

# A deterministic view on explicit data-driven (M)PC

Manuel Klädtke, Dieter Teichrib, Nils Schlüter, and Moritz Schulze Darup

**Abstract.** We show that the explicit realization of data-driven predictive control (DPC) for linear deterministic systems is more tractable than previously thought. To this end, we compare the optimal control problems (OCP) corresponding to deterministic DPC and classical model predictive control (MPC), specify its close relation, and systematically eliminate ambiguity inherent in DPC. As a central result, we find that the explicit solutions to these types of DPC and MPC are of exactly the same complexity. We illustrate our results with two numerical examples highlighting features of our approach.

## I. INTRODUCTION

Data-driven predictive control (DPC), where the prediction of the systems' behavior is carried out based on collected input-output data instead of a model, is becoming more and more popular (see, e.g., [1–4]). Remarkably, assuming perfect data and linear dynamics, Willems' fundamental lemma [5] and variants of it (as, e.g., [6] and [7]) allow establishing the equivalence of the data-driven and model-based approach with respect to the resulting control actions.

However, while strongly related, the two approaches lead to different optimal control problems (OCP). In fact, DPC usually results in an OCP with significantly more decision variables than model-based predictive control (MPC). As a consequence, explicit solutions of the data-driven OCP seem “unattractive” at first sight (especially for noisy setups [8, Sect. IV.B]), even for applications where explicit MPC [9] is tractable. In fact, more decision variables typically result in significantly more complex explicit solutions (in terms of the number of regions etc.). Yet, we show in this paper that the perceived imbalance between MPC and DPC can be completely resolved for the special case of linear deterministic systems. More precisely, we reveal that the larger number of decision variables only results in ambiguous but not more complex solutions in this case. Further, we present a simple method to systematically eliminate this ambiguity. As a central result, we obtain an explicit DPC solution of exactly the same complexity as explicit MPC.

Before detailing our approach, we briefly discuss related works from the literature. First of all, it is already well-known that the optimal input sequences resulting from deterministic DPC and MPC are identical given equivalent initial conditions

[2, Cor. 5.1]. Yet, it is also known that the original OCPs related to MPC are strictly convex while those for DPC are only convex. Thus, optimizers in DPC are typically non-unique, which significantly complicates an explicit solution. Clearly, strict convexity can be enforced through additional regularization [8, 10] (which is also helpful for noisy setups). However, this either destroys the structure we are about to identify or it renders its derivation more difficult. Alternatively, one can consider explicit DPC for fully measurable states. For this simpler case, a result similar to ours has recently been obtained in [11]. Finally, especially since we are dealing with the deterministic case and linear systems, removing ambiguity from the OCP shows many similarities to subspace identification (SID, [12]) and subspace predictive control [13]. In fact, using the data matrices inherent in DPC, one could also identify a state space model and the corresponding MPC formulation would yield another equivalence. However, we provide a simple and direct approach, which can be interpreted as a tailored subspace analysis for DPC.

The remaining paper is organized as follows. In Section II, we summarize classical MPC and fundamentals of DPC. The analysis of explicit solutions of the corresponding OCPs and the central identification of a closer relation between them are carried out in Section III. Finally, we illustrate our findings with two numerical examples in Section IV and we discuss promising directions for future research in Section V.

## II. FUNDAMENTALS OF MPC AND DPC

### A. Classical MPC

We briefly summarize classical MPC in a form that is compatible with the data-driven realization in Section II.B. To this end, we assume that a linear prediction model

$$x(k+1) = Ax(k) + Bu(k) \quad (1a)$$

$$y(k) = Cx(k) + Du(k) \quad (1b)$$

is known. We further assume that input and output constraints are given in terms of convex polyhedral sets

$$\mathcal{U} := \{u \in \mathbb{R}^m \mid M_u u \leq v_u\}, \quad \mathcal{Y} := \{y \in \mathbb{R}^p \mid M_y y \leq v_y\}, \quad (2)$$

which are specified by the matrices  $M_{u/y}$  and vectors  $v_{u/y}$ , respectively. Then, classical MPC (without terminal cost and constraints) can be realized by solving

M. Klädtke, D. Teichrib, N. Schlüter, and M. Schulze Darup are with the Control and Cyberphysical Systems Group, Faculty of Mechanical Engineering, TU Dortmund University, Germany. E-mails: {manuel.klaedtke, dieter.teichrib, nils.schluter, moritz.schulzedarup}@tu-dortmund.de.

\*This paper is a **preprint** of a contribution to the 2022 IEEE 61st Conference on Decision and Control (CDC). The DOI of the original paper is 10.1109/CDC51059.2022.9993384.

$$\begin{aligned}
& \min_{u(k), x(k), y(k)} \sum_{k=0}^{N_f-1} \|y(k)\|_Q^2 + \|u(k)\|_R^2 \quad (3) \\
& \text{s.t.} \quad x(0) = x_0, \\
& \quad x(k+1) = Ax(k) + Bu(k), \quad \forall k \in \{0, \dots, N_f-2\}, \\
& \quad y(k) = Cx(k) + Du(k), \quad \forall k \in \{0, \dots, N_f-1\}, \\
& \quad (y(k), u(k)) \in \mathcal{Y} \times \mathcal{U}, \quad \forall k \in \{0, \dots, N_f-1\}
\end{aligned}$$

in every time-step for the current state  $x_0 \in \mathbb{R}^n$ , where  $Q \in \mathbb{R}^{p \times p}$  and  $R \in \mathbb{R}^{m \times m}$  denote weighting matrices and where  $N_f \in \mathbb{N}$  is the prediction horizon. Now, the OCP (3) is typically condensed into a quadratic program (QP) such that only the inputs remain as decision variables. To this end, one first introduces the sequences

$$u_f := \begin{pmatrix} u(0) \\ \vdots \\ u(N_f-1) \end{pmatrix} \quad \text{and} \quad y_f := \begin{pmatrix} y(0) \\ \vdots \\ y(N_f-1) \end{pmatrix}, \quad (4)$$

and the augmented weighting matrices  $\mathcal{Q} := \text{diag}(Q, \dots, Q)$  and  $\mathcal{R} := \text{diag}(R, \dots, R)$  to rewrite the cost function as

$$\sum_{k=0}^{N_f-1} \|y(k)\|_Q^2 + \|u(k)\|_R^2 = \|y_f\|_{\mathcal{Q}}^2 + \|u_f\|_{\mathcal{R}}^2. \quad (5)$$

We further define the matrices

$$\mathcal{O}_N := \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{N-1} \end{pmatrix} \quad \text{and} \quad \mathcal{T}_N := \begin{pmatrix} D & & & 0 \\ CB & \ddots & & \\ \vdots & \ddots & \ddots & \\ CA^{N-2}B & \dots & CB & D \end{pmatrix},$$

which we will consider for different  $N$  during this note. For  $N = N_f$ , we then obtain the relation

$$y_f = \mathcal{O}_{N_f} x_0 + \mathcal{T}_{N_f} u_f. \quad (6)$$

Finally, substituting (6) into (5) and introducing the augmented matrices  $\mathcal{M}_{u/y} := \text{diag}(M_{u/y}, \dots, M_{u/y})$  leads to

$$\begin{aligned}
u_f^*(x_0) &:= \arg \min_{u_f} \frac{1}{2} u_f^\top H u_f + x_0^\top F^\top u_f \quad (7) \\
&\text{s.t.} \quad G u_f \leq E x_0 + d
\end{aligned}$$

with the parameter  $x_0$  as well as

$$\begin{aligned}
H &:= 2\mathcal{T}_{N_f}^\top \mathcal{Q} \mathcal{T}_{N_f} + 2\mathcal{R}, \quad F := 2\mathcal{T}_{N_f}^\top \mathcal{Q} \mathcal{O}_{N_f}, \\
G &:= \begin{pmatrix} \mathcal{M}_u \\ \mathcal{M}_y \mathcal{T}_{N_f} \end{pmatrix} \quad E := \begin{pmatrix} 0 \\ -\mathcal{M}_y \mathcal{O}_{N_f} \end{pmatrix}, \quad (8) \\
d &:= (v_u^\top \quad \dots \quad v_u^\top \quad v_y^\top \quad \dots \quad v_y^\top)^\top.
\end{aligned}$$

Remarkably,  $H$  is positive definite, i.e., (7) is strictly convex, under the assumption that  $Q$  is positive semi-definite and that  $R$  is positive definite.

## B. DPC using input-output sequences

In contrast to MPC, DPC considers input-output data instead of a model as in (1). More precisely, DPC builds (in its simplest form) on two sequences  $u_d$  and  $y_d$  as in (4) but of length  $N_d \in \mathbb{N}$  that reflect prerecorded system inputs and outputs. We note, at this point, that with slight abuse of notation, we denote both the elements of  $u_f$  and  $u_d$  with  $u(k)$  (and analogously elements in  $y_f$  and  $y_d$  with  $y(k)$ ). However, the specific relationship will always be clear from the context. Now, in order to realize DPC by means of  $u_d$  and  $y_d$ , the sequences have to carry enough information about the systems' dynamics. This holds, for instance, if  $y_d$  is consistent with a persistently exciting and sufficiently long input sequence  $u_d$ . More specifically, for deterministic DPC as considered here, consistency means that there exists a model (1) with initial state  $x_0 \in \mathbb{R}^n$  such that

$$y_d = \mathcal{O}_{N_d} x_0 + \mathcal{T}_{N_d} u_d. \quad (9)$$

Further, according to [5],  $u_d$  is persistently exciting of order  $N_e \in \mathbb{N}$  if the Hankel matrix

$$\mathcal{H}_{N_e}(u_d) := \begin{pmatrix} u(0) & u(1) & \dots & u(N_d - N_e) \\ u(1) & u(2) & \dots & u(N_d - N_e + 1) \\ \vdots & & \ddots & \vdots \\ u(N_e - 1) & u(N_e) & \dots & u(N_d - 1) \end{pmatrix}$$

has full row rank, i.e.,  $\text{rank}(\mathcal{H}_{N_e}(u_d)) = mN_e$ . This requires  $\mathcal{H}_{N_e}(u_d)$  to have as least as many columns as rows, i.e.,

$$N_d - N_e + 1 \geq mN_e \iff N_d \geq (m+1)N_e - 1. \quad (10)$$

Finally, Willems' fundamental lemma [5] allows associating the given sequences  $u_d$  and  $y_d$  with other input-output sequences of the same system. In fact, under the assumption that the underlying system is linear, controllable, and  $u_d$  is persistently exciting of order  $N_e := N_c + n$ , candidate sequences  $(u_c, y_c)$  of length  $N_e \in \mathbb{N}$  belong to the same system as  $(u_d, y_d)$  if and only if

$$\begin{pmatrix} u_c \\ y_c \end{pmatrix} \in \text{im} \begin{pmatrix} \mathcal{H}_{N_e}(u_d) \\ \mathcal{H}_{N_e}(y_d) \end{pmatrix}.$$

At this point, we briefly note that recent extensions of the fundamental lemma in [6] and [7] allow alleviating some of the restrictions above. Now, in order to utilize the previous results for DPC, we proceed similarly to [2]. We choose an integer  $N_p$  equal to (or larger than) the observability index, i.e., such that the corresponding matrix  $\mathcal{O}_{N_p}$  has full column rank (which obviously requires observability). Next, we assume that  $u_d$  is persistently exciting of order

$$N_e := N_p + N_f + n. \quad (11)$$

According to the fundamental lemma, we then find that the concatenated sequences  $(u_p^\top \ u_f^\top)^\top$  and  $(y_p^\top \ y_f^\top)^\top$  with

$$u_p := \begin{pmatrix} u(-N_p) \\ \vdots \\ u(-1) \end{pmatrix} \quad \text{and} \quad y_p := \begin{pmatrix} y(-N_p) \\ \vdots \\ y(-1) \end{pmatrix}$$

and with  $(u_f, y_f)$  as in (4), belong to the same system as  $(u_d, y_d)$  if and only if

$$\begin{pmatrix} u_p \\ u_f \\ y_p \\ y_f \end{pmatrix} = \begin{pmatrix} \mathcal{H}_{N_p+N_f}(u_d) \\ \mathcal{H}_{N_p+N_f}(y_d) \end{pmatrix} a. \quad (12)$$

for some  $a \in \mathbb{R}^l$  with  $l := N_d - N_f - N_p + 1$ . Based on reordering and partitioning, (12) can be rewritten as

$$\xi := \begin{pmatrix} u_p \\ y_p \end{pmatrix} = W_p a, \quad u_f = U_f a, \quad \text{and} \quad y_f = Y_f a \quad (13)$$

with the matrices  $W_p$ ,  $U_f$ , and  $Y_f$  representing blocks of the concatenated Hankel matrices. We are now ready to formulate the OCP associated with DPC. In fact, the combination of (5) and (13) allow expressing the costs

$$\|y_f\|_Q^2 + \|u_f\|_R^2 = \|a\|_{Y_f^\top Q Y_f + U_f^\top R U_f}^2$$

as a function of  $a$ . Taking into account the constraints  $\mathcal{M}_u u_f \leq \mathcal{V}_u$  and  $\mathcal{M}_y u_f \leq \mathcal{V}_y$  and the remaining condition  $\xi = W_p a$  then leads to the QP

$$a^*(\xi) := \arg \min_a \frac{1}{2} a^\top \tilde{H} a \quad \text{s.t.} \quad \tilde{G} a \leq d, \quad W_p a = \xi \quad (14)$$

with the parameter  $\xi$ , the vector  $d$  as in (7), and

$$\tilde{H} := 2Y_f^\top Q Y_f + 2U_f^\top R U_f, \quad \tilde{G} := \begin{pmatrix} \mathcal{M}_u U_f \\ \mathcal{M}_y Y_f \end{pmatrix}.$$

Remarkably, the role of the initial state  $x_0$  in (7) is replaced by  $\xi$ , i.e., the  $N_p$  previous inputs and outputs, in (14). Furthermore,  $a^*(\xi)$  only reflects an intermediate result that is used to compute optimal inputs via  $u_f^*(\xi) := U_f a^*(\xi)$ .

### III. FROM EXPLICIT MPC TO EXPLICIT DPC

The QP (7) or (14) is typically solved for the current state  $x_0$  or the most recent sequences  $\xi$ , respectively, to obtain the optimal input for the current time-step. Subsequently, the procedure is repeated at the next sampling instance. Alternatively, in order to avoid numerical optimization during runtime, (7) can also be solved explicitly using parametric optimization. As a result, we then find the continuous and piecewise affine (PWA) solution

$$u_f^*(x_0) = \begin{cases} L_1 x_0 + c_1 & \text{if } x_0 \in \mathcal{X}_1, \\ \vdots & \vdots \\ L_s x_0 + c_s & \text{if } x_0 \in \mathcal{X}_s, \end{cases} \quad (15)$$

which is defined on a polyhedral partition  $\{\mathcal{X}_i\}_{i=1}^s$  of the state space [9]. Computing this solution offline and evaluating it online is referred to as explicit MPC. While conceptually attractive, explicit MPC can usually only be applied for moderate “sizes” of the underlying QP since it is well-known that the number of regions  $s \in \mathbb{N}$  typically grows exponentially with the number of decisions variables and constraints. As a consequence, solving

(14) parametrically seems unattractive at first sight, since especially the number of decision variables is significantly larger than in (7). In fact, while  $u_f$  is of dimension  $mN_f$ , the dimension  $l$  of  $a$  is lower-bounded by

$$\begin{aligned} l &\geq (m+1)(N_p + N_f + n) - N_f - N_p \\ &= mN_p + mN_f + (m+1)n \end{aligned} \quad (16)$$

according to (10) and (11). Now, while the difference of at least  $mN_p + (m+1)n$  decisions variables is significant especially for  $m > 1$ , we claim that this increase does not result in a more complex explicit solution for the special case of deterministic DPC. In fact, we show that the increase in decision variables only leads to ambiguous solutions and that this ambiguity can be removed by systematically eliminating variables using tools inspired from SID. Remarkably, simultaneously to our work, [14] proposed a conceptually similar way of eliminating decision variables for non-deterministic systems. The focus in [14] is, however, not on explicit DPC.

#### A. Eliminating equality constraints for DPC

Following this claim, we initially eliminate the equality constraints in (14). To this end, we assume that a generalized inverse  $W_p^+$  of  $W_p$  (satisfying the Penrose conditions) and a matrix  $V_p$  characterizing the null-space of  $W_p$  (i.e.,  $\text{im}(V_p) = \ker(W_p)$ ) are known. Then, we can substitute  $a$  in (14) with

$$a := W_p^+ \xi + V_p \alpha, \quad (17)$$

where  $\alpha$  is of dimension

$$\nu := \text{nullity}(W_p) = l - \text{rank}(W_p). \quad (18)$$

Clearly, the equality constraints in (14) are satisfied for every  $\alpha \in \mathbb{R}^\nu$ . Hence, we obtain the transformed QP

$$\alpha^*(\xi) = \arg \min_\alpha \frac{1}{2} \alpha^\top \hat{H} \alpha + \xi^\top \hat{F}^\top \alpha \quad \text{s.t.} \quad \hat{G} \alpha \leq \hat{E} \xi + d \quad (19)$$

with  $\hat{H} := V_p^\top \tilde{H} V_p$ ,  $\hat{G} := \tilde{G} V_p$ ,  $\hat{F} := V_p^\top \tilde{H} W_p^+$ , and  $\hat{E} := -\tilde{G} W_p^+$ . While the elimination of equality constraints is a standard procedure often performed internally by QP solvers, it has a useful interpretation in the case of DPC. Unlike  $a$ , which parametrizes all possible system trajectories of lengths  $N_p + N_f$ , the new variable  $\alpha$  only parametrizes those trajectories that are consistent with  $\xi$ , i.e., the  $N_p$  most recent inputs and outputs. However, it is important to note that (14) is only feasible for  $\xi$  belonging to the system while (19) may also be feasible for other  $\xi$ . This observation will be relevant further below for Theorem 9.

Now, according to (18), the reduction of decision variables when replacing (14) with (19) is determined by  $\text{rank}(W_p)$ . Taking into account that  $W_p$  contains  $mN_p$  rows of the full rank matrix  $\mathcal{H}_{N_p+N_f}(u_d)$ , we immediately find  $\text{rank}(W_p) \geq mN_p$ . A closer investigation reveals the following specification.

**Lemma 1.** *Let  $u_d$  and  $y_d$  be as in Section II.B and consider the partitions (13) of the Hankel matrices in (12). Then,  $\text{rank}(W_p) = mN_p + n$ .*

*Proof.* It is easy to see that  $W_p$  can be written as

$$W_p = \begin{pmatrix} \mathcal{H}_{N_p}(\hat{u}_d) \\ \mathcal{H}_{N_p}(\hat{y}_d) \end{pmatrix},$$

where  $\hat{u}_d$  and  $\hat{y}_d$  refer to the sequences  $u_d$  respectively  $y_d$  shortened by the last  $N_f$  elements. It is further straightforward to show that  $\hat{u}_d$  is persistently exciting of order  $N_p + n$  (i.e., the order of  $u_d$  likewise reduced by  $N_f$ ). As summarized in [7, Sect. I.], Willems' fundamental lemma [5] then implies  $\text{rank}(W_p) = mN_p + n$ . ■

The combination of (16), (18), and Lemma 1 implies

$$\nu = l - \text{rank}(W_p) \geq m(N_f + n). \quad (20)$$

In other words, while the number of decision variables is significantly reduced from (14) to (19), we still find at least  $mn$  more decision variables in (19) than in (7). Fortunately, this deficit can be eliminated as follows.

### B. Eliminating solution candidates in irrelevant null-spaces

As noted in Section II.B, also when applying DPC, we are mainly interested in the optimal control sequence

$$u_f^*(\xi) = U_f a^*(\xi) = U_f W_p^+ \xi + U_f V_p \alpha^*(\xi) \quad (21)$$

(or even only in its first element). As apparent from (21), components of  $\alpha(\xi)$  in the null-space of  $U_f V_p$  will not affect the resulting sequence  $u_f(\xi)$ . As a consequence, it seems promising to parametrize  $\alpha$  by

$$\alpha := K_f \beta + V_f \beta_0, \quad (22)$$

where  $K_f \in \mathbb{R}^{\nu \times \mu}$  and  $V_f \in \mathbb{R}^{\nu \times \nu - \mu}$  with  $\mu := \text{rank}(U_f V_p)$  are such that  $\text{im}(V_f) = \ker(U_f V_p)$  and  $\text{rank}((K_f \ V_f)) = \nu$ . Clearly, the columns of  $K_f$  and  $V_f$  span the subspaces that are relevant and irrelevant for  $u_f$ , respectively. By construction, we thus obtain

$$U_f V_p V_f \beta_0 = 0 \quad \text{for every } \beta_0 \in \mathbb{R}^{\nu - \mu}. \quad (23)$$

Hence,  $\beta_0$  has no effect on the resulting input sequence  $u_f$ . Further, since  $\xi$  determines  $x_0$  in (6) and since  $y_f$  is then determined by  $u_f$ , also  $y_f$  should be independent of  $\beta_0$ . In order to verify this hypothesis, we initially note that  $\xi$  and the assumed observability allow reconstructing  $x(-N_p)$ . This state in combination with  $u_p$  determines  $x_0$ . The relation is formally captured by  $x_0 = \Gamma \xi$ , where

$$\Gamma := \begin{pmatrix} (A^{N_p-1} B \ \dots \ B) - A^{N_p} \mathcal{O}_{N_p}^+ \mathcal{T}_{N_p} & A^{N_p} \mathcal{O}_{N_p}^+ \end{pmatrix}$$

with  $\mathcal{O}_{N_p}^+ := (\mathcal{O}_{N_p}^\top \mathcal{O}_{N_p})^{-1} \mathcal{O}_{N_p}^\top$ . Using this relation in (6) leads to

$$y_f = \mathcal{O}_{N_f} \Gamma \xi + \mathcal{T}_{N_f} u_f. \quad (24)$$

This equation provides the basis for a useful relation between  $W_p$ ,  $U_f$ , and  $Y_f$ . In fact, noting that the columns of these matrices can be interpreted as uniformly shifted sequences  $\xi$ ,  $u_f$ , and  $y_f$ , respectively, one finds

$$Y_f = \mathcal{O}_{N_f} \Gamma W_p + \mathcal{T}_{N_f} U_f \quad (25)$$

as also pointed out in [12, p. 41]. Based on this relation, we can easily derive the analogue to (23) for output sequences.

**Lemma 2.** Let  $Y_f$ ,  $V_p$  and  $V_f$  be defined as in (13), (17) and (22), respectively. Then,

$$Y_f V_p V_f \beta_0 = 0 \quad \text{for every } \beta_0 \in \mathbb{R}^{\nu - \mu}. \quad (26)$$

*Proof.* To prove the claim, we multiply (25) with  $V_p$  as in (17) from the right and obtain

$$Y_f V_p = \mathcal{T}_{N_f} U_f V_p \quad (27)$$

due to  $W_p V_p = 0$ . Substituting (27) in (26) and taking (23) into account completes the proof. ■

The relations (23) and (26) formally show that  $\beta_0$  neither affects input nor output sequences parametrized by  $\alpha$  as in (22). As a consequence, (19) can be replaced by a QP, where only  $\beta \in \mathbb{R}^\mu$  appears as a decision variable. This central observation is formalized in the following theorem.

**Theorem 3.** Let  $U_f$ ,  $W_p^+$ ,  $V_p$  and  $K_f$  be defined as in (13), (17) and (22), respectively. Then, the relation

$$u_f^*(\xi) = U_f W_p^+ \xi + U_f V_p K_f \beta^*(\xi) \quad (28)$$

holds, where

$$\beta^*(\xi) := \arg \min_{\beta} \frac{1}{2} \beta^\top \tilde{H} \beta + \xi^\top \tilde{F} \beta \quad \text{s.t.} \quad \tilde{G} \beta \leq \hat{E} \xi + d \quad (29)$$

with  $\tilde{H} := K_f^\top \hat{H} K_f$ ,  $\tilde{F} := K_f^\top \hat{F}$ , and  $\tilde{G} := \hat{G} K_f$ .

*Proof.* We initially show that

$$\hat{H} V_f \beta_0 = 0, \quad \hat{F}^\top V_f \beta_0 = 0, \quad \text{and} \quad \hat{G} V_f \beta_0 = 0 \quad (30)$$

for every  $\beta_0 \in \mathbb{R}^{\nu - \mu}$ . To see this, we first substitute the expressions for  $\hat{H}$ ,  $\hat{F}$ , as well as  $\hat{G}$  and then insert  $\tilde{H}$ ,  $\tilde{F}$ , as well as  $\tilde{G}$ , respectively. Doing so, we obtain

$$\hat{H} V_f \beta_0 = V_p^\top \tilde{H} V_p V_f \beta_0 = 2V_p^\top (Y_f^\top \mathcal{Q} Y_f + U_f^\top \mathcal{R} U_f) V_p V_f \beta_0$$

for the first expression in (30). Clearly, this expression indeed evaluates to zero for every  $\beta_0 \in \mathbb{R}^{\nu - \mu}$  due to (23) and (26). Analogue observations result for the remaining expressions in (30). Now, the relations in (30) imply that the choice of  $\beta_0$  neither affects the cost function nor the constraints in (19) when  $\alpha$  is parametrized as in (22). Hence, when applying this parametrization to (19), we can omit the variable  $\beta_0$  (or set it to zero) and restrict our attention to the new decision variable  $\beta$ . Formally, this results in the QP (29). ■

Clearly, the number of decision variables in (29) equals  $\mu = \text{rank}(U_f V_p)$ . Since  $U_f V_p$  is of dimension  $mN_f \times \nu$  and since (20) applies, we immediately find

$$\mu = \text{rank}(U_f V_p) \leq \min\{mN_f, \nu\} = mN_f. \quad (31)$$

In other words, while (19) definitely contains more decision variables than (7) according to (20), (29) contains at most as many decision variables as (7) according to (31). At this point, it is important to note that (31) simply reflects the dimensions of  $U_f V_p$ . Recalling that (7) is a strictly convex QP and that (29) provides equivalent solutions according to (28), already excludes the case  $\mu < mN_f$ . In fact, we always have  $\mu = mN_f$  according to the following lemma.

**Lemma 4.** *Let  $U_f$  and  $V_p$  be defined as in (13) and (17), respectively. Then,  $\text{rank}(U_f V_p) = mN_f$ .*

*Proof.* To prove the claim, we consider (12) for the special case  $\xi = 0$ , i.e.,  $(u_p, y_p) = (0, 0)$ . As apparent from (24), any  $u_f \in \mathbb{R}^{mN_p}$  in combination with  $y_f := \mathcal{T}_{N_f} u_f$  leads to consistent sequences for this case. As a consequence, there exists an  $a \in \mathbb{R}^l$  such that

$$\begin{pmatrix} \xi \\ u_f \\ y_f \end{pmatrix} = \begin{pmatrix} 0 \\ u_f \\ \mathcal{T}_{N_f} u_f \end{pmatrix} = \begin{pmatrix} W_p \\ U_f \\ Y_f \end{pmatrix} a$$

for every  $u_f \in \mathbb{R}^{mN_p}$ . Since we have  $0 = W_p a$  by construction, every such  $a$  can be parametrized as  $a = V_p \alpha$  for a suitable  $\alpha \in \mathbb{R}^\nu$  according to (17). Now, since the choice of  $u_f \in \mathbb{R}^{mN_f}$  is not restricted, we find  $\text{im}(U_f V_p) = \mathbb{R}^{mN_f}$ , which immediately completes the proof. ■

Before analyzing implications of  $\mu = mN_f$ , we briefly note that Lemma 4 also allows to specify the choice of  $K_f$ .

**Lemma 5.** *Any  $K_f$  that complies with the parametrization in (22) can be written as*

$$K_f = V_p^\top U_f^\top \Phi \quad (32)$$

for some non-singular matrix  $\Phi \in \mathbb{R}^{mN_f \times mN_f}$ .

*Proof.* By construction of (22), the column space of  $K_f$  has to be equal to the row space of  $U_f V_p$ . Since  $U_f V_p$  has full row rank according to Lemma 4, the row space is, for example, spanned by the  $mN_f$  columns of  $V_p^\top U_f^\top$ . Clearly, any other basis of the row space can be obtained according to (32) by a suitable transition matrix  $\Phi$ . ■

### C. Two sides of the same coin

The parametrizations (17) and (22) reveal a novel relation between MPC and deterministic DPC that goes beyond existing studies of the close relationship (as, e.g., in [2, Sect. V.D]). In fact, the QP (7) associated with MPC is formally related to the DPC variant (29) as follows.

**Lemma 6.** *The cost and constraint specifications of (7) and (29) satisfy the relations*

$$\check{H} = K_f^\top V_p^\top U_f^\top H U_f V_p K_f, \quad (33a)$$

$$\check{F} = K_f^\top V_p^\top U_f^\top F W_p W_p^+ + K_f^\top V_p^\top U_f^\top H U_f W_p^+, \quad (33b)$$

$$\check{G} = G U_f V_p K_f, \quad \text{and} \quad \hat{E} = E W_p W_p^+ - G U_f W_p^+. \quad (33c)$$

*Proof.* In order to prove (33a), we note that

$$\begin{aligned} \check{H} &= 2K_f^\top V_p^\top (Y_f^\top Q Y_f + U_f^\top R U_f) V_p K_f \\ &= 2K_f^\top V_p^\top U_f^\top \mathcal{T}_{N_f}^\top Q \mathcal{T}_{N_f} U_f V_p + 2V_p^\top U_f^\top R U_f V_p K_f \\ &= K_f^\top V_p^\top U_f^\top H U_f V_p K_f \end{aligned}$$

by definition of  $\check{H}$ ,  $\hat{H}$  and  $\tilde{H}$ , due to (27), and by definition of  $H$  in (8), respectively. The remaining relations in (33) can be proven analogously. ■

In principle, we can state a similar result to Lemma 6 for the relation between (7) and (19). However, only (33) involves the terms  $U_f V_p K_f$  with the following useful feature.

**Lemma 7.** *Let  $U_f$ ,  $V_p$ , and  $K_f$  be defined as in (13), (17), and (22), respectively. Then,  $U_f V_p K_f$  is non-singular.*

*Proof.* We initially find  $U_f V_p K_f = U_f V_p V_p^\top U_f^\top \Phi$  for some non-singular  $\Phi$  according to Lemma 5. Further, since  $U_f V_p$  has full row rank by Lemma 4,  $U_f V_p V_p^\top U_f^\top$  is non-singular and, hence, also the product  $U_f V_p V_p^\top U_f^\top \Phi$ . ■

Lemma 7 immediately leads to the following major result.

**Lemma 8.** *The QP (29) is strictly convex.*

*Proof.* Since  $U_f V_p K_f$  is non-singular by Lemma 7,  $\check{H}$  and  $H$  are congruent according to (33a). Hence,  $\check{H}$  inherits the positive definiteness of  $H$ , which proves the claim. ■

We are now ready to address the explicit solutions of (7) and (29). To this end, we recall that (15) can be derived from the parametric Karush-Kuhn-Tucker (KKT) conditions

$$H u_f^*(x_0) + F x_0 + G^\top \lambda^*(x_0) = 0, \quad (34a)$$

$$G u_f^*(x_0) - E x_0 - d \leq 0, \quad (34b)$$

$$\lambda^*(x_0) \geq 0, \quad (34c)$$

$$\text{diag}(\lambda^*(x_0)) (G u_f^*(x_0) - E x_0 - d) = 0 \quad (34d)$$

of (7) [9, Sect. 4.1]. Analogously, the explicit solution of (29) follows from the parametric KKT conditions

$$\check{H} \beta^*(\xi) + \check{F} \xi + \check{G}^\top \check{\lambda}^*(\xi) = 0, \quad (35a)$$

$$\check{G} \beta^*(\xi) - \hat{E} \xi - d \leq 0, \quad (35b)$$

$$\check{\lambda}^*(\xi) \geq 0, \quad (35c)$$

$$\text{diag}(\check{\lambda}^*(\xi)) (\check{G} \beta^*(\xi) - \hat{E} \xi - d) = 0. \quad (35d)$$

A central observation now is that (34) and (35) are equivalent.

**Theorem 9.** *The KKT conditions (34) and (35) are coupled by the relations (28),  $x_0 = \Gamma W_p W_p^+ \xi$ , and  $\lambda^*(x_0) = \check{\lambda}^*(\xi)$ .*

*Proof.* Substituting the coupling relations in (34) and multiplying (34a) with the transpose of  $T := U_f V_p K_f$  from the left, and taking (33) into account, immediately allows us to transform (34) into (35). The inverse transformation follows analogously by noting that  $T$  is invertible according to Lemma 7, which, e.g., allows to derive  $\beta^*(\xi) = T^{-1}(u_f^*(x_0) - U_f W_p^+ \xi)$  from (28). ■

**Remark 1.** The relation between  $x_0$  and  $\xi$  has initially been introduced as  $x_0 = \Gamma\xi$  above (24). However, we require  $x_0 = \Gamma W_p W_p^+ \xi$  in Theorem 9 in order to account for  $\xi$  not belonging to the system but feasible for (19). In fact,  $W_p W_p^+ \xi$  maps such  $\xi$  to belonging ones (and leaves already belonging  $\xi$  unaltered).

Based on the equivalence of the KKT conditions, it is straightforward to see that also the explicit solutions of (7) and (29) are equivalent. Most importantly, we find the following result that we state without a formal proof.

**Corollary 10.** Assume that the explicit solution of (7) can be described based on a continuous PWA function with  $s$  segments as in (15). Then, the same applies to the explicit solution of (29) and vice versa.

**Remark 2.** Note that a similar statement could, in principle, also be formulated for the solution  $a^*(\xi)$  of (14). In fact, by exploiting [2, Cor. 5.1], it immediately follows that  $u_f(\xi) = U_f a^*(\xi)$  can also be described with  $s$  segments. However, even for fixed data matrices,  $a^*(\xi)$  is not unique, which significantly complicates the derivation of an explicit solution (without using the tools leading to (29)).

#### IV. NUMERICAL EXAMPLES

##### A. Illustrating key insights with a 1-dimensional system

As a first example, we consider system (1) with

$$A = 1.2 \quad \text{and} \quad B = C = D = 1$$

subject to the constraints  $\mathcal{U} = [-1, 1]$  and  $\mathcal{Y} = [-4, 4]$ . Further, we choose  $Q = R = 0.5$  and  $N_f = 2$ , which already determines the MPC problem (3). In order to specify (7), we note that  $M_u = M_y = (1 \ -1)^\top$ ,  $v_u = (1 \ 1)^\top$  and  $v_y = (4 \ 4)^\top$  are in line with (2). Explicitly solving (7) then leads to the PWA functions in Figure 1 with  $s = 5$  segments.

Now, to setup and investigate the DPC, we first note that  $N_p = 1$  guarantees full rank of  $\mathcal{O}_{N_p} = C = 1$ . Hence, we choose an input sequence  $u_d$ , which is persistently exciting of order  $N_e = 4$  as in (11). According to (10), this requires at least  $N_d = 7$  elements. It can be easily verified that

$$u_d := (-0.6 \ 0 \ 0 \ 0 \ 0.5 \ 0.5 \ 1)^\top$$

satisfies all conditions. Furthermore,

$$y_d := (-0.1 \ 0 \ 0 \ 0 \ 0.5 \ 1 \ 2.1)^\top$$

is a consistent output sequence since (9) is satisfied for  $x_0 = 0.5$ . According to (13),  $u_d$  and  $y_d$  specify

$$W_p = \begin{pmatrix} -0.6 & 0 & 0 & 0 & 0.5 \\ -0.1 & 0 & 0 & 0 & 0.5 \end{pmatrix}, \quad U_f = \begin{pmatrix} 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 & 1 \end{pmatrix},$$

and  $Y_f$  with  $l = 5$ . In the following, we mainly focus on the transformation to (29) and its explicit solution. To this end, we first require  $W_p^+$  and  $V_p$  as in (17). Taking  $\text{rank}(W_p) = 2$  and, consequently,  $\nu = 3$  into account, suitable choices are

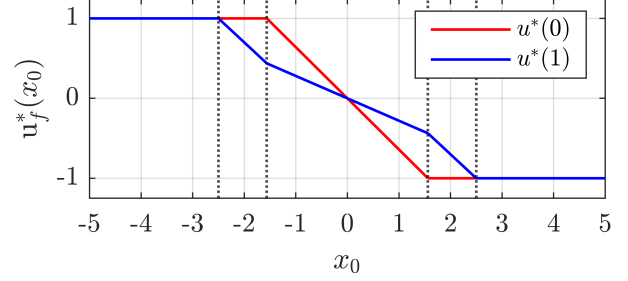


Figure 1: Explicit solution  $u_f^*(x_0)$  for MPC.

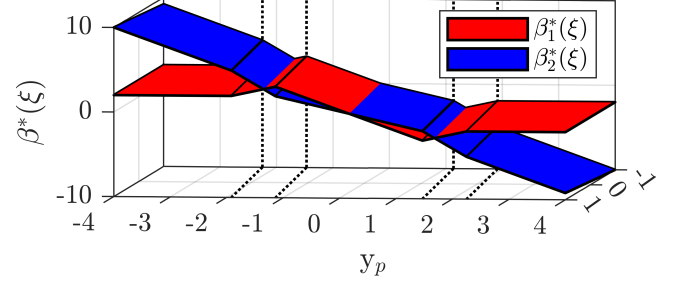


Figure 2: Explicit solution  $\beta^*(\xi)$  for (modified) DPC. Note that  $(y_p, u_p)$  is artificially restricted to  $\mathcal{Y} \times \mathcal{U}$  for visualization.

$$W_p^+ = \begin{pmatrix} -2 & 2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -0.4 & 2.4 \end{pmatrix} \quad \text{and} \quad V_p = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

We next focus on the parametrization in (22) and choose

$$K_f^\top = 2U_f V_p = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

in accordance with (32) for  $\Phi = 2I_2$ . This specifies (29), where we only list

$$\tilde{H} = \begin{pmatrix} 1.75 & 2.5 \\ 2.5 & 3.75 \end{pmatrix} \quad \text{and} \quad \tilde{F} = \begin{pmatrix} -1.34 & 8.04 \\ -1.96 & 11.76 \end{pmatrix}$$

as a reference. Explicitly solving (29) leads to the PWA functions in Figure 2. Obviously,  $\beta^*(\xi)$  likewise consists of  $s = 5$  segments as predicted by Corollary 10.

##### B. Investigating practical features with the double integrator

As a second example, we consider a standard double integrator system with

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}, \quad C = (1 \ 0), \quad \text{and} \quad D = 0$$

subject to the constraints  $\mathcal{U} = [-1, 1]$  and  $\mathcal{Y} = [-25, 25]$ . Further, we choose  $N_f = 5$  as well as  $Q = I_2$  and  $R = 0.01$ . We next reformulate the constraints as in the first example with  $v_y = (25 \ 25)^\top$  and explicitly solve (7) using the multi-parametric toolbox [15]. As a result, we obtain  $mN_f = 5$  PWA functions with  $s = 33$  segments.

The focus of the following analysis of the modified DPC is slightly different to that in Section IV.A. In fact, while the first example aimed for an as simple as possible illustration of the novel approach, this second example addresses more practical implementations. More specifically, we investigate the influence of randomly chosen input sequences  $u_d$  with larger lengths  $N_d$  than theoretically required. In this context, we initially note that full rank of  $\mathcal{O}_{N_p}$  requires  $N_p \geq 2$ . As a consequence, we need at least  $N_d \geq 17$  to achieve persistent excitation of order  $N_e \geq 9$ . Hence, DPC initially results in the QP (14) with  $l \geq 11$  decision variables. Next, by eliminating the equality constraints, we find (19) with  $\nu \geq 7$ . As indicated by (31) and Lemma 4, the final simplification step always leads to the QP (29) with  $\mu = mN_f = 5$  and, hence, as many decision variables as (7) independent of the actual choices of  $N_d$  and  $N_p$ . In addition, also the number of segments of the explicit PWA solution to (29) is identical to that of (7). These observations can be useful in practice since lower bounds for  $N_d$  and  $N_p$  might not always be available.

## V. CONCLUSIONS AND OUTLOOK

By establishing a stricter relation to classical MPC, we have shown that explicit DPC for deterministic linear systems is not as intractable as the “dimensions” of the corresponding OCP suggest. More precisely, through SID-type manipulations of the involved data matrices, we expressed DPC in terms of a strictly convex parametric QP that has exactly as many decisions variables and an exactly as complex explicit solution as its MPC counterpart.

Deterministic DPC for linear systems is of limited use for practical applications, which typically involve uncertainties and nonlinear effects. Hence, future work will address extensions to noisy and uncertain data as well as nonlinear systems. In this context, a promising direction could be the estimation of the “deterministic part” of the system as recently proposed in [14]. Furthermore, we will investigate potential applications of explicit DPC such as, e.g., the extension of the encrypted DPC without constraints in [16] to a realization involving the constraints (2).

## REFERENCES

- [1] H. Yang and S. Li. A new method of direct data-driven predictive controller design. *9th Asian Control Conference*, pp. 1–6, 2013.
- [2] J. Coulson, J. Lygeros, and F. Dörfler. Data-enabled predictive control: In the shallows of the DeePC. *18th European Control Conference*, pp. 307–312, 2019.
- [3] J. Berberich and F. Allgöwer. A trajectory-based framework for data-driven system analysis and control. *2020 European Control Conference*, pp. 1365–1370, 2020.
- [4] F. Dörfler, J. Coulson, and I. Markovsky. Bridging direct&indirect data-driven control formulations via regularizations and relaxations. *IEEE Transactions on Automatic Control*, 2022.
- [5] J. C. Willems, P. Rapisarda, I. Markovsky, and B. De Moor. A note on persistency of excitation. *Syst. Control Lett.*, 54(4):325–329, 2005.
- [6] H. J. van Waarde, J. De Persis, M. K. Çamlıbel, and P. Tesi. Willems’ fundamental lemma for state-space systems and its extension to multiple datasets. *IEEE Control Systems Letters*, 4:602–607, 2020.
- [7] I. Markovsky and F. Dörfler. Identifiability in the behavioral setting. 2020. Available at <http://homepages.vub.ac.be/imarkovs/publications/identifiability.pdf>
- [8] D. Alpag0, F. Dörfler, and J. Lygeros. An extended Kalman filter for data-enabled predictive control. *IEEE Control Systems Letters*, 4:994–999, 2020.
- [9] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1):3–20, 2002.
- [10] V. Breschi, A. Sassella, and S. Formentin. On the design of regularized explicit predictive controllers from input-output data. arXiv:2110.11808v1, 2021.
- [11] A. Sassella, V. Breschi, and S. Formentin. Learning explicit predictive controllers: theory and applications. arXiv:2108.08412v2, 2021.
- [12] P. van Overschee and B. de Moor. *Subspace identification for linear systems*. Kluwer Academic Publishers, 1996.
- [13] F. Fiedler and S. Lucia. On the relationship between data-enabled predictive control and subspace predictive control. *2021 European Control Conference*, pp. 222–229, 2021.
- [14] V. Breschi, A. Chiuso, and S. Formentin. The role of regularization in data-driven predictive control. arXiv:2203.10846v1, 2022.
- [15] M. Herceg, M. Kvasnica, C.N. Jones, and M. Morari. Multi-Parametric Toolbox 3.0. *2013 European Control Conference*, pp. 502–510, 2013.
- [16] A. B. Alexandru, A. Tsiamis, and G. J. Pappas. Towards private data-driven control. *59th IEEE Conference on Decision and Control*, pp. 5449–5456, 2020.