Extremum Seeking-based Iterative Learning Linear MPC

Benosman, M.; Di Cairano, S.; Weiss, A.

TR2014-088 August 2014

Abstract

In this work we study the problem of adaptive MPC for linear time-invariant uncertain models. We assume linear models with parametric uncertainties, and propose an iterative multi-variable extremum seeking (MES)-based learning MPC algorithm to learn online the uncertain parameters and update the MPC model. We show the effectiveness of this algorithm on a DC servo motor control example.

IEEE Multi-Conference on Systems and Control (MSC)

Extremum Seeking-based Iterative Learning Linear MPC

Mouhacine Benosman, Stefano Di Cairano, Avishai Weiss

Abstract— **In this work we study the problem of adaptive MPC for linear time-invariant uncertain models. We assume linear models with parametric uncertainties, and propose an iterative multi-variable extremum seeking (MES)-based learning MPC algorithm to learn online the uncertain parameters and update the MPC model. We show the effectiveness of this algorithm on a DC servo motor control example.**

I. Introduction

Model predictive control (MPC) [1] is a model-based framework for optimal control of constrained multi-variable systems. MPC is based on the repeated, receding horizon solution of a finite-time optimal control problem formulated from the system dynamics, constraints on system states, inputs, outputs, and a cost function describing the control objective. MPC has been applied to several applications such as aerospace [2], [3], automotive [4], [5], and mechatronic systems [6], [7]. Since MPC is a model-based controller, its performance inevitably depends on the quality of the prediction model used in the optimal control computation.

In contrast, extremum seeking (ES) control is a well known approach where the extremum of a cost function associated with a given process performance (under some conditions) is found without the need for detailed modeling information, see, e.g., [8]–[10]. Several ES algorithms (and associated stability analyses) have been proposed, [8]–[12], [12]–[14], and many applications of ES have been reported [15]–[19].

The idea that we introduce in this work, is that the performance of a model-based MPC controller can be combined with the robustness of a model-free ES learning algorithm for simultaneous identification and control of linear time-invariant systems with structural uncertainties. While regulation and identification are seemingly conflicting objectives, by identifying (or re-identifying) the system dynamics online and updating the MPC prediction model, the closed-loop performance may be enhanced relative to a standard MPC scheme that uses an inaccurate (or outdated) model. The optimal solution to this trade-off between identification and control is given by a dynamic program [20], which, for many applications, is computationally intractable. As a result, many suboptimal techniques and heuristics have been developed in recent years – often in a *receding horizon* framework which is suitable for integration with MPC.

In [21], an approximation of the dynamic program is developed for a linear input-output map with no dynamics. Approaches for more complex systems avoid dynamic programming altogether and, instead, sub-optimally trade off between inputs that excite the system and inputs that regulate the state. Excitation signals are often designed to satisfy *persistency of excitation* conditions. For example, a dithering signal may be added on top of the nominal control [22], although difficulties arise in determining the amplitude of the signal, and the dither indiscriminately adds noise to the process. More sophisticated schemes employ *optimal input design*, usually in the frequency domain, where maximizing the Fisher information matrix can be cast as a semi-definite program [23]. However, design in the frequency domain leads to difficulties with constraints that are more naturally addressed in the time domain, e.g., input (and possibly output) amplitude constraints. While the problem formulation in the time domain is highly non-convex, developing such techniques is desirable, and thus the focus of recent work [24]–[32].

In this preliminary work, we aim at proposing an alternative approach to realize an iterative learning-based adaptive MPC. We introduce an approach for a multi-variable extremum seeking (MES)-based iterative learning MPC that merges a model-based linear MPC algorithm with a model-free MES algorithm to realize an iterative learning MPC that adjusts to structured model uncertainties. This approach is an extension to the recent MPC framework reported in [33], where the author proposed to use MES with model-based nonlinear control to design learning-based adaptive controllers for a class of nonlinear systems.

The paper is organized as follows. We start the paper with some preliminaries in Section II. In Section III, we first recall the nominal MPC algorithm and then present the main result of the paper, namely, the iterative learning MPC, with a discussion of the algorithm stability. Section IV is dedicated to a DC servo-motor case study. Finally, we conclude the paper with a brief summary of the results in Section V.

II. Notations and preliminaries

Throughout the paper, $\mathbb{R}$, $\mathbb{Z}_{0+}$, $\mathbb{Z}_{[i,j]}$ denote the set of real numbers, positive integers, and positive integers from i to j, respectively. For $x \in \mathbb{R}^N$ we define $||x|| = \sqrt{x^T x}$, we denote by $A_{ij},\ i = 1,\ldots,n,\ j = 1,\ldots,m$ the elements of the matrix A, and by $[A]_i$ the vector equal to the line i of the matrix A. We denote by $||A||_2$ the spectral matrix norm, and by $||x||_A = \sqrt{x^T A x}$. We denote by $x(i|k)$ the value of x at the time sample i and the MPC cycle k. In the sequel, when we use the term well-posed optimization problem, we mean that the problem admits a unique solution, which is a continuous function of the initial conditions [34].

III. Iterative learning-based adaptive MPC

A. Control objective

We want to design an adaptive controller that solves regulation and tracking problems for linear time-invariant systems with structural model uncertainties under state, input, and output constraints. In what follows, we first present the nominal MPC problem, i.e., without model uncertainties, and then extend this nominal controller to its adaptive form by merging it with an MES algorithm.

B. Constrained linear nominal MPC

Consider a linear MPC, based on the nominal linear prediction model

$$x(k+1) = Ax(k) + Bu(k), \tag{1a}$$
$$y(k) = Cx(k) + Du(k), \tag{1b}$$

M. Benosman (m_benosman@ieee.org), S. Di Cairano (dicairano@ieee.org), and A. Weiss (weiss@merl.com) are with Mitsubishi Electric Research Laboratories, 201 Broadway Street, Cambridge, MA 02139, USA.

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$ are the state, input, and output vectors subject to constraints

$$x_{\min} \le x(k) \le x_{\max}, \tag{2a}$$

$$u_{\min} \le u(k) \le u_{\max}, \tag{2b}$$

$$y_{\min} \le y(k) \le y_{\max}, \tag{2c}$$

where $x_{\min}, x_{\max} \in \mathbb{R}^n$, $u_{\min}, u_{\max} \in \mathbb{R}^m$, and $y_{\min}, y_{\max} \in \mathbb{R}^p$ are the lower and upper bounds on the state, input, and output vectors, respectively. At every control cycle $k \in \mathbb{Z}_{0+}$, MPC solves the finite horizon optimal control problem

$$\min_{U(k)} \quad \sum_{i=0}^{N-1} \|x(i|k)\|^2_{Q_{\mathrm{M}}} + \|u(i|k)\|^2_{R_{\mathrm{M}}} \tag{3a}$$

$$+\|x(N|k)\|^2_{P_{\mathrm{M}}},$$

$$\text{s.t.} \quad x(i+1|k) = Ax(i|k) + Bu(i|k), \tag{3b}$$

$$y(i|k) = Cx(i|k) + Du(i|k), \tag{3c}$$

$$x_{\min} \le x(i|k) \le x_{\max},\ i \in \mathbb{Z}_{[1,N_c]}, \tag{3d}$$

$$u_{\min} \le u(i|k) \le u_{\max},\ i \in \mathbb{Z}_{[0,N_{cu}-1]}, \tag{3e}$$

$$y_{\min} \le y(i|k) \le y_{\max},\ i \in \mathbb{Z}_{[0,N_c]}, \tag{3f}$$

$$u(i|k) = K_f x(i|k),\ i \in \mathbb{Z}_{[N_u,N-1]}, \tag{3g}$$

$$x(0|k) = x(k), \tag{3h}$$

where $Q_{\mathrm{M}} \ge 0$, $P_{\mathrm{M}}, R_{\mathrm{M}} > 0$ are symmetric weight matrices of appropriate dimensions, N is the prediction horizon, $N_u \le N$ is the control horizon (the number of free control moves), $N_{cu} \le N$, $N_c \le N-1$ are the input and output constraint horizons along which the constraints are enforced. The performance criterion is defined by (3a), and (3d)–(3f) enforce the constraints. Equation (3g) defines the pre-assigned terminal controller where $K_f \in \mathbb{R}^{m \times n}$, so that the optimization vector effectively is $U(k) = [u'(0|k) \dots u'(N_u-1|k)]' \in \mathbb{R}^{N_u m}$.

Although the optimal control problem (3) does not explicitly mention a reference, tracking is achieved by including in the state update equation (1a) the reference prediction dynamics

$$r_r(k+1) = A_r r_r(k), \tag{4}$$

and an additional output in (1b) representing the tracking error

$$y_e(k) = Cx(k) - C_r r_r(k), \tag{5}$$

which is then accounted for in the cost function (3a) as later shown in the example (see also [5] for an example in a real world application). At time k, the MPC problem (3) is initialized with the current state value $x(k)$ by (3h) and solved to obtain the optimal sequence $U^*(k)$. Then, the input $u(k) = u_{\mathrm{MPC}}(k) = u^*(0|k) = [I_m\ 0\ \dots\ 0]U(k)$ is applied to the system.

C. Learning-based adaptive MPC algorithm

Consider now, the system (1), with structural uncertainties, such that

$$x(k+1) \ = \ (A+\Delta A)x(k) + (B+\Delta B)u(k) \tag{6a}$$

$$y(k) \ = \ (C+\Delta C)x(k) + (D+\Delta D)u(k), \tag{6b}$$

with the following assumptions.

Assumption 1: The constant uncertainty matrices ΔA, ΔB, ΔC and ΔD, are bounded, s.t. $||\Delta A||_2 \le l_A$, $||\Delta B||_2 \le l_B$, $||\Delta C||_2 \le l_C$, $||\Delta D||_2 \le l_D$, with $l_A,\ l_B,\ l_C,\ l_D > 0$.

Assumption 2: There exists non empty convex sets $\mathcal{K}_a \subset \mathbb{R}^{n \times n}$, $\mathcal{K}_b \subset \mathbb{R}^{n \times m}$, $\mathcal{K}_c \subset \mathbb{R}^{p \times n}$, and $\mathcal{K}_d \subset \mathbb{R}^{p \times m}$, such that $A + \Delta A \in \mathcal{K}_a$ for all ΔA such that $||\Delta A||_2 \le l_A$, $B + \Delta B \in \mathcal{K}_b$ for all ΔB such that $||\Delta B||_2 \le l_B$, $C + \Delta C \in \mathcal{K}_c$ for all ΔC such that $||\Delta C||_2 \le l_C$, $D + \Delta D \in \mathcal{K}_d$ for all ΔD such that $||\Delta D||_2 \le l_D$,.

Assumption 3: The iterative learning MPC problem (3) (and the associated reference tracking extension), where we substitute the model with structural uncertainty (6) for the nominal model (3b) and (3c), is a well-posed optimization problem for any matrices $A + \Delta A \in \mathcal{K}_a$, $B + \Delta B \in \mathcal{K}_b$, $C + \Delta C \in \mathcal{K}_c$, $D + \Delta D \in \mathcal{K}_d$.

Under these assumptions, we postulate the following: If we solve the iterative learning MPC problem (3), where we substitute (6) for (3b) and (3c), iteratively, such that, at each new iteration we update our knowledge of the uncertain matrices ΔA, ΔB, ΔC, and ΔD, using a model-free learning algorithm, in our case the extremum seeking algorithm, we claim that, if we can improve over the iterations the MPC model, i.e., learn over iterations the uncertainties, then we can improve over time the MPC performance, i.e., either in the stabilization or in the tracking. Before formulating this idea in terms of an algorithm, we briefly recall the principle of model-free multi-variable extremum seeking (MES) control.

To use the MES learning algorithm, we define the cost function to be minimized as

$$Q(\hat{\Delta}) = F(y_e(\hat{\Delta})), \tag{7}$$

where $\hat{\Delta}$ is the vector obtained by concatenating all the elements of the estimated uncertainty matrices $\Delta\hat{A}$, $\Delta\hat{B}$, $\Delta\hat{C}$ and $\Delta\hat{D}$, $F : \mathbb{R}^p \to \mathbb{R}$, $F(0) = 0$, $F(y_e) > 0$ for $y_e \neq 0$.

In order to ensure convergence of the MES algorithm, Q need to satisfy the following assumptions.

Assumption 4: The cost function Q has a local minimum at $\hat{\Delta}^* = \Delta$.

Assumption 5: The original parameter estimate vector $\hat{\Delta}$ is close enough to the actual parameters vector Δ.

Assumption 6: The cost function is analytic and its variation with respect to the uncertain variables is bounded in the neighborhood of Δ^*, i.e., there exists $\xi_2 > 0$, s.t. $\|\frac{\partial Q}{\partial \Delta}(\tilde{\Delta})\| \le \xi_2$ for all $\tilde{\Delta} \in \mathcal{V}(\Delta^*)$, where $\mathcal{V}(\Delta^*)$ denotes a compact neighborhood of Δ^*.

Remark 1: Assumption 4 simply means that we can consider that Q has at least a local minimum at the true values of the uncertain parameters.

Remark 2: Assumption 5 indicates that our result will be of local nature, meaning that our analysis holds in a small neighborhood of the actual values of the parameters.

Remark 3: We wrote the cost function (7) as function of the tracking error (5), however, the case of regulation or stabilization can be directly deduced from this formulation by replacing the time-varying reference with a constant reference or an equilibrium point.

Under Assumptions 4, 5, and 6, it has been shown (e.g. [9], [10]), that the MES

$$\begin{array}{l} \dot{z}_i = a_i \sin(\omega_i t + \frac{\pi}{2}) Q(\hat{\Delta}) \\ \hat{\Delta}_i = z_i + a_i \sin(\omega_i t - \frac{\pi}{2}),\ i \in \{1, \dots, N_p\} \end{array} \tag{8}$$

with $N_p \le nn + nm + pn + pm$ is the number of uncertain elements, $\omega_i \neq \omega_j,\ \omega_i + \omega_j \neq \omega_k,\ i, j, k \in \{1, \dots, N_p\}$, and $\omega_i > \omega^*,\ \forall i \in \{1, \dots, N_p\}$, with ω^* large enough, converges to the local minima of Q.

The idea that we want to propose here (refer to [33] where we introduced this concept of learning-based adaptive control for a class of nonlinear systems), is that under Assumptions 1-6, we can merge the MPC algorithm and a discrete-time version of the MES algorithm to obtain an iterative learning MPC algorithm. We formalize this idea in the following iterative algorithm:

ALGORITHM I

```
- Initialize $z_i(0) = 0$, and the uncertainties' vector estimate $\hat{\Delta}(0) = 0$.
- Choose a threshold for the cost function minimization $\epsilon_Q > 0$.
- Choose the parameters, MPC sampling time $\delta T_{\mathrm{mpc}} > 0$, and MES sampling time $\delta T_{\mathrm{mes}} = N_E \delta T_{\mathrm{mpc}}$, $N_E > 0$.
- Choose the MES dither signals' amplitudes and frequencies: $a_i,\ \omega_i,\ i = 1, 2 \ldots, N_p$.
  WHILE(true)
    FOR($\ell = 1, \ell \le N_E, \ell = \ell + 1$)
      - Solve the MPC problem
```

$$
\begin{aligned}
\min_{U(k)} \quad & \sum_{i=0}^{N-1} \|x(i|k)\|^2_{Q_{\mathrm{M}}} + \|u(i|k)\|^2_{R_{\mathrm{M}}} && \text{(9a)}\\
& \quad + \|x(N|k)\|^2_{P_{\mathrm{M}}}, \\
\text{s.t.} \quad & x(i+1|k) = (A + \Delta A)x(i|k) && \text{(9b)}\\
& \qquad\qquad + (B + \Delta B)u(i|k), \\
& y(i|k) = (C + \Delta C)x(i|k) && \text{(9c)}\\
& \qquad\qquad + (D + \Delta D)u(i|k), \\
& x_{\min} \le x(i|k) \le x_{\max},\ i \in \mathbb{Z}_{[1,N_c]}, && \text{(9d)}\\
& u_{\min} \le u(i|k) \le u_{\max},\ i \in \mathbb{Z}_{[0,N_{cu}-1]}, && \text{(9e)}\\
& y_{\min} \le y(i|k) \le y_{\max},\ i \in \mathbb{Z}_{[1,N_c]}, && \text{(9f)}\\
& u(i|k) = K_f x(i|k),\ i \in \mathbb{Z}_{[N_u,N-1]}, && \text{(9g)}\\
& x(0|k) = x(k), && \text{(9h)}
\end{aligned}
$$

```
      - Update $k = k + 1$.
    End
    IF $Q > \epsilon_Q$
      - Evaluate the MES cost function $Q(\hat{\Delta})$
      - Evaluate the new value of the uncertainties $\hat{\Delta}$:
```

$$
\begin{aligned}
& z_i(h+1) = z_i(h) + a_i \delta T_{\mathrm{mes}} \sin(\omega_i h \delta T_{\mathrm{mes}} + \tfrac{\pi}{2}) Q(\hat{\Delta}) \\
& \hat{\Delta}_i(h+1) = z_i(h+1) + a_i \sin(\omega_i h \delta T_{\mathrm{mes}} - \tfrac{\pi}{2}), \\
& i \in \{1, \ldots, N_p\}
\end{aligned}
\tag{10}
$$

```
      - Update $h = h + 1$.
    End
    Reset $\ell = 0$
  End
```

D. Stability discussion

As mentioned, in this preliminary work we aim at presenting an algorithm that merges model-based MPC and model-free MES learning, to obtain an iterative learning MPC algorithm, with encouraging numerical numerical results (see, e.g., the case study in Section IV). A rigorous stability analysis of the combined algorithm is out of the scope of this work, but we want to sketch below an approach to analyze the stability of Algorithm I. We propose here to follow the analysis presented in [33], for the case of learning-based adaptive control for some class of nonlinear systems. By Assumptions 1 -3, the model structural uncertainties $\Delta A, \Delta B, \Delta C$, and ΔD are bounded, the uncertain model matrices $A + \Delta A$, $B + \Delta B$, $C + \Delta C$, and $D + \Delta D$ are elements of convex sets $\mathcal{K}_a$, $\mathcal{K}_b$, $\mathcal{K}_c$, and $\mathcal{K}_d$, and that the MPC problem (9) is well-posed. Based on this, the approach for proving stability is based on establishing a boundedness of the tracking error norm $\|y_e\|$ with the upper-bound being function of the uncertainties estimation error norm $\|\hat{\Delta} - \Delta\|$. One effective way to characterize such a bound is to use an integral Input-to-State Stability (iISS) (or ISS for time-invariant problems) between the input $\|\hat{\Delta} - \Delta\|$ and the augmented state $\|y_e\|$, see, e.g. [33]. If the iISS (or ISS) property is obtained, by reducing the estimation error $\|\hat{\Delta} - \Delta\|$ we also reduce the the tracking error $\|y_e\|$, due to the iISS relation between the two signals. Based on Assumptions 5, 6, and 7, we know, e.g. [9], [10], that the MES algorithm (10) converges to a local minimum of the MES cost Q, which implies (based on Assumption 4), that the estimation error $\|\hat{\Delta} - \Delta\|$ is reducing over the MES iterations. Thus, finally we conclude that the MPC tracking (or regulation) performance is improved over the iterations of Algorithm I. While obviously this discussion is not a rigorous proof of stability of the iterative learning MPC, it provides an interesting guideline to follow for analyzing the controller stability observed during the test case presented next. A more rigorous proof is currently under development and will be presented in future works.

IV. DC SERVO-MOTOR EXAMPLE

The example studied here is about the angular position control of a load connected by a flexible shaft to a voltage actuated DC servo motor, see [35] and references therein. The states are the load angle and angular rate, and the motor angle and angular rate, the control input is the motor voltage, and the outputs are the load angle and the torque acting on the flexible shaft. The model for the system is

$$
\begin{aligned}
\dot{x}_c(t) &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_l}{\mathcal{J}_l} & -\frac{\beta_l}{\mathcal{J}_l} & \frac{k_l}{g\mathcal{J}_l} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_l}{g\mathcal{J}_m} & 0 & -\frac{k_l}{g^2\mathcal{J}_m} & -\frac{\beta_m + R_A^{-1}K_m^2}{\mathcal{J}_m} \end{bmatrix} x_c(t) + \\
& \qquad \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{K_m}{R_A \mathcal{J}_m} \end{bmatrix} u_c(t) \\
y_c(t) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ k_l & 0 & -\frac{k_l}{g} & 0 \end{bmatrix} x_c(t)
\end{aligned}
\tag{11}
$$

where $x_c \in \mathbb{R}^4$ is the state vector, $u_c \in \mathbb{R}$ is the input vector, and $y_c \in \mathbb{R}^2$ is the output vector. In (11) $R_A[\Omega]$ is the armature resistance, K_m[Nm/A] is the motor constant, $\mathcal{J}_l$[kgm^2], β_l[Nms/rad], k_l[Nm/rad], are the inertia, friction and stiffness of load and flexible shaft, $\mathcal{J}_m$[kgm^2], β_m[Nms/rad], are the inertia and friction of the motor, and g is the gear ratio between motor and load. The nominal numerical values used in the simulations are $R_A = 10\Omega$, $K_m = 10$Nm/A, $\mathcal{J}_l = 25$kgm^2, $\beta_l = 25$Nms/rad, $k_l = 1.28 \cdot 10^3$Nm/rad, $\mathcal{J}_m = 0.5$kgm^2, $\beta_m = 0.1$Nms/rad. The system is subject to constraints on motor voltage and shaft torque

$$
\begin{aligned}
-78.5 \le [y_c(t)]_2 \le 78.5, && \text{(12a)}\\
-220 \le u_c(t) \le 220. && \text{(12b)}
\end{aligned}
$$

The control objective is to track a time varying load angle position reference signal $r_l(t)$. Following the nominal MPC presented in Section III-B, the prediction model is obtained by sampling (11) with a period $\delta T_{\mathrm{mpc}} = 0.1s$, and the trajectory tracking problem is solved by augmenting the model with the reference prediction model (4), which in this case represents a constant reference

$$r_r(k+1) = r_r(k),$$

where $r_r \in \mathbb{R}$, and with an incremental formulation of the control input, such that

$$u_c(k+1) = u_c(k) + \Delta v(k)$$

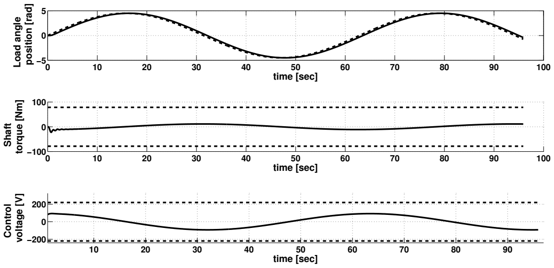

Fig. 1. Outputs and input signals in the nominal case (reference trajectory and constraints limits in dashed-line, obtained signals in solid-line)

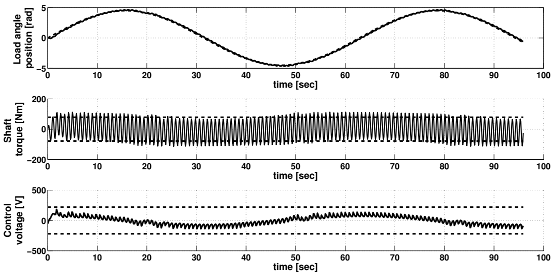

Fig. 2. Outputs and input signals in the uncertain case with nominal MPC control (reference trajectory and constraints limits in dashed-line, obtained signals in solid-line)

Next, the MPC cost function is chosen as

$$\sum_{i=1}^{N} \|[y(i|k)]_1 - r_l(i|k)\|_{Q_y}^2 + \|\Delta v(i|k)\|_{R_v}^2 + \rho\sigma^2, \qquad (13)$$

where $Q_y = 10^3$ and $R_v = 0.05$, prediction, constraints, and control horizons are $N = 20$, $N_c = N_{cu} = N_u = 4$, and $K_f = 0$ in (3g). In this case study output constraints (12a) are considered as soft constraints, which may be (briefly) violated due to the prediction model not being equal to the actual model. Thus, in (13) we add the term $\rho\sigma^2$, where $\rho > 0$ is a (large) cost weight, and σ is an additional variable used to model the (maximum) constraint violation of the softened constraints.

Thus, the MPC problem (3) results in a family of quadratic programs parameterized by the current state $x(k)$ in (3h), with $n_q = 16$ constraints, and $n_u = 4$ variables. In the subsequent simulations, we consider initial state $x(0) = [0\ 0\ 0\ 0]'$ and reference $r_l(t) = 4.5\sin(\frac{2\pi}{T_{ref}}t)$, $T_{ref} = 20\pi\ sec$.

First, to have a base-line performance, we solve the nominal MPC problem, i.e. without model uncertainties. We report the corresponding results on Figure 1, where it is clear that the desired load angular position is precisely tracked, without violating the shaft torque and the control voltage constraints. Next, we introduce the parametric model uncertainty $\delta\beta_l = -70\ [Nms/rad]$. Note that we purposely introduced a very large model uncertainty, i.e. more than 100% of the nominal value, to show clearly the bad effect of this uncertainty on the nominal MPC algorithm and to subsequently test the iterative learning MPC algorithm on a challenging case. We first apply the nominal MPC controller to the uncertain model, we show the obtained performance on Figure 2, where it is clear that the nominal performance is lost, since the second output, i.e. the shaft torque, is oscillating and is violating its upper and lower limits. Furthermore, this oscillations are also present on the control voltage signal.

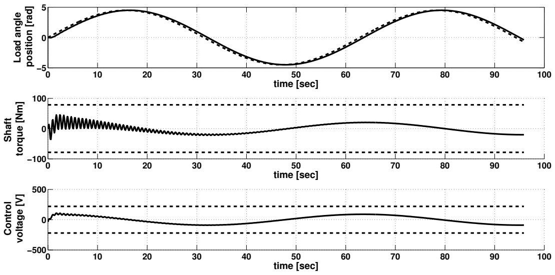

Fig. 3. Outputs and input signals in the uncertain case with iterative learning MPC (reference trajectory and constraints limits in dashed-line, obtained signals in solid-line)

Now, we apply the iterative learning MPC Algorithm I, where we set $\delta T_{\text{mes}} = 1.5T_{ref}$. We choose the MES learning cost function as

$$\begin{aligned} Q &= \sum_{i=0}^{N_E-1} \|[y_e(t - i\delta T_{mpc})]_1\|^2 + \|[\dot{y}_e(t - i\delta T_{mpc})]_1\|^2 \\ &\qquad + \|[y_e(t - i\delta T_{mpc})]_2\|^2, \end{aligned}$$

i.e. the norm of the error in the load angular position and velocity, plus the norm of the error on the shaft torque. To learn the uncertain parameter β_l, we apply the algorithm (10), as

$$\begin{aligned} z_{\beta_l}(k^{'}+1) &= z_{\beta_l}(k^{'}) + a_{\beta_l}\delta T_{mes} sin(\omega_{\beta_l}k^{'}\delta T_{mes} + \tfrac{\pi}{2})Q \\ \delta\hat{\beta}_l(k^{'}+1) &= z_{\beta_l}(k^{'}+1) + a_{\beta_l} sin(\omega_{\beta_l}k^{'}\delta T_{mes} - \tfrac{\pi}{2}), \end{aligned} \qquad (14)$$

with $a_{\beta_l} = 10^{-6}$, $\omega_{\beta_l} = 0.7\ rad/s$. We select ω_{β_l} to be higher than the desired frequency of the closed-loop (around $0.1\ rad/s$), to ensure convergence of the MES algorithm, since the ES algorithms convergence proofs are based on averaging theory, which assumes high dither frequencies, e.g. [9], [10]. We chose a small value of the dither signal amplitude, since we noticed that the MES cost function has large values due to the large simulated uncertainty, so to keep the search excursion amplitude small, and converge to a precise value of the uncertainty $\delta\beta_l$, i.e. to keep $a_{\beta_l}Q$ small, we choose small a_{β_l} (for further explanations on how to tune MES algorithms please refer to [36]). We also set the MES cost function threshold ϵ_Q to $1.5Q_{nominal}$, where $Q_{nominal}$ is the value of the MES cost function obtained in the nominal-model case with the nominal MPC, i.e. the base-line ideal case. In other words, we decide to stop searching for the best estimation of the value of the uncertainty when the uncertain MES cost function, i.e., the value of Q when applying the iterative learning MPC algorithm to the uncertain model, is less or equal to 1.5 of the MES cost function in the case without model uncertainties, which represents the best achievable MES cost function.

The obtained results of the iterative learning MPC algorithm are reported on Figures 3, 4 and 5. First, note on Figure 4, that the uncertain cost function initial value (at the first iteration of the MES learning) is very high, about $4.5 \cdot 10^6$, which is about 50 times the value of the nominal base-line MES cost function value. We see on Figure 4 that this cost function decreases as expected along the MES learning iterations to reach a small value after about 25 iterations. This corresponds to the required number of iterations to learn the actual value of the uncertain parameter as shown on Figure 5. Eventually, after the convergence of the learning algorithm, we see on Figure 3, that the nominal base-line performances of the

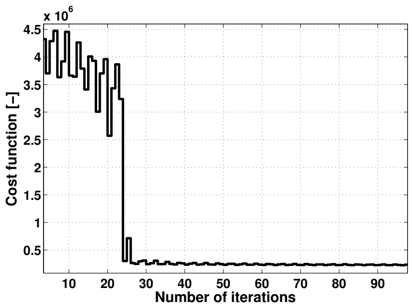

Fig. 4. MES cost function evolution over the learning iterations

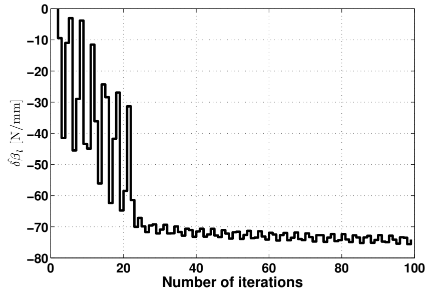

Fig. 5. Uncertain parameter learning evolution over the learning iterations

MPC are recovered and that the output track the desired reference with smooth signals and without violating the desired constraints.

We also tested the case of multiple uncertainties. We assumed the two uncertainties $\delta\beta_l = -70$, $[Nms/rad]$, $\delta J_l = -0.2$, $[kgm^2]$. We first show on Figure 6 the performance of the nominal MPC when applied to this uncertain model. It is clear that the nominal MPC cannot cope with the uncertainties effect on the system's performance, since the shaft torque, is oscillating and is violating its limits. The control voltage signal experiences oscillations, as well. Next, we apply the iterative learning-based MPC, where $\hat{\delta\beta_l}$ is learned using (14), and $\hat{\delta J_l}$ is learned using the ES equations

$$\begin{aligned} z_{J_l}(k'+1) &= z_{J_l}(k') + a_{J_l}\delta T_{mes} sin(\omega_{J_l} k' \delta T_{mes} + \tfrac{\pi}{2})Q \\ \delta\hat{J}_l(k'+1) &= z_{J_l}(k'+1) + a_{J_l} sin(\omega_{J_l} k' \delta T_{mes} - \tfrac{\pi}{2}), \end{aligned} \tag{15}$$

with $a_{J_l} = 10^{-8}$, and $\omega_{J_l} = 0.8\ rad/s$. Note that we choose a smaller dither amplitude for the δJ_l estimation, since the value of the uncertainty on J_l is smaller, so we need a smaller dither signal amplitude for the search of the uncertain value.

The results of the iterative learning MPC algorithm in this case are reported on Figure s 7, 8, 9 and 10. The learning cost function shown on Figure 8, is clearly decreasing and stabilizes after about 30 iterations. The uncertainties are learned and the overall tracking performance is recovered, as shown on Figures 9, 10 and 7, respectively. We notice here that the estimation of the uncertainties has some small residual error, this estimation error can be improved by either fine tuning the MES dither signals' amplitudes, e.g., by using a time-varying dither amplitude [37], or by choosing other type of MES algorithms with larger domain of attraction, e.g. [12], [38].

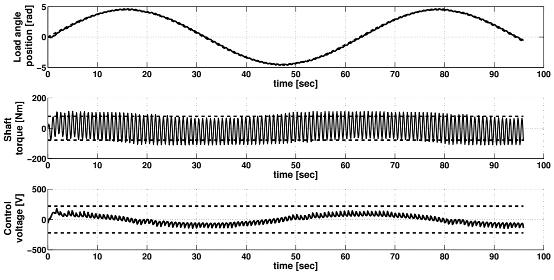

Fig. 6. Outputs and input signals in the uncertain case with nominal MPC control (reference trajectory and constraints limits in dashed-line, obtained signals in solid-line)

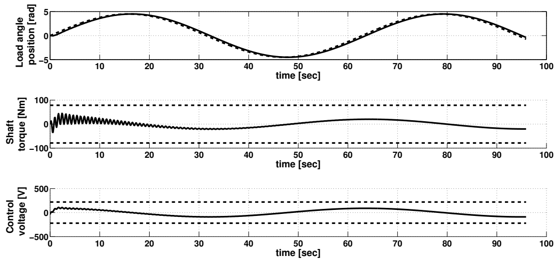

Fig. 7. Outputs and input signals in the uncertain case with iterative learning MPC (reference trajectory and constraints limits in dashed-line, obtained signals in solid-line)

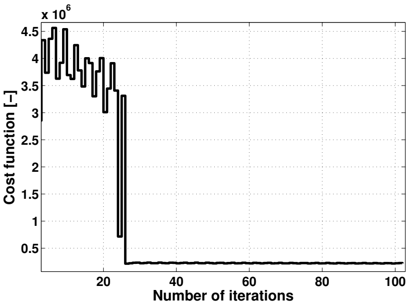

Fig. 8. MES cost function evolution over the learning iterations

V. Conclusion

In this paper, we have reported some preliminary results about an MES-based adaptive MPC algorithm. We have argued that it is possible to merge together a model-based linear MPC algorithm with a model-free MES algorithm to iteratively learn structural model uncertainties and thus improve the overall performance of the MPC controller. We have discussed a possible direction to analyze the stability of such algorithms. However, a more rigorous

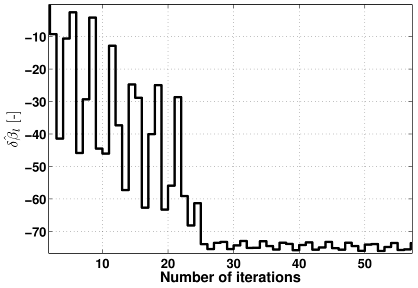

Fig. 9. Uncertain parameter $\delta\beta_l$ learning evolution over the learning iterations

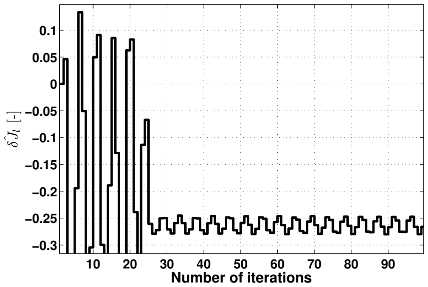

Fig. 10. Uncertain parameter δJ_l learning evolution over the learning iterations

analysis of the stability and convergence of the proposed algorithm is under development, and will be presented in our future reports. We have reported encouraging numerical results obtained on a mechatronics example, namely, a DC servo-motor control example. Future investigations will focus on improving the convergence rate of the iterative learning MPC algorithm, by using different ES algorithms with semi-global convergence properties, e.g. [12], [38], and on extending this work to different types of model-free learning algorithms, e.g. reinforcement learning algorithms, and comparing the learning algorithms in terms of their convergence rate and achievable optimal performances.

REFERENCES

[1] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, pp. 789–814, 2000.

[2] E. Hartley, J. Jerez, A. Suardi, J. Maciejowski, E. Kerrigan, and G. Constantinides, "Predictive control of a boeing 747 aircraft using an fpga," in *Proc. 4th IFAC Nonlinear Model Predictive Control Conference*, Noordwijkerhout, The Netherlands, 2012, pp. 80–85.

[3] S. Di Cairano, H. Park, and I. Kolmanovsky, "Model predictive control approach for guidance of spacecraft rendezvous and proximity maneuvering," 2012, special Issue dedicated to Prof. David W. Clarke. In press.

[4] S. Di Cairano, D. Yanakiev, A. Bemporad, I. Kolmanovsky, and D. Hrovat, "Model predictive idle speed control: Design, analysis, and experimental evaluation," *IEEE, Transactions on Control Systems Technology*, vol. 20, no. 1, pp. 84 –97, 2012.

[5] S. Di Cairano, H. Tseng, D. Bernardini, and A. Bemporad, "Vehicle yaw stability control by coordinated active front steering and differential braking in the tire sideslip angles domain," *IEEE, Transactions on Control Systems Technology*, pp. 1–13, 2012, in press.

[6] S. Di Cairano, A. Bemporad, I. Kolmanovsky, and D. Hrovat, "Model predictive control of magnetically actuated mass spring dampers for automotive applications," *Int. Journal of Control*, vol. 80, no. 11, pp. 1701–1716, 2007.

[7] A. Grancharova and T. Johansen, "Explicit model predictive control of an electropneumtaic clutch actuator using on/off valves and pulsewidth modulation," in *European Control Conference*, Budapest, Hungary, 2009, pp. 4278–4283.

[8] K. Ariyur and M. Krstić, *Real-time optimization by extremum-seeking control*. Wiley-Blackwell, 2003.

[9] K. B. Ariyur and M. Krstic, "Multivariable extremum seeking feedback: Analysis and design," in *Proc. of the Mathematical Theory of Networks and Systems*, South Bend, IN, August 2002.

[10] D. Nesic, "Extremum seeking control: Convergence analysis," *European Journal of Control*, vol. 15, no. 34, pp. 331 – 347, 2009.

[11] M. Krstic, "Performance improvement and limitations in extremum seeking," *Systems & Control Letters*, vol. 39, pp. 313–326, 2000.

[12] Y. Tan, D. Nesic, and I. Mareels, "On non-local stability properties of extremum seeking control," *Automatica*, no. 42, pp. 889–903, 2006.

[13] M. Rotea, "Analysis of multivariable extremum seeking algorithms," in *Proceedings of the American Control Conference*, vol. 1, no. 6. IEEE, 2000, pp. 433–437.

[14] M. Guay, S. Dhaliwal, and D. Dochain, "A time-varying extremum-seeking control approach," in *American Control Conference*, 2013.

[15] T. Zhang, M. Guay, and D. Dochain, "Adaptive extremum seeking control of continuous stirred-tank bioreactors," *AIChE J.*, no. 49, p. 113123., 2003.

[16] N. Hudon, M. Guay, M. Perrier, and D. Dochain, "Adaptive extremum-seeking control of convection-reaction distributed reactor with limited actuation," *Computers & Chemical Engineering*, vol. 32, no. 12, pp. 2994 – 3001, 2008.

[17] C. Zhang and R. Ordez, *Extremum-Seeking Control and Applications*. Springer-Verlag, 2012.

[18] M. Benosman and G. Atinc, "Multi-parametric extremum seeking-based learning control for electromagnetic actuators," in *American Control Conference*, 2013.

[19] ——, "Nonlinear learning-based adaptive control for electromagnetic actuators," in *European Control Conference*, 2013.

[20] A. Feldbaum, "Dual control theory," *Automation and Remote Control*, vol. 21, no. 9, pp. 874–1039, 1960.

[21] M. S. Lobo and S. Boyd, "Policies for simultaneous estimation and optimization," in *American Control Conference*, 1999, pp. 958ñ–964.

[22] O. A. Sotomayor, D. Odloak, and L. F. Moro, "Closed-loop model re-identification of processes under MPC with zone control," *Control Engineering Practice*, vol. 17, no. 5, pp. 551 – 563, 2009.

[23] H. Jansson and H. Hjalmarsson, "Input design via LMIs admitting frequency-wise model specifications in confidence regions," *IEEE, Transactions on Automatic Control*, vol. 50, no. 10, pp. 1534–1549, 2005.

[24] G. Marafioti, R. R. Bitmead, and M. Hovd, "Persistently exciting model predictive control," *International Journal of Adaptive Control and Signal Processing*, 2013.

[25] E. Žáčeková, S. Prívara, and M. Pčolka, "Persistent excitation condition within the dual control framework," *Journal of Process Control*, vol. 23, no. 9, pp. 1270–1280, 2013.

[26] J. Rathouskỳ and V. Havlena, "MPC-based approximate dual controller by information matrix maximization," *International Journal of Adaptive Control and Signal Processing*, vol. 27, no. 11, pp. 974–999, 2013. [Online]. Available: http://dx.doi.org/10.1002/acs.2370

[27] H. Genceli and M. Nikolaou, "New approach to constrained predictive control with simultaneous model identification," *AIChE Journal*, vol. 42, no. 10, pp. 2857–2868, 1996.

[28] C. A. Larsson, M. Annergren, H. Hjalmarsson, C. R. Rojas, X. Bombois, A. Mesbah, and P. E. Modén, "Model predictive control with integrated experiment design for output error systems," in *American Control Conference*, 2013.

[29] T. A. N. Heirung, B. E. Ydstie, and B. Foss, "An adaptive model predictive dual controller," in *Adaptation and Learning in Control and Signal Processing*, vol. 11, no. 1, 2013, pp. 62–67.

[30] V. Adetola, D. DeHaan, and M. Guay, "Adaptive model predictive control for constrained nonlinear systems," *Systems & Control Letters*, vol. 58, no. 5, pp. 320–326, 2009.

[31] A. González, A. Ferramosca, G. Bustos, J. Marchetti, and D. Odloak, "Model predictive control suitable for closed-loop re-identification," in *American Control Conference*. IEEE, 2013, pp. 1709–1714.

[32] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin, "Provably safe and robust learning-based model predictive control," *Automatica*, vol. 49, no. 5, pp. 1216–1226, 2013.

[33] M. Benosman, "Learning-based adaptive control for nonlinear systems," in *European Control Conference*, 2014, pp. 920–925.

[34] J. Hadamard, "Sur les problmes aux drives partielles et leur signification physique," *Princeton University Bulletin*, pp. 49–52, 1902.

[35] S. Di Cairano, M. Brand, and S. Bortoff, "Projection-free parallel quadratic programming for linear model predictive control," *Int. Journal of Control*, vol. 86, no. 8, pp. 1367–1385, 2013.

[36] Y. Tan, D. Nesic, and I. Mareels, "On the dither choice in extremum seeking control," *Automatica*, no. 44, pp. 1446–1450, 2008.

[37] W. Moase, C. Manzie, and M. Brear, "Newton-like extremum seeking part I: Theory," in *IEEE, Conference on Decision and Control*, December 2009, pp. 3839–3844.

[38] W. Noase, Y. Tan, D. Nesic, and C. Manzie, "Non-local stability of a multi-variable extremum-seeking scheme," in *IEEE, Australian Control Conference*, November 2011, pp. 38–43.