Dual Adaptive Model Predictive Control with Disturbances

Zicheng Cai, B. Erik Ydstie

Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: zichengc@andrew.cmu.edu)

Abstract: Model-based control requires good accuracy of model parameters to achieve high performance. Controller design under parametric uncertainties is therefore a challenging topic in control system engineering. One well known control design for uncertain systems is adaptive control. An adaptive controller has two tasks: process regulation and parameter learning. Dual control explores the trade-off between the two seemingly conflicting tasks. The control structure in an adaptive system consists of a model-based controller and a recursive update rule for parameter estimation. In this paper, the adaptive control framework consists of model predictive control for system regulation and recursive least squares for parameter estimation. The requirement for persistent excitation is shown to be necessary for systems with high-dimensional parameter space. Then, a dual formulation is proposed in an attempt to generate excitation signals while maintaining control performance in the adaptive control scheme. The algorithm is implemented and tested on a simulated SISO system.

Keywords: Adaptive Control; Model Predictive Control; Estimation; Identification; Stochastic System

1. INTRODUCTION

Over the past few decades, model predictive control (MPC) has become one of the most effective tools in handling industrial control problems (Qin and Badgwell (2003)). MPC exploits and optimizes trajectories of a plant given a model that simulates the plant dynamics. An approximate linear model is employed in many industrial control examples and the performance can deteriorate over time if the model is not updated to take into account varying operation conditions. An indirect adaptive control system, which consists of a conventional model-based control law with adjustable parameters and an model adaptation loop, is an effective method to handle such problems.

Online parameter adaptation can be considered as a system identification process that uses recursive parametric estimation methods such as the gradient algorithm and the recursive least square (RLS) algorithm. A necessary condition for parameter convergence is persistent excitation (PE), which is satisfied when the input signals have rich information. However, the most common goal in control engineering is set point regulation and the degree of excitation is usually not sufficient under normal operating conditions. If external signals are applied to generate PE, the control performance can be compromised.

Dual control theory (Feldbaum (1960)) discusses the trade-off between the two tasks in adaptive control: trajectory convergence and parameter convergence. The fist goal is to control the system and the learning of parameters comes second. Controllers with active learning components aim to generate more informative signals to satisfy the PE condition. Since active learning with excessive or insufficient excitation compromises performance, a dual adaptive controller is designed to optimally improve the learning while maintaining control performance.

Adaptive model predictive control (AMPC) draws increasing attentions from researchers but many challenges remain open (Mayne (2014)). The majority of adaptive MPC algorithms treat the control and learning as separate tasks. This approach, called certainty equivalence adaptive control, generates input signals that cannot be guaranteed to be rich enough for good parameter estimation (Wittenmark (1995)). One way of approaching this issue is to design a dual controller, which actively explores the system by ensuring a certain level of excitation, either constantly or when needed. The model predictive control and simultaneous identification (MPCI) framework is one of the earliest attempts to formulate dual control based on MPC (Genceli and Nikolaou (1996), Shouche et al. (1998)). Several proposed controllers generate excitation without a specific requirement. Rather, they include a function of information or uncertainty in the MPC cost function and optimize this function together with standard control objectives. Heirung et al. (2015) propose and compare two such formulations that converge to a standard adaptive certainty-equivalence MPC (Åström and Wittenmark (2013)) formulation as the uncertainty is reduced, and show that the excitation can improve closed-loop performance. For a set of finite-impulse-response (FIR) systems, modification of the adaptive MPC formulation can introduce exploratory properties. The approach modifies the nominally optimal input sequence by solving a second optimization problem, the objective of which is to reduce the set of possible models at the next time step

⋆ This work was funded by PITA.

(Tanaskovic et al. (2014)). In these methods, excitation is generated by heuristic modifications of the controller, with the assumption that the resulting excitation improves performance. While this type of algorithm may work well in practice and improves performance over passive-learning approaches, the excitation is not an implicit consequence of optimizing for performance. The algorithm type does, however, illustrate an important distinction: superimposing excitation on a nominally optimal control signal does not generally result in optimal performance, and the inputs are consequently not dual.

Dynamic programming (DP) can be considered as an appropriate solution for dual control by optimally integrating active learning with multistate decision making, which Feldbaum identified in his pioneering papers (Feldbaum (1960)). A scalar dual control problem with one unknown parameter is investigated with dynamic programming, but "the curse of dimensionality" prevents dynamic programming from being a viable solution approach for high order and multivariate adaptive systems. This has motivated the use of approximate methods that directly approximate the dynamic programming equations rather than the problem formulation. It is worth noting that the use of DP provides a direct link to machine learning, in particular reinforcement or Q-learning, which is also referred to as heuristic dynamic programming (HDP) (Werbos (1989)). In the adaptive control community, the Q-learning approach is referred to as direct adaptive control whereas methods that learn a model and then compute control signals are referred to as indirect adaptive control.

Chemical processes are normally nonlinear. However, it cannot be expected that we can develop a good theory for adaptive control of nonlinear systems, before the linear problem is well understood. In this paper, we therefore focus on adaptive control of linear models. Sections 2 and 3 contain a brief review of adaptive MPC, and an introduction to RLS. We derive the proposed Dual formulation in section 4. Simulation results are included in section 5 and section 6 carries out a discussion and lays future plans of interests.

2. MODEL PREDICTIVE CONTROL

We are interested in controlling systems with parameter uncertainties and external noise. The system output is given by the scalar product of a regression vector, a fixed unknown parameter vector, and a disturbance variable so that

$$y(t) = \theta^T \phi(t-1) + v(t) \tag{1}$$

We assume that the regression vector ϕ is generated by data and known at time t. Furthermore, we assume that we have obtained initial estimate $\hat{\theta}(0)$ of the parameters θ and their covariance P_0. The distribution of the noise is unknown, but is assumed to Gaussian for the algorithm development.

In a typical application, the regression vector consists of past outputs and inputs so that

$$\phi(t-1)^T = (y(t-1), ..., y(t-n), u(t-1), ..., u(t-m))$$

and the parameters are organized in a corresponding vector

$$\theta^T = (a_1, a_2, ..., a_n, b_1, b_2, ..., b_m)$$

This provides a linear in the parameter model and we can apply standard estimation techniques such the adaptation of the Kalman filter.

The problem we consider is how to choose input variables so as to solve the predictive control problem

$$\min_u \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - y_{ref})^T Q (y(i) - y_{ref}) + u(i)^T R u(i) | t\right] \tag{2}$$

where Q and R are positive weighting matrices. By conditioning on time step t, we mean that the prediction mechanism utilizes system estimations up to t. The MPC control law from equation 2 is not guaranteed to stabilize the system theoretically. To ensure stability, infinite-horizon prediction and the addition of well-measured terminal cost are common modifications. In practical applications, inequality constraints are added for bounded inputs and outputs. The inequality constraints are not taken into account in the following exposition as they make no difference to the development of the main ideas.

3. CERTAINTY EQUIVALENCE FORMULATION

For an adaptive controller, we have an estimate $\hat{\theta}(t)$ of the parameter vector at time t. This vector can be updated with parameter regression rules such as recursive least squares (RLS) using incoming data.

The advantage of the least squares solution is that it gives the best linear unbiased estimate (BLUE) when $v(t)$ follows a zero-mean Gaussian distribution. Under such assumptions, it makes sense to use the estimated parameter vector to generate predictions. Such an approach is referred to as certainty equivalence and we get to solve the following problem.

$$\begin{aligned} \min_u \quad & \sum_{i=t+1}^{t+T} (\hat{y}(i) - y_{ref})^T Q (\hat{y}(i) - y_{ref}) + u(i)^T R u(i) \\ \text{s.t.} \quad & \hat{y}(i+1) = \hat{\theta}(t)^T \phi(i), \quad i = t, \cdots, t+T-1 \end{aligned} \tag{3}$$

The stochastic MPC problem (2) is now deterministic and can be solved as a quadratic program once we have a suitable algorithm for updating the parameters.

The parameters are usually updated using some version of recursive least squares, which under some conditions, as noted above give the optimal solution. In this case we get the update

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\phi(t-1)e(t) \tag{4}$$

where $e(t) = y(t) - \hat{y}(t)$, with $\hat{y}(t) = \theta(t-1)^T \phi(t-1)$. The following recursive update for P then follows so that

$$P(t) = P(t-1) - \frac{P(t-1)\phi(t-1)\phi(t-1)^T P(t-1)}{1 + \phi(t-1)^T P(t-1)\phi(t-1)} \tag{5}$$

These equations can be derived form the Kalman filter. It follows that under suitable conditions

$$\mathbb{E}\{(\hat{\theta}(t) - \theta)^T (\hat{\theta}(t) - \theta | t\} = P(t) \quad \text{a.s.}$$

4. DUAL FORMULATION

One important drawback of the CE approach is that it does not include an explicit way of introducing excitation into the system. Then, there is no guarantee that the predictions will be optimal or even close to those that would be generated by an optimal model. This problem is compounded when the noise sequence is not Gaussian.

As described in the introduction, the idea behind dual control is to recognize that the estimated parameters are not accurate. To compensate then, it is necessary to generate control actions that at the same time are cautious, meaning robust in some precise sense, and also explorative in the sense that they may deviate from what is thought to be the optimal path in order to gain information about the system. We will now see how this dual approach can be developed from from the stochastic problem (2) with constraints (1) . In practice this means that instead of simply replacing the outputs with estimated values of y in the objective function, we show that the stochastic objective can be replaced with and equivalent deterministic expression. The main disadvantage of the reformulation is that the mathematical programming problem is no longer a classical quadratic program and more intensive numerical optimization is needed.

Theorem. Suppose that the process noise sequence $v(t)$ is Gaussian. The stochastic problem (2) with constraints (1) then can be written as the deterministic optimization problem

$$\begin{aligned}\min_{u} \quad & \sum_{i=t+1}^{t+T} (\hat{y}(i) - y_{ref})^T Q(\hat{y}(i) - y_{ref}) \\ & \qquad + u(i)^T R u(i) + \phi(i-1)^T L P(i-1) L^T \phi(i-1) \\ \text{s.t.} \quad & \hat{y}(i+1) = \hat{\theta}(t)^T \phi(i) \\ & P(i+1) = P(i) - \frac{P(i)\phi(i)\phi(i)^T P(i)}{1 + \phi(i)^T P(i) \phi(i)}, \quad i = t, \cdots, t+T-1\end{aligned} \tag{6}$$

□

The matrix L is obtained via Cholesky decomposition of Q: $Q = LL^T$. This is the key result as it shows that the stochastic problem described above has a deterministic equivalent in the case that the noise process is white. A similar development was provided in Heirung et al. (2017) under more restrictive conditions. In this paper a reformulation was shown that allowed the problem to be solved a quadratically constrained quadratic program (QCQP).

Proof of the Theorem:

At time step t, we have obtained observation $y(t)$ and attempt to compute a stabilizing yet information-rich input signal $u(t)$. We use $\mathbf{y}(t)$ to denote the time series of output signals and $\mathbf{u}(t)$ to denote the time series of input signals , i.e. $\mathbf{y}(t) = [y(0), \cdots, y(t)]$, and $\mathbf{u}(t) = [u(0), \cdots, u(t-1)]$.

To simplify the notations, a squared term is used in the place of the quadratic term in the MPC objective function. Given the prediction horizon T and starting from $t+1$,

$$\begin{aligned}& \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - y_{ref})^2 \mid \mathbf{y}(t), \mathbf{u}(t)\right] \\ =& \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - \hat{y}(i) + \hat{y}(i) - y_{ref})^2 \mid \mathbf{y}(t), \mathbf{u}(t)\right] \\ =& \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - \hat{y}(i))^2 + 2(y(i) - \hat{y}(i))(\hat{y}(i) - y_{ref}) + \right. \\ & \left. (\hat{y}(i) - y_{ref})^2 \mid \mathbf{y}(t), \mathbf{u}(t)\right] \\ =& \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - \hat{y}(i))^2 \mid \mathbf{y}(t), \mathbf{u}(t)\right] + \sum_{i=t+1}^{t+T} (\hat{y}(i) - y_{ref})^2 \\ & + 2 \sum_{i=t+1}^{t+T} (\hat{y}(i) - y_{ref}) \, \mathbb{E}\left[y(i) - \hat{y}(i) \mid \mathbf{y}(t), \mathbf{u}(t)\right]\end{aligned} \tag{7}$$

In appendix A, it is shown that $\hat{\theta}$ is an unbiased estimator. Thus, $\hat{y}$ is an unbiased estimator of y, i.e.$\mathbb{E}[y - \hat{y}, \mid \mathbf{y}(t), \mathbf{u}(t)] = 0$ and the second term is equal to zero. Equivalently, for the quadratic form, we have

$$\begin{aligned}& \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - y_{ref})^T Q (y(i) - y_{ref}) \mid \mathbf{y}(t), \mathbf{u}(t)\right] \\ =& \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - \hat{y}(i))^T Q (y(i) - \hat{y}(i)) \mid \mathbf{y}(t), \mathbf{u}(t)\right] \\ & + \sum_{i=t+1}^{t+T} (\hat{y}(i) - y_{ref})^T Q (\hat{y}(i) - y_{ref})\end{aligned} \tag{8}$$

The quadratic term in $\hat{y}$ is the same term in the CE-MPC objective. Since the weight matrix Q is positive definite, it can take a Cholesky decomposition of the form

$$Q = LL^T \tag{9}$$

where L has the same rank as Q.

For the first expectation term,

$$\begin{aligned}& \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - \hat{y}(i))^T Q (y(i) - \hat{y}(i)) \mid \mathbf{y}(t), \mathbf{u}(t)\right] \\ =& \mathbb{E}\left[\sum_{i=t+1}^{t+T} (y(i) - \hat{y}(i))^T LL^T (y(i) - \hat{y}(i)) \mid \mathbf{y}(t), \mathbf{u}(t)\right] \\ =& \mathbb{E}\left[\sum_{i=t+1}^{t+T} \phi(j-1)^T \tilde{\theta}(j-1) LL^T \tilde{\theta}(j-1)^T \phi(j-1) \mid \mathbf{y}(t), \mathbf{u}(t)\right]\end{aligned} \tag{10}$$

where $\tilde{\theta}$ is the estimation error of θ, i.e. $\tilde{\theta}(t) = \theta - \hat{\theta}(t)$.

The regression variables ϕ are also deterministic, then applying equation B.8 we have

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i=t+1}^{t+T} \phi(j-1)\tilde{\theta}(j-1)LL^T\tilde{\theta}(j-1)^T\phi(j-1) \mid \mathbf{y}(t),\mathbf{u}(t)\right] \\
&= \sum_{i=t+1}^{t+T} \phi(j-1)\,\mathbb{E}[\tilde{\theta}(j-1)LL^T\tilde{\theta}(j-1)^T \mid \mathbf{y}(t),\mathbf{u}(t)]\phi(j-1) \\
&= \sum_{i=t+1}^{t+T} \phi(j-1)\mathrm{Cov}[L^T\theta(j-1)^T \mid \mathbf{y}(t),\mathbf{u}(t)]\phi(j-1) \\
&= \sum_{i=t+1}^{t+T} \phi(j-1)L^T\mathrm{Cov}[\theta(j-1)^T \mid \mathbf{y}(t),\mathbf{u}(t)]L\phi(j-1) \\
&= \sum_{j=t+1}^{t+T} \phi(j-1)L^T P(j-1)L\phi(j-1)
\end{aligned} \tag{11}
$$

Then, if we formulate the Based on the derivation of RLS, matrix P decreases if the regressor vector is more informative. By incorporating the update law for P as a constraint in the optimization formulation, the dual MPC can award exciting input signals. In another word, within the prediction horizon, the P matrices become decision variables and depend on how the system evolve while the estimated parameters are fixed.

5. RESULTS

Control performance of the CE MPC and dual MPC is compared in the following system.

$$y(t+1) = ay(t) + bu(t) + v(t) \tag{12}$$

where true parameters are $a = 1$ and $b = 2$. In the prediction model, the parameter vector $\hat{\theta}(t) = [\hat{a}(t) \quad \hat{b}(t)]$ needs to be estimated. In all simulation cases, $\hat{a}$ is initialized to be 3 and $\hat{b}$ is initialized to be 6, i.e. $\hat{a}(0) = 3$ and $\hat{b}(0) = 6$. The initial value of y is set to 0.5. The weight matrices are applied to both MPC formulations, with $Q = 1$, and $R = 1$. The prediction horizon is set to 5. The performance is also compared with the performance of a MPC with perfect knowledge of the system, i.e., $\hat{a} = 1$ and $\hat{b}$, and no parameter adaptations, which is referred to as the optimal MPC. In all simulations, box constraints are applied to input signals in the optimization problem with $-20 \leq u \leq 20$.

In figure 1 and figure 2, the control performance of dual MPC and CE MPC is compared in the presence of a zero-mean Gaussian noise, i.e. $v(t) \sim N(0, 0.05)$. The CE MPC generates input signals that reach the bounds of the constraints and a spike in the output signals is observed, but the parameters converge close to the true values afterwards. The control performance of the dual MPC is comparable to the optimal MPC in the presence of the zero-mean noise even though the parameters do not converge to the true values.

In figure 3 and figure 3, the control performance of dual MPC and CE MPC is compared in the presence of noise signals sampled from a negative-mean Gaussian distribution, i.e., $v(t) \sim N(-0.1, 0.05)$. In the presence of the negative-mean noise, the CE MPC generates oscillating input signals that reach the bounds of the constraints but the parameters do not converge to the true parameters.

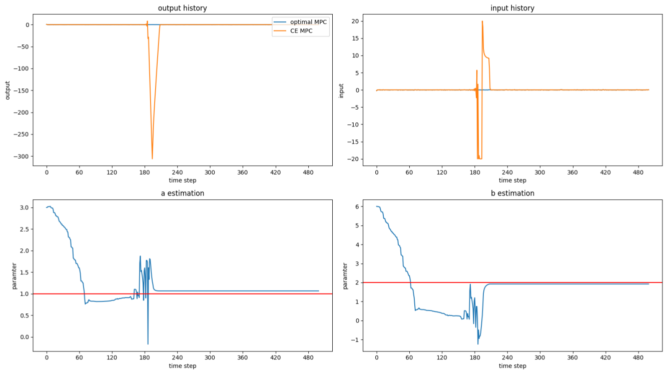

Fig. 1. CE MPC with zero-mean Gaussian process noise: $v \sim N(0, 0.05)$. The oscillations in input signals result in a spike in the output, but parameter estimations converge close to true values.

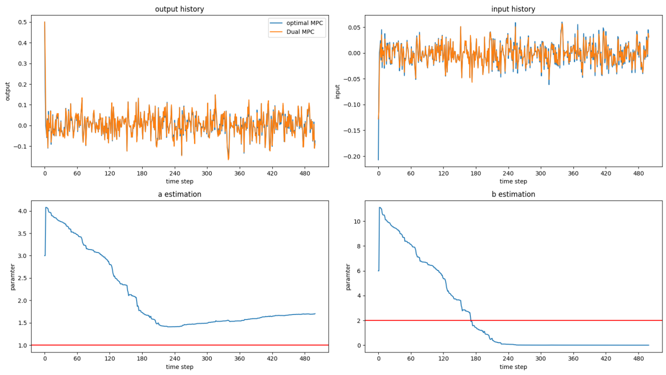

Fig. 2. Dual MPC with zero-mean Gaussian noise: $v \sim N(0, 0.05)$. The parameters did not converge to the true parameters but the performance of the dual MPC was fairly consistent to the controller with true plant parameters.

The estimation of b becomes negative and the CE MPC fails to control the system. The Dual MPC generates stabilizing control signals that are consistent with the optimal MPC. The parameter convergence properties are similar to the zero-mean case with $\hat{a}$ reaching 1.7 and $\hat{b}$ reaching 0.015.

The initial P_0 is set to $1000 * I_2$ and the changes in the entries of $P(t)$ for all four simulation cases are plotted in figure 5. It can be observed that the peaks in input and output signals correspond to large inner product of the regression vector, which leads to the decrease in parameter uncertainties.

6. DISCUSSIONS

We proposed a dual MPC formulation based on the conventional CE MPC. The dual formulation is motivated by utilizing the RLS algorithm in the MPC prediction window and making use of the parameter covariance matrix. In the CE-MPC framework, the past system information are utilized in the update rule for the P matrix (see appendix for a detailed derivation of RLS) but not in the MPC module.

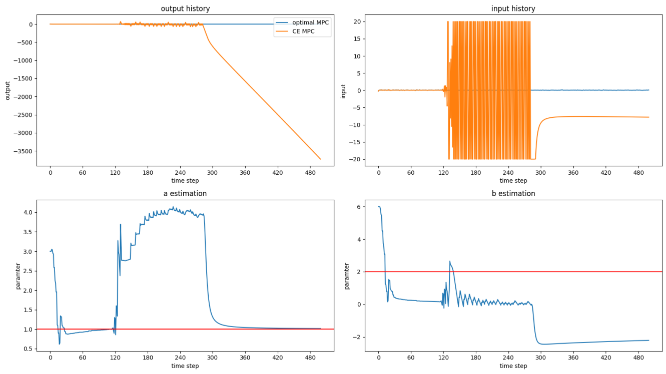

Fig. 3. CE MPC: $v \sim N(-0.1, 0.05)$. The estimation error in b is present and the CE MPC failed to control the plant to the zero set point.

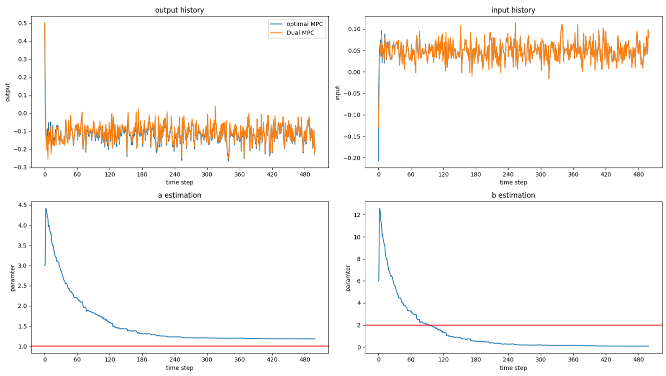

Fig. 4. Dual MPC: $v \sim N(-0.1, 0.05)$. The dual MPC has performance consistent with the optimal MPC.

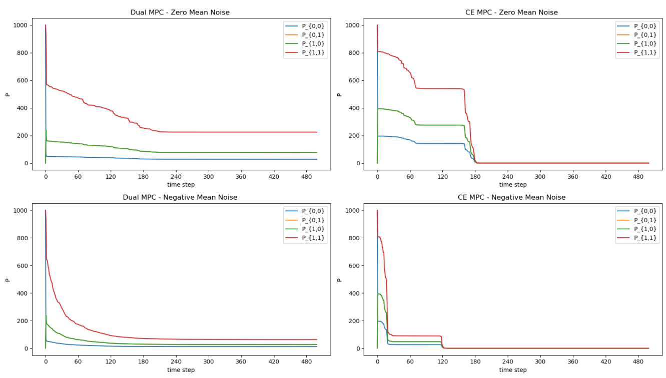

Fig. 5. The history of changes in matrix P for the four simulation cases presented in figure 1 - 4.

Instead of using estimated values of y in the objective function, we aim at minimizing the expectation of $(y - y_{ref})^2$ conditioned on system information from the past embedded in $\mathbf{y}(t)$ and $\mathbf{u}(t)$ in the dual MPC objective function. Then, we show that we can derive an equivalent deterministic expression. We used simulation of a SISO system to demonstrate the improvement of performance. The bursting phenomenon is a common type of failure and is observed in the CE MPC simulation results. It is straightforward to see from the standard RLS derivation in Appendix A that information-rich inputs explicitly decrease P. Input signals with large absolute values have an probing effect and decreases the uncertainties in parameter estimations, but it does not necessarily lead to convergence of parameter estimations to true vales. The dual effect is based on the introduction of the parameter covariance matrix P in the optimization formulation. The quadratic term with respect to P in the dual MPC formulation (6) encourages large input signals to decrease the values of elements in P, but penalizes large values of inputs because u is an element in the regression vector ϕ, which restricts the input signals when the estimation for b is smaller than the true value. This offers an explanation for the stabilizing performance of the dual controller when the parameter estimations do not converge to the true parameters. The simulation results show that the dual MPC maintains stabilizing control performance when the conventional CE MPC fails. We plan to extend the simulations to higher-dimensional cases to examine consistency of the dual MPC formulation. We also plan to study the computation complexity induced by the dual formulation and verify if reformulations of the optimization problem is available.

REFERENCES

Åström, K.J. and Wittenmark, B. (2013). *Adaptive control*. Courier Corporation.

Feldbaum, A. (1960). Dual control theory. i. *Avtomatika i Telemekhanika*, 21(9), 1240–1249.

Genceli, H. and Nikolaou, M. (1996). New approach to constrained predictive control with simultaneous model identification. *AIChE journal*, 42(10), 2857–2868.

Heirung, T.A.N., Foss, B., and Ydstie, B.E. (2015). Mpc-based dual control with online experiment design. *Journal of Process Control*, 32, 64–76.

Heirung, T.A.N., Ydstie, B.E., and Foss, B. (2017). Dual adaptive model predictive control. *Automatica*, 80, 340–348.

Mayne, D.Q. (2014). Model predictive control: Recent developments and future promise. *Automatica*, 50(12), 2967–2986.

Qin, S.J. and Badgwell, T.A. (2003). A survey of industrial model predictive control technology. *Control engineering practice*, 11(7), 733–764.

Shouche, M., Genceli, H., Premkiran, V., and Nikolaou, M. (1998). Simultaneous constrained model predictive control and identification of darx processes. *Automatica*, 34(12), 1521–1530.

Tanaskovic, M., Fagiano, L., Smith, R., and Morari, M. (2014). Adaptive receding horizon control for constrained mimo systems. *Automatica*, 50(12), 3019–3029.

Werbos, P.J. (1989). Neural networks for control and system identification. In *Proceedings of the 28th IEEE Conference on Decision and Control,*, 260–265. IEEE.

Wittenmark, B. (1995). Adaptive dual control methods: An overview. In *Adaptive Systems in Control and Signal Processing 1995*, 67–72. Elsevier.

Appendix A. PARAMETER ESTIMATION: RECURSIVE LEAST SQUARES

The RLS parameter estimator is obtained by minimizing the squared error of the system output and output estimations:

$$\min_{\theta(t)} J(t) = \sum_{i=1}^{t}[y(i) - \theta(t)^T\phi(i-1)]^2 + (\theta(t) - \theta(0))^T P_0^{-1}(\theta(t) - \theta(0)) \tag{A.1}$$

In the second term, P_0 is the covariance matrix for the initial parameter estimation $\theta(0)$ and represents the prior knowledge we have on $\theta(0)$. The value P_0 can either be computed via experimental trials or set as a positive definite matrix of users' choice.
Setting the derivative to zero:

$$\begin{aligned}\frac{\partial J(t)}{\partial \theta(t)} &= -2\sum_{i=1}^{t}[y(i) - \theta(t)^T\phi(i-1)]\phi(i-1) + 2P_0^{-1}(\theta(t) - \theta(0)) \\ &= 0\end{aligned} \tag{A.2}$$

Then,

$$\begin{aligned}&\sum_{i=1}^{t} y(i)\phi(i-1) + P_0^{-1}\theta(0) \\ &= \left[\sum_{i=1}^{t}\phi(i-1)\phi(i-1)^T\right]\theta(t) + P_0^{-1}\theta(t)\end{aligned} \tag{A.3}$$

Define a matrix $P(t)$, the inverse of which is

$$P(t)^{-1} = \sum_{i=1}^{t}\phi(i-1)\phi(i-1)^T + P_0^{-1} \tag{A.4}$$

Then we get

$$\begin{aligned}P(t)^{-1} &= \sum_{i=1}^{t}\phi(i-1)\phi(i-1)^T + P_0^{-1} \\ &= P(t-1)^{-1} + \phi(t-1)\phi(t-1)^T\end{aligned} \tag{A.5}$$

To derive recursive update for $\theta(t)$, From equation (A.3), we get

$$P(t)^{-1}\theta(t) = \sum_{i=1}^{t} y(i)\phi(i-1) + P_0^{-1}\theta(0) \tag{A.6}$$

$$\theta(t) = P(t)\sum_{i=1}^{t} y(i)\phi(i-1) + P(t)P_0^{-1}\theta(0) \tag{A.7}$$

Expand $\sum_{i=1}^{t} y(i)\phi(i-1)$ and get,

$$\sum_{i=1}^{t} y(i)\phi(i-1) = \sum_{i=1}^{t-1} y(i)\phi(i-1) + y(t)\phi(t-1) \tag{A.8}$$

From equation A.6, we have, for $\sum_{i=1}^{t-1} y(i)\phi(i-1)$,

$$\sum_{i=1}^{t-1} y(i)\phi(i-1) = P(t-1)^{-1}\theta(t-1) - P_0^{-1}\theta(0) \tag{A.9}$$

Post-multiply by $\theta(t-1)$ on both sides of equation A.5, we get

$$\begin{aligned}P(t)^{-1}\theta(t-1) &= P(t-1)^{-1}\theta(t-1) + \phi(t-1)\phi(t-1)^T\theta(t-1) \\ &= P(t-1)^{-1}\theta(t-1) + \phi(t-1)\hat{y}(t)\end{aligned} \tag{A.10}$$

and equation A.8 becomes

$$\begin{aligned}\sum_{i=1}^{t} y(i)\phi(i-1) =& P(t)^{-1}\theta(t-1) - \phi(t-1)\hat{y}(t) \\ &+ y(t)\phi(t-1) - P_0^{-1}\theta(0)\end{aligned} \tag{A.11}$$

Substitute into equation A.7, the recursive update rule for $\theta(t)$ becomes,

$$\theta(t) = \theta(t-1) + P(t)\phi(t-1)e(t) \tag{A.12}$$

where $e(t) = y(t) - \hat{y}(t)$.

According to the matrix inversion lemma, we get the following recursive update for P

$$P(t) = P(t-1) - \frac{P(t-1)\phi(t-1)\phi(t-1)^T P(t-1)}{1 + \phi(t-1)^T P(t-1)\phi(t-1)} \tag{A.13}$$

Appendix B. DUAL MPC DETERMINISTIC FORMULATION

the following deterministic form for the expectation of the parameter covariance matrix.
Given the following model

$$y(t) = \theta(t-1)^T\phi(t-1) + v(t) \tag{B.1}$$

and $v(t)$ is a modeling error . To denote information acquired up to time t, we use the following notation.

$$\boldsymbol{y}(t) = [y(1), \cdots, y(t)]^T \tag{B.2}$$

$$\boldsymbol{\phi}(t-1) = \begin{bmatrix}\phi(0)^T \\ \vdots \\ \phi(t-1)^T\end{bmatrix} \tag{B.3}$$

The parameter vector acquired from past information up to time t is the least squares solution to $\boldsymbol{y} = \boldsymbol{\phi}\theta$, which is

$$\hat{\theta} = (\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\boldsymbol{\phi}^T(\boldsymbol{\phi}\theta + \boldsymbol{v}) = \theta + (\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\boldsymbol{\phi}^T\boldsymbol{v} \tag{B.4}$$

For the covariance of estimated parameters, we have

$$\begin{aligned}&E\left[(\hat{\theta}(j) - \theta)(\hat{\theta}(j) - \theta)^T \mid t\right] \\ =&E\left[(\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\boldsymbol{\phi}^T\boldsymbol{v}((\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\boldsymbol{\phi}^T\boldsymbol{v})^T\right] \\ =&E\left[(\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\boldsymbol{\phi}^T\boldsymbol{v}\boldsymbol{v}^T\boldsymbol{\phi}(\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\right] \\ =&E\left[\boldsymbol{v}\boldsymbol{v}^T\right](\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\boldsymbol{\phi}^T\boldsymbol{\phi}(\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1} \\ =&E\left[\boldsymbol{v}\boldsymbol{v}^T\right](\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\end{aligned} \tag{B.5}$$

Assume that the modeling error is a standard normal distribution: $E(v) = 0$ and $Cov(v) = 1$.
From equation B.4, we can show that $\hat{\theta}$ is an unbiased estimation of θ.

$$\begin{aligned}E[\hat{\theta}] = E\left[\theta + (\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\boldsymbol{\phi}^T\boldsymbol{v}\right] &= \theta + (\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1}\boldsymbol{\phi}^T E[\boldsymbol{v}] \\ &= \theta\end{aligned} \tag{B.6}$$

From equation B.5, we get

$$E\left[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\right] = E\left[\boldsymbol{v}\boldsymbol{v}^T\right](\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1} = (\boldsymbol{\phi}^T\boldsymbol{\phi})^{-1} \tag{B.7}$$

In the derivation of RLS, we have , at time step t, $P(t)^{-1} - P(0)^{-1} = \sum_{j=0}^{t-1}\phi(j)\phi(j)^T$. By setting $P(0)$ large enough, e.g. $P(0) = NI$ where N is a large real number (or run random expreiments and collect system information to construct $P(0)$), we have $P(t)^{-1} \approx \sum_{j=0}^{t+-1}\phi(j)\phi(j)^T$. Starting with time step t, we get

$$\text{Cov}[\theta \mid \mathbf{y}(t), \mathbf{u}(t)] = \mathbb{E}[\tilde{\theta}(t)\tilde{\theta}(t)^T \mid \mathbf{y}(t), \mathbf{u}(t)] = P(t) \tag{B.8}$$