

Multivariable MPC Design Based on a Frequency Response Approximation Approach

Gaurang Shah and Sebastian Engell

Abstract—This paper presents an improvement to the MPC tuning approach presented in [1]. The drawbacks of the former approach are discussed and a new tuning approach is proposed to overcome them. The primary aim is to treat linear MPC as a classical controller design problem and to use the tools of linear control theory for determining the tuning parameters. MPC tuning is performed for a desired open-loop frequency response which results from an achievable closed-loop performance determined using the Youla parameterization technique. The tuning approach is performed in two steps and both resulting minimization problems are convex optimization problems. The approach is tested on a challenging example.

I. MOTIVATION

Multivariable control methods have been employed to control many industrial processes. Usually, an industrial process control comprises of either reference tracking or minimizing an economic cost criterion. These industrial processes may require additional constraints on process variables to be satisfied. Model predictive control (MPC) has proven to be an attractive solution to meet various control demands and is being increasingly adopted as a favourable control solution in a wide range of processes. Since MPC is a receding-horizon as well as an optimization-based approach, several parameters, especially the penalty (weighting) matrices, must be appropriately tuned in order to implement it successfully. Tuning these parameters is not straightforward since they behave non-monotonically and no direct relationship exists between them and the process performance. Therefore, usually experience is required to tune an MPC. Most often MPC tuning is performed by fixing all but one tuning parameter which is varied over a certain range and its effect is analyzed, leading to derivation of certain tuning guidelines and simplifying the tuning problem to some extent, see [2], [3], [4], [5], [6], [7]. Shridhar et al. [8] propose analytical expressions for computing the tuning parameters by using approximated first-order plus dead time (FOPDT) models between different pairs of controller output and plant output variables whereas Clarke et al. [9] propose a constrained receding-horizon predictive control (CRHPC) in which the output variable is constrained to be equal to the reference value over a certain constraint range. They derive different conditions on the tuning parameters, i) for guaranteeing asymptotic stability, and ii) for pole-placement control.

G. Shah and S. Engell are with the Dept. of Bio- and Chemical Engineering, Process Dynamics and Operations Group, Technische Universität Dortmund, Germany. Tel.: +49-(0)231-755 5142, Email: {Gaurang.Shah,Sebastian.Engell}@bci.tu-dortmund.de

In an another approach, MPC is treated as a linear feedback control design problem. For example, Yu-Geng et al. [10] design a constrained predictive controller in form of a linear feedback controller. Their approach guarantees closed-loop performance by ensuring that the closed-loop poles are located at the desired locations. Chiou et al. [11] analyze a constrained MPC design problem as designing a set of piecewise linear controllers, each corresponding to different types of predicted active constraint situations, which leads to solving a min-max optimization problem over the frequency domain. If infinite prediction and control horizons are considered, an unconstrained MPC leads to a linear quadratic regulator (LQR) design problem. While the LQR has been well studied in the literature, its design is computationally intractable in the constrained case due to the infinite horizon. With some modifications in the standard quadratic cost function, an unconstrained infinite horizon LQR problem can be mapped to a finite horizon constrained LQR problem, see [12]. Rowe et al. [13] design an H_∞ -based constrained LQR by using loop shaping techniques which enables them to obtain performance trade-offs in the frequency domain. Bemporad et al. [14] develop an explicit MPC algorithm where they design the piecewise state feedback controllers as constrained LQR.

In addition to the nominal stability and the constraint satisfaction requirements, another important requirement in MPC design is that of minimal cross-coupling between various controlled variables. In case of a minimum phase plant, the cross-coupling can be eliminated first by using a suitable decoupler and then tuning MPC's for the resulting SISO loops, see [15], [16]. Alternatively, in case of a non-minimum phase plant, the cross-coupling can be minimized by modifying the cost function, see [17], [18] for details.

The theory of loop shaping in the frequency domain offers a good platform for designing linear multivariable controllers since several design aspects such as robust stability, disturbance rejection, cross-coupling reduction, etc. are well-understood in the frequency domain. This has been considered in two works, i) Vega et al. [19] who solve a mixed sensitivity optimization problem to determine the penalty weights, the prediction and the control horizons, and ii) Fan et al. [20] who implement a two-dimensional frequency domain technique to decouple a cross-directional process model and discuss ways to select penalty weights based on robust control theory. MPC tuning for desired closed-loop performance has been studied by Trierweiler et al. [21]. Their method determines the tuning parameters by making use of a robust performance number (RPN). If a

control horizon of one is chosen, then in the unconstrained case, a linear MPC can be represented as a closed-loop combination of linear controllers which are functions of the penalty matrices used in the standard quadratic cost function of MPC. The closed-loop representation offers an advantage in that the linear controllers can be designed using linear control theory. Further, one can assess the robustness of the loop in the unconstrained case by using the theory of robust loop shaping which is our main motivation. Using the closed-loop analytical approach, the penalty matrices of the quadratic cost function can be designed systematically considering nominal as well as robust performance requirements.

Based on this knowledge, a novel tuning approach was presented in [22] for SISO systems. In this approach, the tuning parameters are determined such that the poles and the zeros of the closed-loop form of MPC are obtained at the desired locations. The tuning parameters are determined in two convex optimization steps. The approach was further extended to MIMO systems, see Shah et al. [1] in which the tuning is performed for obtaining a desired closed-loop performance which is determined based on the steady-state, the overshoot and the rise-time requirements of various input-output channels. Though the latter tuning procedure guarantees desired closed-loop performance, it does not guarantee satisfaction of the robustness criteria which is a main drawback. Secondly, there are numerical problems associated with the optimization step for computing the penalty matrices.

The aim of this paper is to overcome the shortcomings of the above-mentioned tuning approach and to further improve the tuning procedure. The paper is organized as follows. In section II, the closed-loop formulation of MPC is discussed. In section III, the shortcomings of the former tuning approach are briefly discussed and solutions to overcome them are proposed in section IV. The new approach is tested on a challenging example in section V. Section VI concludes with a brief discussion.

Regarding notations, \mathbb{R} and \mathbb{C} denote real and complex spaces respectively. \mathbb{N} is a set of natural numbers and \mathcal{S}^n denotes a set of $n \times n$ symmetric matrices. A vector (or a matrix) is represented by using an underlined variable. $P(z^{-1})$ denotes a discrete-time polynomial and is written as $P(z^{-1}) = p_0 + p_1 z^{-1} + \dots + p_n z^{-n}$, $p_i \in \mathbb{R}$, $z \in \mathbb{C}$, $n \in \mathbb{N}$. The term $1/\Delta$ is a summation operator where $\Delta = (1 - z^{-1})$. $\dim(\underline{M})$ represents the size of a matrix \underline{M} and T_s denotes the sampling time. A Kronecker product between two matrices \underline{X} and \underline{Y} is denoted as $\underline{X} \otimes \underline{Y}$. $\underline{M}(\omega)$ represents the value of a transfer matrix \underline{M} at frequency ω where $\underline{M}(\omega) \in \mathbb{C}$. $\|\cdot\|_F$ denotes a Frobenius norm and $\|\cdot\|_\infty$ denotes an infinity norm.

II. MPC FORMULATION

The linear discrete-time plant is represented as:

$$\begin{aligned} \underline{y}(t+1) &= \underline{A}^{-1}(z^{-1})\underline{B}(z^{-1})\underline{u}(t) \\ &= \underline{\bar{B}}(z^{-1})\underline{\bar{A}}^{-1}(z^{-1})\underline{u}(t) \end{aligned} \quad (1)$$

where \underline{A} and $\underline{\bar{A}}$ are diagonal matrices of polynomials obtained using the least common multiple of the denominators of the corresponding row (or column) of the overall MIMO plant transfer function. Here, $\dim(\underline{A}) = n \times n$, $\dim(\underline{B}) = \dim(\underline{\bar{B}}) = n \times m$, $\dim(\underline{\bar{A}}) = m \times m$ where $n, m \in \mathbb{R}$ are the no. of outputs and inputs respectively.

Generalized Predictive Control (GPC) [23] is a popular form of MPC. It uses a CARIMA (Controlled Auto-Regressive and Integrated Moving Average) prediction model:

$$\underline{A}(z^{-1})\underline{y}(t+k) = \underline{B}(z^{-1})\underline{u}(t+k-1) + \frac{\underline{T}(z^{-1})}{\Delta(z^{-1})}\underline{e}(t+k) \quad (2)$$

where $k = 1 \dots N_p$, $\underline{T}(z^{-1})$ is a matrix of noise filter polynomials, $\underline{e}(t)$ is a vector of white noise, N_p is the prediction horizon and $\dim(\underline{T}) = n \times n$. For the sake of simplicity, a matrix or a polynomial $\underline{M}(z^{-1})$ shall be denoted by \underline{M} henceforth. Considering $\underline{T} = \underline{I}$ for simplicity and assuming the future noise to be 0, the best prediction value is given as

$$\hat{\underline{y}}(t+k) = \underbrace{\underline{G}_k \Delta \underline{u}(t+k-1)}_{\text{free response}} + \underbrace{\underline{H}_k \Delta \underline{u}(t-1) + \underline{F}_k \underline{y}(t)}_{\text{forced response}} \quad (3)$$

where \underline{G}_k , \underline{H}_k , and \underline{F}_k are determined by solving the following Diophantine equations:

$$\underline{T} = \underline{E}_k \underline{A} \Delta + z^{-k} \underline{F}_k, \quad \underline{E}_k \underline{B} = \underline{G}_k \underline{T} + z^{-k} \underline{H}_k. \quad (4)$$

The following quadratic cost function is considered:

$$\begin{aligned} J &= (\hat{\underline{y}} - \underline{w})^T \underline{Q} (\hat{\underline{y}} - \underline{w}) + \Delta \underline{u}^T \underline{\Lambda} \Delta \underline{u} \\ \underline{Q} &\succ 0, \quad \underline{\Lambda} \succeq 0 \end{aligned} \quad (5)$$

the minimization of which gives the optimal control input as

$$\begin{aligned} \Delta \underline{u}(t+k-1) &= \underbrace{(\underline{G}^T \underline{Q} \underline{G} + \underline{\Lambda})^{-1} \underline{G}^T \underline{Q}}_{\underline{K}_0} \cdot \\ &[\underline{w}(t+k) - \underline{H}_k \Delta \underline{u}(t-1) - \underline{F}_k \underline{y}(t)] \end{aligned} \quad (6)$$

where \underline{G} is a matrix containing the coefficients of the elements of \underline{G}_k , $k = N_1 \dots N_2$. N_1, N_2 are the initial and final prediction horizons respectively and $\underline{G} \in \mathbb{R}^{n(N_2 - N_1 + 1) \times m}$. Considering a control horizon $N_u = 1$, the first optimal control input is obtained as

$$\Delta \underline{u}(t) = \underline{k}_0 \underline{w}(t+k) - \underline{k}_0 \underline{H}_k \Delta \underline{u}(t-1) - \underline{k}_0 \underline{F}_k \underline{y}(t) \quad (7)$$

where $\underline{k}_0 \in \underline{K}_0$. Assuming a constant reference input and rearranging the terms, Eq. (7) is obtained as

$$\begin{aligned} \underline{R} \Delta \underline{u}(t) &= \underline{k}_s \underline{w} - \underline{S} \underline{y}(t), \quad \underline{R} = \underline{I} + z^{-1} \underline{k}_0 \underline{H}_k, \quad \underline{S} = \underline{k}_0 \underline{F}_k \\ \underline{k}_s(1, 1) &= \underline{k}_0(1, 1 + 0n) + \underline{k}_0(1, 1 + 1n) + \dots \end{aligned} \quad (8)$$

Note that \underline{k}_0 is introduced as an intermediate degree of freedom which relates the transfer matrices $\underline{R}, \underline{S}, \underline{k}_s$ to the MPC tuning parameters $\underline{Q}, \underline{\Lambda}$. Eq. (8) can be represented as a closed-loop form as shown in Fig. (1) where $\underline{R}, \underline{S}, \underline{k}_s$ represent the controllers the design of which determine the tuning parameters $\underline{Q}, \underline{\Lambda}$.

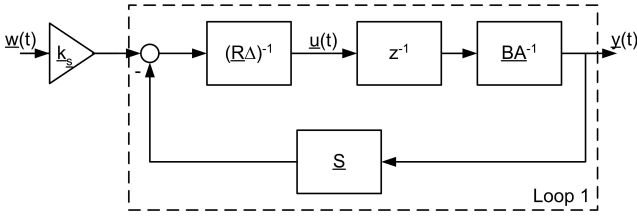


Fig. 1. Closed-loop form of an unconstrained MPC.

III. FORMER TUNING APPROACH

From Fig. (1), the true closed-loop transfer matrix from w to y is given by

$$\begin{aligned} \underline{G}_{cl,true} &= z^{-1} \underline{\bar{B}} (\underline{R} \underline{\bar{A}} \Delta + z^{-1} \underline{S} \underline{\bar{B}})^{-1} \underline{k}_s \Rightarrow \\ \underline{G}_{cl,true}(\omega) &= \dots \\ e^{-j\omega T_s} \underline{\bar{B}}(\omega) (\underline{R}(\omega) \underline{\bar{A}}(\omega) \Delta(\omega) + e^{-j\omega T_s} \underline{S}(\omega) \underline{\bar{B}}(\omega))^{-1} \underline{k}_s \\ \underline{S}(\omega) &= \underline{k}_0 \underline{F}(\omega), \underline{R}(\omega) = \underline{I} + e^{-j\omega T_s} \underline{k}_0 \underline{H}(\omega). \end{aligned}$$

Let $\underline{G}_{cl,desired}$ denote a desired closed-loop transfer matrix. The tuning procedure is performed in two optimization steps where in the first optimization step \underline{k}_0 is determined such that the error between $\underline{G}_{cl,true}$ and $\underline{G}_{cl,desired}$ is minimized and in the second optimization step the penalty matrices \underline{Q} , $\underline{\Lambda}$ are determined using \underline{k}_0 . There are however some shortcomings of this approach which are discussed briefly in the following section.

A. Shortcomings

Let ϵ_{true} represent the true error between $\underline{G}_{cl,true}$ and $\underline{G}_{cl,desired}$. We have

$$\epsilon_{true}(\omega) = \underline{G}_{cl,true}(\omega) - \underline{G}_{cl,desired}(\omega) \quad \forall \omega \in \left[0, \frac{\pi}{T_s}\right]. \quad (9)$$

The primary aim is to minimize $\|\epsilon_{true}\|_F^2$. Since $\underline{G}_{cl,true}(\omega)$ is nonlinear in \underline{k}_0 , Eq. (9) is reformulated into the following convex constraint:

$$\begin{aligned} \underbrace{(\underline{R} \underline{\bar{A}} \Delta + z^{-1} \underline{S} \underline{\bar{B}})^{-1} \underline{\bar{B}}^{-1}}_{\epsilon_k} \epsilon_{true} &= \dots \\ z^{-1} \underline{k}_s - (\underline{R} \underline{\bar{A}} \Delta + z^{-1} \underline{S} \underline{\bar{B}})^{-1} \underline{\bar{B}}^{-1} \underline{G}_{cl,desired} \end{aligned} \quad (10)$$

where the term ω has been dropped in order to give a concise representation and ϵ_k denotes the modified true error. It must be noted that the above reformulation is justified only if ϵ_{true} is small. In the former tuning approach, the following minimization problem is solved in the first optimization step:

$$\begin{aligned} \min_{\underline{k}_0, \epsilon_k} \quad & \|\epsilon'_k\|_F^2 \quad s.t. \\ Eq. (10) \quad & \forall \omega \in \left[0, \frac{\pi}{T_s}\right], \quad \epsilon_k \in \epsilon'_k. \end{aligned} \quad (11)$$

ϵ'_k is a vector containing ϵ_k values computed for every ω . Note that $\|\epsilon'_k\|_F^2$ is quadratic in \underline{k}_0 and is minimized instead of $\|\epsilon_{true}\|_F^2$ in the first minimization step.

The choice of $\underline{G}_{cl,desired}$ influences the solution of the above minimization problem. $\underline{G}_{cl,desired}$ may be chosen to

be (nearly) diagonal or a full matrix depending on the closed-loop requirements such as the amount of cross-coupling and a reasonable closed-loop performance. It is obvious that $\underline{G}_{cl,desired}$ must contain a right-half plane zero or a pure time delay if the plant has one. It may also happen that the performance specified by $\underline{G}_{cl,desired}$ is not realizable by the plant. This may lead to a suboptimal solution of $\underline{R}, \underline{S}, \underline{k}_s$ from the minimization problem (11). This suggests that the optimal value of the cost function may not be close to zero which is necessary especially in the lower frequency region. As a result, the resulting $\underline{G}_{cl,true}$ may specify an undesirable closed-loop behaviour. This implies that an appropriate choice of $\underline{G}_{cl,desired}$ may be critical in some cases.

In a classical control sense, Fig. (1) represents a two degrees-of-freedom control configuration. Assume that $\underline{G}_{cl,desired}$ has been designed considering nominal as well as robust performance requirements. The closed-loop system is robustly stable if and only if the inner loop (shown by a dashed block in Fig. (1)) is stable which depends on the shape of the open-loop frequency response. Though the first minimization step is a convex optimization problem, the set of linear controllers $\{\underline{R}, \underline{S}, \underline{k}_s\}$ corresponding to a given $\underline{G}_{cl,true}$ is not unique. This means that even if $\underline{G}_{cl,true} \approx \underline{G}_{cl,desired}$ is achieved, the obtained inner loop may not be robustly stable. Therefore, satisfaction of the robustness criterion is not handled transparently using the former approach.

IV. NEW TUNING APPROACH

From the previous discussion, it follows that the former tuning approach suffers from two issues:

- a systematic way of determining $\underline{G}_{cl,desired}$ is required, and
- robustness aspects must be considered directly into the optimization problem.

The solution to the above shortcomings are discussed next which subsequently leads to the new tuning approach.

A. Solution to the Shortcomings

The problem of determining a suitable $\underline{G}_{cl,desired}$ is solved by determining an achievable closed-loop performance which shall be denoted as $\underline{G}_{cl,achievable}$ henceforth. $\underline{G}_{cl,achievable}$ is determined for a standard closed-loop, see Fig. (2), by parameterizing the closed-loop performance as a function of the Youla parameter, denoted as \underline{Q}_y , which represents the set of all stable transfer matrices. Owing to the space limitations, the computation of $\underline{G}_{cl,achievable}$ is not shown here but discussions about it can be found in [24], [25]. The computation of $\underline{G}_{cl,achievable}$ can be performed to include external disturbances and various robustness requirements by adding relevant constraints to the involved minimization problem [26]. This way robustness can be achieved systematically. If there are sufficiently many degrees of freedom available in the form of penalty matrices and horizons, $\underline{G}_{cl,true} \approx \underline{G}_{cl,achievable}$, i.e. the closed-loop MPC can

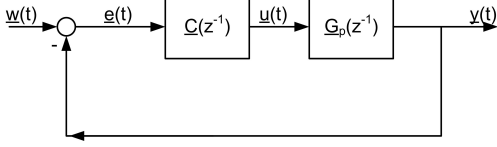


Fig. 2. A standard closed-loop configuration.

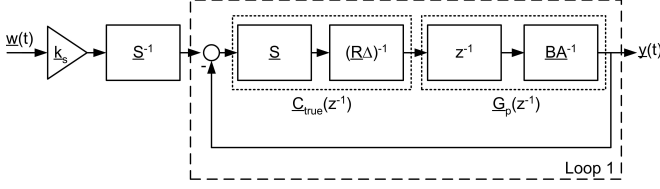


Fig. 3. A rearranged closed-loop form of an unconstrained MPC.

achieve $\underline{G}_{cl,achievable}$.

Rearranging the Closed-loop Form

In order to make the tuning approach transparent, the closed-loop form of MPC is re-arranged as shown in Fig. (3) where Loop 1 (shown by a dashed block) resembles the standard closed-loop form. Let $\underline{C}_{true} = (\underline{R}\underline{\Delta})^{-1}\underline{S}$ denote the true controller and let \underline{C}_a denote the achievable controller which results from $\underline{G}_{cl,achievable}$. Referring to the discussion in Sec. (III-A), the primary aim is to minimize the error between the true and the achievable open-loops which implies that the goal is to obtain $\underline{C}_{true} \approx \underline{C}_a$. In other words, Loop 1 must be tuned. Let

$$\epsilon_c(\omega) = \underline{C}_{true}(\omega) - \underline{C}_a(\omega), \quad \omega \in \left[0, \frac{\pi}{T_s}\right] \quad (12)$$

where $(\cdot)(\omega)$ represent the values at frequency ω . Eq. (12) can be written as

$$\begin{aligned} \epsilon_c(\omega) &= (\underline{R}(\omega)\underline{\Delta}(\omega))^{-1}\underline{S}(\omega) - \underline{C}_a(\omega) \\ &\approx \underline{S}(\omega) - \underline{R}(\omega)\underline{\Delta}(\omega)\underline{C}_a(\omega) \\ &\approx \underline{k}_0\underline{F}(\omega) - (\underline{I} + e^{-j\omega T_s}\underline{k}_0\underline{H}(\omega))\underline{\Delta}(\omega)\underline{C}_a(\omega). \end{aligned} \quad (13)$$

A weighting transfer matrix \underline{W} may be used to emphasize the error over a certain frequency range. The aim is to minimize $\|\underline{W}\epsilon_c\|_F^2$ for all $\omega \in \left[0, \frac{\pi}{T_s}\right]$. Since \underline{k}_0 is nonconvex in \underline{Q} and $\underline{\Lambda}$, determining the penalty matrices by a single step minimization of $\|\underline{W}\epsilon_c\|_F^2$ is a nonconvex problem. In order to solve the minimization problem using convex optimization, the tuning problem is tackled in two steps:

- (i) determine \underline{k}_0 which minimizes $\|\underline{W}\epsilon_c\|_F^2$, and
- (ii) determine $\underline{Q}, \underline{\Lambda}$ which correspond to the \underline{k}_0 values.

The optimality of the result remains unaffected using the above mentioned two steps methodology in comparison to the single step nonconvex minimization problem.

B. Determining the \underline{k}_0 Values

Since $\epsilon_c(\omega)$ is linear in \underline{k}_0 , minimization of $\|\underline{W}\epsilon_c(\omega)\|_F^2$ is a quadratic optimization problem. The latter however has an ill-conditioned objective function since its coefficients widely

differ from each other and vary over a large range. As a result, the optimizer may get stuck in a sub-optimal solution. In order to enable better convergence to the optimal solution, it is proposed that the minimization problem is solved with different initialization values of \underline{k}_0 . To achieve this, \underline{k}_0 is expressed as

$$\underline{k}_0 = \underline{k}_{0,init} + \underline{k}_{0,add}$$

where $\underline{k}_{0,init}$ is an initialization set which represents the fixed part and $\underline{k}_{0,add}$ represents the variable part which is a degree of freedom to be optimized. The minimization problem is repeatedly solved with different $\underline{k}_{0,init}$ as starting values each of which results in a different solution of $\underline{k}_{0,add}$ and respectively \underline{k}_0 . In order to create different sets of $\underline{k}_{0,init}$, one approach would be to simply generate it randomly. Alternatively, $\underline{k}_{0,init}$ can be generated from $(\underline{Q}_{init}, \underline{\Lambda}_{init})$ (initialization sets of $\underline{Q}, \underline{\Lambda}$) since \underline{k}_0 is a function of $\underline{Q}, \underline{\Lambda}$, refer Eq. (6). Different sets of \underline{Q}_{init} and $\underline{\Lambda}_{init}$ can be generated randomly such that the magnitude of each of them is limited by the $\|\cdot\|_\infty$ -norm which can be varied over a certain range. This way different sets of $\underline{k}_{0,init}$ can be systematically constructed from $(\underline{Q}_{init}, \underline{\Lambda}_{init})$. To enable better convergence of the optimizer, only those sets of $\underline{k}_{0,init}$ may be selected which lead to a stable closed-loop configuration although it is not necessary to do so. The complete first step minimization problem can be written as the following:

$$\begin{aligned} \min_{\epsilon_c, \underline{k}_{0,add}} \quad & \|\underline{W}\epsilon_c\|_F^2 \quad s.t. \\ \epsilon_c(\omega) &= \underline{S}(\omega) - \underline{R}(\omega)\underline{\Delta}(\omega)\underline{C}_a(\omega) \\ \underline{S}(\omega) &= \underline{k}_0\underline{F}(\omega), \underline{R}(\omega) = \underline{I} + e^{-j\omega T_s}\underline{k}_0\underline{H}(\omega) \\ \underline{k}_0 &= \underline{k}_{0,init} + \underline{k}_{0,add}, \forall \omega \in \left[0, \frac{\pi}{T_s}\right]. \end{aligned} \quad (14)$$

The minimization problem is repeated with different sets of $\underline{k}_{0,init}$ values and the best solution of \underline{k}_0 , i.e. one for which minimal $\|\underline{W}\epsilon_c\|_F^2$ results, is selected.

C. Determining the Penalty Matrices $\underline{Q}, \underline{\Lambda}$

The optimal \underline{k}_0 obtained in the first minimization step is used to compute the penalty matrices $\underline{Q}, \underline{\Lambda}$. Since $\underline{k}_0 = (\underline{G}^T \underline{Q} \underline{G} + \underline{\Lambda})^{-1} \underline{G}^T \underline{Q}$ (from Eq. (6)), let

$$\epsilon_{k_0, true} = (\underline{G}^T \underline{Q} \underline{G} + \underline{\Lambda})^{-1} \underline{G}^T \underline{Q} - \underline{k}_0$$

where $\epsilon_{k_0, true}$ represents the true error. The primary goal is to minimize $\|\epsilon_{k_0, true}\|_F^2$ but $\epsilon_{k_0, true}$ is nonconvex in $\underline{Q}, \underline{\Lambda}$. It is therefore reformulated into the following convex constraint:

$$\underbrace{(\underline{G}^T \underline{Q} \underline{G} + \underline{\Lambda})\epsilon_{k_0, true}}_{\epsilon_{k_0}} = \underline{G}^T \underline{Q} - (\underline{G}^T \underline{Q} \underline{G} + \underline{\Lambda})\underline{k}_0.$$

Now, $\|\epsilon_{k_0}\|_F^2$ is minimized instead of $\|\epsilon_{k_0, true}\|_F^2$. The corresponding minimization problem is shown below:

$$\begin{aligned} \min_{\underline{Q}, \underline{\Lambda}, \epsilon_{k_0}} \quad & \|\epsilon_{k_0}\|_F^2 \quad s.t. \\ \epsilon_{k_0} &= \underline{G}^T \underline{Q} - (\underline{G}^T \underline{Q} \underline{G} + \underline{\Lambda})\underline{k}_0 \\ \underline{Q} &> 0, \quad \underline{\Lambda} \succeq 0. \end{aligned} \quad (15)$$

This is a convex optimization problem the solution space of which is an intersection of the hyperplane, defined by the linear equality constraint, with the respective positive (semi)definite cones of \underline{Q} and $\underline{\Lambda}$. As a result, there exists a plateau of feasible $\underline{Q}, \underline{\Lambda}$ which result in the same cost value of the objective function.

The solution of the above minimization problem is justified only if $\epsilon_{k0,true}$ is small. This means that a small $\|\epsilon_{k0}\|_F^2$ may not imply a small $\|\epsilon_{k0,true}\|_F^2$. In other words, a large $\|\epsilon_{k0,true}\|_F^2$ may result even if $\|\epsilon_{k0}\|_F^2$ is very small. Therefore, convergence to the required optimal solution is not always guaranteed despite availability of enough degrees of freedom. To overcome this problem, a workaround is proposed as follows. Let $\underline{Q}, \underline{\Lambda}$ be rewritten as

$$\begin{aligned} \underline{Q} &= \underline{Q}_{init} + \delta\underline{Q}, \quad \underline{\Lambda} = \underline{\Lambda}_{init} + \delta\underline{\Lambda} \\ \|\delta\underline{Q}\|_F^2 &\leq \gamma_Q, \quad \|\delta\underline{\Lambda}\|_F^2 \leq \gamma_\Lambda \end{aligned} \quad (16)$$

where $(\underline{Q}_{init}, \underline{\Lambda}_{init})$ are initial values of penalty matrices that were used for the computation of $\underline{k}_0, init$, $\delta\underline{Q}, \delta\underline{\Lambda}$ are perturbation matrices which are the degrees of freedom and whose values are constrained by the constant values of γ_Q, γ_Λ . The optimization problem (15) is repeatedly solved alongwith (16) for different values of γ_Q, γ_Λ , initially beginning with small γ_Q, γ_Λ values and gradually increasing them till a certain limit. If $\|\epsilon_{k0,true}\|_F \approx 0$ is achieved in the above optimization step, it implies that the obtained $\underline{Q}, \underline{\Lambda}$ result in achieving $\underline{C}_{true} \approx \underline{C}_a$.

Remark: The idea is that once Loop 1 has been designed, the resulting overall true closed-loop performance can be fine-tuned with the help of an additional prefilter that is used for manipulating the reference signals. The prefilter is then merely used to provide a faster response for reference steps but not for disturbances. A suitable prefilter can be computed by posing its design as an optimization problem which shall not be discussed in this work.

V. EXAMPLES

The convex optimization problems are solved using the SeDuMi solver¹. In the following example, the discretized equivalent of the continuous transfer function is identified as an ARX transfer function using the MATLAB-based System Identification Toolbox (V7.3).

A. A Shell Standard Control Problem

A shell standard control problem is considered [11]. The discrete-time plant model has sampling time $T_s = 6$ s. and is given as

$$\underline{y}(t) = \begin{bmatrix} \frac{0.4581z^{-5}}{1-0.8869z^{-1}} & \frac{0.1685z^{-5}}{1-0.9048z^{-1}} \\ \frac{0.6096z^{-4}}{1-0.8869z^{-1}} & \frac{0.5445z^{-3}}{1-0.9048z^{-1}} \end{bmatrix} \underline{u}(t).$$

Considering nominal performance requirements, $\underline{C}_{cl,achievable}$ was determined using the Youla parameterization approach wherein 4 individual stable

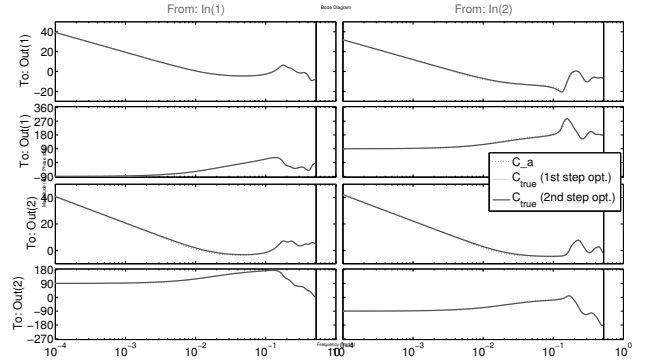


Fig. 4. Bode plots of \underline{C}_a and \underline{C}_{true} .

transfer functions were used for constructing \underline{Q}_y . The achievable (desired) controller \underline{C}_a was then determined from $\underline{C}_{cl,achievable}$ and its frequency response is shown as a dotted line in Fig. (4). Prediction horizons $N_1 = 1$ and $N_2 = 10$ are considered. Tuning of the penalty matrices is performed considering the control horizon $N_u = 1$. From the first optimization step, \underline{k}_0 is obtained as

$$\begin{bmatrix} -0.0385 & 0.0026 & -0.0409 & 0.5170 & -0.9291 & \dots \\ 0.2019 & -0.1568 & -0.1129 & -1.0367 & 0.8911 & \dots \\ -0.0898 & -0.6556 & -0.6818 & 1.0552 & -0.1212 & \dots \\ 0.1039 & 0.7719 & 1.0532 & -1.0459 & 1.1344 & \dots \\ 1.1269 & 0.0529 & 0.2054 & 0.2960 & 1.1119 & \dots \\ -1.1347 & 0.1569 & -1.0170 & -0.3358 & -0.3499 & \dots \\ 0.0803 & -0.8083 & -0.1078 & -0.8917 & 0.0221 & \dots \\ -1.0271 & 0.8471 & 0.1944 & 0.8125 & 0.0533 & \dots \end{bmatrix}$$

and the resulting \underline{C}_{true} is determined whose Bode plot is shown as a solid line in Fig. (4) which indicates a good fit between \underline{C}_{true} and \underline{C}_a . After solving the second optimization problem, \underline{Q} and $\underline{\Lambda}$ are obtained which are not shown here owing to space constraints. The Bode plot of the resulting \underline{C}_{true} fits almost exactly with that from the first optimization and is also shown in Fig. (4).

The obtained Bode plots demonstrate the ability of the tuning approach to achieve a desired open-loop. Further, the overall closed-loop performance can be fine-tuned by manipulating the reference signals using a prefilter. Figs. (5)(a)&(b) show the closed-loop step responses without using a prefilter whereas Figs. (5)(c)&(d) show the same using a prefilter where the following prefilter was used:

$$\underline{W}_{pref} = 0.6667 \frac{(1 - 0.85z^{-1})}{(1 - 0.90z^{-1})} \underline{I}.$$

The prefilter was designed to only influence the closed-loop response for a reference step. The left-half plane zero of the prefilter was chosen to increase the speed of the closed-loop response.

VI. CONCLUSIONS

A new approach to the tuning of MPC parameters has been presented to overcome the drawbacks of the former

¹freely available at <http://sedumi.ie.lehigh.edu/>

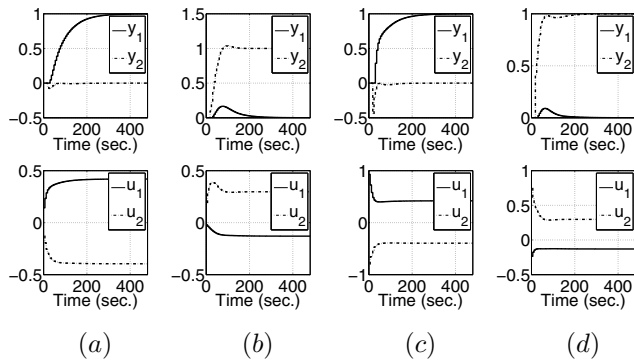


Fig. 5. Closed-loop responses of y_1 & y_2 to a step change in w_1 & w_2 without prefilter ((a), (b)) and with prefilter ((c), (d)).

tuning approach [1]. The successful solution of the former tuning approach depends on an appropriate desired closed-loop transfer matrix whose choice can be critical in some cases. To characterize the possible closed-loop behaviour, an achievable closed-loop performance is determined by parameterizing the true closed-loop transfer matrix as a function of the Youla parameter.

The requirement of robustness is included systematically in the computation of the achievable closed-loop performance which shapes the open-loop frequency response. Hence, in our new approach, the tuning parameters are determined such that the frequency response of the true open-loop is approximated with that of the achievable open-loop. To achieve this, the difference between the respective controllers is minimized in the frequency domain. The new approach addresses solutions to overcome the numerical problems associated with the optimization step for computing the penalty matrices. The obtained results from an example demonstrate the effectiveness of the new tuning approach.

So far, MPC tuning has been discussed only for the case with $N_u = 1$. If $N_u > 1$ is chosen, the second minimization problem becomes nonconvex and hence is difficult to solve. To solve such nonconvex problem, the use of optimization approaches such as bilevel programming [27] or memetic algorithms [28] can be investigated. In order to consider input/output constraints, a proposal would be to modify Fig. (3) such that the limitations are represented as nonlinear saturation blocks [29]. Criteria based on the circle and the Popov criterion may then be used for analyzing the closed-loop stability of linear systems in presence of such nonlinearities.

REFERENCES

- [1] G. Shah and S. Engell, "Tuning MPC for Desired Closed-Loop Performance for MIMO Systems," in *American Control Conference*, 2011, pp. 4404–4409.
- [2] A. R. McIntosh, S. L. Shah, and D. G. Fisher, "Analysis and Tuning of Adaptive Generalized Predictive Control," *The Canadian Journal of Chemical Engineering*, vol. 69, no. 1, pp. 97–110, February 1991.
- [3] J. Ou and R. R. Rhinehart, "CV Damping Versus MV Suppression For MPC Tuning," in *Proceedings of the American Control Conference*, 2002, pp. 2695–2700.
- [4] J. Zhang and X. Yu-Geng, "Some GPC stability results," *International Journal of Control*, vol. 70, no. 5, pp. 831–840, 1998.
- [5] J. L. Garriga and M. Soroush, "Model Predictive Controller Tuning via Eigenvalue Placement," in *Proceedings of the ACC 2008*, June 2008, pp. 429–434.
- [6] B. Kouvaritakis and J. A. Rossiter, "Multivariable Nyquist self-tuning: a general approach," *IEE Proceedings - Pt. D*, vol. 136, no. 5, pp. 189–202, 1989.
- [7] C. Mohtadi, S. L. Shah, and D. G. Fisher, "Frequency response characteristics of MIMO GPC," *International Journal of Control*, vol. 55, no. 4, pp. 877–900, 1992.
- [8] R. Shridhar and D. J. Cooper, "A Tuning Strategy for Unconstrained SISO Model Predictive Control," *Industrial and Engineering Chemistry Research*, vol. 36, pp. 729–746, 1997.
- [9] D. W. Clarke and R. Scattolini, "Constrained Receding-Horizon Predictive Control," in *IEE Proceedings-D*, vol. 138, no. 4, 1991, pp. 347–354.
- [10] X. Yu-Geng and Z. Qun-Liang, "New Approach to Designing Constrained Predictive Controllers," *ACTA AUTOMATICA SINICA*, vol. 31, no. 5, pp. 655–661, 2005.
- [11] H.-W. Chiou and E. Zafiriou, "Frequency Domain Design of Robustly Stable Constrained Model Predictive Controllers," in *American Control Conference*, vol. 3, 1994, pp. 2852–2856.
- [12] P. Scokaert and J. B. Rawlings, "Constrained linear quadratic regulation," *IEEE Transactions on Automatic Control*, vol. 43, no. 8, pp. 1163–1169, 1998.
- [13] C. Rowe and J. Maciejowski, "Tuning MPC using H_∞ Loop Shaping," in *Proceedings of the American Control Conference*, June 2000, pp. 1332–1336.
- [14] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3–20, 2002.
- [15] F. Y. Lin, J. C. Jeng, and H. P. Huang, "Multivariable Control with Generalized Decoupling for Disturbance Rejection," *Ind. Eng. Chem. Res.*, vol. 48, pp. 9175–9185, 2009.
- [16] W. Zenghui, C. Zengqiang, S. Qinglin, and Y. Zhuzhi, "Multivariable Decoupling Predictive Control Based on QFT Theory and Application in CSTR Chemical Process," *Chinese Journal of Chemical Eng.*, vol. 14, no. 6, pp. 765–769, 2006.
- [17] R. H. Middleton and G. J. Adams, "Modification of Model Predictive Control to Reduce Cross-Coupling," in *Proceedings of the 17th IFAC WC*, 2008.
- [18] T. Chai, K. Mao, and X. Qin, "Decoupling design of multivariable generalized predictive control," *IEE Proc.-Control Theory Appl.*, vol. 141, no. 3, pp. 197–201, 1994.
- [19] P. Vega, M. Francisco, and F. Tadeo, "Multiobjective optimization for automatic tuning of robust model based predictive controllers," in *Proceedings of the 17th IFAC WC*, Seoul, South Korea, July 2008, pp. 6980–6985.
- [20] J. Fan, G. E. Stewart, and G. A. Dumont, "Two-dimensional frequency response analysis and insights for weight selection in cross-directional model predictive control," in *42nd IEEE Conference on Decision and Control*, December 2003, pp. 3717–3723.
- [21] J. O. Trierweiler and L. A. Farina, "RPN tuning strategy for model predictive control," *Journal of Process Control*, vol. 13, pp. 591–598, 2003.
- [22] G. Shah and S. Engell, "Tuning MPC for Desired Closed-Loop Performance for SISO Systems," in *Proc. 18th Mediterranean Conference on Control and Automation*, 2010, pp. 628–633.
- [23] D. W. Clarke, C. Mohtadi, and P. S. Tuffs, "Generalized Predictive Control - Part I. The Basic Algorithm," *Automatica*, vol. 23, no. 2, pp. 137–148, 1987.
- [24] S. P. Boyd and C. Barrat, *Linear Controller Design: Limits of Performance*, Second ed. Prentice Hall, Englewood Cliffs, 1991.
- [25] K. Webers and S. Engell, "Comparison of multivariable and decentralized control of a CSTR by convex optimization," in *Proceedings of the 13th IFAC WC*, San Francisco, USA, 1996, pp. 391–396.
- [26] M. Völker and S. Engell, "Multiobjective Controllability Assessment by Finite Dimensional Approximation," in *American Control Conference*, 2005, pp. 39–44.
- [27] S. Dempe, *Foundations of bilevel programming*. Springer, 2002.
- [28] M. Urselmann, S. Barkmann, G. Sand, and S. Engell, "Optimization-based design of reactive distillation columns using a memetic algorithm," *Computers & Chemical Engineering*, vol. 35, no. 5, pp. 787–805, 2011.
- [29] A. H. Glatfelder and W. Schaufelberger, *Control Systems with Input and Output Constraints*. Springer Verlag, 2003.