In 61th IEEE Conference on Decision and Control, Cancun, Mexico, 2022

Data-enabled predictive control with instrumental variables: the direct equivalence with subspace predictive control

Jan-Willem van Wingerden[1], Sebastiaan P. Mulders[1], Rogier Dinkla[1], Tom Oomen[1] and Michel Verhaegen[1]

***Abstract*— Direct data-driven control has attracted substantial interest since it enables optimization-based control without the need for a parametric model. This paper presents a new Instrumental Variable (IV) approach to Data-enabled Predictive Control (DeePC) that results in favorable noise mitigation properties, and demonstrates the direct equivalence between DeePC and Subspace Predictive Control (SPC). The methodology relies on the derivation of the characteristic equation in DeePC along the lines of subspace identification algorithms. A particular choice of IVs is presented that is uncorrelated with future noise, but at the same time highly correlated with the data matrix. A simulation study demonstrates the improved performance of the proposed algorithm in the presence of process and measurement noise.**

I. Introduction

In an era where data is abundantly available and the scientific fields of Machine Learning (ML) and Artificial Intelligence (AI) progress rapidly, the field of data-driven control has consequentially attracted a significant interest [1][2]. In particular in the field of Model Predictive Control (MPC), Data-enabled Predictive Control (DeePC) has been substantially developed. This algorithm has its origin in Willems' fundamental Lemma [3], which states that all future input-output trajectories of a linear system are parameterized by a sufficiently excited past input-output trajectory [4].

Data-driven control methods can be categorized into *direct* and *indirect* approaches. DeePC is a direct data-driven control framework in the sense that it uses past input-output data to make future decisions without an explicit parametric model [5]. In sharp contrast, indirect data-driven control methods perform a separate, and often computational and time expensive system realisation step. That is, in subspace identification a singular value decomposition is performed to find the model order followed by a linear regression problem [6][7]. The DeePC algorithm combines the system realization and control steps in a single optimization using solely data-matrices, making it a potentially efficient and promising method for the control of unknown systems.

In the original DeePC paper [8] the algorithm is derived for deterministic LTI systems without taking noise explicitly into account. Later, in [5] and [9] the algorithm in [8] is extended to address disturbing noise through a regularization term. In practical application, averaging is applied over the past input-output trajectory to mitigate the effect of noise [10]. In [11] a direct data-driven control algorithm called Subspace Predictive Control (SPC) is developed, that does take noise into account in its problem formulation without the need for regularization.

Several recent papers connect the different algorithms in the field of data-driven control. Research in [12] shows that SPC and DeePC are equivalent for the deterministic case, whereas the work of [5] makes a bridge between different data-driven control techniques. In addition, in [5] additional regularization terms are suggested to exploit the underlying causality and low-rank properties for the SPC algorithm. If measurement and process noise are added, the original DeePC algorithm depends on additional regularization and/or data-averaging to mitigate the effect of noise [10]. The open challenge is to optimally mitigate the effect of noise in such data-driven control algorithms.

Although substantial developments have been made in data-driven control algorithms, at present the effects of noise are not completely addressed. The aim of this paper is to present a new algorithm based on the DeePC framework, now including instrumental variables (IVs), which is a well-established technique to mitigate the effect of noise [13]. In addition, full equivalence between SPC and DeePC with IVs is established. To this end, the characteristic equation used in DeePC is first derived from the data-equation used in subspace identification algorithms including measurement and process noise. Then, IVs are introduced to mitigate the negative effect of noise. By formulating the predictive control problem, it becomes evident that the two seemingly different algorithms are equivalent.

The main contributions of this paper are threefold:

- The traditional DeePC algorithm is formulated directly in terms of the data-equations that are encountered in subspace identification.
- Instrumental variables are introduced and shown to result in favourable noise mitigation properties.
- Direct equivalence of IV-based DeePC and SPC is established.

This paper is organized as follows. Section II introduces the model structure and notation used. In Section III, the characteristic equation of the DeePC algorithm is derived and extended with instrumental variables (IVs). These IVs reveal a direct equivalence with SPC. In Section IV, the data-enabled predictive control problem is set up, followed by simulation results in Section V. Finally, conclusions are drawn in Section VI.

This work is part of the research programme closed-loop wake steering for large densely space wind farms" with project number 17512, which is financed by the Dutch Research Council (NWO).). [1]Delft University of Technology, Delft Center for Systems and Control, Mekelweg 2, 2628 CD Delft, The Netherlands. {J.W.vanWingerden, S.P.Mulders, R.T.O.Dinkla, T.A.E.Oomen, M.Verhaegen}@tudelft.nl.

II. Model structure and notation

This section presents the model structure and the notation used for the derivation of the different algorithms considered in this paper.

A. Model structure

The assumed model structure, commonly used in the field of subspace identification, is given by the following Linear Time-Invariant (LTI) discrete-time system in innovation form [14]:

$$\begin{aligned} x_{k+1} =& \quad Ax_k + Bu_k + Ke_k, \\ y_k =& \quad Cx_k + Du_k + e_k, \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^r$, $e_k \in \mathbb{R}^\ell$, $y_k \in \mathbb{R}^\ell$, are the state, input, noise, and output vectors, respectively, and $k \in \mathbb{Z}_{\geq 0}$ denotes the discrete time index. The matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times r}$, $K \in \mathbb{R}^{n\times \ell}$, $C \in \mathbb{R}^{\ell\times n}$, $D \in \mathbb{R}^{\ell\times r}$, are the respective system, input, Kalman, output, and direct feedthrough matrices. By manipulation of (1), the so-called predictor form is obtained:

$$\begin{aligned} x_{k+1} =& \quad \tilde{A}x_k + \tilde{B}u_k + Ky_k, \\ y_k =& \quad Cx_k + Du_k + e_k, \end{aligned} \tag{2}$$

with $\tilde{A} = A - KC$ and $\tilde{B} = B - KD$. The prediction problem inherently present in a data-enabled predictive control setting can now be formulated as:

Problem statement: Given N data instances of an unknown LTI system (1), determine a data-driven prediction of the system trajectory as function of the associated control action over a prediction horizon f.

B. Assumptions and notation

This section introduces the notation and assumptions used throughout this paper, see also [7]. First, the block-Hankel matrix is defined as:

$$Y_{i,s,\bar{N}} = \begin{bmatrix} y_i & y_{i+1} & \cdots & y_{i+\bar{N}-1} \\ y_{i+1} & y_{i+2} & \cdots & y_{i+\bar{N}} \\ \vdots & \vdots & \ddots & \vdots \\ y_{i+s-1} & y_{i+s} & \cdots & y_{i+\bar{N}+s-2} \end{bmatrix}, \tag{3}$$

with $Y_{i,s,\bar{N}} \in \mathbb{R}^{\ell s\times \bar{N}}$, $i \in \mathbb{Z}$, and $\{s,\ \bar{N}\} \in \mathbb{Z}_{>0}$. The block-Hankel matrices $U_{i,s,\bar{N}} \in \mathbb{R}^{rs\times\bar{N}}$ and $E_{i,s,\bar{N}} \in \mathbb{R}^{\ell s\times\bar{N}}$ are defined in a similar fashion. Note that $\bar{N}+s-1$ data samples are used to construct this block-Hankel matrix.

A block-Toeplitz matrix is defined as follows:

$$H_{(B,D)} = \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{f-2}B & CA^{f-3}B & \cdots & CB & D \end{bmatrix}, \tag{4}$$

with $H_{(B,D)} \in \mathbb{R}^{\ell f\times rf}$, and likewise $H_{(K,I)} \in \mathbb{R}^{\ell f\times \ell f}$ with $f \in \mathbb{Z}_{>0}$.

Next, the extended controllability matrix is defined:

$$\mathcal{K}_{(\tilde{B})} = \begin{bmatrix} \tilde{A}^{p-1}\tilde{B} & \tilde{A}^{p-2}\tilde{B} & \cdots & \tilde{B} \end{bmatrix}, \tag{5}$$

with $\mathcal{K}_{(\tilde{B})} \in \mathbb{R}^{n\times rp}$, and similarly $\mathcal{K}_{(K)} \in \mathbb{R}^{n\times \ell p}$ with $p \in \mathbb{Z}_{>0}$. For presentation reasons we also introduce $\mathcal{K} = \begin{bmatrix} \mathcal{K}_{(\tilde{B})} & \mathcal{K}_{(K)} \end{bmatrix} \in \mathbb{R}^{n\times(r+\ell)p}$. The extended observability matrix is defined as:

$$\Gamma = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{f-1} \end{bmatrix}, \tag{6}$$

with $\Gamma \in \mathbb{R}^{\ell f\times n}$. Finally, the variable O indicates a zero matrix, o a zero vector, and I the identity matrix, all of appropriate dimensions.

III. The characteristic equation

In this section, the characteristic equation used in data-enabled predictive control is derived in an alternative way. Section III-A introduces the data equation. Then, the characteristic equation for DeePC is derived in Section III-B. Section III-C introduces the concept of instrumental variables and applies them to the characteristic equation. Finally, Section III-D exhibits the resulting equivalence of DeePC and SPC by the application of IVs.

A. The data equation

An underlying key result that is exploited in both DeePC and subspace identification methods is that the state of a system can be expressed in terms of past input-output data. Forward propagation of the state equation in (2) over p samples leads to:

$$x_{k+p} = \tilde{A}^p x_k + \mathcal{K}\begin{bmatrix} U_{k,p,1} \\ Y_{k,p,1} \end{bmatrix}. \tag{7}$$

With the assumption that $\tilde{A}$ is stable there exist a finite p such that $||\tilde{A}^p||_F \approx 0$, which is an approximation commonly used in many subspace algorithms [15].[1] For large enough p, the following well-known data equation can be constructed using the definitions from the previous subsection [7]:

$$Y_{i_p,f,\bar{N}} = \Gamma\mathcal{K}\begin{bmatrix} U_{i,p,\bar{N}} \\ Y_{i,p,\bar{N}} \end{bmatrix} + H_{(B,D)}U_{i_p,f,\bar{N}} + H_{(K,I)}E_{i_p,f,\bar{N}}, \tag{8}$$

with $\bar{N} = N - p - f + 1$ and $i_p = i + p$. Using the input-output trajectories from time instance i until $i+N-1$. This data equation is used in many subspace identification algorithms (*e.g.* N4SID [17], PO-MOESP [6] or PBSID [7]). In the PO-MOESP algorithm an orthogonal projection is used to remove the effect of $U_{i_p,f,\bar{N}}$ to isolate the low rank matrix $\Gamma\mathcal{K}$. In the PO-MOESP instrumental variables are used to suppress the effect of the noise. The key idea of the DeePC [8] algorithm is to retain $U_{i_p,f,\bar{N}}$ in the data equation, while the effect of noise was not included in the derivation of the original algorithm.

The data equation (8), connects available data from the so-called past (block-Hankel matrices with a p in the subscript)

[1] Please note that in the noiseless case there exists a deadbeat observer, K, such that $||\tilde{A}^n||_F = 0$ [16].

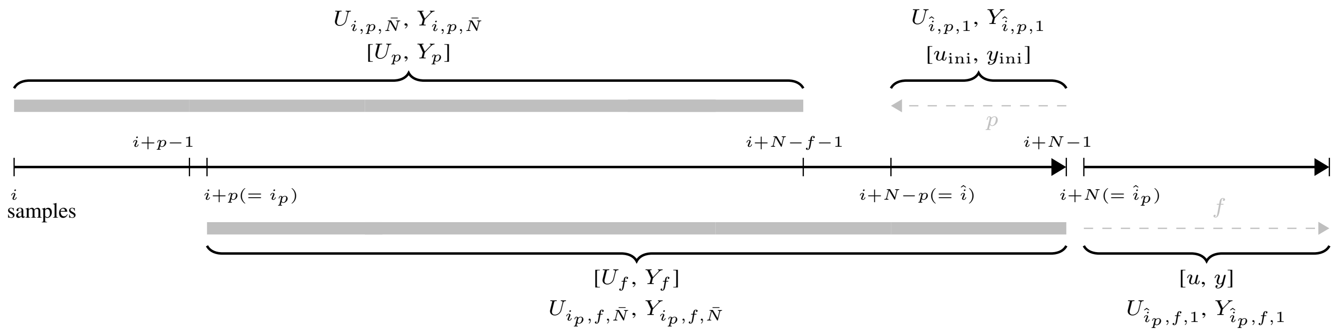


Fig. 1: Visualisation how the data is used within the data-enabled predictive control algorithm. The notation used in this paper is provided, whereas the variables within square brackets correspond to the the notation used in [8].

to the available data in the so-called future (block-Hankel matrices with an f in the subscript). These overlapping datasets are also illustrated in Fig. 1 with the gray bars. In the DeePC algorithm a second similar data equation is defined which connects the last p input-output samples available in the data set to f unknown input-output samples in the future. This is illustrated in Fig. 1 by the dashed gray arrows. This second data-equation is given by:

$$Y_{\hat{i}_p,f,1} = \Gamma\mathcal{K}\begin{bmatrix} U_{\hat{i},p,1} \\ Y_{\hat{i},p,1} \end{bmatrix} + H_{(B,D)} U_{\hat{i}_p,f,1} + H_{(K,I)} E_{\hat{i}_p,f,1}, \quad (9)$$

with $\hat{i} = i{+}N{-}p$ and $\hat{i}_p = i{+}N$. Note that with the current indices $Y_{\hat{i}_p,f,1}$, $U_{\hat{i}_p,f,1}$ are unknown vectors and the variables that will be used in the predictive control problem.

B. The DeePC algorithm characteristic equation

In this section, the data equation, which is the result of the previous subsection, is used to derive the characteristic equation for DeePC. To this end, first the deterministic LTI case is considered, whereas the following subsection incorporates noise in the data equation while also introducing IVs.

Equations (8)-(9) are respectively rewritten for the noise-free case by omitting the noise term as:

$$\begin{bmatrix} \Gamma\mathcal{K} & H_{(B,D)} & -I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} U_{i,p,\bar{N}} \\ Y_{i,p,\bar{N}} \end{bmatrix} \\ U_{i_p,f,\bar{N}} \\ Y_{i_p,f,\bar{N}} \end{bmatrix} = O, \quad (10)$$

$$\begin{bmatrix} \Gamma\mathcal{K} & H_{(B,D)} & -I \end{bmatrix} \begin{bmatrix} \begin{bmatrix} U_{\hat{i},p,1} \\ Y_{\hat{i},p,1} \end{bmatrix} \\ U_{\hat{i}_p,f,1} \\ Y_{\hat{i}_p,f,1} \end{bmatrix} = o. \quad (11)$$

To obtain the characteristic equation of the DeePC algorithm, (10) is at the right hand side multiplied with a vector $g \in \mathbb{R}^{\bar{N}\times 1}$. This basically means that linear combinations of input-output trajectories in the available data set are taken. By subtracting (11) from (10) and introducing g, we obtain:

$$\begin{bmatrix} \Gamma\mathcal{K} & H_{(B,D)} & -I \end{bmatrix} \left(\begin{bmatrix} \begin{bmatrix} U_{i,p,\bar{N}} \\ Y_{i,p,\bar{N}} \end{bmatrix} \\ U_{i_p,f,\bar{N}} \\ Y_{i_p,f,\bar{N}} \end{bmatrix} g - \begin{bmatrix} \begin{bmatrix} U_{\hat{i},p,1} \\ Y_{\hat{i},p,1} \end{bmatrix} \\ U_{\hat{i}_p,f,1} \\ Y_{\hat{i}_p,f,1} \end{bmatrix} \right) = o. \quad (12)$$

This result shows that an unknown input-output trajectory can be embedded as a linear combination of available input-output trajectories. Because the pre-multiplication matrix is full-rank, the following is obtained:

$$\begin{bmatrix} \begin{bmatrix} U_{i,p,\bar{N}} \\ Y_{i,p,\bar{N}} \end{bmatrix} \\ U_{i_p,f,\bar{N}} \\ Y_{i_p,f,\bar{N}} \end{bmatrix} g = \begin{bmatrix} \begin{bmatrix} U_{\hat{i},p,1} \\ Y_{\hat{i},p,1} \end{bmatrix} \\ U_{\hat{i}_p,f,1} \\ Y_{\hat{i}_p,f,1} \end{bmatrix}, \quad (13)$$

which can be used in a data-enabled prediction control problem with g, $U_{\hat{i}_p,f,1}$ and $Y_{\hat{i}_p,f,1}$ as decision variables. Although the derivation is different, the result is in accordance with the results presented in [8].

As already shown in [18], (13) can be split into two equations:

$$\begin{bmatrix} \begin{bmatrix} U_{i,p,\bar{N}} \\ Y_{i,p,\bar{N}} \end{bmatrix} \\ U_{i_p,f,\bar{N}} \end{bmatrix} g = \begin{bmatrix} \begin{bmatrix} U_{\hat{i},p,1} \\ Y_{\hat{i},p,1} \end{bmatrix} \\ U_{\hat{i}_p,f,1} \end{bmatrix}, \quad (14)$$

$$Y_{i_p,f,\bar{N}} g = Y_{\hat{i}_p,f,1}. \quad (15)$$

The data matrix in the left-hand side of (14), has dimensions $(r{+}\ell)\,p{+}rf \times \bar{N}$. Assuming that the conditions on persistency of excitation are fulfilled, it should hold in the deterministic case that $\bar{N} \geq (r{+}\ell)p{+}rf$ to parameterize all possible future trajectories.

In the case of process and measurement noise, the trajectories in the available data set are corrupted by noise and consequently the DeePC algorithm will select linear combinations of these trajectories leading to a reduced performance. As proposed in [8], noise can be mitigated through regularization, or as proposed in [10], mitigated by columnwise data-averaging such that the augmented block-Hankel matrix is square ($\bar{N} = (r{+}\ell)p{+}rf$). The central idea in the present paper is to employ IVs for a systematic mitigation of noise, as is presented in the next section.

C. The DeePC algorithm with Instrumental Variables

In this section, the main result of DeePC with instrumental variables that systematically mitigate noise is presented. Equation (8) is again taken as a starting point for the derivation including noise and is rewritten as:

$$\begin{bmatrix}\Gamma\mathcal{K} & H_{(B,D)} & -I\end{bmatrix}\begin{bmatrix}\begin{bmatrix}U_{i,p,\bar{N}}\\Y_{i,p,\bar{N}}\end{bmatrix}\\U_{i_p,f,\bar{N}}\\Y_{i_p,f,\bar{N}}\end{bmatrix} = -H_{(K,I)}E_{i_p,f,\bar{N}}. \quad (16)$$

Furthermore, Equation (9) is considered for a noise free prediction:

$$\begin{bmatrix}\Gamma\mathcal{K} & H_{(B,D)} & -I\end{bmatrix}\begin{bmatrix}\begin{bmatrix}U_{\hat{i},p,1}\\Y_{\hat{i},p,1}\end{bmatrix}\\U_{\hat{i}_p,f,1}\\Y_{\hat{i}_p,f,1}\end{bmatrix} = o. \quad (17)$$

An instrumental variable $Z_{\bar{N}} \in \mathbb{R}^{q\times\bar{N}}$ is defined such that:

$$\lim_{\bar{N}\to\infty}\frac{1}{\bar{N}}\left(E_{i_p,f,\bar{N}}Z_{\bar{N}}^T\right) = O, \quad (18)$$

and

$$\text{rank}\left(\lim_{\bar{N}\to\infty}\frac{1}{\bar{N}}\begin{bmatrix}U_{i,p,\bar{N}}\\Y_{i,p,\bar{N}}\\U_{i_p,f,\bar{N}}\end{bmatrix}Z_{\bar{N}}^T\right) = \text{rank}\left(\begin{bmatrix}U_{i,p,\bar{N}}\\Y_{i,p,\bar{N}}\\U_{i_p,f,\bar{N}}\end{bmatrix}\right). \quad (19)$$

That is, $Z_{\bar{N}}$ is chosen to be uncorrelated with the block-Hankel matrix containing the noise but highly correlated with the data matrix. A suitable candidate (a variation of the IVs used in PO-MOESP [6]) is given by:

$$Z_{\bar{N}} = \begin{bmatrix}U_{i,p,\bar{N}}\\Y_{i,p,\bar{N}}\\U_{i_p,f,\bar{N}}\end{bmatrix}, \quad (20)$$

which under persistency of excitation conditions [4] satisfies the posed conditions (proof along the lines of IVs for PO-MOESP [6], Chapt. 9.6).

Equation (16) is multiplied with the transpose of the instrumental variable and a vector $\hat{g} \in \mathbb{R}^{q\times 1}$ and (17) is subtracted. This leads to the following result:

$$\lim_{\bar{N}\to\infty}\begin{bmatrix}\Gamma\mathcal{K} & H_{(B,D)} & -I\end{bmatrix}\ldots \quad (21)$$

$$\left(\begin{bmatrix}\begin{bmatrix}U_{i,p,\bar{N}}\\Y_{i,p,\bar{N}}\end{bmatrix}\\U_{i_p,f,\bar{N}}\\Y_{i_p,f,\bar{N}}\end{bmatrix}Z_{\bar{N}}^T\hat{g} - \begin{bmatrix}\begin{bmatrix}U_{\hat{i},p,1}\\Y_{\hat{i},p,1}\end{bmatrix}\\U_{\hat{i}_p,f,1}\\Y_{\hat{i}_p,f,1}\end{bmatrix}\right) = o,$$

and similar as in the noise free case:

$$\lim_{\bar{N}\to\infty}\left(\begin{bmatrix}\begin{bmatrix}U_{i,p,\bar{N}}\\Y_{i,p,\bar{N}}\end{bmatrix}\\U_{i_p,f,\bar{N}}\end{bmatrix}Z_{\bar{N}}^T\hat{g}\right) = \begin{bmatrix}\begin{bmatrix}U_{\hat{i},p,1}\\Y_{\hat{i},p,1}\end{bmatrix}\\U_{\hat{i}_p,f,1}\end{bmatrix}, \quad (22)$$

which can be used in a data-enabled prediction control problem with $\hat{g}$, $U_{\hat{i}_p,f,1}$ as the decision variables, and

$$Y_{\hat{i}_p,f,1} = \lim_{\bar{N}\to\infty}Y_{i_p,f,\bar{N}}Z_{\bar{N}}^T\hat{g}, \quad (23)$$

being an asymptotically noise free prediction of the output. In the nominal DeePC algorithm [8] the predictions are made on linear combinations of noisy output signals. Adding regularization to the objective function of the nominal DeePC [18] or data-averaging [10] can mitigate the effect of the noise.

D. Subspace predictive control

In this section, the direct equivalence between the Subspace Predictive Control (SPC) and the DeePC algorithm is established through the use of the presented instrumental variables. This further strengthens the results in [5] where an initial connection has been established between SPC and the traditional DeePC algorithm.

Since the data matrix in (22) is known and under persistency of excitation conditions invertible, $\hat{g}$ can be explicitly solved for. Substitution of the obtained expression in (23) results in:

$$Y_{\hat{i}_p,f,1} = Y_{i_p,f,\bar{N}}Z_{\bar{N}}^T(Z_{\bar{N}}Z_{\bar{N}}^T)^{-1}\begin{bmatrix}\begin{bmatrix}U_{\hat{i},p,1}\\Y_{\hat{i},p,1}\end{bmatrix}\\U_{\hat{i}_p,f,1}\end{bmatrix}. \quad (24)$$

In the SPC algorithm, a similar solution can be obtained by directly solving (8) for $\begin{bmatrix}\Gamma\mathcal{K} & H_{(B,D)}\end{bmatrix}$ in a least squares sense and putting the available data in a data matrix $Z_{\bar{N}}$:

$$\min_{\begin{bmatrix}\Gamma\mathcal{K} & H_{(B,D)}\end{bmatrix}}\left|\left|Y_{i_p,f,\bar{N}} - \begin{bmatrix}\Gamma\mathcal{K} & H_{(B,D)}\end{bmatrix}Z_{\bar{N}}\right|\right|_F. \quad (25)$$

The solution of this problem is given by the following estimate:

$$\begin{bmatrix}\hat{\Gamma\mathcal{K}} & \hat{H}_{(B,D)}\end{bmatrix} = Y_{i_p,f,\bar{N}}Z_{\bar{N}}^T(Z_{\bar{N}}Z_{\bar{N}}^T)^{-1}. \quad (26)$$

Substitution of this result in (9), and taking the expectation of the future noise, directly leads to (24) which shows the direct equivalence.

The SPC algorithm is also discussed in [5], where it is also observed that the low-rank property of the estimated $\Gamma\mathcal{K}$ matrix, and the lower-triangular block-Toeplitz property of the estimated $H_{(B,D)}$ matrix are not enforced. In [5], methods are proposed to enforce this structure by the employment of regularization techniques.

IV. DeePC

With the equations from the previous section at hand, the original DeePC problem as proposed in [8] is defined as:

$$\begin{aligned}\min_{U_{\hat{i}_p,f,1},g} \quad & Y_{\hat{i}_p,f,1}{}^TQY_{\hat{i}_p,f,1} + U_{\hat{i}_p,f,1}{}^TRU_{\hat{i}_p,f,1}\\ \text{subject to:} \quad & \begin{bmatrix}U_{i,p,\bar{N}}\\Y_{i,p,\bar{N}}\\U_{i_p,f,\bar{N}}\end{bmatrix}g = \begin{bmatrix}U_{\hat{i},p,1}\\Y_{\hat{i},p,1}\\U_{\hat{i}_p,f,1}\end{bmatrix}, \quad (27)\\ & Y_{\hat{i}_p,f,1} = Y_{i_p,f,\bar{N}}g.\end{aligned}$$

The cost-function can be extended with a regularization term (*e.g.*, $\lambda||g||_2$), or input/output constraints. With the addition of IVs to the DeePC problem, a slightly modified optimization problem is defined:

$$\begin{aligned}\min_{U_{\hat{i}_p,f,1},\hat{g}} \quad & Y_{\hat{i}_p,f,1}{}^TQY_{\hat{i}_p,f,1} + U_{\hat{i}_p,f,1}{}^TRU_{\hat{i}_p,f,1}\\ \text{subject to:} \quad & \begin{bmatrix}U_{i,p,\bar{N}}\\Y_{i,p,\bar{N}}\\U_{i_p,f,\bar{N}}\end{bmatrix}Z_{\bar{N}}^T\hat{g} = \begin{bmatrix}U_{\hat{i},p,1}\\Y_{\hat{i},p,1}\\U_{\hat{i}_p,f,1}\end{bmatrix} \quad (28)\\ & Y_{i_p,f,\bar{N}}Z_{\bar{N}}^T\hat{g} = Y_{\hat{i}_p,f,1}\end{aligned}$$

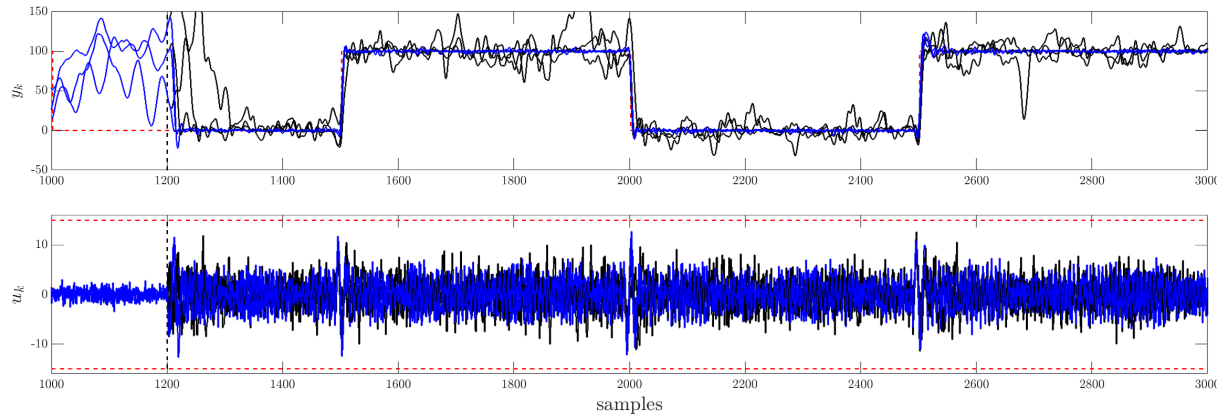


Fig. 2: Simulation results comparing the reference tracking performance of DeePC with averaging (black) and DeePC with IV (blue). For $p = f = 20$ with $\mathrm{var(e_k)} = 0.1^2$ and $\bar{N} = 500$, the proposed IV method outperforms the original method in terms of the tracking error for three distinct realizations.

where $\hat{g}$ can be seen as a dummy variable and explicitly solved for as explained in the previous section. Note that in the noiseless case, the matrix $Z_{\bar{N}} Z_{\bar{N}}^T$ becomes singular and regularization has to be applied or a different data window has to be chosen (different f and/or p).

V. Simulation Study

In this section, the presented IV-based DeePC algorithm is compared with the traditional DeePC algorithm for various levels of noise. The simulation study considers a 5^{th} order system described in [11], representing an actuated laboratory test setup with two circular plates and flexible shafts. The discrete-time system is compatible with the model structure in (1), and the corresponding system matrices are given:

$$A = \begin{bmatrix} 4.4 & 1 & 0 & 0 & 0 \\ -8.09 & 0 & 1 & 0 & 0 \\ 7.83 & 0 & 0 & 1 & 0 \\ -4 & 0 & 0 & 0 & 1 \\ 0.86 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0.00098 \\ 0.01299 \\ 0.01859 \\ 0.0033 \\ -0.00002 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad K = \begin{bmatrix} 2.3 \\ -6.64 \\ 7.515 \\ -4.0146 \\ 0.86336 \end{bmatrix}, \quad D = 0.$$

The objective is to track a square wave with an amplitude of 50 and a DC-offset of 50. During the first 1200 samples the system is excited with a normally distributed white noise signal with a variance of 1. After this period, the excitation is removed and the data-enabled predictive controllers are enabled in a receding horizon setting. The input magnitude and rate of variation are respectively constrained to $|u_k| \leq 15$ and $|u_{k+1} - u_k| \leq 3.75$.

Fig. 2 shows simulation results of the output and input trajectories of DeePC with IV and random averaging. In the case of random averaging, the matrix $Z_{\bar{N}} \in \mathbb{R}^{(r+\ell)p+rf\times\bar{N}}$ is chosen randomly. The prediction window is chosen as $f = 20$ and p is selected to equal f (common choice in subspace identification [7]), the noise variance is taken as $\mathrm{var(e_k)} = 0.1^2$, and the data window $\bar{N} = 500$. It is observed that the tracking performance of the proposed method outperforms the original method for three different realisations.

The simulation studies in the subsequent subsections consider the effect of the data window $(\bar{N})$, noise level $(\mathrm{var(e_k)})$ and the prediction window (f), and are presented in Fig. 3. The performance is measured in terms of the mean squared error of the tracking error divided by the two norm of the reference signal.

A. The data window length $\bar{N}$

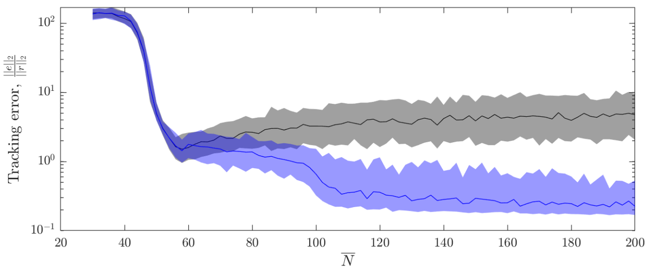
The tracking error performance of both DeePC approaches for different sizes of $\bar{N}$ is shown in Fig. 3a. For each $\bar{N}$, 100 distinct noise realisations are taken. Since $p = f = 20$, the data matrix is square for $\bar{N} = 60$. In the case of $\bar{N} \leq 60$, there is no clear difference between the proposed and the original method. However, increased data windows ($\bar{N} > 60$) clearly demonstrate the strength of the proposed IV approach with increased tracking performance, whereas the performance of the original algorithm deteriorates for larger values of $\bar{N}$.

B. The effect of $\mathrm{var}(e_k)$

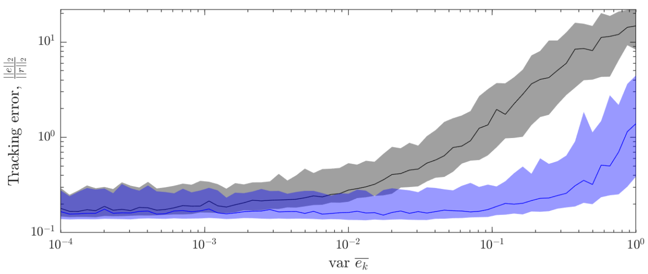
Fig. 3b compares the tracking performance of the two DeePC algorithms for different noise levels. For each $\mathrm{var(e_k)}$, 100 distinct noise realisations are taken. The results clearly demonstrate the beneficial effect of the proposed method by outperforming the original algorithm for higher noise levels. For lower variances of the noise, the performance of both approaches is comparable.

C. The effect of the prediction window f

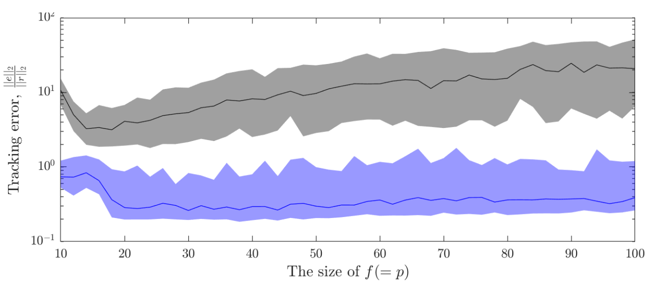
The effect of the prediction window f on the tracking performance for the proposed and original method are presented in Fig. 3c. To mitigate the effect of $\bar{N}$ on the presented results,



(a) Varying data window $\bar{N}$, and $p = f = 20$ with $\text{var}(e_k) = 0.5^2$



(b) Varying noise level $\text{var}(e_k)$, and $p = f = 20$ with $\bar{N} = 200$.



(c) Varying prediction window $f(= p)$, and $\text{var}(e_k) = 0.5^2$ with $\bar{N} = 6\times f$.

Fig. 3: Performance in terms of tracking error for the DeePC with instrumental variables (blue), and the original DeePC method with random averaging (black). Simulations are repeated 100 times for different noise realisations. The solid lines represent the median of the tracking error while the shaded area covers 80% of the realisations.

the data window is chosen as $\bar{N} = 6\times f$, and for simplicity $p = f$. For each f, 100 new noise realisations are taken.

The results demonstrate that there is a minor overall effect of the prediction window on the performance. It is concluded that the gained performance of the IV method with respect to random averaging is independent from the prediction window f.

VI. CONCLUSIONS

In this paper, a new instrumental variable approach to DeePC-based data-driven control is presented, that leads to a direct equivalence between two existing classes of direct data-driven control methods: Data-enabled Predictive Control (DeePC) and Subspace Predictive Control (SPC). To end up at this result, first the characteristic equation used in DeePC is derived from the data-equation typically used in subspace identification algorithms. Instrumental variables (IVs) are included in the DeePC framework, with the intent to asymptotically remove the effect of process and measurement noise. A particular choice of IVs is made that is uncorrelated with future noise, but at the same time highly correlated with the data matrix. Simulation studies showcase the improved performance in terms of the tracking error of the proposed algorithm in the case of process and measurement noise.

REFERENCES

[1] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2019.

[2] H. J. van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel, "Data informativity: A new perspective on data-driven analysis and control," *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4753–4768, 2020.

[3] 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.

[4] H. J. van Waarde, C. De Persis, M. K. Camlibel, and P. Tesi, "Willems' fundamental lemma for state-space systems and its extension to multiple datasets," *IEEE Control Systems Letters*, vol. 4, no. 3, pp. 602–607, 2020.

[5] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct and indirect data-driven control formulations via regularizations and relaxations," *IEEE Transactions on Automatic Control*, pp. 1–1, 2022.

[6] M. Verhaegen and V. Verdult, *Filtering and system identification: a least squares approach*. Cambridge university press, 2007.

[7] G. van der Veen, J. W. van Wingerden, M. Bergamasco, M. Lovera, and M. Verhaegen, "Closed-loop subspace identification methods: an overview," *IET Control Theory & Applications*, vol. 7, no. 10, pp. 1339–1358, 2013.

[8] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 307–312.

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

[10] N. H. Jo and H. Shim, "Data-driven output-feedback predictive control: Unknown plant's order and measurement noise," *arXiv preprint arXiv:2201.03136*, 2022.

[11] W. Favoreel, B. De Moor, and M. Gevers, "SPC: Subspace predictive control," *IFAC Proceedings Volumes*, vol. 32, no. 2, pp. 4004–4009, 1999.

[12] F. Fiedler and S. Lucia, "On the relationship between data-enabled predictive control and subspace predictive control," in *2021 European Control Conference (ECC)*. IEEE, 2021, pp. 222–229.

[13] T. Söderström and P. Stoica, "Instrumental variable methods for system identification," *Circuits, Systems and Signal Processing*, vol. 21, no. 1, pp. 1–9, 2002.

[14] L. Ljung, "System identification," in *Signal analysis and prediction*. Springer, 1998, pp. 163–173.

[15] A. Chiuso, "The role of vector autoregressive modeling in predictor-based subspace identification," *Automatica*, vol. 43, no. 6, pp. 1034–1048, 2007.

[16] I. Houtzager, J. W. van Wingerden, and M. Verhaegen, "VARMAX-based closed-loop subspace model identification," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*. IEEE, 2009, pp. 3370–3375.

[17] P. Van Overschee and B. De Moor, "N4SID: subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.

[18] J. Coulson, J. Lygeros, and F. Dorfler, "Distributionally robust chance constrained data-enabled predictive control," *IEEE Transactions on Automatic Control*, 2021.