A performance monitoring algorithm for sustained optimal operation with economic MPC

Marco Vaccari[1] and Gabriele Pannocchia[1]

Abstract— This paper addresses the problem of performance monitoring for Economic Model Predictive Control (EMPC) in the presence of plant parameter changes. In order to cope with plant-model mismatch, we adopt a recently developed offset-free EMPC algorithm which requires the gradient of the plant input-output steady-state map. A subspace identification method is used in order to estimate this plant gradient from transient measurements. However, when the plant parameters change, this method may fail unless re-identification is performed. Hence, to start a new data collection for the identification an event-triggered mechanism is proposed, based on a suitable performance monitoring strategy. In this case this mechanism investigates a possible, more profitable, steady-state equilibrium and, if convenient, it re-identifies the plant gradient. The proposed monitoring technique is then successfully tested over an illustrative example of a chemical reactor.

I. INTRODUCTION

Monitoring of model predictive control (MPC) performance has been debated for the last two decades and essentially can be summarized to two basic problems: evaluation of the MPC objective function and model quality validation [1]. The usual logic steps for control performance monitoring are evaluating potential performances and related statistics, followed by designing diagnosis methodologies to avoid performance degradation [2]. Statistical analysis has been used to model validation and error estimation, via implementation of quality indexes. [3]. The prediction error has also been used as the primary variable for performance monitoring of MPC systems [4]. Pannocchia et al. [5] proposed a prediction error analysis exploiting the structure of a model augmented by an integrating disturbance. A subspace identification method is then used to discriminate between plant-model mismatch or erroneous disturbance estimation.

On the other hand, performance monitoring for economic MPC (EMPC) [6] it is still an open field. The EMPC performance assessment has been deeply studied in the last decade. It has been shown that, for certain combinations of system and cost functions, steady-state solution may be less economically profitable than oscillating ones [6], [7]. As for MPC, also EMPC performance can suffers from bad model quality causing plant-model mismatch. EMPC can reliably reach steady-state optimal performance exploiting a combination of the traditional offset-free tracking methodology [8] and modifier-adaptation (MA) ideas [9], as recently shown in [10], [11]. A strong drawback of this technique is

[1]Marco Vaccari and Gabriele Pannocchia are with Department of Civil and Industrial Engineering, University of Pisa, Largo Lazzarino 2, 56126 Pisa, Italy marco.vaccari@ing.unipi.it, gabriele.pannocchia@unipi.it

requiring knowledge of the plant gradient, hence, a system identification method to estimate its approximation has been recently proposed [11], [12]. Results have shown how the identified gradient, used into the proposed technique, compensates satisfactorily the plant-model mismatch.

The purpose of this work is to formulate a performance monitoring strategy able to overcome changes in process or cost functions parameters. To this aim, we recall the concept of *event-triggered identification*. An identification or estimation procedure is activated by a specific event happened, and this finds many applications in various fields (see e.g. [13], [14]). Specifically, we aim at applying an event-triggered method to test the quality of the estimated plant gradient. In particular, the data collection for the identification procedure shown in [11], [12] is executed only once. This does not guarantee a "good quality" plant gradient in case of plant parameter changes. Hence, a specific event-triggering mechanism promoting a new data collection for the gradient identification procedure can be a good monitoring technique.

The rest of this paper is organized as follows. A review of the problem definition is presented in Section II. The proposed method, with a detailed analysis and description of the event-triggered monitoring technique is presented in Section III. The methodology is tested over a case study in Section IV, where numerical results are reported and discussed. Finally, Section V summarizes the paper and gives possible future developments of this work.

II. PLANT, COST AND EMPC ALGORITHM

A. Plant and cost specification

The discrete-time dynamical system considered in this work has the following form:

$$\begin{aligned} x_p^+ &= F_\theta(x_p, u) \\ y &= H_\theta(x_p) \end{aligned} \tag{1}$$

in which $x_p \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$ are the plant state, control input and output at a given time, respectively, x_p^+ is the successor state, and the subscript θ denotes a vector an unknown process parameters, which may change over time. The plant output is measured at each time $k \in \mathbb{Z}$ and it is denoted by y_k. Functions $F_\theta : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $H_\theta : \mathbb{R}^n \to \mathbb{R}^p$ are supposed not precisely known but differentiable. The economic optimal equilibrium of system (1) is defined by:

$$(x_s^0, u_s^0, y_s^0) = \arg\min_{x,u,y} \ell_e(y, u) \qquad \text{s.t.} \tag{2a}$$

$$x = F_\theta(x, u) \tag{2b}$$

$$y = H_\theta(x) \tag{2c}$$

$$u_{\min} \le u \le u_{\max}, \quad y_{\min} \le y \le y_{\max} \tag{2d}$$

in which $u_{\min}, u_{\max} \in \mathbb{R}^m$ and $y_{\min}, y_{\max} \in \mathbb{R}^p$ are the bound vectors. Since (1) is poorly represented, (x_s^0, u_s^0, y_s^0) is unknown, albeit assumed unique and problem (2) feasible.

Assumption 1: The cost function $\ell_e(y, u) : \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}$ is assumed continuously differentiable.

B. Model and Offset-free augmented system

The process *model* used to design the EMPC algorithm is:

$$\begin{aligned} x^+ &= f(x, u) \\ y &= h(x) \end{aligned} \tag{3}$$

in which $x, x^+ \in \mathbb{R}^n$ denote the current and successor model states. The (known) functions $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^p$ are assumed to be differentiable.

To cope with plant-model mismatch, an offset-free algorithm is based on an augmented model [8]:

$$\begin{aligned} x^+ &= F(x, u, d) \\ d^+ &= d \\ y &= H(x, d) \end{aligned} \tag{4}$$

in which $d \in \mathbb{R}^{n_d}$ is the so-called *disturbance*. The functions $F : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{n_d} \to \mathbb{R}^n$ and $H : \mathbb{R}^n \times \mathbb{R}^{n_d} \to \mathbb{R}^p$ are assumed to be continuous and *consistent* with (3), i.e. $F(x, u, 0) = f(x, u)$ and $H(x, 0) = h(x)$.

Assumption 2: The augmented system (4) is observable.

C. State and disturbance estimation

Given the system (4), an observer is defined to estimate the augmented state $\xi = [x^T, d^T]^T$. The selected estimator is the Moving Horizon Estimator (MHE). At every sample time k, an optimization problem based on N_T past output data $\{y_{k-N_T+1}, \ldots, y_k\}$ gives the new ξ estimate.

Let us define $\boldsymbol{\xi} = \{\xi_0, \ldots, \xi_{N_T}\}$ as the augmented state sequence and $\mathbf{w} = \{w_0, \ldots, w_{N_T-1}\}$, $\mathbf{v} = \{v_0, \ldots, v_{N_T-1}\}$ for the augmented process and measurement noise sequences respectively, where $w_j = [w_j^{x\,T}, w_j^{d\,T}]^T$. The MHE problem to solve, at the generic time k, is:

$$(\boldsymbol{\xi}^*, \mathbf{w}^*, \mathbf{v}^*) = \arg\min_{\boldsymbol{\xi}, \mathbf{w}, \mathbf{v}} \Gamma_k(\gamma) + \sum_{j=0}^{N_T-1} \ell_{MHE}(w_j, v_j) \tag{5a}$$

subject to:

$$\gamma = \xi_0 - \bar{\xi}_{k-N_T+1} \tag{5b}$$

$$x_{j+1} = F(x_j, u_j, \hat{d}_j) + w_j^x \tag{5c}$$

$$d_{j+1} = d_j + w_j^d \tag{5d}$$

$$y_{j+k-N_T+1} = H(x_j, d_j) + v_j \tag{5e}$$

$$y_{\min} \le H(x_j, d_j) \le y_{\max} \tag{5f}$$

in which $\bar{\xi}_{k-N_T+1}$ represents the *a priori* estimate of ξ at the beginning of the horizon and N_T is the horizon length. The filtered estimates of $x(k)$ and $d(k)$ in (4), $\hat{x}_k$ and $\hat{d}_k$ are obtained from $\xi_{N_T-1}^*$. The term $\Gamma_k(\gamma)$ approximates the arrival cost of the full estimation problem before $k - N_T + 1$.

The cost function in (5a) is defined as follows:

$$\ell_{MHE}(w_j, v_j) = \|w_j^x\|_{Q_x^{-1}}^2 + \|w_j^d\|_{Q_d^{-1}}^2 + \|v_j\|_{R^{-1}}^2 \tag{6}$$

in which $Q_x \in \mathbb{R}^{n \times n}$, $Q_d \in \mathbb{R}^{n_d \times n_d}$, $R \in \mathbb{R}^{p \times p}$ represent the process and measurement noise covariance matrices.

D. Target calculation with MA technique

The offset-free EMPC algorithm solves the target problem:

$$(x_{s,k}^*, u_{s,k}^*, y_{s,k}^*) = \arg\min_{x, u, y} \ell_e(y, u) \qquad \text{s.t.} \tag{7a}$$

$$x = F(x, u, \hat{d}_k) \tag{7b}$$

$$y = H(x, \hat{d}_k) + (\lambda_{k-1}^G)^T (u - u_{s,k-1}^*) \tag{7c}$$

$$u_{\min} \le u \le u_{\max}, \quad y_{\min} \le y \le y_{\max} \tag{7d}$$

in which $u_{s,k-1}^*$ is the steady-state input target found at the previous sampling time $k-1$, and λ_{k-1}^G is the modifier matrix calculated at the previous iteration. We define $G : \mathbb{R}^{m+n_d} \to \mathbb{R}^p$ the model steady-state input-to-output map, i.e.:

$$\begin{cases} x_s = F(x_s, \hat{d}_k, u_s) \\ y_s = H(x_s, \hat{d}_k) \end{cases} \Rightarrow y_s = G(u_s, \hat{d}_k) \tag{8}$$

A steady-state input-output map $y_{p,s} = G_p(u_s)$ can be similarly defined for the actual plant (1). For the sake of clarity, the dependance of $G_p(\cdot)$ to the time-varying parameter θ is omitted albeit present. The filtering update for λ_{k-1}^G is:

$$\lambda_k^G = (1 - \alpha_s)\lambda_{k-1}^G + \alpha_s \left(\nabla_u G_p(u_{s,k}^*) - \nabla_u G(u_{s,k}^*, \hat{d}_k) \right) \tag{9}$$

in which $\alpha_s \in (0, 1]$, the operator $\nabla_u(\cdot)$ is the gradient of the considered function respect to the variable u and $\lambda_0^G = 0$.

E. Optimal control problem with MA technique

Let $\boldsymbol{x} = \{x_0, x_1, \ldots, x_N\}$ and $\boldsymbol{u} = \{u_0, u_1, \ldots, u_{N-1}\}$ be, respectively, the state and the input sequences. The Finite Horizon Optimal Control Problem (FHOCP) to solve is:

$$(\boldsymbol{x}_k^*, \boldsymbol{u}_k^*) = \arg\min_{\boldsymbol{x}, \boldsymbol{u}} \sum_{i=0}^{N-1} \ell_e(y_i, u_i) \qquad \text{s.t.} \tag{10a}$$

$$x_0 = \hat{x}_k, \tag{10b}$$

$$x_{i+1} = F(x_i, u_i, \hat{d}_k) \tag{10c}$$

$$y_i = H(x_i, \hat{d}_k) + (\lambda_{k-1}^G)^T (u_i - u_{s,k}^*) \tag{10d}$$

$$u_{\min} \le u_i \le u_{\max}, \quad y_{\min} \le y_i \le y_{\max} \tag{10e}$$

$$x_N = x_{s,k}^* \tag{10f}$$

Assuming problem (10) feasible, the associated receding horizon implementation is given by:

$$u_k = u_{0,k}^* \tag{11}$$

F. Plant gradient estimation

Since the steady-state plant-map gradient $\nabla_u G_p(\cdot)$ used in (9) is generally unknown, a gradient approximation $\nabla_u \tilde{G}_p(\cdot)$ is calculated via identification of a linear system. The identification algorithm needs a collection of sufficiently informative N_{id} data points. In such case, for $k \in [k_{id}, k_{id} + N_{id}]$, (11) is replaced with:

$$u_k = u_{0,k}^* + \sigma_k \tag{12}$$

in which $\sigma_k \in \mathbb{R}^m$ is a random vector. Feasibility problems are avoided by checking the *excited* input u_k and projecting

it, when necessary, to the nearest feasible point. During this window, an input sequence named $U_{id} \in \mathbb{R}^{m \times N_{id}}$ and an output sequence $Y_{id} \in \mathbb{R}^{p \times N_{id}}$ are collected and then used in the identification algorithm at time $k = k_{id} + N_{id}$. Given A, B and C the locally linearized state-space matrices, the identified system gain is $\nabla_u \tilde{G}_p(\cdot)^T = C(I - A)^{-1}B$. Both input and output sequences are centered with respect to the model steady-state triple evaluated with (7) corresponding to the first data in the sequences. For $k > k_{id} + N_{id}$ the original receding horizon law (11) is restored, U_{id} and Y_{id} slide over, and the identification is repeated every time that $\|U_{id}\| \geq \tau_{id}$ where τ_{id} is a threshold value. This helps preventing numerical instability for non-informative input sequences.

III. PROPOSED TECHNIQUE

Since the plant gradient approximation $\nabla_u \tilde{G}_p(\cdot)^T$ may differ from the true plant gradient, once an equilibrium value is reached, we cannot know if this value is the process optimum or not. In particular, we cannot evaluate how much the reached equilibrium is far from the actual optimal triple defined in (2). It is also desirable to understand if a new data collection for the system identification algorithm is required, because $\nabla_u \tilde{G}_p(\cdot)$ is outdated. Therefore, a methodology able to evaluate a neighborhood of the reached optimum and establish a direction in which the cost function can possibly be lower, is needed. In such case, the identification procedure has to be repeated to obtain a new gradient approximation.

A. KKT analysis with plant gradient error

Let us first rewrite the plant stationary problem (2) as:

$$\min_u \ell_e(G_p(u), u) \qquad \text{s.t.} \tag{13a}$$

$$u_{\min} \leq u \leq u_{\max} \tag{13b}$$

$$y_{\min} \leq G_p(u) \leq y_{\max} \tag{13c}$$

If u^* is a (local) solution to (13), there exist vectors $\pi_1^*, \pi_2^*, \pi_3^*, \pi_4^*$ satisfying the following first-order necessary optimality KKT conditions:

$$\Gamma_p(u^*, \pi_1^*, \pi_2^*, \pi_3^*, \pi_4^*) = 0 \tag{14a}$$

$$u^* - u_{\max} \leq 0 \tag{14b}$$

$$u_{\min} - u^* \leq 0 \tag{14c}$$

$$G_p(u^*) - y_{\max} \leq 0 \tag{14d}$$

$$y_{\min} - G_p(u^*) \leq 0 \tag{14e}$$

$$\pi_1^*, \pi_2^*, \pi_3^*, \pi_4^* \geq 0 \tag{14f}$$

$$(u^* - u_{\max})_j \pi_{1,j}^* = 0 \qquad j = 1, \ldots, m \tag{14g}$$

$$(u_{\min} - u^*)_j \pi_{2,j}^* = 0 \qquad j = 1, \ldots, m \tag{14h}$$

$$(G_p(u^*) - y_{\max})_j \pi_{3,j}^* = 0 \qquad j = 1, \ldots, p \tag{14i}$$

$$(y_{\min} - G_p(u^*))_j \pi_{4,j}^* = 0 \qquad j = 1, \ldots, p. \tag{14j}$$

in which

$$\Gamma_p(u^*, \pi_1^*, \pi_2^*, \pi_3^*, \pi_4^*) = \nabla_u \ell_e(u^*, G_p(u^*)) + \pi_1^* - \pi_2^* + \nabla_u G_p(u^*)\pi_3^* - \nabla_u G_p(u^*)\pi_4^*$$

and $\pi_1, \pi_2 \in \mathbb{R}^m$ and $\pi_3, \pi_4 \in \mathbb{R}^p$ are the multiplier vectors for the input bound (13b), output bound (13c) constraints.

As a KKT matching consequence, at the optimum u^* the active constraints, and the corresponding Lagrange multiplier, are the same for plant and model [9]. When there are m active constraints, then the first-order MA technique described in Section II-D becomes unnecessary since there are no degrees of freedom. On the other hand, when the number of active constraints is less than m, the MA formulation is still needed to attain the plant optimum value for the "unconstrained variables". For this reason, let us consider the vector of all the inequality constraints stacked vertically:

$$c(u) = \begin{bmatrix} u - u_{\max} \\ u_{\min} - u \\ G_p(u) - y_{\max} \\ y_{\min} - G_p(u) \end{bmatrix} \tag{15}$$

Then, the set of active inequality constraints is defined as:

$$\mathscr{A}(u) = \{j \in N_{in} \text{ such that: } c_j(u) = 0\}, \tag{16}$$

in which $N_{in} = \{1, \ldots, 2m + 2p\}$ and c_j is the j-th inequality constraint. It has to be noted that $\mathscr{A}(u)$ contains the components of c representing the active inequality constraints. For the sake of simplicity, let us consider the case where $\mathscr{A}(u^*) \equiv \emptyset$, that is the KKT problem (14) can be reduced only to the condition (14a) as follows:

$$\Gamma_p(u^*, \pi_1^*, \pi_2^*, \pi_3^*, \pi_4^*) = \nabla_u \ell_e(G_p(u), u)\,|_{u^*} = 0 \tag{17}$$

that can be expanded into:

$$\left[\mathrm{D}_u \ell_e(\cdot, u)|_{u^*} + \mathrm{D}_{G_p} \ell_e(G_p(\cdot), \cdot)\, \mathrm{D}_u G_p(u)|_{u^*} \right]^T = 0 \tag{18}$$

where the symbol $\mathrm{D}_x z(y, x)$ denotes the derivative (i.e. transpose of the gradient) of $z(\cdot)$ with respect to x. It has to be noted that the term $\mathrm{D}_u G_p(u)|_{u^*}$ is supposed to be unknown, but with the procedure described in Section II-F, an approximated calculation of this derivative, is available. Let us, for notation simplicity, define $K_p^{id} = \nabla_u \tilde{G}_p(\cdot)^T$. Hence, the true plant derivate at a generic point, can be rewritten as follows:

$$\mathrm{D}_u G_p(u) = K_p^{id} + \delta_{id} \tag{19}$$

in which $\delta_{id} \in \mathbb{R}^{p \times m}$ is the error associated with the identified gain matrix K_p^{id}. As detailed in Section II-F, the current $K_{p,k}^{id}$ value, is the last one calculated by the identification algorithm the last time it run. Hence, $K_{p,k}^{id}$ is for sure the plant gradient approximation used by the model to calculate $(\lambda_{k-1}^G)^T$ into (9) and, consequently, $u_{s,k}^*$. Given that, we can now inspect the neighborhood of the current steady-state solution $u_{s,k}^*$ to find a possible lower value of the cost function $\ell_e(\cdot)$. That is, we look for a Δu such that:

$$\ell_e^* = \ell_e(G_p(u_{s,k}^* + \Delta u), u_{s,k}^* + \Delta u^*) < \ell_e(G_p(u_{s,k}^*), u_{s,k}^*) \tag{20}$$

Δu is correlated to the error δ_{id} through a relation $\Delta u = \psi(\delta_{id})$. In order to establish this relation we define

$$\bar{\Gamma}_p(\bar{u}, \delta_{id}) = \left[\mathrm{D}_u \ell_e(\cdot, u)|_{\bar{u}} + \mathrm{D}_{G_p} \ell_e(G_p(\cdot), \cdot)\left(K_p^{id} + \delta_{id}\right) \right]^T \tag{21}$$

where $\bar{u} = u^*_{s,k} + \Delta u$, and we note that $\bar{\Gamma}_p(\bar{u}, \delta_{id}) = 0$ must be fulfilled. To this aim, a first order Taylor expansion of $\bar{\Gamma}_p(\bar{u}, \delta_{id})$ is performed, leading to:

$$\bar{\Gamma}_p(u^*_{s,k}, 0) + D_u\bar{\Gamma}_p(\cdot)\big|_{(u^*_{s,k}, 0)}\Delta u + D_\delta\bar{\Gamma}_p(\cdot)\big|_{(u^*_{s,k}, 0)}\delta_{id} = 0$$

From the KKT matching described in [10], the KKT problem (17) can be rewritten as:

$$\bar{\Gamma}_p(u^*_{s,k}, 0) = \left[D_u\ell_e(\cdot, u)\big|_{u^*_{s,k}} + D_{G_p}\ell_e(G_p(\cdot), \cdot)K^{id}_{p,k}\right]^T = 0 \tag{22}$$

Hence, the relation $\psi(\delta_{id})$ is calculated as follows:

$$D_u\bar{\Gamma}_p(\cdot)\big|_{(u^*_{s,k}, 0)}\Delta u = -D_\delta\bar{\Gamma}_p(\cdot)\big|_{(u^*_{s,k}, 0)}\delta_{id} \tag{23}$$

which defines $\Delta u = \psi(\delta_{id})$. A proper value for δ_{id}, and hence a suitable value for Δu^*, has to be calculated so that the active constraints do not change, i.e. $\mathscr{A}(u^*_{s,k} + \Delta u^*) \equiv \mathscr{A}(u^*_{s,k})$.

Remark 1: Testing the neighborhood of the current steady-state equilibrium, identified by $\ell_e(G_p(u^*_{s,k}), u^*_{s,k})$, to find a better steady-state equilibrium, identified by ℓ^*_e, any generic point in this search $\bar{u}$ must satisfy (14), i.e. the relation $\psi(\delta_{id})$ is defined only for steady-state solutions.

Remark 2: The restriction of inactive constraints, $\mathscr{A}(u^*_{s,k}) \equiv \emptyset$, can be overcome with a more general analysis of the KKT problem (14) when $\dim(\mathscr{A}(u^*_{s,k})) \leq m$, where dim is the dimension of active set, by restricting that the active set does not change, i.e. $\mathscr{A}(u^*_{s,k} + \Delta u^*) \equiv \mathscr{A}(u^*_{s,k})$.

B. Evaluating the sub-optimality of the current equilibrium

Recalling the convergence results obtained with the EMPC algorithm from [12] and the offset-free technique results, the steady-state values $G_p(u^*_{s,k})$ are known since they correspond to the augmented model ones. On the opposite hand, $G_p(u)$ cannot be evaluated in any generic point. For this reason another first order Taylor expansion has to be defined:

$$\bar{G}_p(u^*_{s,k}, \Delta u) = G_p(u^*_{s,k}) + K^{id}_{p,k}\Delta u \tag{24}$$

To assess the sub-optimality of the current equilibrium, $u^*_{s,k}$, the minimum cost ℓ^* is evaluated as a function of the plant gradient error δ_{id}, via the following optimization problem:

$$\ell^*_e = \min_{\delta_{id}, \Delta u} \ell_e(\bar{G}_p(u^*_{s,k}, \Delta u), u^*_{s,k} + \Delta u) \qquad \text{s.t.} \tag{25a}$$

$$\Delta u = \psi(\delta_{id}) \tag{25b}$$

$$u_{\min} + \varepsilon_u \leq u^*_{s,k} + \Delta u \leq u_{\max} - \varepsilon_u \tag{25c}$$

$$y_{\min} + \varepsilon_y \leq \bar{G}_p(u^*_{s,k}, \Delta u) \leq y_{\max} - \varepsilon_y \tag{25d}$$

$$-\delta_{\max} \leq \delta_{id} \leq \delta_{\max} \tag{25e}$$

in which ε_u and ε_y are threshold fixed values assuring that bounds constraints are conservatively satisfied.

The bound $\delta_{\max}$ represents the maximum error done evaluating $K^{id}_{p,k}$ with the identification. For instance, $\delta_{\max}$ can be calculated with the *asymptotic method* described by [15]. This method gives an upper bound on the difference between true transfer function and the identified one. For the sake of simplicity, instead, we can suppose that: $\delta_{\max} = a\left|K^{id}_{p,k}\right|$ where the symbol $\left|K^{id}_{p,k}\right|$ stands for the element-wise absolute
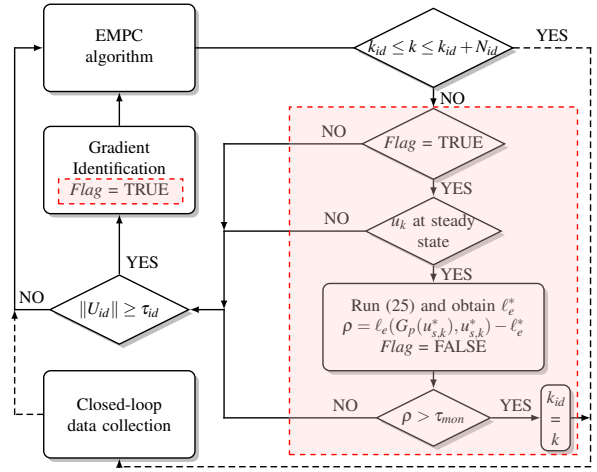

Fig. 1. Scheme of the event-triggering mechanism (inside the red areas) inserted into the offset-free EMPC algorithm.

value of $K^{id}_{p,k}$, that is each component of matrix $\delta_{\max}$ represents a relative error a on the corresponding $K^{id}_{p,k}$ component taken in its absolute value. Clearly, $\ell^*_e \leq \ell_e(G_p(u^*_{s,k}), u^*_{s,k})$, since $(\delta_{id}, \Delta u) = (0,0)$ is feasible. When calculated, the expected value of the cost function, ℓ^*_e, is compared with the actual one. Defining the sub-optimality performance index $\rho \geq 0$ in the following way:

$$\rho = \ell^*_e - \ell_e(G_p(u^*_{s,k}), u^*_{s,k}) \tag{26}$$

two cases are distinguished: if $\rho \geq \tau_{mon}$, where τ_{mon} is a threshold value, the data collection is repeated and the closed-loop control law (12) is enabled, otherwise the current steady-state value is near enough to the real optimum one.

C. Summary: the event-triggering mechanism

Figure 1 depicts the event-triggering mechanism inside the offset-free EMPC algorithm as defined in Section II.

Some notes have to be added to this figure. First of all, there are conditions for the event-triggering mechanism to be tested. The first one, represented by the boolean variable *Flag*, is that a plant gradient has been identified and it has to be tested (*Flag* =TRUE), that is the system identification has to be run at least once before testing the gradient quality, i.e. the first data collection and identification are imposed (see Section II-F). When *Flag* = FALSE no gradient has to be tested since no estimation is present or it has already been tested. Its initial value is *Flag* = FALSE. Then, the system has to be at steady-state, and, in order to avoid unnecessary computational costs, this steady-state has to be different from the one previously tested. When both conditions apply, the optimization problem (25) is run and quantity ρ calculated. In addition, the boolean variable *Flag* is set to its new values. Finally, the black dashed line represents the data collection window, hence, during this time period, neither the gradient identification nor the monitoring test can be applied. If the monitoring technique is not switched on, this path is followed only once, i.e. variable *Flag* is never updated to "TRUE".

TABLE I

DIFFERENT ACTUAL KINETIC CONSTANTS AND CORRESPONDING
CALCULATED ECONOMIC OPTIMUM.

Case	Kinetic constants min^{-1}		Economic optimum triple		
	k_1	k_2	u_s^0	$x_{1,s}^0$	$x_{2,s}^0$
1	1	0.05	1.043	0.511	0.467
2	0.7	0.035	0.831	0.509	0.471
3	1.25	0.071	1.179	0.512	0.463

IV. CASE STUDY

A. Process and optimal economic performance

The application chosen is a Continuous Stirred Tank Reactor (CSTR), where two consecutive reactions take place: $A \xrightarrow{k_1} B \xrightarrow{k_2} C$. The reactor is described by:

$$\begin{aligned}
\dot{x}_1 &= \frac{u}{V}(c_{A0} - x_1) - k_1 x_1 \\
\dot{x}_2 &= \frac{u}{V}(c_{B0} - x_2) + k_1 x_1 - k_2 x_2
\end{aligned} \quad (27)$$

in which x_1 and x_2 are the molar concentrations of A and B in the reactor, c_{A0} and c_{B0} are the corresponding concentrations in the feed, u is the feed flow rate regulated through a valve, V is the constant reactor volume, k_1 and k_2 are the kinetic constants. The reactor is assumed to be isothermal and the state measurables. The parameters of the actual system are taken from [12] and the process economics is expressed by:

$$\tilde{\ell}(u, x_2) = \beta_A u c_{A0} - \beta_B u x_2. \quad (28)$$

with β_A, β_B prices for the species A and B, respectively. A variation of the reactor temperature is supposed to occur and consequently the true actual kinetic constants are supposed to vary along with the economic optimum computed. The different values for kinetic constants and economic optimum are reported in Table I. Obviously, every optimum point in each case is supposed to be unknown since the controller is designed with wrong parameters as explained below.

B. Model and controllers

The designed controllers are affected by the same plant-model mismatch, that is $\bar{k}_1 = 1.2$ min^{-1} and $\bar{k}_2 = 0$ min^{-1}, meaning that side reaction is always not modeled. The same nominal model cost function and sampling time of $T_s = 1$ min are used to compare the EMPC algorithms designed according to Section II.

Remark 3: As described in [10] and shown in [11], [12], the economic cost function (28) must be integrated to be used in (10), since its point-wise evaluation would not lead to a dissipative system [7], i.e. the closed-loop system equilibrium would not be stable.

Hence, the dynamics (27) is discretized using an implicit Euler method and the state disturbance model is used, i.e. the considered augmented system (4) is $F(x, u) = f(x, u) + d$ and $H(x) = x$, while the integrated cost function is:

$$\ell(u, x_2) = (\beta_A u c_{A0} - \beta_B u x_2^+) T_s. \quad (29)$$

The estimator type distinguishes two formulations [16]: EMPC1 adopts the *filtering* updating for the arrival cost, whereas EMPC2 uses the *smoothing* arrival cost updating The estimator tuning for both the EMPC configurations is
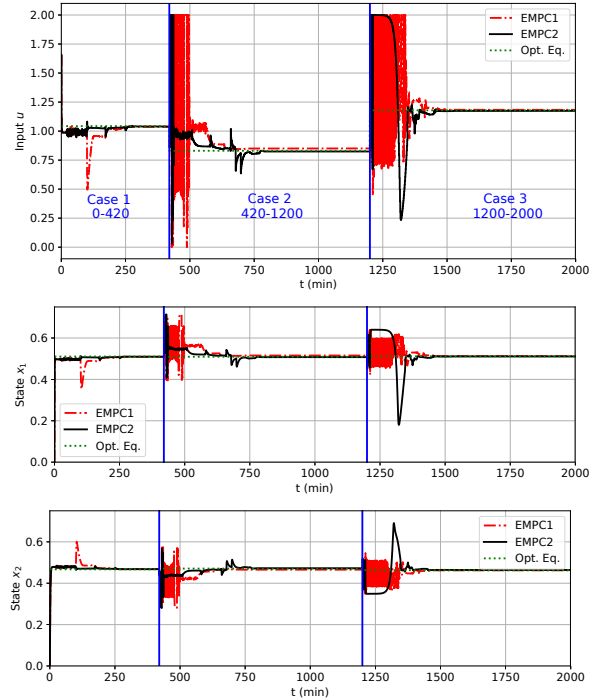

Fig. 2. Closed-loop results for the two selected controllers: input (top) and states (middle and bottom).

the following: $Q_x = 10^{-4}I$, $Q_d = I$, $R = 10^{-4}I$, $P_0 = I$. The identification algorithm adopted is the N4SID method using an open-source package[1]. The identification parameters are: $N_{id} = 80$, model order $n = 2$, future and past horizons $SS_f = SS_p = 10$ and the threshold is $\tau_{id} = 0.1$.

C. Monitoring parameters

As described in Section III, the relation $\Delta u = \psi(\delta_{id})$ is needed to solve the optimization problem (25). Considering the cost function (29) the various terms in (22) are:

- $D_u \ell_e(\cdot, u) = (\beta_A c_{A0} - \beta_B x_2) T_s$
- $D_{G_p} \ell_e(G_p(\cdot), \cdot) = \begin{bmatrix} 0 & -\beta_B u T_s \end{bmatrix}$

From (23) we obtain: $\Delta u = \psi(\delta_{id}) = -\dfrac{u_{s,k}^* \begin{bmatrix} 0 & 1 \end{bmatrix} \delta_{id}}{2 \begin{bmatrix} 0 & 1 \end{bmatrix} K_{p,k}^{id}}$. The parameters ε_u and ε_y are set to be equal to 5% of u_{max} and y_{max} respectively, whereas the uncertainty on plant gradient is $a = 1\%$. Finally, the threshold value for deciding if the expected cost (ℓ_e^*) is sufficiently lower than the actual one ($\ell_e(G_p(u_{s,k}^*), u_{s,k}^*)$) is: $\tau_{mon} = 0.02 |\ell_e(G_p(u_{s,k}^*), u_{s,k}^*)|$.

D. Results

The closed-loop behavior obtained with the above described controllers is depicted in Figure 2. It can be noticed how both algorithms accomplish the goal to reach the economic optimum even under plant-model mismatch and under process parameter changes. The monitoring techniques can be clearly seen in action two times for EMPC2 and one for EMPC1. Controller EMPC2 seems to give the best performance in terms of oscillations and being the one nearest to the economic optimum values. The second

[1]SIPPY (Systems Identification Package for PYthon) is available at GITHUB https://github.com/CPCLAB-UNIPI/SIPPY
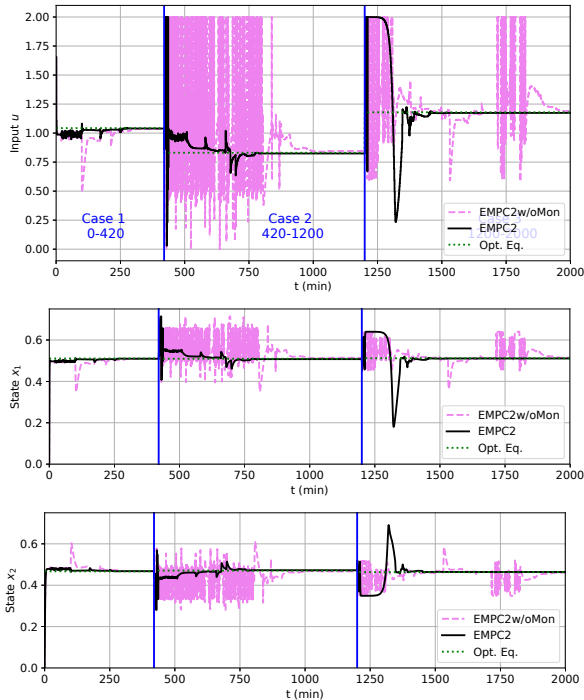
Fig. 3. Closed-loop results for the EMPC2 with and without monitoring techniques: input (top) and states (middle and bottom).

monitoring action is applied for EMPC1 at $k = 1210$ while the first one is at $k = 429$ for EMPC2, but only at $k = 487$ for EMPC1. This is because of the filtering nature of EMPC1, that makes the controller more prone to oscillate rather than EMPC2. Anyway, Figure 2 allows us to understand how the monitoring methodology works. When the system goes into oscillation due to the process parameter change, as soon as two consecutive input values are equals and the corresponding steady-state triple is into the feasible region, the data collection starts again. That is because $\rho \geq \tau_{mon}$ holds true, i.e. $\ell_e^* - \ell_e(G_p(u_{s,k}^*), u_{s,k}^*) \geq 0.02|\ell_e(G_p(u_{s,k}^*), u_{s,k}^*)|$.

Let us now define EMPC2w/oMon as the controller EMPC2 without the monitoring technique, i.e. the plant gradient identification will be updated only if the data window is sufficiently unsteady. The closed-loop behaviors for EMPC2w/oMon are reported in Figure 3. After the fist data collection and identification, as how it should be, both cases have a comparable behavior. Once the first parameter change occurs (*Case 2* of Table I), both of them begin to oscillate between the input bounds, but only EMPC2 has been able to recover thanks to the monitoring technique. The oscillation is due to an outdated plant gradient not able to approximate the real one with sufficient accuracy anymore. The monitoring technique is able to diagnose the problem and, hence, to start a new data collection to re-identify the gradient. The EMPC2w/oMon, instead, when entering *Case 2*, continues using the gradient calculated at *Case 1* with some updates generated by identification with weakly informative data, oscillating between two bounds.

V. CONCLUSIONS

In this work a performance monitoring technique for an economic model predictive control (EMPC) algorithm has

been proposed. This technique relies upon a previously defined EMPC implementation, that in presence of plant-model mismatch is able to asymptotically achieve the optimal performance [12]. When deemed appropriate, an event-triggered data collection followed by an identification is induced by the proposed monitoring algorithm to update the plant gradient approximation. This approximation is then applied to offset-free EMPC algorithm detailed in [12]. The monitoring technique has been described, detailing the event-triggering mechanism based on a KKT alteration analysis. We applied the monitoring methodology to a CSTR example, testing various state and disturbance estimation methods. Results show how the proposed methodology can overcome plant-model mismatch and plant parameter changes, converging to the best economic equilibrium. Results also show that the Moving Horizon Estimation (MHE) method with smoothing update is particularly efficient versus the filtering update MHE counterpart.

REFERENCES

[1] B. Huang and R. Kadali, *Dynamic modeling, predictive control and performance monitoring: a data-driven subspace approach.* Springer, 2008.

[2] T. J. Harris, C. Seppala, and L. Desborough, "A review of performance monitoring and assessment techniques for univariate and multivariate control systems," *Journal of Process Control*, vol. 9, no. 1, pp. 1–17, 1999.

[3] Z. Sun, S. J. Qin, A. Singhal, and L. Megan, "Performance monitoring of model-predictive controllers via model residual assessment," *Journal of Process Control*, vol. 23, no. 4, pp. 473–482, 2013.

[4] C. A. Harrison and S. J. Qin, "Discriminating between disturbance and process model mismatch in model predictive control," *Journal of Process Control*, vol. 19, no. 10, pp. 1610–1616, 2009.

[5] G. Pannocchia, A. De Luca, and M. Bottai, "Prediction error based performance monitoring, degradation diagnosis and remedies in offset-free MPC: Theory and applications," *Asian Journal of Control*, vol. 16, no. 4, pp. 995–1005, 2014.

[6] J. B. Rawlings, D. Angeli, and C. N. Bates, "Fundamentals of economic model predictive control," in *51st IEEE Conference on Decision and Control (CDC)*, 2012, pp. 3851–3861.

[7] D. Angeli, R. Amrit, and J. B. Rawlings, "On average performance and stability of economic model predictive control," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1615–1626, 2012.

[8] G. Pannocchia and J. B. Rawlings, "Disturbance models for offset-free model-predictive control," *AIChE Journal*, vol. 49, no. 2, pp. 426–437, 2003.

[9] A. G. Marchetti, B. Chachuat, and D. Bonvin, "Modifier-adaptation methodology for real-time optimization," *Industrial & Engineering Chemistry Research*, vol. 48, no. 13, pp. 6022–6033, 2009.

[10] M. Vaccari and G. Pannocchia, "A modifier-adaptation strategy towards offset-free economic MPC," *Processes*, vol. 5, no. 1, p. 2, 2016.

[11] G. Pannocchia, "An economic MPC formulation with offset-free asymptotic performance," in *10th IFAC Symposium on Advanced Control of Chemical Processes (ADCHEM)*, 2018, pp. 387–392.

[12] M. Vaccari and G. Pannocchia, "Implementation of an economic MPC with robustly optimal steady-state behavior," in *6th IFAC Conference on Nonlinear Model Predictive Control (NMPC)*, 2018, pp. 94–99.

[13] A. J. Isaksson, A. Horch, and G. Dumont, "Event-triggered deadtime estimation from closed-loop data," in *American Control Conference, 2001. Proceedings of the 2001*, vol. 4. IEEE, 2001, pp. 3280–3285.

[14] S. Carter, M. Moeglein, J. D. DeLoach Jr, W. T. Riley, D. H. Agre, and L. Sheynblat, "Event-triggered data collection," June 20 2006, uS Patent 7,065,351.

[15] Y. Zhu, *Multivariable system identification for process control.* Elsevier, 2001.

[16] C. V. Rao, J. B. Rawlings, and J. H. Lee, "Constrained linear state estimation– a moving horizon approach," *Automatica*, vol. 37, no. 10, pp. 1619–1628, 2001.