LEARNING MPC FOR REPETITIVE OPERATION PROCESSES

A. H. GONZÁLEZ[1], E. J. ADAM[1], D. ODLOAK[2] and J. L. MARCHETTI[1]

[1] Institute of Technological Development for the Chemical Industry (INTEC), CONICET - Universidad Nacional del Litoral, Santa Fe – Argentina.
[2] Department of Chemical Engineering, Polytechnic School, University of São Paulo, São Paulo – SP, Brazil.

ABSTRACT- Several industrial systems work by executing a sequence of similar or almost identical finite-time operations, which allows learning from previous runs to improve control performance. The operation of batch and semi-batch reactors are typical examples of these systems in the chemical industry. In this work, a model predictive controller (MPC) based on the concept of shrinking horizon is proposed to deal with finite-time operations. In addition, the proposed controller includes a learning mechanism leading to improve the performance as the sequence of operations progresses. When a control strategy is developed for these systems, two convergence analyses are necessary: the convergence of the closed-loop responses during the time-interval the single operation lasts and the convergence from operation to operation, i.e., the convergence of the learning mechanism that provides high performance when the number of trials increases. All these features found with the proposed MPC are illustrated by simulating a case study.

KEYWORDS: Predictive Control, Finite Time Operations, Learning procedures

1. INTRODUCTION

Repetitive dynamic system can be found in several industrial fields such as robot manipulation [1], injection molding [2], batch processes [3-4] and semiconductor processes [5]. The operation of repetitive dynamic systems frequently presents two well defined control objectives: one addresses the final target condition which is associated to the product quality; the other refers to economic aspects of the operation and is mainly defined through a recipe or output variable trajectory to be accomplished. Because of the repetitive characteristic, we analyze these systems using two count indexes or time scales: one is for the time running within the interval the operation lasts, and the other is associated to the number of operations in the sequence [6-7]. Consequently, a control strategy for repetitive systems requires accounting for two different objectives: the first one is the disturbance rejection during a single operation in the sequence; this frequently implies the tracking of a predetermined trajectory. The other one is the control of the sequence of operations, which recognizes disturbances or model mismatches that remain from one operation to the next. The important property pursued when designing this control system is the ability to use the information from previous operations to improving the tracking performance as the sequence progresses.

ISSN 2178-3659

The paper is organized as follows. In Section 2 the single operation problem is presented, and the shrinking horizon MPC formulation is presented. Section 3 describes how the sequence of operation problems is accounted by the proposed strategy. Then, the learning properties of the strategy (to control a operation sequence) is illustrated by a simulation example in Section 4 and the conclusions are presented in Sections 5.

2. THE SINGLE OPERATION PROBLEM

The augmented dynamic model used to forecast the output variables is given by:

$$x_{k+1} = Ax_k + Bu_k \tag{1}$$

$$d_{k+1} = d_k \tag{2}$$

$$y_k = Cx_k + d_k \tag{3}$$

where A, B and C are matrices of appropriate dimension, x are the standard states and the augmented state d_k is an estimated disturbance included for updating the output predictions Cx_k and providing offset elimination [8-9]. The sub index k stands for the real time during a single operation, and since every operation is assumed to last T_f time instants, it goes from 0 to T_f .

Let us assume that the objective is to follow an output reference trajectory, y_k^r , and to do that an input reference candidate, u_k^r , is available. Given that the control algorithm is executed along the time-interval the operation lasts only, a shrinking output horizon is defined as the difference between the current time k and the final time T_f, that is, $H(k) := T_f - k$. Then, the optimization problem to be solved at time k, as part of the single operation, is as follows:

Problem P1

$$V_k = \min_{\left\{\bar{u}_{k|k},\cdots,\bar{u}_{k+N_s-1|k}\right\}} \left\{ \sum_{j=0}^{H-1} \ell\left(e_{k+j|k}, \bar{u}_{k+j|k}\right) + F\left(e_{T_f|k}\right) \right\}$$

subject to:

$$e_{k+j|k} = Cx_{k+j|k} + d_{k+j} - y_{k+j}^r, \quad j = 0,\cdots,H , \tag{4}$$

$$x_{k+j+1|k} = Ax_{k+j|k} + Bu_{k+j|k}, \quad j = 0,\cdots,H-1, \tag{5}$$

$$d_{k+j} = \hat{d}_k, \qquad j = 0,\cdots,H \qquad ..(6)$$

$$u_{k+j|k} \in U, \quad j = 0,1,\cdots,H-1, \tag{7}$$

$$u_{k+j|k} = u_{k+j}^r + \bar{u}_{k+j|k}, \quad j = 0,1,\cdots,H-1 , \tag{8}$$

$$\bar{u}_{k+j|k} = 0, \quad j \geq N_s , \tag{9}$$

where $\ell(e,u) := \|e\|_Q^2 + \|u\|_R^2$ is the step cost, $F\left(e_{T_f|k}\right) := \left\|Cx_{T_f|k} + d_{T_f} - y_{T_f}^r\right\|_P^2 = \left\|e_{T_f}\right\|_P^2$ is a final cost associated to the final targeting error, and U = {u : $u_{min} \leq u \leq u_{max}$; $|\Delta u| \leq \Delta u_{\max}$ }. In this formulation, N_s is the control horizon, which is equal to N most of the time but, over the end, it becomes a shrinking control horizon given by $N_s = min(H,N)$. The matrices Q, R and P are such that Q, P>0 and R≥0; the current state and disturbance are estimated by an observer, then, $x_{k|k} = \hat{x}_k$, and $\hat{d}_k$ is used as indicated in (6). Because this formulation contains some new concepts, a few remarks are needed to clarify the key points: i) in any single operation, T_f optimization problems must be solved (from k=0 to k=T_f-1). Each problem gives a sequence of optimal input values $u_{k+j|k}^{opt}$, for j=0,…,H-1, but only the first one in the sequence, $u_{k|k}^{opt}$, is applied to the system. Notice this is not an application of the typical receding horizon policy since the end point remains fix during every single operation. ii) The decision variables $\bar{u}_{k+j|k}$ are corrections, to the open-loop reference sequence u^r_{k+j} (see Eq. (8)), for improving the closed-loop performance. Notice that because of constraint (9), $\bar{u}_{k+j|k}$ is different from zero only in the first N_s steps and therefore, Problem P1 has Ns decision variables.

Theorem 1 (single-operation cost boundeness): Using the control law resulting from the successive solutions to Problem P1 on the system (1)-(3), the obtained cost function is decreasing, that is, $V_k - V_{k-1} + \ell(e_{k-1}, \bar{u}_{k-1}) \le 0$, for $1 \le k \le T_f$-1.

Proof.

Let $\bar{\mathbf{u}}^{opt}_{k-1} := \left\{\bar{u}^{opt}_{k-1|k-1}, \cdots, \bar{u}^{opt}_{k+N_s-2|k-1}, 0, \cdots, 0\right\}$ be the optimal solution of Problem P1 at time k-1. At this time instant $H = T_f - k+1$ and the cost function can be written as

$$V_{k-1} = \sum_{j=0}^{H-1} \ell\left(e^{opt}_{k+j-1|k-1}, \bar{u}^{opt}_{k+j-1|k-1}\right) + F\left(e^{opt}_{T_f|k-1}\right) \quad (10)$$

Assume the input resultant from the first solution, $u_{k-1|k-1} = u^r_{k-1} + \bar{u}^{opt}_{k-1|k-1}$ is implemented, and shift the remaining solution values to compose a feasible solution $\bar{\mathbf{u}}^{feas}_k = \left\{\bar{u}^{opt}_{k|k-1}, \cdots, \bar{u}^{opt}_{k+N_s-2|k-1}, 0, \cdots, 0, 0\right\}$ to Problem P1 at time k. If there are no unknown disturbances, the cost at time k corresponding to the feasible solution $\bar{\mathbf{u}}^{feas}_k$ is as follows:

$$V^{feas}_k = \sum_{j=0}^{H-1} \ell\left(e^{opt}_{k+j|k-1}, \bar{u}^{opt}_{k+j|k-1}\right) + F\left(e^{opt}_{T_f|k-1}\right), \quad (11)$$

where now $H = T_f - k$. Subtracting (10) from (11), we have

$$V^{feas}_k - V_{k-1} = -\ell\left(e^{opt}_{k-1|k-1}, \bar{u}^{opt}_{k-1|k-1}\right). \quad (12)$$

Notice this equality determines the "natural" decreasing rate of the cost function due to a one-step shorter horizon H. However, the actual solution of Problem P1 at time k, $\bar{\mathbf{u}}^{opt}_k := \left\{\bar{u}^{opt}_{k|k}, \cdots, \bar{u}^{opt}_{k+N_s-1|k}, 0, \cdots, 0\right\}$ should be better than simply using the remaining of the feasible solution obtained at k-1. Then, if no new disturbance appears in the interval [k-1,k], the optimal cost at time k, should be lower or at least equal to the one obtained with the feasible solution. In other words,

$$V_k \le V^{feas}_k = V_{k-1} - \ell\left(e^{opt}_{k-1|k-1}, \bar{u}^{opt}_{k-1|k-1}\right). \quad (13)$$

This is because $\bar{\mathbf{u}}^{opt}_k$ provides one more degree of freedom than $\bar{\mathbf{u}}^{feas}_k$ (the first null value in $\bar{\mathbf{u}}^{feas}_k$ is substituted by $\bar{u}^{opt}_{k+N_s-1|k}$ in $\bar{\mathbf{u}}^{opt}_k$). Finally, notice that $e^{opt}_{k-1|k-1}$ and $\bar{u}^{opt}_{k-1|k-1}$ represent actual already known values. Thus, we may write

$$V_k - V_{k-1} + \ell\left(e_{k-1}, \bar{u}_{k-1}\right) \le 0 \qquad 1 \le k \le T_f \quad (14)$$

This shows that, if the output error is different from zero, the cost function with finite horizon decreases more than what is due by simply reducing the distance from k to T_f in one sampling time interval. Though this result ensures bounded behavior and supports the stability issue, it does not imply a definite assessment about closed loop performance [8].

3. SEQUENCE OF OPERATION PROBLEM

The systematic repetition of single similar operations motivates the use of previous experiences to improve the control performance such to obtain a closer tracking of economic trajectories and to reach with more accuracy the desired target over the end. These performance improvements are the main reason for extending the above control strategy such to include learning capabilities. This learning capability can be developed by creating memory reservoirs for the inputs injected to both the actual process and the model, and for the correspondent output mismatch. For instance, these memory reservoirs can be incorporated into Problem P1 through the following additional constraints:

$$u^{r\,i}_{k+j} = u^{i-1}_{k+j}, \qquad j = 0,\cdots,H-1 \tag{15}$$

$$d^{i}_{k+j} = d^{i-1}_{k+j}, \qquad j = 0,\cdots,H-1 \tag{16}$$

where d^{i-1}_{k+j} represents the model mismatch or disturbance found during the operation *i*-1. The first constraint updates the input reference for operation *i* with the last optimal control sequence executed in operation *i*-1 (i.e. $\boldsymbol{u}^{r\,i} = \boldsymbol{u}^{i-1}$, for i =1,2,…). The second constraint updates the disturbance profile for operation *i* with the last estimated sequence in operation *i*-1 (i.e. $\boldsymbol{d}^{i} = \boldsymbol{d}^{i-1}$, for *i*=1,2,…, with an initial value given by $\boldsymbol{d}^{0} := [0 \quad \dots \quad 0]$).

The learning capacity is achieved by two reasons: one is the successive update of the input sequence taken as reference; the other is the update of the systematic model disturbance ***d*** appearing along the operating trajectory. As the number of operations progresses, both, the input and the disturbance sequences evolve to stable profiles. The input sequence changes from iteration to iteration because the control horizon in Problem P1 is much shorter than the prediction horizon (mostly due to computational limitations) and a single optimization does not suffice to reach the final complete optimal input profile. On the other hand, the disturbance sequence changes as a consequence of the input profile evolution and provides information about the system future behavior. For instance, once the disturbance sequence accumulates knowledge about the difference between plant and model for every sampling instant of the desired trajectory, better output forecast are made because the predictive algorithm anticipates these differences. This is an unusual feature for feedback control systems (since traditional designs require realizable conditions) and it is the reason that explains the algorithmic capacity of reaching extraordinary performances.

However, notice that the disturbance update in (16) is not an online update but an operation to operation one. This is equivalent to ignore changes due to unmeasured variables and to delay model mismatch adjustments until the next operation. To overcome this problem, an additional correction to account for "current" disturbances is proposed, this is given by $\Delta d^{i}_{k+j} = \hat{d}^{i}_{k} - d^{i-1}_{k}$, $j = 0,\cdots,H$ where $\hat{d}^{i}_{k}$ is estimated by the observer. This correction modifies the disturbance constraint in (6), as follows:

$$d^{i}_{k+j} = d^{i-1}_{k+j} + \Delta d^{i}_{k+j} = d^{i-1}_{k+j} + \hat{d}^{i}_{k} - d^{i-1}_{k}, \quad j = 0,..,H-1$$

Based on the later formulation, it is possible to establish the following theorem:

Theorem 2: Let us assume that no disturbance enter the system. For the system (1)-(3), by using the control law derived from the on-line execution of problem P1 in a receding horizon manner, together with the learning updating (15) and (16), and assuming that a feasible perfect control input trajectory there exists (i.e., a control trajectory for which the output trajectory is the output reference trajectory), then the output error trajectory converges to zero as $i \to \infty$.

Proof.

The proof of this theorem is similar to the one shown in [10].

Remark: In real systems a perfect control input trajectory is not possible to reach, mainly for the case of abrupt references trajectories. In these cases, the costs of the proposed MPC controller converge to a non-null finite value as $i \to \infty$. Furthermore, if an operation cost is defined as $J_i := \sum_{k=0}^{T_f-1} V_k^{i\,opt}$, then, this cost will be decreasing and it will converge to the smallest possible value taking into account the input constraints.

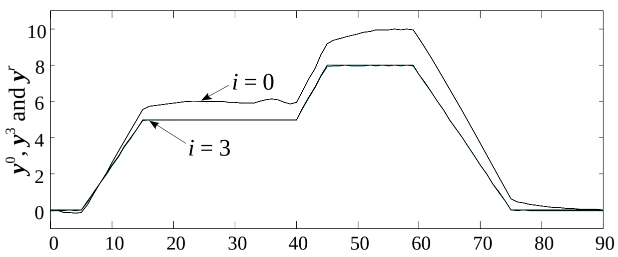
Given that the impossibility to reach perfect control is exclusively related to the input and/or states limits (which should be consistent with the control problem constraints), the proposed strategy will find the best approximation to the perfect control, which constitutes an important advantage of the method.

4. APPLICATION EXAMPLE

In order to evaluate the performance of the proposed controller we choose a simple 2nd order SISO system. The true and nominal processes are represented by $G(s)$=1/15s^2+8s+1 and $G(s)$=0.8/12s^2+7s+1, respectively [3]. The sampling time adopted to develop the discrete state space model is T=1 and the final batch time is given by T_f = 90T. The idea is to show that the proposed strategy achieves a good control performance in the first two or three operations, with a rather reduced control horizon. The controller parameters are as follows: Q=1500, R=0.05, N=5. Figure 1 shows the output response together with the output reference, and the inputs $\boldsymbol{u}^i$ and $\bar{\boldsymbol{u}}^i$, for the first and third operations. At the first operation, since the input reference is a constant value (u_s=0), $\boldsymbol{u}^i$ and $\bar{\boldsymbol{u}}^i$ are the same, and the output performance is quite poor (mainly because of the model mismatch). At the third operation, however, given that a disturbance state is estimated from the previous run, the output response and the output reference are undistinguishable.

As expected, the batch error is reduced drastically from operation 1 to operation 3, while the MPC cost is decreasing for each run (Figure 2). Notice that the MPC cost is normalized taking into account the maximal value ($V_k^i / V_{\max}^i$), where $V_{\max}^1 \approx 1 \cdot 10^6$ and $V_{\max}^1 \approx 286.5$. This shows that the MPC output error decrease from one run to the next, as was stated in Theorem 2. Finally, Figure 3 shows the normalized norm of the error corresponding to each operation, $J_i := \sum_{k=0}^{T_f-1} V_k^{i\,opt}$.

Notice that to achieve an acceptable tracking both, the input values and the input increments are quite aggressive. If realistic input constraints are considered, they would become active when the performance tends to perfect control condition. This is why it is usually impossible to achieve a perfect control in real systems. In this example no input constraints were considered in order to show the operation-to-operation convergence to perfect control. The rigorous convergence analysis was not included in this report due to space limitations.



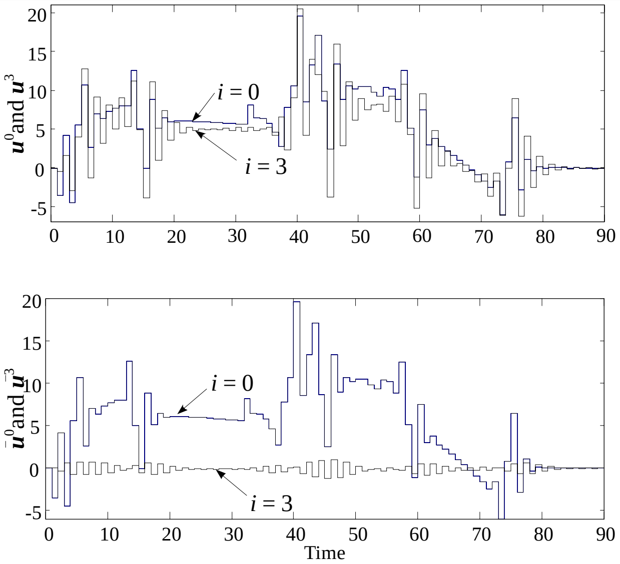


Figure 1 – Output and input responses.

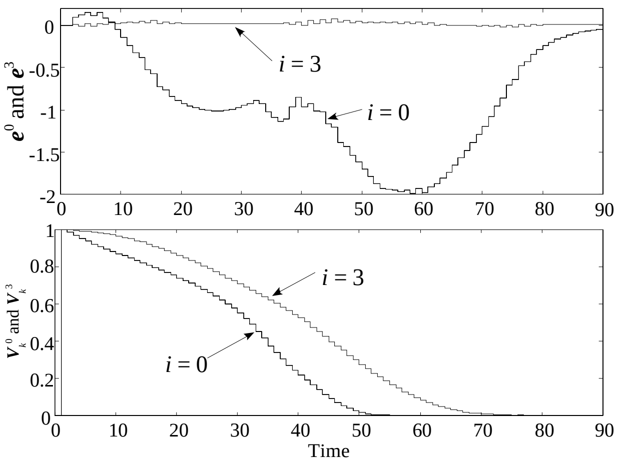


Figure 2 – Error and Normalized MPC cost.

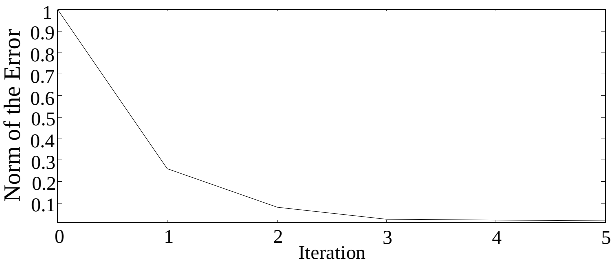


Figure 3 – Normalized norm of the iteration error.

5. CONCLUSIONS

A model predictive controller with learning properties to be applied to repetitive dynamic systems is presented. The proposed strategy integrates in a single MPC formulation the single operation and the sequence of operations control. Two memory reservoirs are used to construct the learning procedure: an open-loop control trajectory and a model mismatch history accumulated as disturbance information. The analysis of the finite-time control problem shows not only the decreasing property of the cost function but also that the achievable performance can increase as the operation approaches to the final running time. Besides, a convergence analysis for the sequence of operations shows that, unless system limitations arise, the proposed learning scheme is capable to reach perfect tracking performance.

6. REFERENCES

[1] TAN, K.K.; HUANG, S. N.; LEE, T. H.; TAY, A. Disturbance compensation incorporated in predictive control system using a repetitive learning approach. *System and Control Letters*, v. 56, p. 75-82, 2007.

[2] YAO, K.; GAO, F.; ALLGOWER, F. Barrel temperature control during operation transition in injection molding. *Control Engineering Practice*, v. 16, No. 11, p. 1259-1264, 2008.

[3] LEE, K.S.; LEE, J.H. Convergence of constrained model predictive control for batch processes. *IEEE Trans A.C.*, v. 45, p. 1928-1932, 2000.

[4] BONVIN, D.; SRINIVASAN, B.; HUNKELER, D. Control and optimization of bath processes – improvement of process operation in the production of specialty chemicals. *IEEE Trans. Control Syst. Mag.* V. 26, No.6, p. 34-45, 2006.

[5] MOYNE, J.E.; del CASTILLO, E.; HURWITZ, A.M. Run to run control in

semiconductor manufacturing. CRC Press, Boca Raton, Florida, 2001.

[6] SRINIVASAN, B.; BONVIN, D. Controllability and stability of repetitive batch processes. *Journal of Process Control*, v. 17 p. 285-295, 2007.

[7] LEE, K.S.; LEE, J.H. Model predictive control technique combined with iterative learning for batch processes. *AIChE Journal*, v. 45, p. 2175-2187, 1999.

[8] MUSKE, K.R.; BADGWELL, T.M. Disturbance modeling for offset free linear model predictive control. *Journal of Process Control*, v. 12, p. 617–632, 2002.

[9] GONZALEZ, A.H.; ADAM, E.J.; MARCHETTI, J.L. Conditions for offset elimination in receding horizon controllers: a tutorial analysis. *Chemical Engineering and Processing*, v. 47, p. 2184-2194, 2008.

[10] GONZALEZ, A.H.; ADAM, E.J.; MARCHETTI, J.L. Infinite horizon MPC applied to batch process. *XIIIº Reunión de Trabajo en Procesamiento de la Información y Control (RPIC), Rosario, Argentina,* 2009.