, \dots, u_{N-1})$, $x = (x_0, \dots, x_N)$, $y = (y_0, \dots, y_{N-1})$ are the decision

variables, and $r_{t+k} \in \mathbb{R}^p$ is the desired reference at time $t+k$, where $t \in \mathbb{Z}_{\geq 0}$ is the time at which the optimization problem should be solved. The norm $\|u_k\|_R$ denotes the quadratic form $u_k^T R u_k$ (similarly for $\|\cdot\|_Q$), where $R \in \mathbb{R}^{m \times m}$ is the control cost matrix and $Q \in \mathbb{R}^{p \times p}$ is the output cost matrix. The estimated state at time t is denoted by $\hat{x}(t)$ and the predicted state and output at time $t+k$ are denoted by x_k and y_k , respectively. If the entire state is measured then $\hat{x}(t) = x(t)$. When the state measurement is not available, but system (1) is observable, an observer is typically used to estimate the state based on knowledge of the system (1) and the measured output y [15].

The classical MPC algorithm involves solving optimization problem (2) in a receding horizon manner.

Algorithm 1 MPC

Input: (A, B, C, D) , reference trajectory r , past input/output data (u, y) , constraint sets \mathcal{U} and \mathcal{Y} , and performance matrices Q and R

- 1) Generate state estimate $\hat{x}(t)$ using past input/output data.
 - 2) Solve (2) for $u^* = (u_0^*, \dots, u_{N-1}^*)$.
 - 3) Apply inputs $(u(t), \dots, u(t+s)) = (u_0^*, \dots, u_s^*)$ for some $s \leq N-1$.
 - 4) Set t to $t+s$ and update past input/output data.
 - 5) Return to 1.
-

Under standard assumptions, it can be shown that the MPC algorithm is recursively feasible and stabilizing [16]. One crucial ingredient for MPC is an accurate model of the system; this is needed both to formulate problem (2) and, in cases where the state is not measured exactly, to generate the initial state estimates $\hat{x}(t)$. The need for a model is traditionally addressed through system identification, where observations of the system are collected offline before the online control operation begins and are used to estimate a model of the form (1) that matches the observed data in an appropriate sense. For complex systems, this can be a cumbersome and expensive process [3]. For this reason, we propose a Data-enabled Predictive Control (DeePC) algorithm which does not require an explicit model, system identification, or state estimation (see Algorithm 2).

IV. PRELIMINARIES

A. Non-parametric system representation

Behavioural system theory is a natural way of viewing a dynamical system when one is not concerned with a particular system representation, but rather the subspace of the signal space in which trajectories of the system live. This is in contrast with classical system theory, where a particular parametric system representation is used to describe the input/output behaviour, and properties of the system are derived by studying the chosen system representation. Following [12], we define a dynamical system and its properties in terms of its behaviour.

Definition 4.1: A *dynamical system* is a 3-tuple $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathcal{B})$ where $\mathbb{Z}_{\geq 0}$ is the discrete-time axis, \mathbb{W} is a signal space, and $\mathcal{B} \subseteq \mathbb{W}^{\mathbb{Z}_{\geq 0}}$ is the behaviour.

Definition 4.2: Let $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathcal{B})$ be a dynamical system.

- (i) $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathcal{B})$ is *linear* if \mathbb{W} is a vector space and \mathcal{B} is a linear subspace of $\mathbb{W}^{\mathbb{Z}_{\geq 0}}$.
- (ii) $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathcal{B})$ is *time invariant* if $\mathcal{B} \subseteq \sigma\mathcal{B}$ where $\sigma: \mathbb{W}^{\mathbb{Z}_{\geq 0}} \rightarrow \mathbb{W}^{\mathbb{Z}_{\geq 0}}$ is the forward time shift defined by $(\sigma w)(t) = w(t+1)$ and $\sigma\mathcal{B} = \{\sigma w \mid w \in \mathcal{B}\}$.
- (iii) $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathcal{B})$ is *complete* if \mathcal{B} is closed in the topology of pointwise convergence.

Note that if a dynamical system satisfies (i)-(ii) then (iii) is equivalent to finite dimensionality of \mathbb{W} (see [12, Section 7.1]). We denote the class of systems $(\mathbb{Z}_{\geq 0}, \mathbb{R}^{m+p}, \mathcal{B})$ satisfying (i)-(iii) by \mathcal{L}^{m+p} , where $m, p \in \mathbb{Z}_{\geq 0}$. With slight abuse of notation and terminology, we denote a dynamical system in \mathcal{L}^{m+p} only by its behaviour \mathcal{B} .

Next, we define the set $\mathcal{B}_T = \{w \in (\mathbb{R}^{m+p})^T \mid \exists v \in \mathcal{B} \text{ s.t. } w_t = v_t, 1 \leq t \leq T\}$ of trajectories truncated to a window of length T . Without loss of generality, we assume that \mathcal{B} can be written as the product space of two sub-behaviours \mathcal{B}^u and \mathcal{B}^y , where $\mathcal{B}^u = (\mathbb{R}^m)^{\mathbb{Z}_{\geq 0}}$ and $\mathcal{B}^y \subseteq (\mathbb{R}^p)^{\mathbb{Z}_{\geq 0}}$ are the spaces of input and output signals, respectively (see [17, Theorem 2]), that is, any trajectory $w \in \mathcal{B}$ can be written as $w = \text{col}(u, y)$, where $\text{col}(u, y) := (u^T, y^T)^T$. We now present two concepts: controllability, and persistency of excitation.

Definition 4.3: A system $\mathcal{B} \in \mathcal{L}^{m+p}$ is *controllable* if for every $T \in \mathbb{Z}_{>0}$, $w^1 \in \mathcal{B}_T$, $w^2 \in \mathcal{B}$ there exists $w \in \mathcal{B}$ and $T' \in \mathbb{Z}_{>0}$ such that $w_t = w_t^1$ for $1 \leq t \leq T$ and $w_t = w_t^2$ for $t > T + T'$.

In other words, a behavioural system is controllable if any two trajectories can be patched together in finite time.

Definition 4.4: Let $L, T \in \mathbb{Z}_{>0}$ such that $T \geq L$. The signal $u = \text{col}(u_1, \dots, u_T) \in \mathbb{R}^{Tm}$ is *persistently exciting of order L* if the Hankel matrix

$$\mathcal{H}_L(u) := \begin{pmatrix} u_1 & \cdots & u_{T-L+1} \\ \vdots & \ddots & \vdots \\ u_L & \cdots & u_T \end{pmatrix}$$

is of *full row rank*.

The term persistently exciting describes an input signal sufficiently rich and long as to *excite* the system yielding an output sequence that is representative for the system's behaviour.

B. Parametric system representation

There are several equivalent ways of representing a behavioural system $\mathcal{B} \in \mathcal{L}^{m+p}$, including the classical input/output/state representation (1) denoted by $\mathcal{B}(A, B, C, D) = \{\text{col}(u, y) \in (\mathbb{R}^{m+p})^{\mathbb{Z}_{\geq 0}} \mid \exists x \in (\mathbb{R}^n)^{\mathbb{Z}_{\geq 0}} \text{ s.t. } \sigma x = Ax + Bu, y = Cx + Du\}$. The input/output/state representation of smallest order (i.e., smallest state dimension) is called a *minimal representation*, and we denote its order by $n(\mathcal{B})$. Another important property of

a system $\mathcal{B} \in \mathcal{L}^{m+p}$ is the *lag* defined by the smallest integer $\ell \in \mathbb{Z}_{>0}$ such that the observability matrix $\mathcal{O}_\ell(A, C) := \text{col}(C, CA, \dots, CA^{\ell-1})$ has rank $n(\mathcal{B})$. We denote the lag by $\ell(\mathcal{B})$ (see [12, Section 7.2] for equivalent input/output/state representation free definitions of lag). We can now present a uniqueness result.

Lemma 4.1: ([10, Lemma 1]): Let $\mathcal{B} \in \mathcal{L}^{m+p}$ and $\mathcal{B}(A, B, C, D)$ a minimal input/output/state representation. Let $T_{\text{ini}}, N \in \mathbb{Z}_{>0}$ with $T_{\text{ini}} \geq \ell(\mathcal{B})$ and $\text{col}(u_{\text{ini}}, u, y_{\text{ini}}, y) \in \mathcal{B}_{T_{\text{ini}}+N}$. Then there exists a unique $x_{\text{ini}} \in \mathbb{R}^{n(\mathcal{B})}$ such that

$$y = \mathcal{O}_N(A, C)x_{\text{ini}} + \begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{N-2}B & \cdots & CB & D \end{pmatrix} u. \quad (3)$$

In other words, given a sufficiently long window of initial system data $\text{col}(u_{\text{ini}}, y_{\text{ini}})$, the state to which the system is driven by the sequence of inputs u_{ini} is unique. Furthermore, if the matrices A, B, C, D are known, the state x_{ini} can be computed. We now present a result known in behavioural systems theory as the *Fundamental Lemma* [10].

Lemma 4.2: ([18, Theorem 1]): Consider a controllable system $\mathcal{B} \in \mathcal{L}^{m+p}$. Let $T, t \in \mathbb{Z}_{>0}$, and $w = \text{col}(u, y) \in \mathcal{B}_T$. Assume u to be persistently exciting of order $t + n(\mathcal{B})$. Then $\text{colspan}(\mathcal{H}_t(w)) = \mathcal{B}_t$.

Note that the number of data points T must be at least $(m+1)(t + n(\mathcal{B})) - 1$ in order to satisfy the persistency of excitation condition. Lemma 4.2 replaces the need for a model or system identification process and allows for any trajectory of a controllable LTI system to be constructed using a finite number of data samples generated by a persistently exciting input signal. In a sense, the *Hankel matrix is itself a non-parametric predictive model* based on raw data. It allows one to implicitly estimate the state of an LTI system, predict its future behaviour, and design optimal feedforward control inputs [10]. Lemma 4.1 and Lemma 4.2 can be used together in a predictive control algorithm similar to Algorithm 1 by replacing the state observer (respectively, the system model) with the data $\text{col}(u_{\text{ini}}, y_{\text{ini}})$ (respectively, $\mathcal{H}_t(w)$).

V. DEEPC: DATA-ENABLED PREDICTIVE CONTROL

A. Data collection

We begin by assuming that the data is generated by an unknown controllable LTI system $\mathcal{B} \in \mathcal{L}^{m+p}$ with a minimal input/output/state representation $\mathcal{B}(A, B, C, D)$. Let $T, T_{\text{ini}}, N \in \mathbb{Z}_{>0}$ such that $T \geq (m+1)(T_{\text{ini}}+N+n(\mathcal{B}))-1$. Let $u^d = \text{col}(u_1^d, \dots, u_T^d) \in \mathbb{R}^{Tm}$ be a sequence of T inputs applied to \mathcal{B} , and $y^d = (y_1^d, \dots, y_T^d) \in \mathbb{R}^{Tp}$ the corresponding outputs. Furthermore, assume u^d is *persistently exciting of order $T_{\text{ini}}+N+n(\mathcal{B})$* . The superscript d is used to indicate that these are sequences of data samples collected during an offline procedure from the unknown system. Note that the data $\text{col}(u^d, y^d) \in \mathcal{B}_T$ can be equivalently thought of as coming from the minimal input/output/state representation $\mathcal{B}(A, B, C, D)$. Next, we partition the input/output data into

two parts which we call *past data* and *future data*. More formally, define

$$\begin{pmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{pmatrix} := \mathcal{H}_{T_{\text{ini}}+N}(u^d), \quad \begin{pmatrix} Y_p \\ Y_f \end{pmatrix} := \mathcal{H}_{T_{\text{ini}}+N}(y^d), \quad (4)$$

where U_p consists of the first T_{ini} block rows of $\mathcal{H}_{T_{\text{ini}}+N}(u^d)$ and U_f consists of the last N block rows of $\mathcal{H}_{T_{\text{ini}}+N}(u^d)$ (similarly for Y_p and Y_f). In the sequel, past data denoted by the subscript p will be used to estimate the initial condition of the underlying state, whereas the future data denoted by the subscript f will be used to predict the future trajectories.

B. State estimation and trajectory prediction

By Lemma 4.2, we can construct any $T_{\text{ini}} + N$ length trajectory of $\mathcal{B}_{T_{\text{ini}}+N}$ using the data collected. Indeed, a trajectory $\text{col}(u_{\text{ini}}, u, y_{\text{ini}}, y)$ belongs to $\mathcal{B}_{T_{\text{ini}}+N}$ if and only if there exists $g \in \mathbb{R}^{T-T_{\text{ini}}-N+1}$ such that

$$\begin{pmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{pmatrix} g = \begin{pmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \\ y \end{pmatrix}. \quad (5)$$

If $T_{\text{ini}} \geq \ell(\mathcal{B})$, then Lemma 4.1 implies that there exists a unique $x_{\text{ini}} \in \mathbb{R}^{n(\mathcal{B})}$ such that the output y is uniquely determined by (3). Intuitively, the trajectory $\text{col}(u_{\text{ini}}, y_{\text{ini}})$ fixes the underlying initial state x_{ini} from which the trajectory $\text{col}(u, y)$ evolves. Note, however, that (5) does not require the input/output/state representation of the system to be known. The state x_{ini} is only “fixed” implicitly by $\text{col}(u_{\text{ini}}, y_{\text{ini}})$. As first shown in [10], this allows one to predict future trajectories. Indeed, given an initial trajectory $\text{col}(u_{\text{ini}}, y_{\text{ini}}) \in \mathcal{B}_{T_{\text{ini}}}$ of length $T_{\text{ini}} \geq \ell(\mathcal{B})$ and a sequence of future inputs $u \in \mathbb{R}^{Nm}$, the first three block equations of (5) can be solved for g . The sequence of future outputs are then given by $y = Y_f g$. By Lemma 4.1 the vector y contains the unique sequence of outputs corresponding to the inputs u .

C. DeePC algorithm

Given a time horizon $N \in \mathbb{Z}_{>0}$, a reference trajectory $r = (r_0, r_1, \dots) \in (\mathbb{R}^p)^{\mathbb{Z}_{\geq 0}}$, past input/output data $\text{col}(u_{\text{ini}}, y_{\text{ini}}) \in \mathcal{B}_{T_{\text{ini}}}$, input constraint set $\mathcal{U} \subseteq \mathbb{R}^m$, output constraint set $\mathcal{Y} \subseteq \mathbb{R}^p$, output cost matrix $Q \in \mathbb{R}^{p \times p}$, and control cost matrix $R \in \mathbb{R}^{m \times m}$, we formulate the following optimization problem:

$$\begin{aligned} & \underset{g, u, y}{\text{minimize}} && \sum_{k=0}^{N-1} \left(\|y_k - r_{t+k}\|_Q^2 + \|u_k\|_R^2 \right) \\ & \text{subject to} && \begin{pmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{pmatrix} g = \begin{pmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \\ y \end{pmatrix}, \\ & && u_k \in \mathcal{U}, \forall k \in \{0, \dots, N-1\}, \\ & && y_k \in \mathcal{Y}, \forall k \in \{0, \dots, N-1\}. \end{aligned} \quad (6)$$

Note here, that u and y are not independent decision variables of the optimization problem. Rather they are described completely by the fixed data matrices U_f and Y_f and the

decision variable g . A comparison of the two optimal control problems (2) and (6) yields only a single (though key) difference; the model and the state estimate in (2) are replaced completely with input/output data samples in (6). We now present the DeePC algorithm.

Algorithm 2 DeePC

Input: $\text{col}(u^d, y^d) \in \mathcal{B}_T$, reference trajectory $r \in \mathbb{R}^{Np}$, past input/output data $\text{col}(u_{\text{ini}}, y_{\text{ini}}) \in \mathcal{B}_{T_{\text{ini}}}$, constraint sets \mathcal{U} and \mathcal{Y} , and performance matrices Q and R

- 1) Solve (6) for g^* .
- 2) Compute the optimal input sequence $u^* = U_f g^*$.
- 3) Apply input $(u(t), \dots, u(t+s)) = (u_0^*, \dots, u_s^*)$ for some $s \leq N-1$.
- 4) Set t to $t+s$ and update u_{ini} and y_{ini} to the T_{ini} most recent input/output measurements.
- 5) Return to 1.

D. Equivalence of DeePC and MPC

It can be shown that Algorithm 1 and Algorithm 2 yield equivalent closed-loop trajectories under certain assumptions.

Theorem 5.1: (Feasible Set Equivalence): Consider a controllable LTI system $\mathcal{B} \in \mathcal{L}^{m+p}$ with minimal input/output/state representation $\mathcal{B}(A, B, C, D)$ given as in (1). Consider the MPC and DeePC optimization problems (2) and (6). Let $T_{\text{ini}} \geq \ell(\mathcal{B})$ and $\text{col}(u_{\text{ini}}, y_{\text{ini}}) \in \mathcal{B}_{T_{\text{ini}}}$ be the most recent input/output measurements from system (1). Assume that the data $\text{col}(u^d, y^d) \in \mathcal{B}_T$ in $\text{col}(U_p, Y_p, U_f, Y_f)$ is such that u^d is persistently exciting of order $T_{\text{ini}} + N + n(\mathcal{B})$, where $T \geq (m+1)(T_{\text{ini}} + N + n(\mathcal{B})) - 1$. Then there exists a state estimate $\hat{x}(t)$ in (2) such that the feasible sets of (2) and (6) are equal.

Corollary 5.1: (Equivalent Closed Loop Behaviour): Consider the MPC Algorithm 1 and the DeePC Algorithm 2 with $Q \succeq 0$, $R \succ 0$, and \mathcal{U} , \mathcal{Y} convex and non-empty. Under the assumptions of Theorem 5.1, Algorithm 1 and Algorithm 2 result in equivalent closed-loop behaviour, i.e., the optimal control sequence u^* and corresponding system output y^* at every iteration is identical.

One notable feature of the DeePC algorithm is its simplicity. The DeePC algorithm achieves system ID, state estimation, and trajectory prediction with one linear equation resulting in a quadratic program with $T - T_{\text{ini}} - N + 1$ number of decision variables where $T \in \mathbb{Z}_{>0}$ is the amount of data collected. In order to satisfy the persistency of excitation assumption in Theorem 5.1, one must collect a minimum of $(m+1)(T_{\text{ini}} + N + n(\mathcal{B})) - 1$ data samples implying that the number of decision variables in (6) is at least $m(T_{\text{ini}} + N) + (m+1)n(\mathcal{B})$.

Note that the persistency of excitation assumption assumes knowledge of $n(\mathcal{B})$ and $\ell(\mathcal{B})$, which are properties that are a priori unknown. We know that $\ell(\mathcal{B}) \leq n(\mathcal{B})$ by the Cayley-Hamilton theorem. Hence an upper bound for $n(\mathcal{B})$ is sufficient. In practice, one would simply collect a sufficiently large amount of data to exceed the necessary amount for persistency of excitation.

VI. BEYOND DETERMINISTIC LTI SYSTEMS

In this section we provide a promising extension of the DeePC algorithm beyond deterministic LTI systems. We offer insightful algorithmic extensions by means of salient regularizations, show their utility through a numerical study, and provide plausible reasoning for the regularizations.

A. Regularized DeePC Algorithm

Consider now the nonlinear discrete-time system whose states and outputs are given by $x(t+1) = f(x(t), u(t))$ and $y(t) = h(x(t), u(t), \eta(t))$, respectively, where $\eta(t) \in \mathbb{R}^p$ is white measurement noise, and $f: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ and $h: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ are not necessarily linear. One may also consider a system affected by process noise. However, we focus on systems **only affected by measurement noise** in order to isolate its effect on the DeePC algorithm. To apply the DeePC algorithm to the system above, we propose two regularizations to the optimal control problem (6). In particular, we introduce the following *regularized* optimization problem:

$$\begin{aligned} & \underset{g, u, y, \sigma_y}{\text{minimize}} && \sum_{k=0}^{N-1} \left(\|y_k - r_{t+k}\|_Q^2 + \|u_k\|_R^2 \right) \\ & && + \lambda_g \|g\|_1 + \lambda_y \|\sigma_y\|_1 \\ & \text{subject to} && \begin{pmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{pmatrix} g = \begin{pmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ \sigma_y \\ 0 \\ 0 \end{pmatrix}, \quad (7) \\ & && u_k \in \mathcal{U}, \quad \forall k \in \{0, \dots, N-1\}, \\ & && y_k \in \mathcal{Y}, \quad \forall k \in \{0, \dots, N-1\}, \end{aligned}$$

where $\sigma_y \in \mathbb{R}^{T_{\text{ini}} p}$ is an auxiliary slack variable, $\lambda_y, \lambda_g \in \mathbb{R}_{>0}$ are regularization parameters. Let us explain these regularizations. We offer convincing numerical evidence in Section VI-B.

Slack variable: When the output measurements are corrupted **by noise**, the constraint equation in (6) may become inconsistent. Hence, in (7) we include the **slack variable σ_y in the constraint to ensure feasibility of the constraint at all times**. We penalize the slack variable with a weighted one-norm penalty function. Choosing λ_y sufficiently large gives the desired property that $\sigma_y \neq 0$ only if the constraint is infeasible, that is, only if the data is inconsistent.

One-norm regularization on g : The cost includes a one-norm penalty on g . We conjecture that this regularization is related to **distributionally robust optimization problems**, in which similar regularization terms arise [19].

B. Aerial Robotics Case Study

We illustrate the performance of the regularized DeePC algorithm, i.e., Algorithm 2 with (7), by simulating it on a high-fidelity nonlinear quadcopter model, and compare the performance to system identification (ID) followed by MPC using the identified model. The states of the quadcopter model are given by the 3 spatial coordinates (x, y, z) and their velocities, and the 3 angular

coordinates (α, β, γ) and their velocities, i.e., the state is $(x, y, z, \dot{x}, \dot{y}, \dot{z}, \alpha, \beta, \gamma, \dot{\alpha}, \dot{\beta}, \dot{\gamma})$. The inputs are given by the thrusts from the 4 rotors, (u_1, u_2, u_3, u_4) . We assume full state measurement to facilitate the comparison to standard MPC. Data was collected from the nonlinear model subject to additive white measurement noise. We collected 214 input/output measurements with a sample time of 0.1 seconds. Drawing the input sequence from a **uniformly distributed random variable** ensured that the data was persistently exciting. For the model-based MPC, the data was used to identify the system parameters through the least square prediction error method with an assumed state dimension of 12. The following parameters were chosen for the optimization problems (2) and (7): $N = 30$, $T_{\text{ini}} = 1$, $R = I$, $Q = \text{diag}(200, 200, 300, 1, \dots, 1)$, $\lambda_g = 30$, $\lambda_y = 10^5$. The thrust from each rotor was constrained between 0 and 1, and the (x, y, z) coordinates were constrained between -3 and 3 .

We simulated the system ID followed by the MPC algorithm and the regularized DeePC algorithm on the nonlinear and stochastic quadcopter model in which the quadcopter was commanded to follow a series of figure-eight trajectories for a duration of 60 seconds. Figure 1 shows that DeePC performs better than sequential ID and MPC in terms of reference tracking and constraint satisfaction.

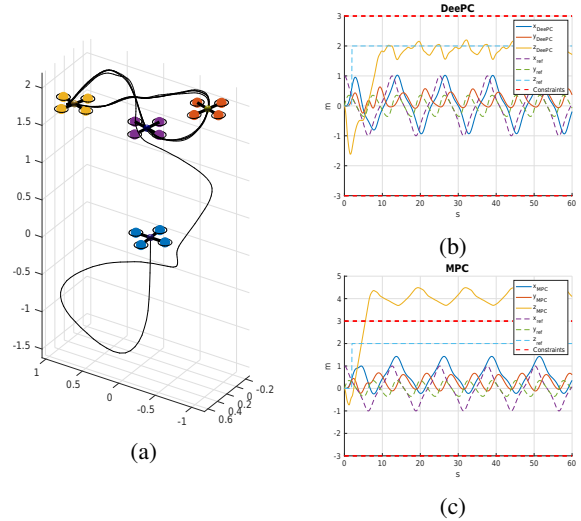


Fig. 1: Figure (a): three dimensional plot of the trajectory of the quadcopter at different instances of time controlled with DeePC. Figure (b) and (c): trajectories of the spatial coordinates when controlled by DeePC and MPC, respectively. The horizontal red dashed lines represent constraints.

Another simulation was performed in which the quadcopter was commanded to perform a simple step trajectory in the (x, y, z) coordinates with the same constraints listed above. The duration of constraint violations and the cost were measured. This was repeated 30 times with different data sets for constructing $\text{col}(U_p, Y_p, U_f, Y_f)$, and different random seeds for the measurement noise. The results are displayed in Figure 2 and show that DeePC consistently outperforms identification-based MPC in terms of cost and

constraint satisfaction. While these observations should be made cautiously, an intuitive explanation for the superior performance of DeePC is that (7) simultaneously optimizes for the best system model, state estimation, and control policy, whereas conventional MPC requires fixing a system model and performs these tasks independently.

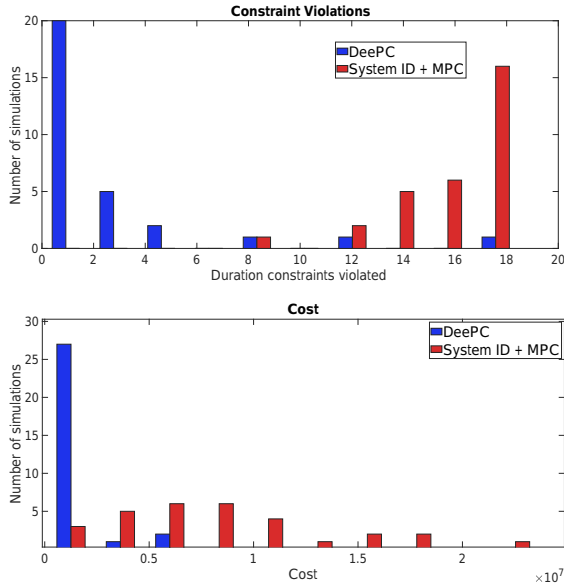


Fig. 2: Constraint violations and cost comparison of DeePC and system ID with MPC.

To study the effect of the regularizations on the performance of the DeePC algorithm, we performed a sensitivity analysis on the regularization parameters λ_y and λ_g . The quadcopter was commanded to follow the same step trajectory. The duration of constraint violations and cost were measured. This was repeated 8 times with different data sets, and the duration of constraint violations and cost were averaged over these 8 data sets. The results in Figure 3 show that the regularizations improve performance.

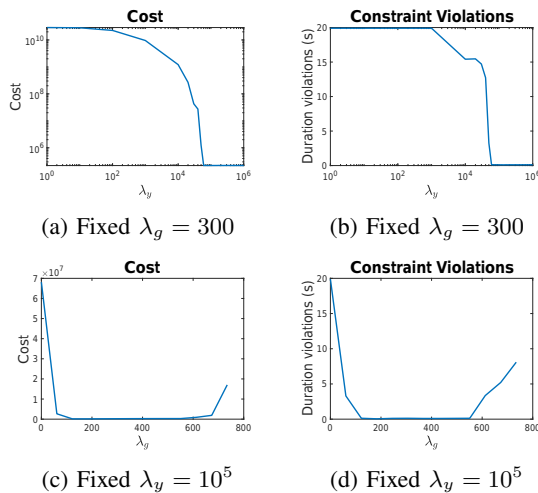


Fig. 3: Performance of quadcopter controlled by the DeePC algorithm with different regularization parameters.

VII. CONCLUSION

We presented a DeePC algorithm that can be applied to unknown LTI systems and formally showed its equivalence to the classical MPC algorithm. The DeePC algorithm uses data to compute optimal controls that drive the system along a desired trajectory while respecting system constraints. Furthermore, we simulated a regularized version of the algorithm on stochastic nonlinear quadcopter dynamics illustrating its capabilities beyond deterministic LTI systems. Future work focuses on the robustness of DeePC when applied to stochastic and nonlinear dynamics.

Acknowledgements: We thank Bassam Bamieh, Ben Recht, Ashish Cherukuri, and Manfred Morari for useful discussions, and Melanie Zeilinger and Colin Jones for providing the simulation model.

REFERENCES

- [1] Z. S. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Information Sciences*, vol. 235, pp. 3–35, 2013.
- [2] J. B. Rawlings and D. Q. Mayne, *Model Predictive Control: Theory and Design*. Nob Hill Publishing, 2009.
- [3] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, no. 3, pp. 393–438, 2005.
- [4] B. Recht, "A tour of reinforcement learning: The view from continuous control," *arXiv preprint arXiv:1806.09460*, 2018.
- [5] M. C. Campi and E. Weyer, "Finite sample properties of system identification methods," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1329–1334, 2002.
- [6] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *arXiv preprint arXiv:1710.01688*, 2017.
- [7] F. Berkenkamp, M. Turchetta, A. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," in *Advances in Neural Information Processing Systems*, 2017, pp. 908–918.
- [8] R. W. Freedman and A. Bhatia, "Adaptive dynamic matrix control: Online evaluation of the DMC model coefficients," in *American Control Conference*. IEEE, 1985, pp. 220–225.
- [9] P. Lundström, J. H. Lee, M. Morari, and S. Skogestad, "Limitations of dynamic matrix control," *Computers & Chemical Engineering*, vol. 19, no. 4, pp. 409–421, 1995.
- [10] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.
- [11] W. Favoreel, B. De Moor, P. Van Overschee, and M. Gevers, "Model-free subspace-based LQG-design," in *American Control Conference*, vol. 5. IEEE, 1999, pp. 3372–3376.
- [12] I. Markovsky, J. C. Willems, S. Van Huffel, and B. De Moor, *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. SIAM, 2006.
- [13] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," *arXiv preprint arXiv:1811.05890*, 2018.
- [14] A. Ferramosca, D. Limón, I. Alvarado, T. Alamo, and E. F. Camacho, "MPC for tracking with optimal closed-loop performance," *Automatica*, vol. 45, no. 8, pp. 1975 – 1978, 2009.
- [15] D. Q. Mayne, "Model predictive control: Recent developments and future promise," *Automatica*, vol. 50, no. 12, pp. 2967–2986, 2014.
- [16] E. F. Camacho and C. B. Alba, *Model Predictive Control*. Springer Science & Business Media, 2013.
- [17] J. C. Willems, "From time series to linear system – Part I. Finite dimensional linear time invariant systems," *Automatica*, vol. 22, no. 5, pp. 561–580, 1986.
- [18] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.
- [19] S. Shafieezadeh-Abadeh, D. Kuhn, and P. M. Esfahani, "Regularization via mass transportation," *arXiv preprint arXiv:1710.10016*, 2017.