2025 23rd European Control Conference (ECC)
June 24-27, 2025. Thessaloniki, Greece

An Innovation-Based Approach to Adaptive Data-Driven Predictive Control

Yibo Wang, Hao Ye, Dexian Huang, Xiang Yin, and Chao Shang, *Member, IEEE*

***Abstract*— Recently data-driven predictive control (DDPC) methods have attracted considerable interest, primarily due to the benefit that optimal control inputs can be obtained directly from raw data without identifying an explicit model. However, time-varying characteristics of complex systems can substantially degrade the control performance of existing DDPC methods. In this work, we propose a new adaptive DDPC scheme in which innovations play a critical role. By leveraging the (approximate) white-noise and Gaussian properties of innovations, we develop a data-driven scheme that can effectively detect changes in system dynamics. This naturally motivates a statistically grounded adaptive DDPC framework in which the online adaptation is triggered by the detection of changes in system dynamics. Numerical examples demonstrate the efficacy of the proposed detection scheme and the enhanced control performance of the adaptive DDPC method in contrast to existing heuristic methods.**

I. Introduction

As a cutting-edge control strategy, Model Predictive Control (MPC) has been widely employed in practice due to its capacity to handle constraints and multiple objectives [1]. Typically, MPC utilizes an explicit state-space model to predict future outputs. In recent years, there has been a paradigm shift towards data-driven predictive control (DDPC) methods, thanks to the availability of massive data and advancements in computing power [2], [3]. A mainstream of DDPC methods builds upon the famous Willems' fundamental lemma, a formula rooted in the behavioral theory [4]. Based on the fundamental lemma, any trajectory of a linear-invariant (LTI) system can be expressed as a linear combination of trajectories from a trajectory dictionary. This enables DDPC methods to obtain optimal control design directly based on raw input-output data without the need for an explicit model. Moreover, DDPC methods retain the traits of conventional MPC, including the capability of handling constraints and the robustness of receding horizon implementation [5].

The generic DDPC method, known as data-enabled predictive control (DeePC) [6], is built upon deterministic LTI systems, where data-based equations instead of an identified parameterized model are used for output prediction. However, when the raw data are corrupted by process disturbances and measurement noise in stochastic systems, the data-driven output predictor (OP) in DeePC becomes susceptible to uncertainties thereby considerably compromising the reliability of the control design. To address these issues, various extensions of the generic DeePC are proposed to mitigate the effect of noise in stochastic systems. The most common approaches are based on regularizations [7]–[9]. Meanwhile, the use of instrumental variables is shown to perform well [10], and this is further extended to the case under feedback control [11]. Min-max robust formulations of DeePC have been investigated in [12], [13], which offer probabilistic performance guarantees. Recently, a new data-driven OP based on innovations was put forward in [14] to effectively address additive uncertainty in stochastic systems. The data-driven OP is shown to yield the same predictions as the model-based Kalman Filter (KF)-like OP provided that the innovations of offline data are exactly known. Based on this OP, a new innovation-based DDPC method was developed, which showcases remarkably improved control performance compared with generic DDPC methods.

Besides uncertainty, some complex systems may also exhibit time-varying characteristics in practice. These changes in system dynamics can substantially reduce the prediction accuracy of the data-driven OP and consequently jeopardize the control performance of DDPC. In [15], [16], the generic DeePC method is made adaptive for linear parameter varying systems by leveraging prior information of possible variations in system dynamics, and new input-output measurements are appended to capture the new dynamics. In [17], a fast singular value decomposition (SVD) method is employed to alleviate the cumbersome computations due to recursive updating, while the earliest data are continually discarded in order to ensure effective adaptation in [18].

In this paper, we consider the scenarios where there exist unknown time-varying system dynamics during the receding horizon implementation of DDPC. Building upon the innovation-based DDPC in [14], we develop a new adaptive DDPC framework to effectively adapt to changes in system dynamics by elucidating the role that innovations and their uncertainty play. First, a data-driven scheme for detecting changes in system dynamics is proposed. Specifically, a hypothesis test procedure is designed, where the innovations are the core ingredient thanks to their (approximate) whiteness and Gaussianity. This inspires a simple yet efficient strategy to detect the changes in system behaviors. Based on this, a statistically grounded adaptive DDPC framework is

This work was supported by National Natural Science Foundation of China under Grants 62373211 and 62327807. *(Corresponding author: Chao Shang.)*

Y. Wang is with Department of Automation, Tsinghua University, Beijing 100084, China (e-mail: wyb21@mails.tsinghua.edu.cn).

H. Ye, D. Huang and C. Shang are with Department of Automation, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (e-mail: haoye@tsinghua.edu.cn, huangdx@tsinghua.edu.cn, c-shang@tsinghua.edu.cn).

X. Yin is with School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: yinxiang@sjtu.edu.cn).

developed, where the data-driven OP is updated to accommodate the new system dynamics upon detecting changes. Numerical examples demonstrate the efficacy of the proposed detection method and the superiority of the adaptive DDPC over existing heuristic adaptation schemes.

The rest of this work is organized as follows. Section II gives a brief introduction to the generic adaptive DDPC and the innovation-based DDPC method. In Section III, the data-driven detection method for changes in system dynamics is proposed, followed by the adaptive DDPC framework based on innovations. The results of numerical examples are presented in Section IV, followed by final conclusions.

Notation: The zero vector of size s and zero matrix of size $s_1 \times s_2$ are denoted by $0_s \in \mathbb{R}^s$ and $0_{s_1\times s_2} \in \mathbb{R}^{s_1\times s_2}$, respectively. For a vector x and a matrix X, $x_{[i:j]}$ and $X_{[i:j]}$ denote the subvector and submatrix with the entries from the ith row to jth row of x and X, respectively. We use $X^\dagger$ to denote the pseudo-inverse of X. The submatrix of X with entries from the ith column to jth column is denoted by $X_{[:,i:j]}$. $\|X\|_F$, $X \succ 0$ denotes the Frobenius norm and positive semidefiniteness of the matrix X, respectively. Given a sequence $\{x(k)\}_{k=1}^N$, $x_{(i:j)}$ denotes the restriction of x to the interval $[i,j]$ as $\mathrm{col}(x(i),...,x(j)) = \begin{bmatrix} x(i)^\top & \cdots & x(j)^\top \end{bmatrix}^\top$. A block Hankel matrix of depth $s \in \mathbb{N}^*$ can be constructed from $x_{(i:j)}$ via the Hankel operator $\mathcal{H}_s(x_{(i:j)})$.

II. Preliminaries

A. DDPC and its adaptation scheme

Consider the discrete LTI system in state-space form

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{1}$$

where $u(t) \in \mathbb{R}^{n_u}$, $y(t) \in \mathbb{R}^{n_y}$, and $x(t) \in \mathbb{R}^{n_x}$ denote input, output, and unmeasurable state, respectively. It is assumed that the system (6) is minimal. Given a long-enough input sequence $u^d_{(1:N)}$ with persistent excitation of order $L + n_x$ and its corresponding output $y^d_{(1:N)}$ from the system (1), the generic DDPC method, namely DeePC, can be formulated as the following optimization problem at time instance t with the past horizon $L_p \geq n_x$ and the future horizon $L_f \geq 1$ [6]:

$$\min_{u_f(t),y_f(t),g(t)} \mathcal{J}(u_f(t), y_f(t)) \tag{2a}$$

$$\text{s.t.} \quad \begin{bmatrix} U_p \\ U_f \\ Y_p \\ Y_f \end{bmatrix} g(t) = \begin{bmatrix} u_p(t) \\ u_f(t) \\ y_p(t) \\ y_f(t) \end{bmatrix}, \tag{2b}$$

$$u_f(t) \in \mathbb{U}, \; y_f(t) \in \mathbb{Y}, \tag{2c}$$

where

$$\begin{aligned} &U_d = \mathcal{H}_L(u^d_{(1:N)}), \; U_p = U_{d,[1:n_uL_p]}, \; U_f = U_{d,[n_uL_p+1:n_uL]}, \\ &u_p(t) = u_{(t-L_p:t-1)}, \; u_f(t) = u_{(t:t+L_f-1)}, \; L = L_p + L_f, \end{aligned} \tag{3}$$

and similarly for $Y_p, Y_f, E_p, E_f, y_p(t), y_f(t), e_p(t)$, $g(t)$ is a decision variable to be optimized, $\mathbb{U}$ and $\mathbb{Y}$ are input and output constraint sets. In DeePC (2), the data-based equations (2b) serve as a data-driven counterpart to the state-space model (1). Specifically, any input-output trajectory can be represented as a linear combination of columns in the so-called "trajectory dictionary", which is constructed by using Hankel matrices in (2b). The cost function in (2) can be defined as $\mathcal{J}(u_f(t), y_f(t)) = \sum_{k=t}^{t+L_f-1} \|y(k) - r(k)\|_Q^2 + \|u(k)\|_R^2$, where $Q, R \succ 0$ are weighting matrices and $r(\cdot)$ is the future output reference. Solving the problem (2) can yield the equivalent closed-loop control performance to that of MPC based on the state-space model (1) with system matrices $\{A, B, C, D\}$.

A variety of adaptive schemes have been designed to account for time-varying system dynamics. A simple and intuitive heuristic is to recursively update the trajectory dictionary in (2b) online. For example, [17] proposed to continually append the latest L input-output data as a new trajectory to the dictionary:

$$\begin{bmatrix} U_p \\ U_f \\ Y_p \\ Y_f \end{bmatrix} \leftarrow \begin{bmatrix} U_p & u_{(t-L+1:t-L_f)} \\ U_f & u_{(t-L_f+1:t)} \\ Y_p & y_{(t-L+1:t-L_f)} \\ Y_f & y_{(t-L_f:t)} \end{bmatrix}. \tag{4}$$

Although this strategy can better capture the latest dynamics, retaining past samples in the trajectory dictionary may have a detrimental effect on prediction accuracy and control performance. On account of this, [18] proposed to concurrently add new data and discard the earliest trajectory:

$$\begin{bmatrix} U_p \\ U_f \\ Y_p \\ Y_f \end{bmatrix} \leftarrow \begin{bmatrix} U_{p,[:,2:N-L+1]} & u_{(t-L+1:t-L_f)} \\ U_{f,[:,2:N-L+1]} & u_{(t-L_f+1:t)} \\ Y_{p,[:,2:N-L+1]} & y_{(t-L+1:t-L_f)} \\ Y_{f,[:,2:N-L+1]} & y_{(t-L_f:t)} \end{bmatrix}. \tag{5}$$

This heuristic not only captures time-varying dynamics better than (4) but also keeps the computational complexity under control. Rather, these adaptation schemes bear no statistical interpretation, and it remains challenging to maintain the informativity of the trajectory dictionary updated with online input-output data.

B. DDPC based on innovations

Next, we turn to the stochastic LTI system expressed in innovation form [19]:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + Ke(t) \\ y(t) &= Cx(t) + Du(t) + e(t), \end{aligned} \tag{6}$$

where the innovation $e(t) \in \mathbb{R}^{n_y}$ is zero-mean white noise with Gaussian distribution, i.e., $e(t) \sim \mathcal{N}(0, \Sigma_e)$. The steady-state Kalman gain K can render all eigenvalues of $\bar{A} = A - KC$ strictly inside the unit circle. Indeed, expressing stochastic systems as (6) does not cause much loss of generality since a generic stochastic system with additive process disturbance and measurement noise can be represented in this form under mild assumptions [19].

When handling stochastic system (6), the control performance of the generic DDPC (2) can be significantly impaired due to the presence of uncertainty. To address this, the

DDPC method based on innovations has been proposed to mitigate the effect of noise. It is formulated as the following optimization problem at time t [14]:

$$\min_{u_f(t),\hat{y}_f(t)} \mathcal{J}(u_f(t),\hat{y}_f(t)) \tag{7a}$$

$$\text{s.t.} \quad \hat{y}_f(t) = Y_f \begin{bmatrix} U_p \\ U_f \\ Y_p \\ \hat{E}_p \\ \hat{E}_f \end{bmatrix}^{\dagger} \begin{bmatrix} u_p(t) \\ u_f(t) \\ y_p(t) \\ \hat{e}_p(t) \\ 0_{pL_f} \end{bmatrix}, \tag{7b}$$

$$u_f(t) \in \mathbb{U},\ \hat{y}_f(t) \in \mathbb{Y}, \tag{7c}$$

where $\hat{E}_p$ and $\hat{E}_f$ are constructed from the innovation estimates $\hat{e}^d_{(1:N)}$. As to be detailed in the sequel, the innovation estimates $\hat{e}^d_{(1:N)}$ can be obtained from input-output data $\{u^d(k), y^d(k)\}_{k=1}^N$ by fitting a nonparametric vector autoregressive with exogenous input (VARX) model without knowing (A,B,C,D,K).

As long as the true value of innovations $e^d_{(1:N)}$ are used in (7), the data-driven OP (7b) can generate identical output predictions to those from the following KF-based multi-step output prediction with $k \geq 0$:

$$\begin{aligned} \hat{x}(t+k+1) &= A\hat{x}(t+k) + Bu(t) \\ \hat{y}(t+k) &= C\hat{x}(t+k) + Du(t), \end{aligned} \tag{8}$$

whose initial state $\hat{x}(t)$ in (8) is derived from a steady-state Kalman filter (SSKF):

$$\begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + Bu(t) + K[y(t) - \hat{y}(t)] \\ \hat{y}(t) &= C\hat{x}(t) + Du(t). \end{aligned} \tag{9}$$

Indeed, the so-called true values of the innovations $e^d_{(1:N)}$ corresponding to the input-output data $\{u^d(k), y^d(k)\}_{k=1}^N$ is the one-step ahead prediction error of the SSKF with the system matrices (A,B,C,D,K).

After solving the problem (7) at time t in the receding horizon implementation of the DDPC method (7), the innovation vector $\hat{e}_p(t+1)$ required at the next instance can be updated in a moving window fashion:

$$\begin{aligned} \hat{e}_p(t+1) &= \mathrm{col}(\hat{e}(t-L_p+1), \cdots, \hat{e}(t)) \\ &= \mathrm{col}(\hat{e}_{p,[n_y+1:n_yL_p]}, y(t) - \hat{y}_{f,[1:n_y]}(t)), \end{aligned} \tag{10}$$

where the one-step ahead predicted output $\hat{y}_{f,[1:n_y]}(t)$ is obtained from (7).

Data-driven innovation estimation: Due to the stability of matrix $\bar{A}$, we assume throughout this paper that there exists a $\rho \in \mathbb{N}^*$ such that $\bar{A}^\rho = 0$ [20]. In this case, the VARX model can be formulated as:

$$\mathcal{H}_1(y^d_{(1:N)}) = \begin{bmatrix} \Phi_y & \Phi_u & D \end{bmatrix} \underbrace{\begin{bmatrix} \mathcal{H}_\rho(y^d_{(1-\rho:N-1)}) \\ \mathcal{H}_\rho(u^d_{(1-\rho:N-1)}) \\ \mathcal{H}_1(u^d_{(1:N)}) \end{bmatrix}}_{\Gamma} + \mathcal{H}_1(e^d_{(1:N)}), \tag{11}$$

where $\Phi_y \triangleq C\begin{bmatrix} \bar{A}^{\rho-1}K & \cdots & \bar{A}K & K \end{bmatrix}$, and $\Phi_u \triangleq C\begin{bmatrix} \bar{A}^{\rho-1}\bar{B} & \cdots & \bar{A}\bar{B} & \bar{B} \end{bmatrix}$ with $\bar{B} = B - KD$. An approximation of innovations $\mathcal{H}_1(e^d_{(1:N)})$, namely $\hat{e}^d_{(1:N)}$, can be obtained as the residuals of the following least-squares regression:

$$\min_{\Phi_y,\Phi_u,D} \left\| \mathcal{H}_1(y^d_{(1:N)}) - \begin{bmatrix} \Phi_y & \Phi_u & D \end{bmatrix} \Gamma \right\|_F^2. \tag{12}$$

III. Adaptive DDPC based on innovations

In this section, we first design a data-driven method for effectively detecting changes in system dynamics, with a key focus on innovations and their statistical properties. Then, the data-driven OP is updated online to adapt to new system dynamics once they are detected.

A. Data-driven detection of changes in system dynamics

Since the data-driven OP (7b) is equal to the KF-like predictor (8) with the true value of $e^d_{(1:N)}$, we have the following asymptotic results regarding the online innovations.

Theorem 1: If the true value of $\{e^d(k)\}_{k=1}^N$ is used in the innovation-based DDPC (7) and the initial online innovation $e_p(0)$ is obtained by solving the following optimization problem

$$\begin{aligned} \min_{e_p(0),g} \quad & \|e_p(0)\|_2^2 \\ \text{s.t.} \quad & \begin{bmatrix} U_p \\ Y_p \\ E_p \end{bmatrix} g = \begin{bmatrix} u_p(0) \\ y_p(0) \\ e_p(0) \end{bmatrix}, \end{aligned} \tag{13}$$

then the online innovations $e(t) = y(t) - \hat{y}_{f,[1:n_y]}(t)$ with $\hat{y}_{f,[1:n_y]}(t)$ from (7) are white and satisfy the asymptotic properties $\lim_{t\to\infty} \mathbb{E}[e(t)] = 0$ and $\lim_{t\to\infty} \mathbb{E}[e(t)e(t)^\top] = \Sigma_e$.

Proof: Under the conditions in Theorem 1 and given $L_p \geq n$, the initial trajectory $\{u_p(0), y_p(0), e_p(0)\}$ uniquely corresponds to a specific initial state $x(0)$. The data-driven OP (7b) produces identical output prediction to the KF-like predictor (8) with the same initial state. Thus, the online one-step ahead prediction error $e(t)$ from (7b) coincides with the innovation in the model-based SSKF (9). As a consequence, the innovation $e(t)$ from the innovation-based DDPC is a Gaussian distributed white noise process and satisfies the asymptotic properties [21, Theorem 5.1]. ■

Theorem 1 indicates that innovations may play a central role in the online detection of changes in system dynamics. Specifically, when system dynamics remain unchanged, the online innovation $e(t)$ derived from the innovation-based DDPC method (7) is expected to ideally follow the Gaussian distribution $\mathcal{N}(0, \Sigma_e)$. When system dynamics change, the distribution of innovations also shifts. Henceforth, the changes in system dynamics can be statistically detected by monitoring whether the online innovations adhere to the normal distribution $\mathcal{N}(0, \Sigma_e)$. This motivates a hypothesis test procedure for detection purposes. Even if innovation estimates are used in DDPC and the resulting OP no longer coincides with KF, this idea still holds promise since the innovation-based OP is known to closely resemble KF [14].

The key to performing the hypothesis test is to estimate the mean and covariance of innovations from $\hat{e}^d_{(1:N)}$:

$$\hat{\mu}_e = \frac{1}{N}\sum_{k=1}^{N}\hat{e}^d(k),\ \hat{\Sigma}_e = \frac{1}{N}\sum_{k=1}^{N}[\hat{e}^d(k)-\hat{\mu}_e][\hat{e}^d(k)-\hat{\mu}_e]^\top, \tag{14}$$

whose rationale is as follows.

Theorem 2: If the estimated parameters $\{\hat{\Phi}_y, \hat{\Phi}_u, \hat{D}\}$ are derived by solving (12) and the estimated mean and covariance $\{\hat{\mu}_e, \hat{\Sigma}_e\}$ of the innovations are given by (14), then the following statements hold.

(a) The estimated parameters $\{\hat{\Phi}_y, \hat{\Phi}_u, \hat{D}\}$ are unbiased, i.e., $\mathbb{E}\left(\begin{bmatrix}\hat{\Phi}_y & \hat{\Phi}_u & \hat{D}\end{bmatrix}\right) = \begin{bmatrix}\Phi_y & \Phi_u & D\end{bmatrix}$.

(b) The estimates $\{\hat{\Phi}_y, \hat{\Phi}_u, \hat{D}\}$ are consistent, i.e., $\{\hat{\Phi}_y, \hat{\Phi}_u, \hat{D}\} \rightarrow \{\Phi_y, \Phi_u, D\}$ with probability 1 for $N \rightarrow \infty$.

(c) The estimated mean and covariance of innovations, satisfies $\lim_{N\rightarrow\infty}\hat{\mu}_e = 0$ and $\lim_{N\rightarrow\infty}\hat{\Sigma}_e = \Sigma_e$ with probability 1.

Proof: The true innovation sequence $\{e^d(k)\}_{k=1}^N$ is a Gaussian distributed white noise process and is statistically independent from the offline data Γ in (11). Thus, the conclusions (a) and (b) naturally follow from [22]. The statement (c) can be directly derived from the conclusions (a) and (b). ■

According to Theorem 2(c), it is reasonable to use the innovation estimations $\{\hat{e}^d(k)\}_{k=1}^N$, obtained as the residuals of (12), to estimate $\{\hat{\mu}_e, \hat{\Sigma}_e\}$ through (14). As the size N of offline data in (12) increases, the estimated mean and covariance $\{\hat{\mu}_e, \hat{\Sigma}_e\}$ become closer to the true values $\{0, \Sigma_e\}$. The estimation procedure related to the innovations can be summarized in Algorithm 1.

Algorithm 1 Estimation of innovations and their statistics

Input: Input-output data $\{u^d(k), y^d(k)\}_{k=-\rho}^N$

Output: Innovation estimates $\hat{e}^d_{(1:N)}$, Hankel matrices $\{U_d, Y_d, \hat{E}_d\}$, mean and covariance estimates $\{\hat{\mu}_e, \hat{\Sigma}_e\}$

1: Estimate $\hat{e}^{\mathrm{d}}_{(1:N)}$ by (12);

2: Construct $\{U_d, Y_d, \hat{E}_d\}$ from $\{u^d(i), y^d(i), \hat{e}^d(i)\}_{i=1}^N$, estimate $\hat{\mu}_e$ and $\hat{\Sigma}_e$ by (14) from $\hat{e}^d_{(1:N)}$;

To improve the robustness of the detection scheme, we carry out the normality test on the sample innovation segment instead of a single point. The Mardia's kurtosis test is used since the test is known to be consistent for all alternative distributions with a nonsingular covariance matrix [23]. The null hypothesis H_0 and the alternative hypothesis H_1 at time t can be formulated using the empirical estimates in (14):

$$\begin{aligned} &H_0 : \hat{e}_{(t-N_w+1:t)} \sim \mathcal{N}(\hat{\mu}_e, \hat{\Sigma}_e), \\ &H_1 : \hat{e}_{(t-N_w+1:t)} \nsim \mathcal{N}(\hat{\mu}_e, \hat{\Sigma}_e), \end{aligned} \tag{15}$$

where N_w denotes the window length of past data. The test statistic is calculated based on the measure of multivariate kurtosis [24]:

$$b_k = \frac{1}{N_w}\sum_{k=-N_w+1}^{0}[\hat{e}(t+k)-\hat{\mu}_e]^\top\hat{\Sigma}_e^{-1}[\hat{e}(t+k)-\hat{\mu}_e], \tag{16}$$

which is based on the Mahalanobis distance. Specifically, the test statistic b_k bears the following asymptotic property [24]:

$$\beta_k \triangleq \frac{b_k - n_y(n_y+2)}{\sqrt{8n_y(n_y+2)/N_w}} \sim \mathcal{N}(0,1),\ N_w \rightarrow \infty. \tag{17}$$

Based on this, the p-value can be calculated as $p = 2(1-\Phi(\beta_k))$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. When the calculated p-value is less than a predetermined significance level α (e.g., 0.05), the null hypothesis is rejected, indicating that the changes in system dynamics have occurred.

B. Adaptive DDPC scheme

By applying the proposed data-driven detection scheme to adapt to underlying changes in system dynamics, we arrive at a new adaptive DDPC scheme based on innovations summarized in Algorithm 2. Upon detecting the changes in system dynamics at a certain time instance t_d, the adaptation is then triggered and then data collection is initiated. To ensure informativity of input-output data collected online, it is suggested to use a new output reference sequence of length $N+\rho$ that is sufficiently exciting. In this way, the input-output data $\{u(t_d+k), y(t_d+k)\}_{k=1}^{N+\rho}$ capturing the new system dynamics are collected online and then used to construct a new data-driven OP as a substitution of the old one. Specifically, the dictionary is reframed using $\{u(t_d+k), y(t_d+k)\}_{k=\rho+1}^{N+\rho}$, and their corresponding innovations $\{\hat{e}(t_d+k)\}_{k=\rho+1}^{N+\rho}$ shall be also estimated and used to initialize $\hat{e}_p(t_d+N+\rho)$ in DDPC. Meanwhile, the statistics of the innovations, namely $\{\hat{\mu}_e, \hat{\Sigma}_e\}$, shall also be updated for the subsequent implementation of the detection scheme.

Algorithm 2 Adaptive DDPC based on innovations

Input: $N_w \in \mathbb{N}^*$, $\alpha \in [0,1]$

1: Collect data $\{u^d(i), y^d(i)\}_{i=-\rho}^N$, obtain $\{U_d, Y_d, \hat{E}_d\}$ and $\{\hat{\mu}_e, \hat{\Sigma}_e\}$ from Algorithm 1;

2: Initialize $\hat{e}_p(0)$ by (13) based on $\{u(k), y(k)\}_{k=-L_p}^0$;

3: Obtain $u_f^*(t)$ and $\hat{y}_f^*(t)$ from (7) and apply $u^*_{f,[1:n_u]}(t)$ into the plant;

4: Collect $y(t)$ and calculate p-value by (16), (17) and $p = 2(1-\Phi(\beta_k))$;

5: If $p < \alpha$ at time instance t_d, implement a new output reference sequence and collect an extra input-output trajectory $\{u(t_d+k), y(t_d+k)\}_{k=1}^{N+\rho}$. Then obtain new $\{U_d, Y_d, \hat{E}_d\}$, $\{\hat{\mu}_e, \hat{\Sigma}_e\}$ and $\{\hat{e}(t_d+k)\}_{k=\rho+1}^{N+\rho}$ from Algorithm 1, and replace the online past innovations with $\{\hat{e}(t_d+k)\}_{k=\rho+1}^{N+\rho}$;

6: Update $\hat{e}_p(t+1)$ based on (10) and update $\{u_p(t+1), y_p(t+1)\}$;

7: $t \leftarrow t+1$, go to Step 2.

IV. Simulation Example

We apply the proposed adaptive DDPC to the control of the longitudinal dynamics of the Boeing 747 aircraft, which is a benchmark in the predictive control literature [25]. The perturbation equations for the aircraft cruising at an altitude of 40,000 ft and a horizontal velocity of $V = 774\,\text{ft/s}$ can be represented in a continuous 4-state system as follows [26]:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + B_w \begin{bmatrix} w_h(t) \\ w_v(t) \end{bmatrix} \\ \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} &= Cx(t) + v(t), \end{aligned} \tag{18}$$

where

$$\begin{aligned} A &= \begin{bmatrix} -0.003 & 0.039 & 0 & -0.322 \\ -0.065 & -0.319 & 7.74 & 0 \\ 0.02 & -0.101 & -0.429 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\ B &= \begin{bmatrix} 0.010 & 1 \\ -0.18 & -0.04 \\ -1.16 & 0.598 \\ 0 & 0 \end{bmatrix}, B_w = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \\ C &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 7.74 \end{bmatrix}, \end{aligned} \tag{19}$$

$w_h \in \mathbb{R}$ and $w_v \in \mathbb{R}$ denote the horizontal and vertical wind velocity disturbances, and $v \in \mathbb{R}^2$ are the output measurement noise. The system (18) has two inputs, i.e., the throttle u_1 and the elevator angle (u_2, deg), and two outputs, i.e., the longitudinal velocity (y_1, ft/s) and climb rate (y_2, ft/s). The process disturbance w_h and w_v can be modeled as the following transfer functions [27]:

$$\begin{aligned} w_h &= \sigma_u \sqrt{2L_u/(V\pi)} \frac{1}{1 + (L_u/V)s} w_1, \\ w_v &= \sigma_v \sqrt{2L_v/(V\pi)} \frac{1 + (2\sqrt{3}L_v/V)s}{[1 + 2(L_v/V)s]^2} w_2, \end{aligned} \tag{20}$$

where the intensities are chosen as $\sigma_u = 10$ and $\sigma_v = 10$, and the turbulence length scales are $L_u = 1750\,\text{ft}$ and $L_v = L_u/2$.

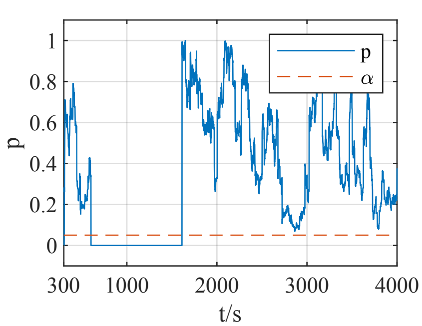


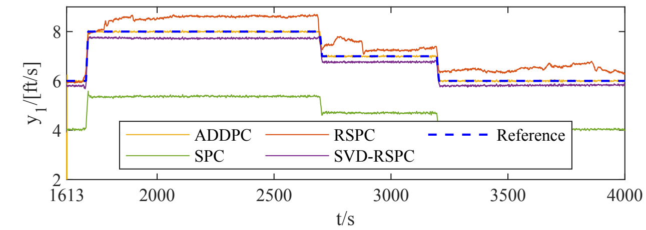
Fig. 1. Online p-value from the proposed data-driven detection method with significance level 0.025.

In the simulations, the continuous system (18) is discretized with a zero-order holder for a sampling time $t_s = 1\,\text{s}$. The process disturbance $w = \text{col}(w_1, w_2)$ and the measurement noise v are zero-mean white Gaussian distributed noise with covariances $\Sigma_w^2 = 1 \times 10^{-5} I_2$ and $\Sigma_v^2 = 1 \times 10^{-6} I_2$. For offline data collection, a pseudo-random binary signal with amplitude levels of $[-1, 1]$ is generated as the offline input with the data length $N = 250$. The past horizon L_p is chosen as 10 and the future L_f as 15. The cost weighting matrices are set as $Q = I_2$ and $R = 0.01 I_2$. Input and output constraint sets are set as $\mathbb{U} = \mathbb{R}^{n_u L_f}$ and $\mathbb{Y} = \mathbb{R}^{n_y L_f}$. For a comprehensive comparison, the following control strategies are implemented in a receding horizon fashion.

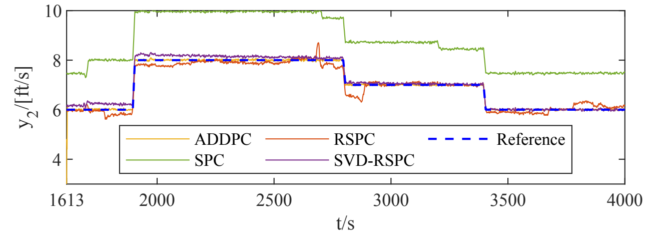
- **ADDPC**: The proposed Adaptive DDPC in Algorithm 2, where $N_w = 300$ and $\alpha = 0.025$. Moreover, the new output reference used in online data collection is designed as a pseudo-random binary signal with amplitude levels of $[-3, 3]$.
- **SPC** [28]: The well-known subspace predictive control (SPC) method without adaptation.
- **RSPC** [18]: The generic SPC with the adaptive scheme based on (5).
- **SVD-RSPC** [17]: The generic SPC with the adaptive scheme based on (4), where the fast SVD algorithm [17, Algorithm 2] is utilized to recursively update the trajectory dictionary.

In the ADDPC, the sample length N_w is chosen as 300 to alleviate the effect of outliers and improve the robustness of the plant-mismatch monitoring scheme.

The offline data is collected at the horizontal velocity of $V = 774\,\text{ft/s}$. In the online experiment, the horizontal velocity varies from $774\,\text{ft/s}$ to $800\,\text{ft/s}$ at time instance $t = 600\,\text{s}$ to simulate the change in operating point. This induces the changes in system matrices $\{A, B, C, D\}$ of (18) and noise statistics of $\{w_h, w_v\}$. We carry out a simulation



(a) Longitudinal velocity y_1



(b) Climb rate y_2

Fig. 2. Performance of the different control algorithms after the update of data-driven OP in the ADDPC.

TABLE I

CONTROL PERFORMANCE METRICS OF DIFFERENT CONTROL ALGORITHMS AFTER THE UPDATE OF DATA-DRIVEN OP IN THE ADDPC

	$\mathbb{E}\{\|y_1(t)-r_1(t)\|^2\}$	$\mathbb{E}\{\|y_2(t)-r_2(t)\|^2\}$	$\mathbb{E}\{\|u(t)\|_2^2\}$
ADDPC	0.001	0.002	5.482
SPC	5.441	3.051	3.072
RSPC	0.254	0.028	6.118
SVD-RSPC	0.055	0.017	5.196

to investigate the efficiency of the data-driven detection scheme for changes in system dynamics and compare the control performance of different control algorithms. Firstly, the results related to the detection method are presented in Fig. 1. The p-value is obtained after $t = 300\,\mathrm{s}$ since sufficient samples of innovations are needed to calculate β_k. The obtained p-value is less than α at $t = 602\,\mathrm{s}$ due to the change in the operating point. Then, new input-output data with length 1010 is collected to construct the new Hankel matrices in the interval of $[602\,\mathrm{s}, 1612\,\mathrm{s}]$. The p-value is larger than α from $t = 1613\,\mathrm{s}$ owing to the implementation of the appropriate data-driven OP. We can observe that the proposed data-driven detection scheme can effectively detect the changes in system dynamics as well as maintain stable detection when the system dynamics remain constant.

The results of control performance after $1613\,\mathrm{s}$, namely when the data-driven OP in ADDPC is updated, are presented in Fig. 2 and Table I. Significant biases can be observed in the outputs of SPC. In contrast, both RSPC and SVD-RSPC outperform SPC owing to the benefits of the recursive updates to the Hankel matrices. Notably, the proposed ADDPC method achieves the best control performance among the four methods.

V. CONCLUSIONS

In this paper, we proposed a new adaptive DDPC framework based on innovations that can adapt to changes in system dynamics. By harnessing the (approximate) whiteness and Gaussianity of innovations, a data-driven detection scheme was proposed, which enables effective detection of changes in system dynamics. This gave rise to an adaptive DDPC scheme, where the new system dynamics can be accommodated by updating the data-driven OP. Numerical simulations showcased the usefulness of the proposed data-driven detection scheme and the superior control performance of the proposed adaptive DDPC over the generic DDPC methods.

REFERENCES

[1] F. Borrelli, A. Bemporad, and M. Morari, *Predictive Control for Linear and Hybrid Systems*. Cambridge University Press, 2017.

[2] Z.-S. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Information Sciences*, vol. 235, pp. 3–35, 2013.

[3] C. Shang and F. You, "Data analytics and machine learning for smart process manufacturing: Recent advances and perspectives in the big data era," *Engineering*, vol. 5, no. 6, pp. 1010–1016, Dec. 2019.

[4] 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.

[5] V. Breschi, A. Chiuso, and S. Formentin, "Data-driven predictive control in a stochastic setting: a unified framework," *Automatica*, vol. 152, p. 110961, 2023.

[6] 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.

[7] I. Markovsky and F. Dörfler, "Data-driven dynamic interpolation and approximation," *Automatica*, vol. 135, p. 110008, 2022.

[8] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1702–1717, 2021.

[9] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct & indirect data-driven control formulations via regularizations and relaxations," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 883–897, Feb. 2023.

[10] J.-W. van Wingerden, S. P. Mulders, R. Dinkla, T. Oomen, and M. Verhaegen, "Data-enabled predictive control with instrumental variables: the direct equivalence with subspace predictive control," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 2111–2116.

[11] Y. Wang, Y. Qiu, M. Sader, D. Huang, and C. Shang, "Data-driven predictive control using closed-loop data: An instrumental variable approach," *IEEE Control Systems Letters*, vol. 7, pp. 3639–3644, 2023.

[12] J. Coulson, J. Lygeros, and F. Dörfler, "Distributionally robust chance constrained data-enabled predictive control," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3289–3304, 2021.

[13] L. Huang, J. Zhen, J. Lygeros, and F. Dörfler, "Robust data-enabled predictive control: Tractable formulations and performance guarantees," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 3163–3170, 2023.

[14] Y. Wang, K. You, D. Huang, and C. Shang, "Data-driven output prediction and control of stochastic systems: An innovation-based approach," *Automatica*, vol. 171, p. 111897, 2025.

[15] C. Verhoek, H. S. Abbas, R. Tóth, and S. Haesaert, "Data-driven predictive control for linear parameter-varying systems," *IFAC-PapersOnLine*, vol. 54, no. 8, pp. 101–108, 2021.

[16] T. B. Hamdan, P. Coirault, G. Mercère, and T. Dairay, "Data enabled predictive control of LPV systems," *Control Engineering Practice*, vol. 149, p. 105969, 2024.

[17] J. Shi, Y. Lian, and C. N. Jones, "Efficient recursive data-enabled predictive control," in *2024 European Control Conference (ECC)*. IEEE, 2024, pp. 880–887.

[18] Y. Lian, J. Shi, M. Koch, and C. N. Jones, "Adaptive robust data-driven building control via bilevel reformulation: An experimental result," *IEEE Transactions on Control Systems Technology*, vol. 31, no. 6, pp. 2420–2436, 2023.

[19] B. Huang and R. Kadali, *Dynamic Modeling, Predictive Control and Performance Monitoring: A Data-Driven Subspace Approach*. Springer, 2008.

[20] 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.

[21] E. W. Kamen and J. K. Su, *Introduction to Optimal Estimation*. Springer Science & Business Media, 1999.

[22] K. J. Åström, "Lectures on the identification problem: The least squares method," 1968.

[23] N. Henze, "On Mardia's kurtosis test for multivariate normality," *Communications in Statistics - Theory and Methods*, vol. 23, no. 4, pp. 1031–1045, 1994.

[24] K. V. Mardia, "Measures of multivariate skewness and kurtosis with applications," *Biometrika*, vol. 57, no. 3, pp. 519–530, 1970.

[25] P. Verheijen, V. Breschi, and M. Lazar, "Handbook of linear data-driven predictive control: Theory, implementation and design," *Annual Reviews in Control*, vol. 56, p. 100914, 2023.

[26] A. E. Bryson, *Control of Spacecraft and Aircraft*. Princeton University Press, 1993, vol. 41.

[27] R. Smith, M. Abdalmoaty, and M. Yin, "Data-driven formulation of the Kalman filter and its application to predictive control," in *63rd IEEE Conference on Decision and Control (CDC 2024)*, 2024.

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