

An Extended Kalman Filter for Data-Enabled Predictive Control

Journal Article**Author(s):**

Alpago, Daniele; [Dörfler, Florian](#) ; [Lygeros, John](#) 

Publication date:

2020-10

Permanent link:

<https://doi.org/10.3929/ethz-b-000425063>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

IEEE Control Systems Letters 4(4), <https://doi.org/10.1109/LCSYS.2020.2998296>

Funding acknowledgement:

787845 - Optimal control at large (EC)

An Extended Kalman Filter for Data-enabled Predictive Control

Daniele Alpagò, Florian Dörfler and John Lygeros

Abstract—The literature dealing with data-driven analysis and control problems has significantly grown in the recent years. Most of the recent literature deals with linear time-invariant systems in which the uncertainty (if any) is assumed to be deterministic and bounded; relatively little attention has been devoted to stochastic linear time-invariant systems. As a first step in this direction, we propose to equip the recently introduced Data-enabled Predictive Control algorithm with a data-based Extended Kalman Filter to make use of additional available input-output data for reducing the effect of noise, without increasing the computational load of the optimization procedure.

Index Terms—Optimal control, Stochastic systems, Kalman filtering.

I. INTRODUCTION

WITH the increasing complexity of applications in science and engineering, the use of model-based control techniques is becoming more and more challenging as they usually require accurate descriptions of model and uncertainties, often difficult to obtain. The need of data-driven approaches was already perceived in the 1980s with the advent of system identification [1] and adaptive control [2], and received new impetus in the 1990s: cornerstones such as Iterative Feedback Tuning, Correlation-based Tuning, Virtual Reference Feedback Tuning [3]–[5] are among the most effective approaches and currently represent active research topics. Another successful stream of ongoing work originated in those years was that of Reinforcement Learning (RL) [6]. In particular, model-free RL approaches such as Policy Search [7] and Approximate Dynamic Programming [8] represent flexible frameworks for phrasing relevant data-driven control problems, see [9], [10] and references therein. In recent years, the rediscovery of a result originally formulated in the context of behavioral system theory [11], [12] by J. C. Willems and coauthors, known as the *Fundamental Lemma* [13], shed new light on data-driven analysis and control approaches. The result states that if the input signal is sufficiently rich, all possible trajectories of a deterministic, linear time-invariant (LTI) system can be generated from linear combinations of past trajectories of the same system. This allows one to use a Hankel matrix constructed from input-output data as an implicit representation of the underlying dynamics. This representation has been first exploited for data-driven simulation

and linear-quadratic tracking for deterministic LTI systems in a behavioral setting [14], [15]. This has resulted in a growing stream of literature dealing with data-driven analysis [16], [17] and control problems [18]–[21].

In this context, particular attention has been devoted to the problem of optimal trajectory tracking [22]–[26], widely studied in model-based control. The aim is to compute an optimal control policy based on output feedback that drives the system along an output reference while minimizing a stage cost and satisfying input and output constraints. Model Predictive Control (MPC) has been one of the most effective methods to tackle such problems. MPC requires an accurate model of the system and an accurate description of the uncertainty (if any), which might be challenging and expensive to obtain in many applications [27], [28]. The success of MPC and the difficulties arising from the requirement for models have led to the introduction of a Data-enabled Predictive Control (DeePC) algorithm. The latter does not rely on a particular parametric system representation, but rather on a Hankel-matrix representation of the underlying dynamics, constructed from the system's input-output trajectories [23]–[25]. Unlike classical model-based predictive control methods, DeePC performs the control computation without identifying the system by solving a (parametric) convex optimization problem that includes the Hankel matrix in the constraints. The complexity of such optimization depends on the system dimensions, the MPC horizon and the amount of available data. Therefore, it is not obvious how to efficiently include additional data (for example, past measurements of the output of the closed-loop system) in the procedure, without increasing the computational burden of the optimization that has to be solved online.

Unlike earlier methods, recent data-driven approaches based on the Fundamental Lemma have devoted little attention to stochastic systems. Some approaches adopt a “robust control” perspective and treat the uncertainty as a deterministic and bounded sequence [18], [20], [26], sometimes affecting just the output. The focus of the present paper is to extend the DeePC algorithm to stochastic LTI systems and design methods for including more data to improve closed-loop performance, without increasing the computational load of the optimization. This is accomplished through a combination of an offline averaging of Hankel matrix predictors and an online, data-driven Extended Kalman Filter (EKF). The latter is based on an implicit model constructed from the parametric solution of the DeePC optimization program. The combination of off-line averaging with online EKF dramatically improves the closed-loop performance, as evidenced by numerical experiments.

In Section II we introduce the notation and recall some

Research supported by the ERC under the project OCAL, the SEMP and the Gini Foundation. D. Alpagò is with the Department of Information Engineering at University of Padova, Italy, dalpagò@dei.unipd.it; F. Dörfler and J. Lygeros are with the Department of Information Technology and Electrical Engineering at ETH Zürich, Switzerland, dorfler@ethz.ch, jlygeros@ethz.ch.

preliminary results. In Section III the DeePC algorithm is reviewed and the issues that motivate our approach are highlighted. Section IV-B introduces the proposed approach and Section V presents numerical experiments. Finally, in Section VI we draw conclusions and outline future lines of research.

II. NOTATION AND PRELIMINARIES

We recall the definition of persistently exciting signals and a state-space version of a result from behavioral system theory [11], [12], known as the *Fundamental Lemma* [13], [14].

Notation. Given a matrix $A \in \mathbb{R}^{m \times n}$, A^\top denotes its transpose and A^{-1} its inverse (if $m = n$). The notation $A(i: j, :)$, $i \leq j \leq m$, stands for the sub-matrix of A that goes from the i -th row of A to the j -th row of A , included. If $i = j$ we simply write $A(i, :)$. We denote with I_m the identity matrix of dimension $m \times m$. With $\text{diag}(A_1, \dots, A_n)$ we denote the block-diagonal matrix formed with the matrices A_1, \dots, A_n . Given a sequence of matrices A_h, \dots, A_{h+k} in $\mathbb{R}^{m \times n}$, with $k > 0$, we denote by $A_{h,h+k} := \text{col}(A_h, \dots, A_{h+k}) := [A_h^\top \dots A_{h+k}^\top]^\top \in \mathbb{R}^{(k+1)m \times n}$. Given a vector $w \in \mathbb{R}^m$ we introduce the quadratic form $\|w\|_P^2 = w^\top P w$ which coincides with the squared Euclidean norm $\|w\|_2^2$ when $P = I_m$. The symbol $\mathbb{E}[\cdot]$ denotes the expectation operator.

Persistency of Excitation. Let w_0, \dots, w_{N-1} be $N \in \mathbb{N}$ samples of a given sequence $(w_k)_{k \in \mathbb{Z}}$ taking values in \mathbb{R}^q . For $M \in \mathbb{N}$, $M \leq N$, let

$$\mathcal{H}_M(w_{0,N-1}) := \begin{bmatrix} w_0 & w_1 & \dots & w_{N-M} \\ w_1 & w_2 & \dots & w_{N-M+1} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M-1} & w_M & \dots & w_{N-1} \end{bmatrix}$$

be the block-Hankel matrix associated to the trajectory $w_{0,N-1}$, with M block-rows and the maximal number of columns $N - M + 1$. We say that the signal $w_{0,N-1}$ is *persistently exciting* of order M if the Hankel matrix $\mathcal{H}_M(w_{0,N-1})$ has full row-rank qM . This requires the sequence $w_{0,N-1}$ to be rich and long enough, in particular $N \geq (q+1)M - 1$.

Fundamental Lemma. Consider a state-space representation of an n -dimensional LTI system

$$\begin{cases} x_{k+1} = Ax_k + Bu_k \\ y_k = Cx_k + Du_k \end{cases} \quad (1)$$

where $(x_k)_{k \in \mathbb{Z}}$ is the n -dimensional state-process, $(u_k)_{k \in \mathbb{Z}}$ the m -dimensional input process, and $(y_k)_{k \in \mathbb{Z}}$ is the p -dimensional output process. Given an initial condition $x_0 \in \mathbb{R}^n$ and a sequence of inputs $u_{0,k-1} \in \mathbb{R}^{mk}$, the output of the system can be written as

$$y_{k-1} = CA^{k-1}x_0 + [CA^{k-2}B \dots CBD]u_{0,k-1}, \quad k \geq 2, \quad (2)$$

and $y_0 = Cx_0 + Du_0$. Let $T \in \mathbb{N}$ and collect T -long input-output trajectories $u_{0,T-1}, y_{0,T-1}$ of the system (1). For fixed positive integers $N_p, N_f \in \mathbb{N}$ and $k \geq N_p$, we can associate to the vectors

$$\begin{aligned} u_p^{(k)} &:= u_{k-N_p, k-1}, & y_p^{(k)} &:= y_{k-N_p, k-1}, \\ u_f^{(k)} &:= u_{k, k+N_f-1}, & y_f^{(k)} &:= y_{k, k+N_f-1}, \end{aligned} \quad (3)$$

the block-Hankel matrices

$$\begin{aligned} \begin{bmatrix} U_p \\ U_f \end{bmatrix} &= \begin{bmatrix} u_p^{(N_p)} & \dots & u_p^{(T-N_f-1)} \\ u_f^{(N_p)} & \dots & u_f^{(T-N_f-1)} \end{bmatrix} =: \mathcal{H}_{N_p+N_f}(u_{0,T-1}), \\ \begin{bmatrix} Y_p \\ Y_f \end{bmatrix} &= \begin{bmatrix} y_p^{(N_p)} & \dots & y_p^{(T-N_f-1)} \\ y_f^{(N_p)} & \dots & y_f^{(T-N_f-1)} \end{bmatrix} =: \mathcal{H}_{N_p+N_f}(y_{0,T-1}). \end{aligned}$$

From (2) we can then construct the $(m+p)(N_p+N_f) \times (T - N_p - N_f + 1)$ data matrix

$$\mathcal{H} := \begin{bmatrix} U_p \\ U_f \\ Y_p \\ Y_f \end{bmatrix} = \begin{bmatrix} 0 & I_{m(N_p+N_f)} \\ 0 & I_{m(N_p+N_f)} \\ \mathcal{O}_p(A, C) & \mathcal{T}_p(B, D) \\ \mathcal{O}_f(A, C) & \mathcal{T}_f(B, D) \end{bmatrix} \begin{bmatrix} X \\ X \end{bmatrix}, \quad (4)$$

where $U := [u_0 \dots u_{T-1}]$ and $X := [x_0 \dots x_{T-1}]$ are the block-Hankel matrices containing the inputs and the corresponding sequence of states, respectively. Here,

$$\begin{aligned} \mathcal{O}_p(A, C) &:= \text{col}(C, CA, \dots, CA^{N_p-1}), \\ \mathcal{O}_f(A, C) &:= \text{col}(CA^{N_p}, CA^{N_p+1}, \dots, CA^{N_f-1}), \end{aligned}$$

are observability matrices,

$$\mathcal{T}_p(B, D) := \begin{bmatrix} D & 0 & \dots & 0 & 0 & \dots & 0 \\ CB & D & \dots & 0 & 0 & \dots & 0 \\ CAB & CB & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & 0 \\ CA^{N_p-2}B & CA^{N_p-3}B & \dots & D & 0 & \dots & 0 \end{bmatrix}, \quad (5)$$

and $\mathcal{T}_f(B, D)$ is defined similarly. The following fundamental result from [13] guarantees that the matrix (4) can be used in place of the parametric representation (1), as long as the input $u_{0,T-1}$ is persistently exciting.

Lemma 1 (Fundamental Lemma): Assume the system (1) to be controllable and the input trajectory $u_{0,T-1}$ to be persistently exciting of order $N_p + N_f + n$. Then, a sequence $\text{col}(u_p, u_f, y_p, y_f)$ is an input-output trajectory of the system (1) if and only if it is in the range space of \mathcal{H} .

Recalling the persistency of excitation condition, $T \geq (m+1)(N_p + N_f + n) + 1$ is a necessary condition for $u_{0,T-1}$ to be persistently exciting of order $N_p + N_f + n$. Lemma 1 has been originally proven using the behavioral language in [13, Theorem 1]. For an equivalent recent state-space proof see [16] and [20, Appendix A].

III. DATA-ENABLED PREDICTIVE CONTROL REVIEW

We briefly introduce the recently proposed Data-enabled Predictive Control (DeePC) method [23], and highlight some related issues when dealing with stochastic LTI systems. This will serve as a motivation for what follows.

Consider the stochastic version of (1)

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + Ev_k \\ y_k = Cx_k + Du_k + Fv_k \end{cases} \quad (6)$$

where $(v_k)_{k \in \mathbb{Z}}$ is p -dimensional white noise (zero mean and unit-variance). We collect sufficiently long input-output trajectories $u_{0,T-1}, y_{0,T-1}$, i.e. $T \geq (m+1)(N_p + N_f + n) + 1$.

The same computations leading to (4), lead to the matrix corresponding to the stochastic model (6):

$$\mathcal{H} := \begin{bmatrix} U_p \\ U_f \\ Y_p \\ Y_f \end{bmatrix} = \begin{bmatrix} 0 & I_{m(N_p+N_f)} & 0 \\ 0 & & 0 \\ \mathcal{O}_p(A,C) & \mathcal{T}_p(B,D) & \mathcal{T}_p(E,F) \\ \mathcal{O}_f(A,C) & \mathcal{T}_f(B,D) & \mathcal{T}_f(E,F) \end{bmatrix} \begin{bmatrix} X \\ U \\ V \end{bmatrix}, \quad (7)$$

where $V := [v_0 \cdots v_{T-1}]$ is the block-Hankel matrix gathering the noise samples and $\mathcal{T}_p(E,F)$, $\mathcal{T}_f(E,F)$ are block-Toeplitz matrices similar to $\mathcal{T}_p(B,D)$, $\mathcal{T}_f(B,D)$ in (5).

The DeePC algorithm proposed in [23] attempts to compute an optimal-control action based on past input-output data coming from the unknown system (6), without previous identification. The control action is selected through an MPC-like optimization problem that allows one to enforce constraints ensuring safety and performance requirements. The previously collected data is directly used on-line in the MPC optimization problem; the predictor is therefore implicit and arises as the outcome of the optimization problem. In particular, at the generic iteration $k \geq N_p$, the DeePC computes optimal control actions by solving

$$\begin{aligned} \min_{g, u_i, y_i} \quad & \sum_{i=1}^{N_f} \|y_i - r_{k+i}\|_Q^2 + \|u_i\|_R^2 + \lambda_y \|Y_p g - y_p^{(k)}\|_2^2 + \lambda_g \|g\|_2^2 \\ \text{subject to} \quad & \begin{bmatrix} U_p \\ U_f \\ Y_f \end{bmatrix} g = \begin{bmatrix} u_p^{(k)} \\ u_i \\ y_i \end{bmatrix}, \quad u_i \in \mathcal{U}, \quad i = 1, \dots, N_f, \\ & \quad \quad \quad y_i \in \mathcal{Y}, \quad i = 1, \dots, N_f, \end{aligned} \quad (8)$$

where N_f is the prediction horizon, $r \in \mathbb{R}^{pN_f}$ is the output reference signal we want to track, $\mathcal{U} \subseteq \mathbb{R}^m$, $\mathcal{Y} \subseteq \mathbb{R}^p$ are the input and output constraint sets, respectively, $Q \in \mathbb{R}^{p \times p}$ is the output cost matrix (positive semidefinite), $R \in \mathbb{R}^{m \times m}$ is the control cost matrix (positive definite), $\lambda_y \geq 0$ and $\lambda_g \geq 0$ are the regularization parameters, and $u_p^{(k)}$, $y_p^{(k)}$ are the most recent N_p input-output measurements from (6), according to the notation (3). For simplicity we consider input and output box constraints of the form $\mathcal{U} = [u_{\min}, u_{\max}]$ and $\mathcal{Y} = [y_{\min}, y_{\max}]$, respectively. Note that the block-Hankel matrices U_p, U_f, Y_p, Y_f are fixed throughout the online iterations.

If we let g_k^* be the optimal solution of problem (8) at iteration k , DeePC provides an implicit predictor $y_f^{(k)} = Y_f g_k^*$ whose model is never derived explicitly but whose predictions are implicitly used for obtaining the optimal control actions $u_f^{(k)} = U_f g_k^*$. Problem (8) is solved in a receding-horizon fashion: of the N_f -long optimal control sequence $u_f^{(k)} = \text{col}(u_k^*, \dots, u_{k+N_f-1}^*)$ solution of (8), we apply a subsequence $u_k^*, \dots, u_{k+N_c}^*$, for some $N_c \leq N_f - 1$, to the system, update $u_p^{(k)}, y_p^{(k)}$ to the most recent input-output measurements and set k to $k + N_c + 1$.

Equation (8) is a relaxation of the corresponding problem enforcing the constraint $Y_p g = y_p^{(k)}$, which was proven to be equivalent to the classical receding-horizon MPC in the case of deterministic LTI systems [23]. Here, to cope with potential infeasibility due to the disturbances, the constraint is substituted with the least-squares regularization term $\|Y_p g - y_p^{(k)}\|_2^2$. The two-norm regularization on g has been introduced to

avoid overfitting and it has been shown to relate to distributional robustness of the method with respect to a range of uncertainties [24]. One can also see that, besides infeasibility issues, considering stochastic models makes both the implicit predictor model $y_f^{(k)} = Y_f g_k^*$ and the control actions $u_f^{(k)} = U_f g_k^*$ defined by the optimization (8) to be non-linear in the past-data $(y_p^{(k)}, u_p^{(k)})$ (Section IV).

We conclude this section with a note on the effect of acquiring more data. Classical LTI system identification methods use historical data off-line to compute estimates of the matrices in the system dynamics. The resulting matrices are then used on-line to generate state estimates and perform the prediction [1]. Other prediction methods estimate the linear relation from $u_p^{(k)}, u_f^{(k)}, y_p^{(k)}$ to $y_f^{(k)}$ off-line, and then use it online to carry on the predictions, [29] and references therein. This kind of prediction architectures naturally possesses a data-compression mechanism. For the classical system identification methods only storage of the estimates of the system matrices is required. This depends on input and output dimensions m and p and on the (guessed) state dimension, but not on the amount T of data nor on the prediction horizon N_f . Likewise, the storage requirements for linear predictors depends on input and output dimensions m and p , and the horizons N_p, N_f , but not on the amount T of data. If more data become available one can then use it to improve predictions with no need to store it or use it online. The DeePC approach, on the other hand, requires one to carry all the data at every iteration and the size of the optimization problem (8) increases both with the amount T of available data and with the parameters m, p, N_p , and N_f . Though additional data would in principle also be beneficial for a DeePC controller, it is not clear how one can incorporate it without increasing the on-line computational burden.

IV. METHOD DESCRIPTION

We introduce a possible way to effectively incorporate more data in the DeePC framework to reduce the effect of noise in the solution of problem (8). We propose to use additional data that may be available off-line to de-noise the block-Hankel matrices Y_p and Y_f and to equip the DeePC with an EKF based on data to handle the noise in the on-line measurements $y_p^{(k)}$.

A. Averaging Data Matrices

The output matrices Y_p and Y_f are constructed off-line from the trajectory $y_{0,T-1}$. Let $(u_{0,T-1}^{(1)}, y_{0,T-1}^{(1)}), \dots, (u_{0,T-1}^{(N)}, y_{0,T-1}^{(N)})$ be available T -long input-output trajectories and denote with $x_0^{(i)}, i = 1, \dots, N$, the corresponding initial conditions. Using those additional data on-line to improve the prediction will lead to an intractable optimization problem. However, we can make use of additional trajectories off-line to construct N different data matrices $\mathcal{H}^{(1)}, \dots, \mathcal{H}^{(N)}$ defined analogously to (7), and average those matrices to obtain

$$\bar{\mathcal{H}}_N := \frac{1}{N} \sum_{i=1}^N \mathcal{H}^{(i)} = \begin{bmatrix} 0 & I_{m(N_p+N_f)} & 0 \\ 0 & & 0 \\ \mathcal{O}_p(A,C) & \mathcal{T}_p(B,D) & \mathcal{T}_p(E,F) \\ \mathcal{O}_f(A,C) & \mathcal{T}_f(B,D) & \mathcal{T}_f(E,F) \end{bmatrix} \begin{bmatrix} \bar{X}_N \\ \bar{U}_N \\ \bar{V}_N \end{bmatrix} \quad (9)$$

where $\bar{V}_N := \frac{1}{N} \sum_{i=1}^N V^{(i)}$ is the average of the matrices $V^{(i)} = [v_0^{(i)} \cdots v_{T-1}^{(i)}]$, $i = 1, \dots, N$, constructed from the different noise realizations $v_{0,T-1}^{(i)}$ affecting the trajectories $y_{0,T-1}^{(i)}$, and similarly \bar{X}_N and \bar{U}_N contain the averaged state trajectories corresponding to the different initial conditions $x_0^{(i)}$ and input trajectories $u_{0,T-1}^{(i)}$, respectively. Since $(v_k)_{k \in \mathbb{Z}}$ is assumed to be white-noise, the Law of Large Numbers guarantees that $\bar{V}_N \rightarrow 0$ as $N \rightarrow \infty$ [30]. The averaging procedure makes use of additional data to mitigate the effect of noise in the data-driven model, hence reducing the risk of overfitting that would be present if the data was used directly in DeePC. Accordingly, the more matrices are involved in the average, the smaller the value of the regularization parameter λ_g that gives the optimal closed-loop cost (see Figure 3, Section V).

We recognize that such a method heavily exploits the underlying linear structure of the problem. However, this is meant to be a first attempt to exploit additional data for improving the performance of the algorithm when dealing with stochastic systems, without increasing the dimension of the optimization problem (11) to be solved online. Indeed, thanks to the linear structure and superposition, the sub-matrices \bar{X}_N and \bar{U}_N in the averaged data matrix (9) still represent valid system trajectories corresponding to the average of the initial states $x_0 = \frac{1}{N} \sum_i x_0^{(i)}$ and the average of the input sequences $u = \frac{1}{N} \sum_i u^{(i)}$, used in the experiments that generated the data for each $\mathcal{H}^{(i)}$. The only point that requires attention is ensuring that the resulting average input sequence respects the persistence of excitation requirement. The simplest way to ensure this is to assume that the same persistently exciting input sequence is applied in all cases.

B. An EKF for DeePC

The averaging procedure represents a simple way to make use of multiple T -long trajectories to denoise off-line the data matrix representing the dynamics in the constraints of (8). Numerical evidence suggests that additional denoising of the on-line data $y_p^{(k)}$ that enters the cost of (8) can lead to a further, significant improvement in performance (Section V). In a model based setting, such on-line denoising could be performed by a Kalman filter. In our data-driven setting, however, the classical Kalman filter cannot be applied as it requires a model of the system dynamics. Here we show how the non-linear one-step predictor implicit in (8) can be used to derive an EKF for integrating past measurements into an implicit “state estimate”; this can in turn be used to improve the asymptotic performance of the algorithm.

We introduce a fictitious state vector $z_k := \text{col}(y_{k-N_p+1}, \dots, y_k) \in \mathbb{R}^{pN_p}$ and keep track of a filtered state-estimate $\hat{z}_{k|k}$ and the corresponding error covariance $P_{k|k}$. At the generic iteration $k \geq N_p$, the EKF-DeePC algorithm uses the estimate $\hat{z}_{k|k}$ by solving the following optimization

problem

$$\begin{aligned} \min_{g \in \mathbb{R}^d} \quad & \sum_{i=1}^{N_f} \|(Y_f g)_i - r_{k+i}\|_Q^2 + \|(U_f g)_i\|_R^2 \\ & + \lambda_y \|Y_p g - \hat{z}_{k|k}\|_2^2 + \lambda_g \|g\|_2^2 \\ \text{subject to} \quad & U_p g = u_p^{(k)}, \\ & u_{\min} \leq (U_f g)_i \leq u_{\max}, \quad i = 1, \dots, N_f, \\ & y_{\min} \leq (Y_f g)_i \leq y_{\max}, \quad i = 1, \dots, N_f. \end{aligned} \quad (10)$$

The formulation (10) is obtained from (8) by substituting the constraints $u_i = U_f g$ and $y_i = Y_f g$ in the cost, leaving g as the only decision variable. The crucial difference with respect to (8) is that the past data $Y_p g$ used in the implicit predictor is now required to fit the state estimate $\hat{z}_{k|k}$ instead of the N_p most-recent measurements $y_p^{(k)}$. Rewriting (10) as a multi-parametric quadratic program (mp-QP) in the parameter $\theta_k := \text{col}(\hat{z}_{k|k}, u_p^{(k)})$ leads to an explicit relation between the optimum g_k^* and the parameter θ_k ; we exploit this relation to build the EKF. The mp-QP form of (10) is

$$\begin{aligned} \min_{g \in \mathbb{R}^d} \quad & \frac{1}{2} g^\top P g + (G\theta_k + q_k)^\top g + \theta_k^\top H \theta_k + \frac{1}{2} \mathbf{r}_k^\top \mathbf{Q} \mathbf{r}_k \\ \text{subject to} \quad & U_p g = B_{\text{eq}} \theta_k, \quad A_{\text{in}} g \leq b_{\text{in}}, \end{aligned} \quad (11)$$

where the inequality constraint defined by

$$\begin{aligned} b_{\text{in}} &:= \text{col}(\mathbb{1}_{mN_f} u_{\max}, -\mathbb{1}_{mN_f} u_{\min}, \mathbb{1}_{pN_f} y_{\max}, -\mathbb{1}_{pN_f} y_{\min}), \\ A_{\text{in}} &:= \text{col}(U_f, -U_f, Y_f, -Y_f), \end{aligned}$$

has to be understood *component-wise*, i.e. $(A_{\text{in}} g)_i \leq s_i$ for $i = 1, \dots, 2(m+p)N_f$. The cost is then defined by the reference signal $\mathbf{r}_k := \text{col}(r_{k+1}, \dots, r_{k+N_f})$ and the matrices

$$\begin{aligned} P &:= Y_f^\top \mathbf{Q} Y_f + U_f^\top \mathbf{R} U_f + \lambda_y Y_p^\top Y_p + \lambda_g I_d, \quad q_k := -Y_f^\top \mathbf{Q} \mathbf{r}_k \\ G &:= \begin{bmatrix} -\lambda_y Y_p^\top & 0 \\ 0 & I_{mN_p} \end{bmatrix}, \quad H := \begin{bmatrix} (\lambda_y/2) \cdot I_{pN_p} & 0 \\ 0 & 0 \end{bmatrix}. \\ B_{\text{eq}} &:= \begin{bmatrix} 0 & I_{mN_p} \end{bmatrix} \end{aligned}$$

Here $\mathbb{1}_N := \text{col}(1, 1, \dots, 1) \in \mathbb{R}^N$, $\mathbf{Q} = \text{diag}(Q, \dots, Q)$ and $\mathbf{R} = \text{diag}(R, \dots, R)$. Assuming $P = P^\top > 0$ and the KKT-matrix for problem (11) to be positive semi-definite (which is always the case for an MPC problem with input weighting matrix $R > 0$ [31]), the optimizer g_k^* is a piecewise affine function of the parameters, and can be written as $g_k^* = \tilde{\mathcal{A}}_k \hat{z}_{k|k} + \tilde{\mathcal{B}}_k u_p^{(k)} + \tilde{h}_k$. In particular, the equality constraints in (11) and the noise in the model are responsible for the affine structure while the inequality constraints implies this affine relation to hold just locally, i.e. in a neighborhood of the parameter θ_k , known as the *critical region*. The coefficients $\tilde{\mathcal{A}}_k$, $\tilde{\mathcal{B}}_k$ and \tilde{h}_k coming from the KKT conditions for problem (11), are therefore region-dependent themselves [31]: we consider the affine expansion of g_k^* pertaining to a specific value for the parameter θ_k , fixed by the previous iteration.

The fact that, under suitable assumptions, the piecewise affine relation between the optimizer g_k^* of (11) and the parameter θ_k can be derived from the KKT conditions for problem (11), might suggest that the predictor implicit in (10) could be made explicit and be constructed off-line. Doing this would require one to construct all the regions on which

the affine expansion is defined, for all the possible values of the parameters. As the number of these regions scales exponentially with the QP size (parameters plus constraints), hence with the amount T of available data and the horizons N_p and N_f , this approach is likely to be computationally intractable. This is the main reason for keeping the predictor implicit.

Let $\mathcal{M} := \text{col}(Y_p(2 : pN_p, :), Y_f(1, :))$ be the prediction map. We can exploit the piece-wise affine form of the optimizer g_k^* to incorporate the implicit predictor provided by (11) in an EKF-like architecture. To fix ideas, suppose we start running the algorithm at $k = N_p$. From an initial guess of the mean $\hat{z}_{N_p|N_p} = \mathbb{E}[z_{N_p}]$ and the covariance matrix $P_{N_p|N_p} = \mathbb{E}[(z_{N_p} - \hat{z}_{N_p|N_p})(z_{N_p} - \hat{z}_{N_p|N_p})^\top]$, we then compute recursively (at every point in time) the standard Kalman filter update steps in Figure 1. We note that, because of its dependence on g_k^* , the implicit predictor is piecewise affine (the matrices $\mathcal{A}_k, \mathcal{B}_k, h_k$ at the current iteration, depend on the critical region) making this a data-driven analogue to an *Extended* Kalman Filter.

Prediction	Update
$P_{k+1 k} = \mathcal{A}_k P_{k k} \mathcal{A}_k^\top + \Omega_k$	$P_{k+1 k+1} = (I - K_{k+1} \mathcal{C}) P_{k+1 k}$
$\hat{z}_{k+1 k} = \mathcal{M} g_k^* = \mathcal{A}_k \hat{z}_{k k} + \mathcal{B}_k u_p^{(k)} + h_k$	$\hat{z}_{k+1 k+1} = \hat{z}_{k+1 k} + K_{k+1} (y_{k+1} - \mathcal{C} \hat{z}_{k+1 k})$
$\hat{y}_{k+1} = \mathcal{C} \hat{z}_{k+1 k}$	

Fig. 1: EKF equations. Here, $\mathcal{C} := [0 \dots 0 I_p]$ by construction, $\mathcal{A}_k := \mathcal{M} \tilde{\mathcal{A}}_k$, $\mathcal{B}_k := \mathcal{M} \tilde{\mathcal{B}}_k$, $h_k := \mathcal{M} \tilde{h}_k$ and $S_{k+1} = \mathcal{C} P_{k+1|k} \mathcal{C}^\top + \mathcal{R}_k$, $K_{k+1} = P_{k+1|k} \mathcal{C}^\top S_{k+1}^{-1}$, are the variance of the innovation process and the filter gain, respectively.

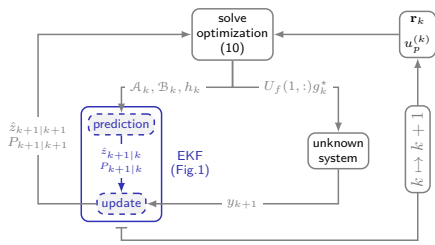


Fig. 2: EKF-DeePC Algorithm.

The matrices $\Omega_k \in \mathbb{R}^{pN_p \times pN_p}$ and $\mathcal{R}_k \in \mathbb{R}^{p \times p}$ are the variances of the process and measurement noise implicitly affecting the state and output dynamics $z_{k+1} = \mathcal{A}_k z_k + \mathcal{B}_k u_p^{(k)} + h_k$ and $y_k = \mathcal{C} z_k$, respectively. In practice, to implement the filter in Figure 1 one needs to address the same issues as for any Kalman filter, i.e. choosing the initial conditions $\hat{z}_{N_p|N_p}$ and $P_{N_p|N_p}$, and obtaining an estimate (or guess) of the variances Ω_k and \mathcal{R}_k . For a model-based Kalman filter, one can use the residuals of the system identification to obtain estimates for these quantities; we speculate that something similar would be possible using the residuals between the average and individual data matrices in Section IV-A, though we do not pursue this direction here, due to space limitations. The steps outlined above are summarized in Figure 2.

V. NUMERICAL VALIDATION

We present numerical experiments assessing the effects of the proposed solution in tackling the noise acting in the optimization (8). Consider the stochastic system

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + w_k \\ y_k = Cx_k + v_k \end{cases}, \quad A = \begin{bmatrix} 0.8 & 1 \\ 0 & 0.8 \end{bmatrix}, \quad B = [0 \ 1]^\top, \\ C = [1 \ 1]. \quad (12)$$

where $(w_k)_{k \in \mathbb{Z}}, (v_k)_{k \in \mathbb{Z}}$ are independent zero-mean Gaussian white noises with covariance $\Sigma_w := EE^\top$ and $\Sigma_v := FF^\top$ respectively; below we consider $\Sigma_w = \sigma_w^2 I_n$ and $\Sigma_v = \sigma_v^2 I_p$ and report results for different values of σ_w^2 and σ_v^2 . Model (12) is reachable and observable. The performance metric we will consider throughout is the closed-loop cost $J(u, y) = \sum_{k=1}^{N_{\text{sim}}} \|y_k - r_k\|_Q^2 + \|u_k\|_R^2$, where $r_k = 5 \sin(0.3k)$ is the reference signal. To isolate the effects of averaging and the EKF, for each numerical experiment we tuned the regularization parameters λ_y and λ_g through exhaustive search to minimize the closed-loop cost for the standard DeePC, the averaged DeePC and the averaged DeePC with EKF. All numerical evidence comes from 100 repetitions for different data-sets. Unless otherwise stated, we set $T = 100$, $N_{\text{sim}} = 100$, $N_p = 3$, $N_f = 5$, $Q = I_p$, $R = I_m$, $\sigma_w^2 = \sigma_v^2 = 0.5$ and $N = 40$ data matrices in the average (9).

First we show the effect of the averaging introduced in Section IV-A. To isolate the averaging effect, we consider noisy Y_p, Y_f but $y_p^{(k)}$ generated from the model with neither process noise, nor measurement noise. The result, shown in Figure 3, is as expected: the cost decreases towards the ideal cost of MPC with the same horizon but perfect model and full state measurement while on the right, the optimal (numerically found) value of the regularization parameter λ_g decreases to zero, as we expect from standard results in optimization. Fig-

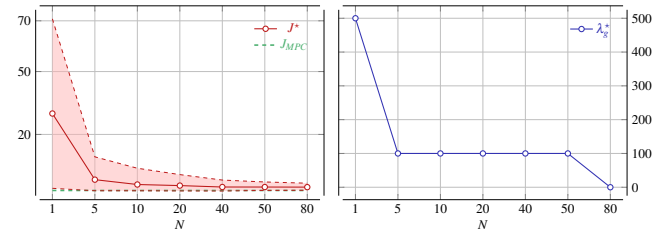


Fig. 3: Closed-loop cost (left) and optimal λ_g (right) when averaging N data matrices. The dashed-green line shows the MPC cost with the same horizon, but perfect model and full state measurement.

ure 4 shows a comparison between the closed-loop costs of the standard DeePC and the averaged DeePC with EKF algorithms when varying the noise variances σ_w^2 and σ_v^2 . Figure 4 shows that introducing averaging and the EKF substantially improves the performance of the DeePC algorithm. Further experiments (data not shown) with averaged DeePC without EKF and EKF without averaging confirmed that DeePC with a combination of averaging and EKF substantially outperformed all other combinations; indeed this was the case across a broad range of values of the regularizes λ_y and λ_g , suggesting that averaged DeePC with EKF is easier to tune. Figure 5 shows how the

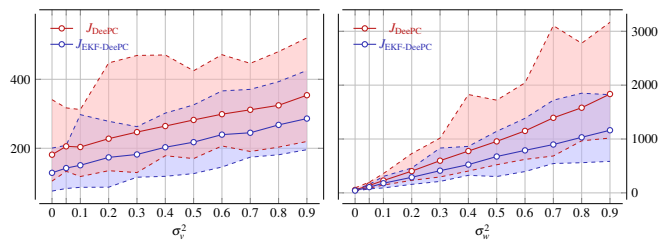


Fig. 4: Closed-loop cost of standard DeePC and averaged DeePC with EKF for different σ_w^2 while keeping $\sigma_v^2 = 0.1$ (left), and different σ_v^2 while keeping $\sigma_w^2 = 0.2$ (right).

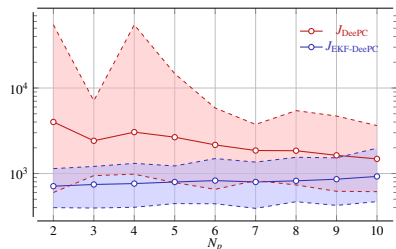


Fig. 5: Closed-loop costs while varying the past horizon N_p .

closed-loop cost of the two algorithms behaves with respect to the past horizon N_p . The improvement in the performance is due to the introduction of the averaging plus the EKF, highlighting the difference between fixed horizon and recursive filtering.

VI. CONCLUSION AND OUTLOOK

We presented an extension of the data-enabled predictive control (DeePC) algorithm introduced in [23] to tackle control problems on unknown stochastic LTI systems, by making use of additional data without increasing the dimension of the on-line optimization problem. The procedure features an on-line data-driven EKF that filters out the noise in the measurements, and an (off-line) averaging of multiple data matrices to get a cleaner data-driven model. The performance of the proposed method has been validated experimentally on an LTI stochastic system showing improvements with respect to standard DeePC. Future work includes testing the proposed method on a real-case scenario and comparing it to algorithms such as [3]–[5].

ACKNOWLEDGMENTS

The authors are grateful to Colin Jones and Francesco Micheli for detailed discussions and the anonymous reviewers for insightful suggestions.

REFERENCES

- [1] L. Ljung, *System Identification: Theory for the User*. Prentice Hall information and system sciences series, Prentice Hall PTR, 1999.
- [2] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*. USA: Dover Publications, Inc., 2009.
- [3] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin, "Iterative feedback tuning: theory and applications," *IEEE Control Systems Magazine*, vol. 18, no. 4, pp. 26–41, 1998.
- [4] A. Karimi, L. Mišković, and D. Bonvin, "Convergence analysis of an iterative correlation-based controller tuning method," *IFAC Proceedings Volumes*, vol. 35, no. 1, pp. 413–418, 2002. 15th IFAC World Congress.
- [5] M. Campi, A. Lecchini, and S. Savaresi, "Virtual reference feedback tuning: a direct method for the design of feedback controllers," *Automatica*, vol. 38, no. 8, pp. 1337–1346, 2002.
- [6] D. P. Bertsekas, *Reinforcement learning and optimal control*. Athena Scientific Belmont, MA, 2019.
- [7] M. G. Lagoudakis and R. Parr, "Least-squares policy iteration," *Journal of machine learning research*, vol. 4, no. Dec, pp. 1107–1149, 2003.
- [8] W. B. Powell, *Approximate Dynamic Programming: Solving the curses of dimensionality*, vol. 703. John Wiley & Sons, 2007.
- [9] B. Recht, "A tour of reinforcement learning: The view from continuous control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, no. 1, pp. 253–279, 2019.
- [10] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *Int. J. Rob. Res.*, vol. 32, p. 1238–1274, Sept. 2013.
- [11] J. C. Willems, "Paradigms and puzzles in the theory of dynamical systems," *IEEE Transactions on Automatic Control*, vol. 36, pp. 259–294, March 1991.
- [12] J. C. Willems and J. W. Polderman, *Introduction to mathematical systems theory: a behavioral approach*, vol. 26. Springer Science & Business Media, 2013.
- [13] J. C. Willems, P. Rapisarda, I. Markovskiy, and B. L. D. Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.
- [14] I. Markovskiy and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.
- [15] I. Markovskiy and P. Rapisarda, "On the linear quadratic data-driven control," in *2007 European Control Conference (ECC)*, pp. 5313–5318, IEEE, 2007.
- [16] H. J. van Waarde, C. D. Persis, M. K. Camlibel, and P. Tesi, "Willems' fundamental lemma for state-space systems and its extension to multiple datasets," 2020.
- [17] A. Romer, J. Berberich, J. Köhler, and F. Allgöwer, "One-shot verification of dissipativity properties from input–output data," *IEEE Control Systems Letters*, vol. 3, pp. 709–714, July 2019.
- [18] J. Berberich, A. Romer, C. W. Scherer, and F. Allgöwer, "Robust data-driven state-feedback design," *arXiv preprint arXiv:1909.04314*, 2019.
- [19] J. Berberich and F. Allgöwer, "A trajectory-based framework for data-driven system analysis and control," *arXiv preprint arXiv:1903.10723*, 2019.
- [20] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Transactions on Automatic Control*, vol. 65, pp. 909–924, March 2020.
- [21] H. J. van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel, "Data informativity: a new perspective on data-driven analysis and control," 2019.
- [22] W. Favoreel, B. D. Moor, and M. Gevers, "SpC: Subspace predictive control," *IFAC Proceedings Volumes*, vol. 32, no. 2, pp. 4004–4009, 1999. 14th IFAC World Congress 1999, Beijing, Chia, 5-9 July.
- [23] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the deepc," in *2019 18th European Control Conference (ECC)*, pp. 307–312, IEEE, 2019.
- [24] J. Coulson, J. Lygeros, and F. Dörfler, "Regularized and distributionally robust data-enabled predictive control," *arXiv preprint arXiv:1903.06804*, 2019.
- [25] L. Huang, J. Coulson, J. Lygeros, and F. Dorfler, "Data-enabled predictive control for grid-connected power converters," *arXiv preprint arXiv:1903.07339*, 2019.
- [26] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *arXiv preprint arXiv:1906.04679*, 2019.
- [27] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, no. 3, pp. 393–438, 2005. Data-Based Modelling and System Identification.
- [28] B. Bulut, M. R. Katebi, and M. A. Johnson, "Industrial application of model based predictive control as a supervisory system," in *Proceedings of the 2000 American Control Conference*, vol. 6, pp. 3763–3767 vol.6, June 2000.
- [29] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory—Implementation—Applications*. Springer Science & Business Media, 2012.
- [30] P. Billingsley, *Probability and Measure*. Wiley Series in Probability and Statistics, Wiley, 1995.
- [31] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3–20, 2002.