Robust multi-rate predictive control using multi-step prediction models learned from data

Enrico Terzi[1], Marcello Farina[1], Lorenzo Fagiano[1], and Riccardo Scattolini[1]

Abstract— This note extends a recently proposed algorithm for model identification and robust model predictive control (MPC) of asymptotically stable, linear time-invariant systems subject to process and measurement disturbances. Independent output predictors for different horizon values are estimated with Set Membership methods. It is shown that the corresponding prediction error bounds are the least conservative in the considered model class. Then, a new multi-rate robust MPC algorithm is developed, employing said multi-step predictors to robustly enforce constraints and stability against disturbances and model uncertainty, and to reduce conservativeness. A simulation example illustrates the effectiveness of the approach.

I. INTRODUCTION

In a recent paper [6], we presented a unitary approach to model identification and robust Model Predictive Control (MPC) design for linear, asymptotically stable, discrete time systems subject to process and measurement disturbances. A Set Membership (SM) identification approach was used to obtain multi-step prediction models used in the cost function definition, while state and control constraints were tightened by propagating the uncertainty bound of a simulation model, tuned using the knowledge of the multi-step models and the associated error intervals. Being the multi-step predictors linear in their parameters, it was possible to derive tight uncertainty bounds in a tractable way. However, these bounds were not directly exploited to deal robustly with constraints, with a consequent limited advantage in terms of conservativeness reduction in the constraint tightening procedure.

In the present paper, we develop this line of research with two main contributions: first, we prove that the prediction error bounds obtained with the SM approach proposed in [6] are smaller than those of *any* linear simulation model iterated p times. This further motivates the use of such predictors both in the cost function and for constraint tightening. We do so in our second contribution, since we propose a new robust MPC scheme that explicitly relies on the optimal SM multi-step models, thus dramatically reducing conservativeness. To deal with the particular structure of the multi-step predictors, which prevents the use of a standard robust MPC approach, we adopt a novel multi-rate receding horizon strategy, for which we prove guaranteed constraint satisfaction and convergence properties. Many multirate schemes have been proposed in the literature for predictive control design, see for example [5], [8], [7] and the references therein, usually to cope with different sampling rates in outputs sampling,

state update, and control implementation. On the contrary, here the multirate implementation stems from the particular form of the predictors.

In the last section of the paper, the new approach is compared with that of [6] in a simulation example. The proofs of the main results are reported in Appendix.

Notation: I_n is the identity matrix of dimension n, $\bar{I}_n$ is the matrix with zero entries except for those on the anti-diagonal, which are equal to 1, $0_{m,n}$ is the null matrix of dimensions m and n. The Cartesian product between n sets $\mathbf{T}_1,\ldots,\mathbf{T}_n$ is $\prod_{i=1}^{n} \mathbf{T}_i$. For a generic vector x, $\|x\|^2 \doteq x^T x$ and $\|x\|_Q^2 \doteq x^T Q x$ with Q being a given square matrix of suitable dimension. For a matrix A, $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$ is its induced 2-norm and $\rho(M)$ its spectral radius, i.e. the maximum absolute value of its eigenvalues. Given sets $\mathbf{A}, \mathbf{B} \subset \mathbb{R}^n$, $\mathbf{A} \oplus \mathbf{B} = \{a+b : a \in \mathbf{A}, b \in \mathbf{B}\}$ and $\mathbf{A} \ominus \mathbf{B} = \{a \in \mathbf{A} : \forall b \in \mathbf{B}, a+b \in \mathbf{A}\}$.

II. PROBLEM STATEMENT, IDENTIFICATION ALGORITHM, AND ERROR BOUNDS

Consider a linear and time-invariant (LTI) discrete-time system of order n with input $u(k) \in \mathbb{R}$, output $z(k) \in \mathbb{R}$, measured output $y(k) \in \mathbb{R}$, process disturbance $v(k) \in \mathbb{R}$, and measurement disturbance $d(k) \in \mathbb{R}$, where $k \in \mathbb{N}$ is the discrete time variable. We define $\varphi_z^{(p)}(k) \in \mathbb{R}^{2n+p-1}$ as:

$$\varphi_z^{(p)}(k) = [z(k),\ldots,z(k-n+1), \tag{1}$$
$$u(k-1),\ldots,u(k-n+1),u(k),\ldots,u(k+p-1)]^T,$$

with $p \in \mathbb{N}$. The system can be expressed in ARX (autoregressive-exogenous) form as

$$\begin{cases} z(k+1) = \bar{\theta}^{(1)^T} \varphi_z^{(1)}(k) + v(k) \\ y(k) = z(k) + d(k), \end{cases} \tag{2}$$

where $\bar{\theta}^{(1)} \in \mathbb{R}^{2n}$ is the vector of unknown parameters.

Assumption 1: (Disturbance boundedness). $|v(k)| \leq \bar{v}$, $|d(k)| \leq \bar{d}$, $\forall k \in \mathbb{N}$ with $\bar{d}$ known. $\square$
The value of $\bar{d}$ is assumed to be available from prior knowledge, and/or it can also be estimated from data, see e.g. [1], whereas $\bar{v}$ is not necessarily known.
Using the SM method presented in [6], the following predictors of order o can be obtained for all p values up to a finite horizon $\bar{p}$:

$$\hat{z}(k+p) = \hat{\theta}^{(p)*^T} \varphi_y^{(p)}(k), \tag{3}$$

where $\hat{\theta}^{(p)*} = [\hat{\theta}_{AR}^{(p)*^T} \quad \hat{\theta}_U^{(p)*^T} \quad \hat{\theta}_{\bar{U}}^{(p)*^T}]^T$ and $\hat{\theta}_{AR}^{(p)*^T} \in \mathbb{R}^o$, $\hat{\theta}_U^{(p)*^T} \in \mathbb{R}^{o-1}, \hat{\theta}_{\bar{U}}^{(p)*^T} \in \mathbb{R}^p$ are vectors of known parameters

[1] The authors are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133, Milano, Italy. E-mail: `name.surname@polimi.it`

resulting from the identification phase. We refer to these predictors as *multi-step* in the remainder. The derivation of $\hat{\theta}^{(p)*}$ for a given value of p is recalled later on in this section. Moreover, in (3)

$$\varphi_y^{(p)}(k) = \begin{aligned}[t] &[y(k),\ldots,y(k-o+1), \\ &u(k-1),\ldots,u(k-o+1),u(k),\ldots u(k+p-1)]^T \end{aligned}$$

Assumption 2: (Model order) The order of the models (3) is $o \geq n$ □

An algorithm to estimate o is described in [1]. The SM learning phase also returns an estimate of the bound on the worst-case prediction error:

$$|z(k+p) - \hat{z}(k+p)| \leq \hat{\tau}_p(\hat{\theta}^{(p)*}) \qquad (4)$$

In fact, for each step $p \leq \bar{p}$ one can derive a guaranteed upper bound $\hat{\tau}_p$ of the difference between the nominal output and its prediction obtained with a generic predictor

$$\hat{z}(k+p) = \hat{\theta}^{(p)^T} \varphi_y^{(p)}(k) \qquad (5)$$

For the identification of $\hat{\theta}^{(p)*}$ and $\hat{\tau}_p$ a finite number N of measured data is available, composed of pairs $(\varphi_y^{(p)}(k), y(k+p))$, $k = 1,\ldots,N$. We first estimate an error bound $\hat{\bar{\varepsilon}}_p = \alpha \underline{\lambda}_p$, $\forall p = 1,\ldots,\bar{p}$, through

$$\underline{\lambda}_p = \min_{\theta^{(p)},\lambda} \lambda, \quad \text{subject to}$$
$$|y(k+p) - \theta^{(p)^T} \varphi_y^{(p)}(k)| \leq \lambda + \bar{d}, \ k = 1,\ldots,N,$$

The latter value is inflated by a scalar $\alpha > 1$ to account for the fact that the available dataset is finite. The Feasible Parameter Sets (FPSs) are then defined as

$$\Theta^{(p)} = \Big\{ \hat{\theta}^{(p)} : |y(k+p) - \hat{\theta}^{(p)^T} \varphi_y^{(p)}(k)| \leq \hat{\bar{\varepsilon}}_p + \bar{d}, \\ k = 1,\ldots,N \Big\} \qquad (6)$$

For each p, $\Theta^{(p)}$ is a convex set and, if the data are informative enough, it is also compact. This property can be checked easily by linear programming; if the set $\Theta^{(p)}$ is not bounded then this is a sign that more informative data should be collected. In the remainder, we consider that $\Theta^{(p)}$ is compact for any p. Let us further denote with $\Phi^{(p)} \subseteq \mathbb{R}^{2o-1+p}$ a compact set containing all possible values of $\varphi_y^{(p)}(k)$. In practice, this means that we restrict our analysis and results to a set of system trajectories of interest, which contains the available data points. This is a reasonable assumption in practice. Since $\bar{\theta}^{(p)}$ in (2) belongs to $\Theta^{(p)}$, the smallest bound on the error $|z(k+p) - \hat{z}(k+p)|$ (see (4)) $\forall p = 1,\ldots,\bar{p}$ is:

$$|z(k+p) - \hat{z}(k+p)|$$
$$\leq \max_{\varphi_y^{(p)} \in \Phi^{(p)}} \max_{\theta^{(p)} \in \Theta^{(p)}} |(\theta^{(p)} - \hat{\theta}^{(p)})^T \varphi_y^{(p)}| + \hat{\bar{\varepsilon}}_p = \tau_p(\hat{\theta}^{(p)})$$
$$(7)$$

The bound (7) is global, since it holds for any regressor value inside $\Phi^{(p)}$ and for any model compatible with the data, i.e. contained in the set $\Theta^{(p)}$. However it cannot be computed in practice since the set $\Phi^{(p)}$ is not available. On the other

hand, an approximation $\hat{\tau}_p(\hat{\theta}^{(p)}) \approx \tau_p(\hat{\theta}^{(p)})$ can be easily computed as $\hat{\tau}_p(\hat{\theta}^{(p)}) = \gamma \underline{\tau}_p(\hat{\theta}^{(p)})$ with

$$\underline{\tau}_p(\hat{\theta}^{(p)}) = \max_{k=1,\ldots,N} \max_{\theta^{(p)} \in \Theta^{(p)}} |(\theta^{(p)} - \hat{\theta}^{(p)})^T \varphi_y^{(p)}(k)| + \hat{\bar{\varepsilon}}_p \quad (8)$$

i.e. by computing the worst-case prediction error with respect to the available data. This approximation includes a second scaling factor $\gamma \geq 1$, again to account for the finite available dataset. The nominal predictor (3), for each step p, is chosen as the minimizer of this worst case error $\underline{\tau}_p(\hat{\theta}^{(p)})$, i.e.

$$\hat{\theta}^{(p)*} = \arg \min_{\hat{\theta}^{(p)} \in \Theta^{(p)}} \underline{\tau}_p(\hat{\theta}^{(p)}) \qquad (9)$$

The following theorem is concerned with the optimality (in terms of size of the uncertainty bound) of the multistep prediction models.

Theorem 1: Consider any 1-step-ahead LTI system model (i.e. of the form (5) with $p = 1$) with coefficient vector $\hat{\theta}^{(1)} \in \mathbb{R}^{2o}$. Let $\hat{\theta}^{(p),1} \in \mathbb{R}^{2o-1+p}$ be the corresponding vector of multi-step predictor coefficients, obtained by iterating p times such a 1-step-ahead model. Then, for all $p = 1,\ldots,\bar{p}$ it holds:

$$\tau_p(\hat{\theta}^{(p)*}) \leq \tau_p(\hat{\theta}^{(p),1}). \qquad (10)$$

Proof: See the Appendix. ∎

Theorem 1 justifies the use of multi-step models for robust MPC design, since in general they yield smaller error bounds.

III. MPC DESIGN AND PROPERTIES

The multi-step models previously introduced can not be directly used in existing robust MPC schemes. Therefore we propose a new multirate MPC approach where the predicted behavior of the system is optimized by considering a prediction/control horizon of N_p "long" steps, with index $j \in \mathbb{N}$, each one consisting of $\bar{p}$ "short" sampling times with index k. Note that the "short" sampling interval is the one assumed for the true system (2). The optimal control problem is thus solved at every long step j (i.e. every $\bar{p}$ short steps) and the solution provides the values of the control input to be applied at each step k in the interval $\{j\bar{p},\ldots,(j+1)\bar{p}-1\}$ according to a standard receding horizon formulation. For clarity, we represent the long and short sampling times on a common time-scale in Figure 1. Also, in the remainder we will use the upper-case letters to denote variables defined at a long sampling time. Assume $\bar{p} > o$ for simplicity, although it is not necessary, and define the system state, the input, and the disturbance at time j as $X(j) = [y(j\bar{p}),\ldots y(j\bar{p}-o+1),u(j\bar{p}-1),\ldots,u(j\bar{p}-o+1)]^T$, $U(j) = [u(j\bar{p}),\ldots,u(j\bar{p}+\bar{p}-1)]^T$, $W(j) = [w_1(j\bar{p}),\ldots,w_{\bar{p}}(j\bar{p})]^T$, respectively. Denote with $\bar{w}_p$ a value such that $|w_p(j\bar{p})| \leq \bar{w}_p$, for all $p = 1,\ldots,\bar{p}$, which accounts for the error stemming from the identification procedure, the process noise, and the measurement disturbance. Given the bound (8), since the state $X(j)$ comprises samples of the measured output y affected by measurement noise d, it is possible to obtain $\bar{w}_p$ as

$$\bar{w}_p = \hat{\tau}_p(\hat{\theta}^{(p)*}) + \bar{d} \qquad (11)$$
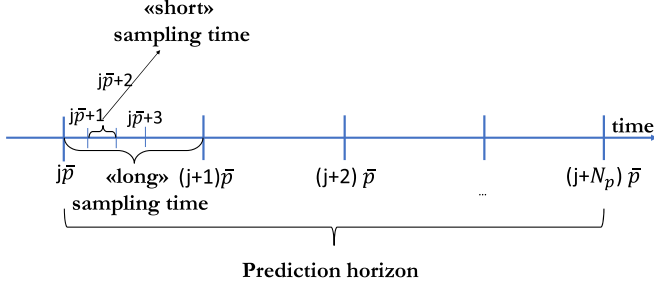
Fig. 1. Sketch of the time-scales involved in the simulation, with "long" and "short" sampling times.

thus directly exploiting the multi-step error bounds previously obtained. The state transition equation, that maps the current state $X(j)$ into the $\bar{p}$ steps ahead state $X(j+1)$, is:

$$X(j+1) = \bar{A}X(j) + \bar{B}U(j) + \bar{M}W(j) \tag{12}$$

where:

$$\bar{A} = \begin{bmatrix} \hat{\theta}_{AR}^{(\bar{p})*T} & \hat{\theta}_{U}^{(\bar{p})*T} \\ \vdots & \vdots \\ \hat{\theta}_{AR}^{(\bar{p}-o+1)*T} & \hat{\theta}_{U}^{(\bar{p}-o+1)*T} \\ 0_{o-1,o} & 0_{o-1,o-1} \end{bmatrix}, \bar{B} = \begin{bmatrix} \hat{\theta}_{\bar{U}}^{(\bar{p})*T} \\ \vdots \\ \begin{bmatrix} \hat{\theta}_{\bar{U}}^{(\bar{p}-o+1)*T} & 0_{1,o-1} \end{bmatrix} \\ \begin{bmatrix} 0_{o-1,\bar{p}-o+1} & \bar{I}_{o-1} \end{bmatrix} \end{bmatrix}$$

$$\bar{M} = \begin{bmatrix} 0_{o,\bar{p}-o} & \bar{I}_o \\ 0_{o-1,\bar{p}-o} & 0_{o-1,o} \end{bmatrix} \tag{13}$$

The following assumption is introduced.

Assumption 3: The pair $(\bar{A},\bar{B})$ is stabilizable. $\qquad\square$
Since the model is obtained from input-output data, Assumption 3 is usually satisfied in practice and is thus not restrictive. We rewrite models (4) as system output equations:

$$\hat{z}(j\bar{p}+p) = C_p X(j) + D_p U(j) \tag{14}$$

where $C_p = \begin{bmatrix} \hat{\theta}_{AR}^{(p)*T} & \hat{\theta}_{U}^{(p)*T} \end{bmatrix}$, $D_p = \begin{bmatrix} \hat{\theta}_{\bar{U}}^{(p)*T} & 0_{1,\bar{p}-p} \end{bmatrix}$.
Consistently with (4), we can write

$$z(j\bar{p}+p) = C_p X(j) + D_p U(j) + w_p(j\bar{p}) \tag{15}$$

For notational convenience let us stack matrices C_p and D_p, for all $p = 1,\dots\bar{p}$, as

$$\bar{C} = \begin{bmatrix} C_1^T & \dots & C_{\bar{p}}^T \end{bmatrix}^T, \quad \bar{D} = \begin{bmatrix} D_1^T & \dots & D_{\bar{p}}^T \end{bmatrix}^T \tag{16}$$

so that we can define the predictions of outputs in the long sampling time, but at a short sampling period basis, as $\hat{Z}(j) = \begin{bmatrix} \hat{z}(j\bar{p}+1) & \dots & \hat{z}(j\bar{p}+\bar{p}) \end{bmatrix}^T$. Thanks to the predictors in (14), we write

$$\hat{Z}(j) = \begin{bmatrix} \hat{z}(j\bar{p}+1) \\ \vdots \\ \hat{z}(j\bar{p}+\bar{p}) \end{bmatrix} = \bar{C}X(j) + \bar{D}U(j) \tag{17}$$

In the control design phase a tube-based robust control approach is used [2] and the input $U(j)$ is defined as

$$U(j) = \bar{U}(j) + K(X(j) - \bar{X}(j)) \tag{18}$$

The input $\bar{U}(j)$ will be computed by MPC, while the term $K(X(j) - \bar{X}(j))$ aims to reduce the error between the state $\bar{X}(j)$ of a suitably defined nominal dynamic system and the actual value of $X(j)$, available at time $k = j\bar{p}$. The gain K is chosen such that $\bar{F} = \bar{A} + \bar{B}K$ is Schur stable, which is possible thanks to Assumption 3.
The nominal dynamic system is defined based on (12):

$$\bar{X}(j+1) = \bar{A}\bar{X}(j) + \bar{B}\bar{U}(j) \tag{19}$$

The p steps ahead nominal output predictor corresponding to (14) is computed as:

$$\hat{\bar{Z}}(j) = \bar{C}\bar{X}(j) + \bar{D}\bar{U}(j) \tag{20}$$

The difference between the real available data vector $X(j)$ and the state of the nominal system is defined as $E(j) = X(j) - \bar{X}(j)$. From (12) and (19), it evolves according to:

$$E(j+1) = (\bar{A} + \bar{B}K)E(j) + \bar{M}W(j) \tag{21}$$

Let $\mathbb{E}$ be a robust positively invariant (RPI) [3] set for the system (21). Similarly to [2], the constraints and the optimization problem will be defined with reference to the nominal model (19). This will require to define suitable tightened state and input constraints, that allow one to account for the difference between $\bar{X}(j)$ and $X(j)$.

Remark 1: In (21) only the last o components of $W(j)$ are involved in the computation of $\mathbb{E}$, and they depend on the estimates $\hat{\tau}_p(\hat{\theta}^{(p)*})$ of the bounds proved to be optimal in Theorem 1, see (11). Moreover, since $\bar{A} + \bar{B}K$ is Schur stable and evolves over a (possibly long) $\bar{p}$-steps-ahead period, it is prone to have a smaller spectral radius and norm with respect to the one corresponding to a 1-step state space model, e.g. the one considered in [6]. Thus, this results in a smaller set $\mathbb{E}$ and less conservative constraint tightening, as also illustrated in the example of Section IV. $\qquad\square$
The MPC controller must guarantee the fulfillment of input and output constraints for all $k \geq 0$:

$$u(k) \in \mathbb{U} \quad, \quad z(k) \in \mathbb{Z} \tag{22}$$

where $\mathbb{U}$ and $\mathbb{Z}$ are suitable convex sets containing the origin in their interior. For ease of notation, let us introduce the higher-dimensional convex sets $\mathbf{U} = \mathbb{U}^{\bar{p}}$ and $\mathbf{Z} = \mathbb{Z}^{\bar{p}}$. Similarly to [2], it is first necessary to constrain $\bar{X}(j)$ at time $j\bar{p}$ to lie in the neighborhood of $X(j)$, i.e

$$X(j) - \bar{X}(j) \in \mathbb{E} \tag{23a}$$

Regarding the input variable, to guarantee that (22) holds from time $j\bar{p}$ to $(j+N_p-1)\bar{p}$, it is enough to enforce the following tightened constraints, for all $i = 0,\dots,N_p - 1$.

$$\bar{U}(j+i) \in \mathbf{U} \ominus K\mathbb{E} \tag{23b}$$

As for the output, to guarantee that (22) holds at time $j\bar{p} + 1,\dots,(j+N_p)\bar{p}$, we define $\forall p = 1,\dots,\bar{p}$

$$\mathbb{T}_p = \{t \in \mathbb{R} : |t| \leq \hat{\tau}_p(\hat{\theta}^{(p)*})\}$$

and the tightened set $\hat{\mathbf{Z}}$ as

$$\hat{\mathbf{Z}} = \mathbf{Z} \ominus \prod_{p=1}^{\bar{p}} \mathbb{T}_p \qquad (23c)$$

This set is such that, by construction, if $\hat{Z}(j+i) \in \hat{\mathbf{Z}}$, then $Z(j+i) \in \mathbf{Z}$, $i = 0, \ldots, N_p - 1$. We thus enforce the following tightened constraint, again related to the nominal system (19), for all $i = 0, \ldots, N_p - 1$.

$$\hat{\bar{Z}}(j+i) \in \hat{\mathbf{Z}} \ominus (\bar{C} + \bar{D}K)\mathbb{E} \qquad (23d)$$

Finally, to guarantee recursive feasibility, we also need to enforce a terminal constraint of the type

$$\bar{X}(j+N_p) \in \mathbb{X}_F \qquad (23e)$$

where $\mathbb{X}_F$ is defined as a positively invariant set for the system $\hat{X}(j+1) = (\bar{A} + \bar{B}K)\hat{X}(j)$ that verifies
- $(\bar{C} + \bar{D}K)\mathbb{X}_F \subseteq \hat{\mathbf{Z}} \ominus (\bar{C} + \bar{D}K)\mathbb{E}$
- $K\mathbb{X}_F \subseteq \mathbf{U} \ominus K\mathbb{E}$

For consistency, the following assumption is required.

Assumption 4: There exists a ball $\mathscr{B}$ in space $\mathbb{R}^{\bar{p}}$, centered at the origin and with radius ε, such that

$$(\bar{C} + \bar{D}K)\mathbb{E} \oplus \mathscr{B} \subseteq \hat{\mathbf{Z}} \qquad (24a)$$
$$K\mathbb{E} \oplus \mathscr{B} \subseteq \mathbf{U} \qquad (24b)$$

$\square$

The cost function to be minimized at time step k is

$$J(j) = \sum_{i=0}^{N_p-1} \|\hat{\bar{Z}}(j+i)\|_Q^2 + \|\bar{U}(j+i)\|_R^2 + \|\bar{X}(j+N_p)\|_P^2$$

where $Q = \mathrm{diag}(q_1, \ldots, q_{\bar{p}}) > 0$, $R = \mathrm{diag}(r_0, \ldots, r_{\bar{p}-1}) > 0$, N_p is the prediction horizon, and P is the unique positive definite solution to the Riccati equation (see Assumption 3)

$$\bar{F}^T P \bar{F} - P = -\left(\bar{G}^T Q \bar{G} + K^T R K\right) \qquad (25)$$

where $\bar{G} = (\bar{C} + \bar{D}K)$. Note that Q and R can be chosen freely while in [6] they were selected according to the solution to an LMI problem, so limiting the possible trade-offs between bandwidth and control activity of the closed-loop system. Now, denoting the vector of decision variables with

$$\bar{\mathbf{U}}(j) = \begin{bmatrix} \bar{U}(j)^T & \ldots & \bar{U}(j+N_p-1)^T \end{bmatrix}^T,$$

the optimization problem to be solved at each "long" sampling time $j \geq 0$, reads

$$J^*(j) = \min_{\bar{X}(j), \bar{\mathbf{U}}(j)} J(j) \text{ s.t. constraints (23)} \qquad (26)$$

If problem (26) is feasible, its solution is denoted with $\bar{X}^*(j), \bar{\mathbf{U}}^*(j) = [\bar{U}^*(j)^T, \ldots, \bar{U}^*(j+N_p-1)^T]^T$, and the input sequence $U^*(j) = \bar{U}^*(j) + K(X(j) - \bar{X}^*(j))$ in (18) is applied to the system according to the Receding Horizon principle. Also, we denote with $\bar{X}^*(j+i)$ the future nominal state predictions generated using (19) with input $\bar{\mathbf{U}}^*(j)$, as well as all the other derived quantities, such as $\hat{\bar{Z}}^*(j)$ (see (20)).

Theorem 2: If (26) is feasible at time step $j = 0$ then it is feasible at all time steps $j > 0$ and, for all $j \geq 0$, the constraints (22) are satisfied. Moreover, $\hat{\bar{Z}}^*(j) \to 0$ as $j \to \infty$. Finally, $\delta(Z(j), (\bar{C} + \bar{D}K)\mathbb{E}) \to 0$ as $j \to \infty$, where $\delta(\alpha, \beta)$ denotes the distance between point α and set β $\qquad \square$.

Proof: See the Appendix. ∎

IV. SIMULATION EXAMPLE

Consider the system employed in [6], obtained by discretizing, with sampling time $T_s = 0.1$, the continuous-time transfer function

$$G(s) = \frac{160}{(s+10)(s^2 + 1.6s + 16)}$$

A dataset of 1000 pairs (u, y) has been collected by exciting the system with a signal u taking value in $\{-1, 0, 1\}$ randomly each 5 units of time, and adding the disturbance $v(k)$ and $d(k)$, with $\bar{v} = 0.01$ and $\bar{d} = 0.1$, respectively, consistently with (2). The multi-step bounds estimates $\hat{\tau}_p(\hat{\theta}^{(p)*})$ have been computed according to the algorithm described in [6], with $\bar{p} = 10$ (resulting in a "long" sampling time equal to $T_s\bar{p} = 1s$) and model order $o = 4$. In Figure 2 they are plotted and compared with the bounds computed by simply iterating the simulation model (i.e., the 1-step ahead predictor) and propagating its uncertainty bound accordingly.

In the control design phase, the matrix K has been computed with Linear Quadratic (LQ) control, while the prediction horizon for the MPC controller is $N_p = 3$. The weighting matrices are defined as $Q = 100I_{\bar{p}}$ and $R = 1I_{\bar{p}}$, while matrix P is obtained thanks to (25). Both the input u and the output z have been enforced to belong to the set $[-10, 10]$ for each time instant.

The input and output trajectories, comparing the closed-loop with the open-loop response of the system, are plotted in Figures 3 and 4 together with the relevant bounds. The controller, based on the identified model, is able to regulate the real system (2) to zero with a much faster time constant and sensibly damping the oscillations. In Table I we also report, for the same tuning of the LQ problem, the spectral radius and norm of the state transition matrix of the nominal system (19) subject to the auxiliary law K, see also Remark 1. Note that the norm of such matrix directly affects the computation of the invariant set $\mathbb{E}$. Moreover, by comparing the effect on the constraint tightening, we note that, while in [6] the tightened output constraints correspond to the interval $[-7.7, 7.7]$ for each prediction step and the input constraints to the interval $[-9.05, 9.05]$, with the new algorithm proposed. here we obtain the following box-inequalities, to be intended entry-wise, $i = 0, \ldots, N_p - 1$: Specifically, define

$$Z_m = \begin{bmatrix} 8.3 & 7.4 & 7.8 & 8.2 & 8.8 & 9.0 & 9.0 & 9.3 & 9.1 & 8.9 \end{bmatrix}$$
$$U_m = \begin{bmatrix} 9.77 & 9.68 & 9.72 & 9.60 & 9.53 & 9.60 & 9.67 & 9.88 & 9.87 & 9.92 \end{bmatrix}$$

and the constraints

$$-Z_m^T \leq \hat{\bar{Z}}(j+i) \leq Z_m^T \text{ and } -U_m^T \leq \bar{U}(j+i) \leq U_m^T$$

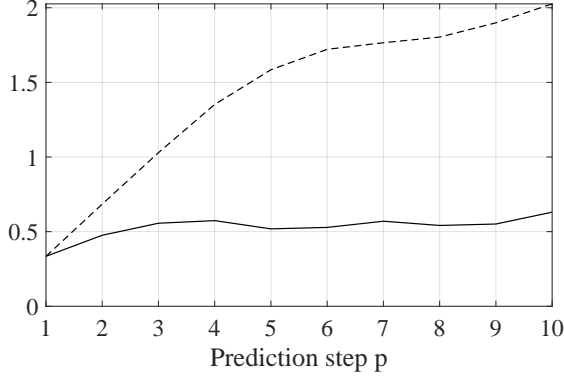which confirm a conservativeness reduction.

Fig. 2. Computed bounds. Dashed line: bound obtained by iterating $\bar{w}_1$ with the one-step model, solid line: bounds $\bar{w}_p, p = 1, \ldots, \bar{p}$
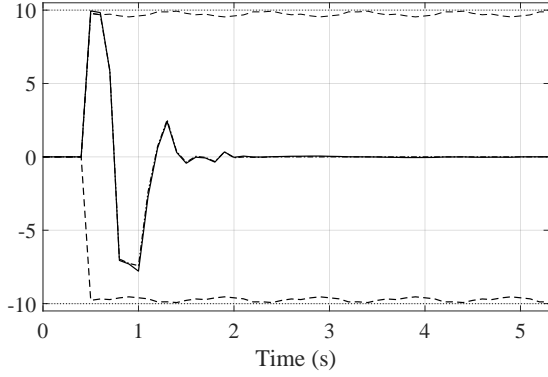

Fig. 3. Input variable. Dash-dotted line: $\bar{U}(k)$, solid line: $U(k)$, dashed lines: tightened constraints (23b), dotted lines: absolute constraints.
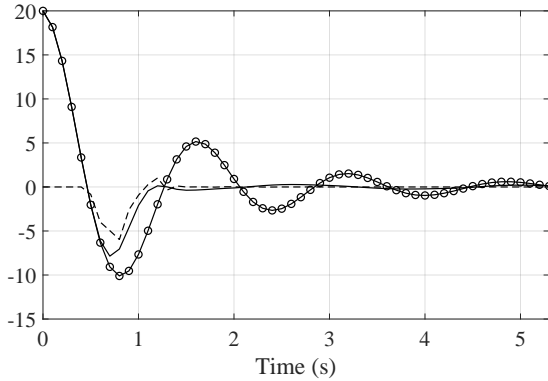

Fig. 4. Output variable. Solid line: $z(k)$, dashed line: $\hat{\bar{Z}}(j)$, line with circles: open loop response.

Algorithm in [6]	Proposed algorithm
$\rho(A + B_1 K) = 0.78$	$\rho(\bar{A} + \bar{B}K) = 0.2974$
$\|A + B_1 K\| = 1.77$	$\|\bar{A} + \bar{B}K\| = 0.455$

APPENDIX

Proof of Theorem 1.

We derive a multi-step predictor by iterating a generic (simulation, i.e. 1-step ahead) model with coefficient vector $\hat{\theta}^{(1)}$ and focusing on the function linking prediction steps $p = 1$ and $p = 2$. The extension up to $\bar{p}$ is straightforward. In the following we will use $'$ on the variables predicted with the (possibly iterated) simulation model.

First recall the definition of the one-step predictor regressor vector $\varphi_y^{(1)}(k) = \begin{bmatrix} y(k) & \ldots & y(k - o + 1) & u(k - 1) & \ldots & u(k - o + 1) & u(k) \end{bmatrix}^T$ and note that $\varphi_y^{(2)}(k) = \begin{bmatrix} \varphi_y^{(1)}(k) \\ u(k+1) \end{bmatrix}$.

Assuming to be at time k+1, to proceed 1-step ahead we would need $y(k+1)$ to compute

$$\hat{z}'(k+2) = \hat{\theta}^{(1)^T} \varphi_y^{(1)}(k+1) \tag{27}$$

If we are at time k, the measured value of $y(k+1)$ is not available, hence its nominal prediction computed with the simulation model is used in its place, i.e. $\hat{z}'(k+1) = \hat{\theta}^{(1)^T} \varphi_y^{(1)}(k)$. This results in

$$\hat{z}'(k+2) = \hat{\theta}^{(1)^T} \varphi_y^{(1)'}(k+1) \tag{28}$$

where $\varphi_y^{(1)'}(k+1) = [\hat{z}'(k+1) \ldots y(k - o + 2) u(k - 1) \ldots u(k - o + 2), u(k), u(k+1)]^T$, which can be expressed as a function of $\varphi_y^{(1)}(k)$ as $\varphi_y^{(1)'}(k+1) =$

$$= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ u(k+1) \end{bmatrix} + \begin{bmatrix} \hat{\theta}^{(1)*^T} \\ I_{o-1} & 0_{o-1,1} & 0_{o-1,o} \\ 0_{1,o-1} & 0 & [0_{1,o-1} \quad 1] \\ 0_{o-2,o-1} & 0_{o-2,1} & [I_{o-2} \quad 0_{o-2,2}] \\ 0_{1,o-1} & 0 & 0_{1,o} \end{bmatrix} \varphi_y^{(1)}(k) \tag{29}$$

or, in shorter notation $\varphi_y^{(1)'}(k+1) = \mathcal{U}_1 + S(\hat{\theta}^{(1)})\varphi_y^{(1)}(k)$, where $\mathcal{U}_1$ and $S(\hat{\theta}^{(1)})$ are are implicitly defined in (29).

By replacing (29) in (28) we get

$$\begin{aligned} \hat{z}'(k+2) &= \hat{\theta}^{(1)^T} \left(\mathcal{U}_1 + S(\hat{\theta}^{(1)})\varphi_y^{(1)}(k) \right) \\ &= \underbrace{\begin{bmatrix} \hat{\theta}^{(1)^T} S(\hat{\theta}^{(1)}) & \hat{\theta}_{last}^{(1)} \end{bmatrix}}_{\hat{\theta}^{(2),1}} \begin{bmatrix} \varphi_y^{(1)}(k) \\ u(k+1) \end{bmatrix} \\ &= \hat{\theta}^{(2),1} \varphi_y^{(2)}(k) \end{aligned} \tag{30}$$

where $\hat{\theta}_{last}^{(1)}$ is the last element of $\hat{\theta}^{(1)}$, and $\hat{\theta}^{(2),1}$ is introduced. Note that the entries of $\hat{\theta}^{(2),1}$ are polynomial combinations of the ones of $\hat{\theta}^{(1)}$, see (30). With similar

manipulations, for any step p the predictor obtained by iterating the simulation model with parameters $\hat{\theta}^{(1)}$ reads

$$\hat{z}'(k+p) = \hat{\theta}^{(p),1^T} \varphi_y^{(p)}(k) \qquad (31)$$

where $\hat{\theta}^{(p),1}(\hat{\theta}^{(1)}) : \mathbb{R}^{2o} \to \mathbb{R}^{2o+p-1}$ is a vector function of polynomials of degree up to p of the elements of $\hat{\theta}^{(1)}$.

Let us now focus on the worst case prediction error, with arguments similar to (7):

$$|z(k+p) - \hat{z}'(k+p)| \leq$$
$$\leq |\varphi_y^{(p)^T}(\bar{\theta}^{(p)} - \hat{\theta}^{(p),1})| + \bar{\varepsilon}_p$$
$$\leq \max_{\theta \in \Theta^{(p)}} \max_{\varphi_y^{(p)} \in \Phi^{(p)}} |\varphi_y^{(p)^T}(\theta - \hat{\theta}^{(p),1})| + \bar{\varepsilon}_p = \tau_p(\hat{\theta}^{(p),1})$$
$$(32)$$

where the last inequality holds thanks to the fact that $\bar{\theta}^{(p)} \subseteq \Theta^{(p)}$. We now aim to show that $\tau_p(\hat{\theta}^{(p),1}) \geq \tau_p(\hat{\theta}^{(p)*})$. For a given vector $\hat{\theta}^{(p),1}$, one of these two cases occur:

- If $\hat{\theta}^{(p),1} \in \Theta^{(p)}$, then from (32), it follows

$$\tau_p(\hat{\theta}^{(p)*}) = \min_{\bar{\theta}^{(p)} \in \Theta^{(p)}} \max_{\theta \in \Theta^{(p)}} \max_{\varphi_y^{(p)} \in \Phi^{(p)}} |\varphi_y^{(p)^T}(\theta - \bar{\theta}^{(p)})| + \bar{\varepsilon}_p$$

$$\leq \tau_p(\hat{\theta}^{(p),1})$$

hence proving the claim.

- If $\hat{\theta}^{(p),1} \notin \Theta^{(p)}$, let us consider a generic element $\theta \in \Theta^{(p)}$ and the convex combination

$$(1 - \bar{\alpha})\theta + \bar{\alpha}\hat{\theta}^{(p),1} = \theta^{(p)''} \qquad (33)$$

where

$$\bar{\alpha} = \max_{\alpha} \alpha$$
$$s.t. \quad \alpha \in [0,1]$$
$$(1 - \alpha)\theta + \alpha\hat{\theta}^{(p),1} \in \Theta^{(p)}$$

The point $\theta^{(p)''}$, belongs to the boundary of $\Theta^{(p)}$ along the direction connecting the chosen $\theta \in \Theta^{(p)}$ to $\hat{\theta}^{(p),1} \notin \Theta^{(p)}$. Consider (32), omitting the dependence on time k for brevity we compute

$$|\varphi_y^{(p)^T}(\theta - \hat{\theta}^{(p),1})| =$$
$$= |\varphi_y^{(p)^T}(1 - \bar{\alpha})(\theta - \hat{\theta}^{(p),1})| + |\varphi_y^{(p)^T}\bar{\alpha}(\theta - \hat{\theta}^{(p),1})|$$
$$= |\varphi_y^{(p)^T}(\theta^{(p)''} - \hat{\theta}^{(p),1})| + |\varphi_y^{(p)^T}(\theta - \theta^{(p)''})|$$
$$(34)$$

where in the last equality the term $\theta^{(p)''}$ as defined in (33) has been substituted. The latter expression allows us to split the value

$$|\varphi_y^{(p)^T}(\theta - \hat{\theta}^{(p),1})|$$

in a contribution given by a predictor θ inside $\Theta^{(p)}$, namely $|\varphi_y^{(p)^T}(\theta - \theta^{(p)''})|$, plus a contribution outside $\Theta^{(p)}$, that is $|\varphi_y^{(p)^T}(\theta^{(p)''} - \hat{\theta}^{(p),1})|$. Therefore we can write $\tau_p(\hat{\theta}^{(p),1}) =$

$$= \max_{\theta \in \Theta^{(p)}} \max_{\varphi_y^{(p)} \in \Phi^{(p)}} |\varphi_y^{(p)^T}(\theta - \hat{\theta}^{(p),1})| + \bar{\varepsilon}_p$$

$$\geq \min_{\theta^{(p)''} \in \Theta^{(p)}} \max_{\theta \in \Theta^{(p)}} \max_{\varphi_y^{(p)} \in \Phi^{(p)}} |\varphi_y^{(p)^T}(\theta - \theta^{(p)''})| + \bar{\varepsilon}_p = \tau_p(\hat{\theta}^{(p)*})$$
$$(35)$$

that completes the proof. Note that in (35), for a given $\hat{\theta}^{(p),1}$, $\theta^{(p)''}$ depends only on θ. $\qquad \square$

We first prove recursive feasibility by induction. Assume that, at $k = j\bar{p}$, a solution to the optimization problem (26) exists and denote it with $\bar{X}^*(j|j)$, $\bar{U}^*(j|j)$. All constraints (23) are therefore verified by the nominal state trajectories associated with the optimal solution $\bar{X}^*(j+i|j)$ and $\bar{U}^*(j|j)$:

$$X(j) - \bar{X}^*(j|j) \in \mathbb{E} \qquad (36a)$$
$$\bar{U}^*(j+i|j) \in \mathbf{U} \ominus K\mathbb{E} \qquad (36b)$$
$$\hat{Z}^*(j+i|j) = \bar{C}\bar{X}^*(j+i|j) + \bar{D}\bar{U}^*(j+i|j)$$
$$\in \hat{\mathbf{Z}} \ominus (\bar{C} + \bar{D}K)\mathbb{E} \qquad (36c)$$
$$\bar{X}^*(j+N_p|j) \in \mathbb{X}_F \qquad (36d)$$

with $i = 0, \ldots, N_p - 1$. Finally, the input $U(j)$ is defined according to (18) with $\bar{U}(j) = \bar{U}^*(j|j)$ and $\bar{X}(j) = \bar{X}^*(j|j)$. Let us call this quantity $U^*(j|j)$. At yime $k = (j+1)\bar{p}$,

$$X(j+1) = \bar{A}X(j) + \bar{B}U^*(j|j) + \bar{M}W(j)$$

We can show that a feasible, although possibly suboptimal, solution to (26) can be defined, i.e., as $\bar{X}^*(j+1|j), \tilde{\bar{U}}^*(j+1|j) = (\bar{U}^*(j+1|j), \ldots, \bar{U}^*(j+N_p-1|j), K\bar{X}^*(j+N_p|j))$.

First of all, we have

$$X(j+1) - \bar{X}^*(j+1|j) =$$
$$(\bar{A} + \bar{B}K)(X(j) - \bar{X}^*(j|j)) + \bar{M}W(j) \in \mathbb{E}$$

in view of (36a) and of the fact that $\mathbb{E}$ is RPI. Moreover, $\bar{U}^*(j+i|j) \in \mathbf{U} \ominus K\mathbb{E}$ in view of (36b), for all $i = 1, \ldots, N_p - 1$, and $K\bar{X}^*(j+N_p|j) \in K\mathbb{X}_F \subseteq \mathbf{U} \ominus K\mathbb{E}$ in view of (36d) and of (23b). In addition,

$$(\bar{C} + \bar{D}K)\bar{X}^*(j+i|j) \in \hat{\mathbf{Z}} \ominus (\bar{C} + \bar{D}K)\mathbb{E}$$

for all $i = 1, \ldots, N_p - 1$ in view of (36c) and

$$(\bar{C} + \bar{D}K)\bar{X}^*(j+N_p|j) \in (\bar{C} + \bar{D}K)\mathbb{X}_F \subseteq \hat{\mathbf{Z}} \ominus (\bar{C} + \bar{D}K)\mathbb{E}$$

in view of (36d) and of (23b).

Finally, it holds that

$$\bar{X}^*(j+N_p+1|j) = (\bar{A} + \bar{B}K)\bar{X}^*(j+N_p|j) \in \mathbb{X}_F$$

in view of (36d) and of the positive invariance of $\mathbb{X}_F$. Since feasibility holds by assumption at time $j\bar{p}$, $j = 0$ then, by induction, it is guaranteed also for all $j > 0$.

Constraint satisfaction.
Constraint satisfaction is now proven. In view of the feasibility of the problem (26) at any time instant $j\bar{p}$, $j \geq 0$, it results that constraints (36) are verified. Therefore, from (18), (36a), and (36b),

$$U^*(j|j) = \bar{U}^*(j|j) + K(X(j) - \bar{X}^*(j|j)) \in (\mathbf{U} \ominus K\mathbb{E}) \oplus K\mathbb{E} \subseteq \mathbf{U}.$$

Then, by definition of $U(j)$ and the set $\mathbf{U}$, input constraints satisfaction in (22) follows. Also, from (36a) and (36c),

$$\hat{Z}(j|j) = \bar{C}\bar{X}^*(j|j) + \bar{D}U^*(j|j) + (\bar{C} + \bar{D}K)(X(j) - \bar{X}^*(j|j))$$
$$\in (\hat{\mathbf{Z}} \ominus (\bar{C} + \bar{D}K)\mathbb{E}) \oplus (\bar{C} + \bar{D}K)\mathbb{E} \subseteq \hat{\mathbf{Z}}.$$

Indeed, if $\hat{Z}(j|j) \in \hat{\mathbf{Z}}$, then $Z(j|j) \in \mathbf{Z}$ follows, and eventually by definition of the latter ones, (22) is satisfied.

Convergence is proven with standard arguments (see [4]) by showing that the optimal cost function is decreasing in time, i.e.

$$J^*(j+1|j+1) - J^*(j|j) \leq -\left(\|\hat{\bar{Z}}^*(j|j)\|_Q^2 + \|\hat{\bar{U}}^*(j|j)\|_R^2 \right) \tag{37}$$

Since $J(j)$ is positive by definition, and decreasing in view of (37), then $\hat{\bar{Z}}^*(j|j)$ and $\bar{U}^*(j|j) \to 0$ as $j \to +\infty$. Also, recalling (36a), it holds

$$\bar{C}(X(j) - \bar{X}^*(j|j)) + \bar{D}(U^*(j|j) - \bar{U}^*(j|j)) = \hat{Z}(j|j) - \hat{\bar{Z}}^*(j|j)$$
$$\in (\bar{C} + \bar{D}K)\mathbb{E},$$

for all j which concludes the proof. $\qquad\square$

REFERENCES

[1] M. Lauricella and L. Fagiano. Set membership identification of linear systems with guaranteed simulation accuracy. *IEEE Transactions on Automatic Control*, 65(12):5189–5204, 2020.

[2] D.Q. Mayne, M.M. Seron, and S.V. Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, 2005.

[3] S. V. Rakovic and K. I. Kouramas. Invariant approximations of the minimal robust positively invariant set via finite time aumann integrals. In *46th IEEE Conference on Decision and Control*, pages 194–199, 2007.

[4] J.B. Rawlings and D.Q. Mayne. Model predictive control: Theory and design. 2009.

[5] R. Scattolini and N Schiavoni. A multirate model based predictive controller. In *Proceedings of 1994 33rd IEEE Conference on Decision and Control*, volume 1, pages 243–248. IEEE, 1994.

[6] E. Terzi, L. Fagiano, M. Farina, and R. Scattolini. Learning-based predictive control for linear systems: A unitary approach. *Automatica*, 108:108473, 2019.

[7] J. Wang, Y. Song, and G. Wei. Robust model predictive control for multirate systems with model uncertainties and circular scheduling. *International Journal of Robust and Nonlinear Control*, 30(18):8206–8227, 2020.

[8] X. Wang, B. Huang, and T. Chen. Multirate minimum variance control design and control performance assessment: A data-driven subspace approach. *IEEE Transactions on Control Systems Technology*, 15(1):65–74, 2006.