Preprints of the
9th International Symposium on Advanced Control of Chemical Processes
The International Federation of Automatic Control
June 7-10, 2015, Whistler, British Columbia, Canada

Dual MPC for FIR Systems: Information Anticipation ⋆

Tor Aksel N. Heirung * **B. Erik Ydstie** ** **Bjarne Foss** *

* *Department of Engineering Cybernetics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway (e-mail: heirung@itk.ntnu.no).*
** *Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA.*

Abstract: Dual model predictive control (DMPC) optimally combines plant excitation and control based on current and predicted parameter estimation errors. Exact solution of dual control problems with constraints is in general computationally prohibitive. Our deterministic equivalent of the stochastic optimal control problem enables convergence toward optimality for a specific class of finite-horizon problems. The cost function shows that the optimal controls are functions of the current and future parameter-estimate error covariances. Our proposed objective-function reformulation provides the optimal combination of caution, probing, and nominal control. We show that the nonconvex optimization problem can be solved as a quadratic program with bilinear constraints. This type of problem can be efficiently solved with existing algorithms based on branch and bound with McCormick-type estimators. We demonstrate the application of DMPC to a singe-input single-output (SISO) finite impulse response (FIR) system. In the simulation example the parameter estimates converge quickly, and accurate and precise estimates are obtained even though the excitation vanishes.

Keywords: Dual control, model predictive control, adaptive control, optimal control, stochastic control, parameter estimation, system identification, excitation, exploration, endogenous learning.

1. INTRODUCTION

Optimal control of plants under parametric uncertainty is a challenging task for which a number of approaches are available. Robust control for systems under worst-case uncertainties was long a major focus in literature on uncertainty in control, with a high degree of maturity reached in the 1990s (Doyle et al., 1990). Adaptive control (Åström and Wittenmark, 1995) also achieved a state of considerable maturity by the early 1990s, and the field includes many well-studied control approaches for systems with unknown and possibly changing dynamics. Plant uncertainty and disturbances are in general accounted for through adjustable control parameters and a mechanism for adjusting those parameters. Optimal control of systems with probabilistic uncertainties has received less attention, and as a result there are fewer strong results in the field. Minimizing the expected value of some function of the output error is a challenging problem when the process is uncertain, and the optimal control must direct the output toward the reference while also exploring the plant to generate information on system characteristics.

This dual nature of an optimal control signal for an unknown plant was first recognized by Feldbaum (1961), who described the twofold effect of a dual control signal as investigating as well as directing. Feldbaum also identified stochastic dynamic programming as an appropriate solution method in his series of papers analyzing the dual control problem. This initial work on dual control was pioneering in its integration of active learning with multistage decision making under uncertainty, and provided a foundation for much of the later work in stochastic adaptive control.

Dual control is an intuitive and appealing concept to the control designer, but the development of practical algorithms in the 1960s and 1970s was limited by the available computational power and the lack of efficient algorithms for nonlinear programming. Åström and Helmersson (1986) were among the first to numerically solve a dual control problem using dynamic programming. They considered a scalar integrator with one unknown parameter and spent 180 CPU hours to obtain the solution using a time horizon of 30 samples (Åström, 1983). The "curse of dimensionality" renders dynamic programming impractical for solving even moderately large problems, despite the superior computers available today. The computational complexity associated with obtaining dual control laws has led to a very limited number of industrial implementations. One of the first applications of suboptimal dual control in process control was reported by Allison et al. (1995). They used several heuristics to simplify the control design and their dual controller improved performance by better identifying gain changes and preventing turn-off.

⋆ We gratefully acknowledge the funding provided by the Center of Integrated Operations in the Petroleum Industry at NTNU, Trondheim, Norway (www.iocenter.no).

Several strategies to break the curse of dimensionality in dual control design have been studied in the literature. A recent approach by Lee and Lee (2009) relies on approximate dynamic programming. Their design uses Monte Carlo simulations with multiple known suboptimal controllers to define a limited region of the hyper-state space and then obtains approximate solutions with dynamic programming within that region. Bayard and Schumitzky (2010) used forward dynamic programming with particle filtering and sampling to develop a related method. Both of these studies demonstrate that even fairly simple dual control problems require advanced solution approaches.

In this article we consider finite impulse response (FIR) systems with probabilistic parametric uncertainty subjected to white noise. One advantage of this model class is that it allows for a compact, exact, and computationally attractive problem reformulation that corresponds to a well-known bilinear formulation in the field of nonconvex optimization. The covariance predictions are explicit functions of the decision variables. In our proposed deterministic reformulation of the stochastic control problem, the current covariance matrix provides a rationale for caution, whereas the future covariance matrices induce probing.

The algorithm we present here is related to our earlier work (Heirung et al., 2013, 2015) in that we previously developed algorithms that improve performance by exciting the plant when parameter uncertainty is large. However, those approaches are based on (partly) heuristic terms in the objective function that reward uncertainty reduction, and not on an exact reformulation of the objective function like the one we present here. Furthermore, the proposed reformulations presented in this paper lead to an optimal control problem with a specific structure, as opposed to a general nonlinear programming (NLP) problem.

The FIR control formulation we present in this paper ensures that the system is sufficiently excited for accurate and precise parameter estimation but does not require a persistently exciting input. Our deterministic equivalent is formulated such that the excitation term in the objective function corresponds to a time-varying L-optimal experiment design criterion (see Gevers et al. (2011)).

This article is organized as follows: we formulate the stochastic control problem in Section 2. In Section 3 we state and prove a theorem and a corollary that we use to formulate the equivalent deterministic control problem. Section 4 contains the dual control algorithm, followed by a simulation example in Section 5. We provide a brief discussion of the results in Section 6. In Section 7 we conclude the paper along with some thoughts for future work.

2. PROBLEM STATEMENT

We consider the single input single output (SISO) output tracking problem for FIR systems of the form

$$y(t) = \sum_{j=1}^{n_b} b_j u(t-j) + v(t) + d(t-1) \tag{1}$$

where $y(t)$ is the plant output, $u(t)$ the control input, $v(t)$ an additive process disturbance, and $d(t)$ can be a constant or time-varying bias term, all at the discrete time t. The difference between $v(t)$ and $d(t)$ is that $d(t)$ is $\mathcal{Y}(t)$ measurable (defined below), where $\mathcal{Y}(t)$ is the available information at time t. Often, $d(t-1) = y(t-1)$ in a more general system formulation. For simplicity of exposition we let $d(t) \equiv 0$ in the following. The unknown parameters $\{b_j\}_{j=1}^{n_b}$ with $b_1 \neq 0$ can be independently drawn from Gaussian distributions; that is, $b_j \sim \mathcal{N}(\hat{b}_j, P_{jj})$ with $P_{ij} = 0$ for $i \neq j$. The independent and identically distributed random variables $v(t)$ are Gaussian with zero mean and variance r. We collect the parameters in the vector

$$\theta = [b_1 \ b_2 \ \cdots \ b_{n_b}]^\top \tag{2}$$

and the inputs in a regressor vector such that

$$\varphi(t-1) = [u(t-1) \ u(t-2) \ \cdots \ u(t-n_b)]^\top \tag{3}$$

and write the plant (1) as

$$y(t+1) = \theta^\top \varphi(t) + v(t+1) \tag{4}$$

A standard definition of information recorded up to and including time t is the set of all past decisions and measurements:

$$\mathcal{Y}(t) = \{u(t), u(t-1), \dots, y(t), y(t-1), \dots\} \tag{5}$$

In this article we find it convenient to extend this definition to include future predicted measurements and future decisions and define

$$\begin{aligned}\mathcal{Y}(k \mid t) = &\\ &\{\underbrace{u(k), \dots, u(t+1), \hat{y}(k \mid t), \dots, \hat{y}(t+1 \mid t),}_{\text{anticipated information, } k \geq t+1} \\ &\quad \underbrace{u(t), u(t-1), \dots, y(t), y(t-1), \dots}_{\text{past information}}\}\end{aligned} \tag{6}$$

Note that the future inputs are decisions and not subject to uncertainty, whereas the future outputs are defined based on $\mathcal{Y}(t)$ in (7). Thus, $\mathcal{Y}(k \mid t)$ does not include information from the plant beyond time t. However, it can be used to characterize anticipated information using the future decisions, the current model, and the uncertainty description. With the mean $\hat{\theta}(t) = \mathrm{E}[\theta \mid \mathcal{Y}(t)] = [\hat{b}_1(t) \ \cdots \ \hat{b}_{n_b}(t)]^\top$ we define the output predictor

$$\begin{aligned}\hat{y}(k+1 \mid t) &= \mathrm{E}\left[y(k+1) \mid \mathcal{Y}(t)\right], \qquad k \geq t \\ &= \hat{\theta}^\top(t)\varphi(k)\end{aligned} \tag{7}$$

Accordingly, $P(t) := \mathrm{E}[(\theta - \hat{\theta}(t))(\theta - \hat{\theta}(t))^\top]$ is the covariance matrix for the estimates at time t.

We define the moving-horizon control cost

$$\begin{aligned}J_N(t) = \sum_{k=t}^{t+N-1} \Big\{ &\mathrm{E}\left[(y(k+1) - y^*(k+1))^2 \mid \mathcal{Y}(k \mid t)\right] \\ &+ w_2 u^2(k) + w_3(\Delta u(k))^2 \Big\}\end{aligned} \tag{8}$$

for MPC (see Mayne et al. (2000)), where $1 \leq N \leq \infty$ is the length of the prediction horizon, $y^*(k+1)$ is the output reference, $w_2 \geq 0$, $w_3 \geq 0$ are tuning weights, and $\Delta u(k) := u(k) - u(k-1)$ is the control input change.

The moving-horizon stochastic optimal control problem we want to solve at each time t is then

$$\min J_N(t) \tag{9a}$$

subject to

$$\hat{y}(k+1\,|\,t) = \hat{\theta}^\top(t)\varphi(k) \tag{9b}$$
$$y_{\min} \le \hat{y}(k+1\,|\,t) \le y_{\max} \tag{9c}$$
$$u_{\min} \le u(k) \le u_{\max} \tag{9d}$$
$$\Delta u_{\min} \le \Delta u(k) \le \Delta u_{\max} \tag{9e}$$
$$k \in [t,\, t+N-1] \tag{9f}$$
$$\hat{\theta}(t),\ \{u(k)\}_{k=t-1}^{t-n_b+1},\ P(t) \text{ given} \tag{9g}$$

Note that we take the expected value of the squared output error with respect to $\mathcal{Y}(k\,|\,t)$ in $J_N(t)$, which has interesting consequences; we discuss these in Section 3. If we instead use the expectation of the output values with respect to the current information $\mathcal{Y}(t)$ we get the classic MPC output cost

$$\sum_{k=t}^{t+N-1} \big(\mathrm{E}\left[y(k+1)\,\middle|\,\mathcal{Y}(t)\right] - y^*(k+1)\big)^2 = \sum_{k=t}^{t+N-1} (\hat{y}(k+1\,|\,t) - y^*(k+1))^2 \tag{10}$$

which results in a standard certainty-equivalence type MPC where predictions are based on the most recent parameter estimate $\hat{\theta}(t)$ as defined by (7). This objective rewards neither caution nor probing (see Bar-Shalom (1981)), which means the resulting controller is not risk averse in the face of large uncertainty and does not excite the system to reduce uncertainty. Combining this idea with system identification then gives the indirect adaptive predictive controller.

We write $y(k+1\,|\,t)$ in (9) since the parameter estimate $\hat{\theta}(t)$ appears directly in the definition of this variable; it would be consistent to also note the dependence on t explicitly in the other predicted optimization variables. However, since it is clear from context whether a variable is an optimization variable on the prediction horizon in a problem constructed at time t or a physical realization of that variable in the process, we omit the explicit dependence on t to simplify the notation.

We now describe the method for updating the parameter-estimate statistics, and then discuss a how we evaluate $J_N(t)$ to transform (9) into a tractable problem for $N < \infty$.

2.1 Parameter estimation with past information

We estimate the parameters in (7) online by minimizing a least-squares criterion (Ljung, 1999). Let $R(t)$ be the information matrix

$$R(t) = \sum_{k=t_0}^{t} r^{-1}\lambda^{t-k}\varphi(k-1)\varphi^\top(k-1) \tag{11}$$

with the forgetting factor $\lambda \in (0,1]$. $R(t)$ can alternatively be expressed recursively,

$$R(t) = \lambda R(t-1) + r^{-1}\varphi(t-1)\varphi^\top(t-1), \quad t > t_0 \tag{12}$$

with $R(t_0) = P^{-1}(t_0)$ given. Instead of calculating the inverse $P(t) = R^{-1}(t)$ at every time t we update the least squares parameter estimate recursively using

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)\big(y(t) - \hat{\theta}^\top(t-1)\varphi(t-1)\big) \tag{13a}$$
$$K(t) = P(t-1)\varphi(t-1) \times \big(r\lambda + \varphi^\top(t-1)P(t-1)\varphi(t-1)\big)^{-1} \tag{13b}$$
$$P(t) = \big(I - K(t)\varphi^\top(t-1)\big)P(t-1)(1/\lambda) \tag{13c}$$

This algorithm can be derived with the matrix-inversion lemma (see Ljung (1999)). When $\lambda = 1$ the conditional distribution of θ given the information $\mathcal{Y}(t)$ is Gaussian with mean $\hat{\theta}(t)$ and covariance $P(t)$ as described by the equation set (13) (Åström and Wittenmark, 1995). Note that $r = 1$ in (13) when there is no process noise in (4).

3. PROPOSED REFORMULATION

FIR processes belong to a class of systems where the output is dependent on only past inputs and not past outputs. Since we have the freedom to choose the future decisions, we effectively decide $\mathcal{Y}(k\,|\,t)$ for any $k \ge t$. This means that the future covariances $P(t)$ can be predicted exactly and are determined by our choice of future process inputs $u(t)$. The following theorem, which we use to reformulate the optimal control problem (9), is a consequence of this.

Theorem 1. For a stochastic FIR process (4),

$$\mathrm{E}[y^2(k+1)\,|\,\mathcal{Y}(k\,|\,t)] = \hat{y}^2(k+1\,|\,t) + \varphi^\top(k)P(k)\varphi(k) + r \tag{14}$$

for all $k \ge t$.

Proof. For any $k \ge t$ we take $\mathrm{E}[y^2(k+1)\,|\,\mathcal{Y}(k\,|\,t)]$ and add and subtract the model $\hat{y}(k+1\,|\,k) = \hat{\theta}^\top(k)\varphi(k)$ such that

$$\mathrm{E}[y^2(k+1)\,|\,\mathcal{Y}(k\,|\,t)] = \mathrm{E}\left[\big(\hat{y}(k+1\,|\,k) + y(k+1) - \hat{y}(k+1\,|\,k)\big)^2\,\middle|\,\mathcal{Y}(k\,|\,t)\right] \tag{15}$$

since $\mathcal{Y}(k\,|\,t)$ is known. We let $\tilde{\theta}(k) := \theta - \hat{\theta}(k)$ and then have

$$\mathrm{E}[y^2(k+1)\,|\,\mathcal{Y}(k\,|\,t)] = \mathrm{E}\left[\big(\hat{y}(k+1\,|\,k) + \tilde{\theta}^\top(k)\varphi(k) + v(k+1)\big)^2\,\middle|\,\mathcal{Y}(k\,|\,t)\right] \tag{16}$$

Expanding the square gives

$$\begin{aligned}\mathrm{E}[y^2(k+1)\,|\,\mathcal{Y}(k\,|\,t)] = \mathrm{E}\big[&\hat{y}^2(k+1\,|\,k) + \varphi^\top(k)\tilde{\theta}(k)\tilde{\theta}^\top(k)\varphi(k) + v^2(k+1)\\ &+ 2\hat{y}(k+1\,|\,k)\varphi^\top(k)\tilde{\theta}(k) + 2\hat{y}(k+1\,|\,k)v(k+1)\\ &+ 2\varphi^\top(k)\tilde{\theta}(k)v(k+1)\,|\,\mathcal{Y}(k\,|\,t)\big]\end{aligned} \tag{17}$$

The three last terms are all zero because $\mathrm{E}[\tilde{\theta}(k)\,|\,\mathcal{Y}(k\,|\,t)] = 0$, $\mathrm{E}[v(k)] = 0$, and $\tilde{\theta}(k)$ and $v(k+1)$ are independent. Using

$$\mathrm{E}[\theta\,|\,\mathcal{Y}(k\,|\,t)] = \mathrm{E}[\theta\,|\,\mathcal{Y}(t)] = \hat{\theta}(t), \qquad k \ge t \tag{18}$$

and the definition of $P(t)$ leaves

$$\mathrm{E}[y^2(k+1)\,|\,\mathcal{Y}(k\,|\,t)] = \hat{y}^2(k+1\,|\,t) + \varphi^\top(k)P(k)\varphi(k) + r \tag{19}$$

which completes the proof. □

The following corollary extends the above result to the case where we want to track a time-varying output reference $y^*(t)$.

Corollary 2. For a stochastic FIR process (4),

$$\begin{aligned}&\mathrm{E}[(y(k+1)-y^*(k+1))^2 \mid \mathcal{Y}(k \mid t)] = \\ &\quad (\hat{y}(k+1 \mid t) - y^*(k+1))^2 + \varphi^\top(k)P(t)\varphi(k) + r \end{aligned} \tag{20}$$

for all $k \geq t$.

Proof. We have that

$$\begin{aligned}&\mathrm{E}[(y(k+1)-y^*(k+1))^2 \mid \mathcal{Y}(k \mid t)] \\ &\quad = \mathrm{E}[y^2(k+1) - 2y(k+1)y^*(k+1) \\ &\qquad + (y^*(k+1))^2 \mid \mathcal{Y}(k \mid t)] \\ &\quad = \mathrm{E}[y^2(k+1) \mid \mathcal{Y}(k \mid t)] \\ &\qquad - 2\,\mathrm{E}[y(k+1) \mid \mathcal{Y}(k \mid t)]y^*(k+1) + (y^*(k+1))^2\end{aligned}$$

Using Theorem 1, we can write

$$\begin{aligned}&\mathrm{E}[(y(k+1)-y^*(k+1))^2 \mid \mathcal{Y}(k \mid t)] \\ &\quad = \hat{y}^2(k+1 \mid t) + \varphi^\top(k)P(k)\varphi(k) + r \\ &\qquad - 2\hat{y}(k+1 \mid t)y^*(k+1) + (y^*(k+1))^2 \\ &\quad = (\hat{y}(k+1 \mid t) - y^*(k+1))^2 + \varphi^\top(k)P(k)\varphi(k) + r\end{aligned}$$

which is the desired result. □

Corollary 2 allows the stochastic objective (8) to be reformulated into the equivalent deterministic function

$$\begin{aligned}J_N(t) = \sum_{k=t}^{t+N-1} \Big\{ &(\hat{y}(k+1 \mid t) - y^*(k+1))^2 + \varphi^\top(k)P(k)\varphi(k) \\ &+ r + w_2u^2(k) + w_3(\Delta u(k))^2 \Big\}\end{aligned} \tag{21}$$

which can be minimized by augmenting the constraint set of (9) to arrive at a deterministic optimal control problem that is equivalent to (9). This problem is

$$\min J_N(t) \tag{22a}$$

subject to

$$\hat{y}(k+1 \mid t) = \hat{\theta}^\top(t)\varphi(k) \tag{22b}$$
$$K(k+1) = P(k)\varphi(k)\big(r + \varphi^\top(k)P(k)\varphi(k)\big)^{-1} \tag{22c}$$
$$P(k+1) = \big(I - K(k+1)\varphi^\top(k)\big)P(k) \tag{22d}$$
$$y_{\min} \leq \hat{y}(k+1 \mid t) \leq y_{\max} \tag{22e}$$
$$u_{\min} \leq u(k) \leq u_{\max} \tag{22f}$$
$$\Delta u_{\min} \leq \Delta u(k) \leq \Delta u_{\max} \tag{22g}$$
$$k \in [t,\, t+N-1] \tag{22h}$$
$$\hat{\theta}(t),\ \{u(k)\}_{k=t-1}^{t-n_b+1},\ P(t) \text{ given} \tag{22i}$$

The solution to this nonlinear programming problem includes an optimal sequence of predicted control inputs $\{u^{\mathrm{o}}(k \mid t)\}_{k=t}^{N-1}$, the first element of which is applied to the plant: $u(t) = u^{\mathrm{o}}(t \mid t)$.

Note that the formulation of the objective $J_N(t)$ in (21) converges to the certainty-equivalence objective (10) as the uncertainty represented by $P(t)$ goes to zero. This implies that the excitation induced by the parameter uncertainty vanishes as the uncertainty is resolved. This property highlights the main idea of certainty equivalence control, which is that we assume that the parameter estimates actually give the correct representation of the plant dynamics.

Although the solution to (22) exactly minimizes

$$\begin{aligned}J_N(t) = \sum_{k=t}^{t+N-1} \Big\{ &\mathrm{E}\left[(y(k+1)-y^*(k+1))^2 \mid \mathcal{Y}(k \mid t)\right] \\ &+ w_2u^2(k) + w_3(\Delta u(k))^2 \Big\}\end{aligned} \tag{8}$$

over the finite horizon $k \in [t,\, t+N-1]$, the constraint set is nonconvex because of the inclusion of the nonlinear covariance-prediction equality constraints (22c)–(22d). This motivates investigation of reformulation approaches that facilitate solving the optimal control problem. We consider this reformulation the main contribution of the paper.

3.1 Reformulation

In order to simplify the formulation (22) we define the vector $z(k) := P(k)\varphi(k)$ for $k \in [t,\, t+N-1]$, or equivalently

$$R(k)z(k) = \varphi(k), \qquad k \in [t,\, t+N-1] \tag{23}$$

The objective function (22) is then equivalent to

$$\begin{aligned}J_N(t) = \sum_{k=t}^{t+N-1} \Big\{ &(\hat{y}(k+1 \mid t) - y^*(k+1))^2 + \varphi^\top(k)z(k) \\ &+ r + w_2u^2(k) + w_3(\Delta u(k))^2 \Big\}\end{aligned} \tag{24}$$

Accordingly, the optimal control problem (22) is equivalent to

$$\min J_N(t) \tag{25a}$$

subject to

$$\hat{y}(k+1 \mid t) = \hat{\theta}^\top(t)\varphi(k) \tag{25b}$$
$$R(k+1) = R(k) + r^{-1}\varphi(k)\varphi^\top(k) \tag{25c}$$
$$R(k+1)z(k+1) = \varphi(k+1) \tag{25d}$$
$$y_{\min} \leq \hat{y}(k+1 \mid t) \leq y_{\max} \tag{25e}$$
$$u_{\min} \leq u(k) \leq u_{\max} \tag{25f}$$
$$\Delta u_{\min} \leq \Delta u(k) \leq \Delta u_{\max} \tag{25g}$$
$$k \in [t,\, t+N-1] \tag{25h}$$
$$\hat{\theta}(t),\ \{u(k)\}_{k=t-1}^{t-n_b+1},\ R(t),\ z(t) \text{ given} \tag{25i}$$

Note that r can be removed from $J_N(t)$ without changing the solution of the optimization problems. This formulation is still nonlinear, but all nonlinearities are now either bilinear or quadratic. In fact, (25) is a quadratically-constrained quadratic programming (QCQP) problem and can be written in the modified standard form (cf. Misener and Floudas (2012))

$$\min_x x^\top Q_0(t)x + \alpha_0^\top(t)x \tag{26a}$$

subject to

$$0 \leq x^\top Q_i x + \alpha_i^\top x \leq 0, \qquad i = 1, \ldots, n_q \tag{26b}$$
$$\beta_{\min,i} \leq \alpha_i^\top(t)x \leq \beta_{\max,i}, \quad i = n_q+1, \ldots, n_q+n_\ell \tag{26c}$$
$$\text{some elements of } x \text{ given} \tag{26d}$$

where the quadratic and the linear constraints are separated for clarity. Here, n_q is the number of quadratic constraints and n_ℓ is the number of linear constraints, x is a vector containing all variables, linear equality constraints are formulated by letting $\beta_{\min,i} = \beta_{\max,i}$, and the constant terms in the objective (24) have been left out. The structures of the matrices Q_i ($i = 0, \ldots, n_q$) and the vectors α_i ($i = 0, \ldots, n_q + n_\ell$), $\beta_{\min,i}$, and $\beta_{\max,i}$

($i = n_q + 1, \ldots, n_q + n_\ell$), depend on how the variables are organized in the vector x.

There are several algorithms that solve QCQP problems to ϵ-global optimality (Tawarmalani and Sahinidis, 2002). See Misener and Floudas (2012) for a presentation of one such algorithm and a good overview of the QCQP problem class.

4. DUAL CONTROL

We now propose the following algorithm for DMPC with a receding finite prediction horizon.

Algorithm for dual control

0) Initialize at t_0: specify $\hat{\theta}(t_0)$, $\{u(k)\}_{k=t_0-1}^{t_0-1-n_b}$, $P(t_0)$.
1) Collect plant data: measure $y(t)$ and $u(t-1)$.
2) Estimate parameters: update $\hat{\theta}(t)$ and $P(t)$ using (13).
3) Solve (25) to obtain the solution $\{u^{\mathrm{o}}(k\,|\,t)\}_{k=t}^{N-1}$.
4) Implement $u(t) = u^{\mathrm{o}}(t\,|\,t)$.
5) Set $t \leftarrow t+1$ and go to step 1.

Note that we also measure $u(t-1)$ since the actual control sometimes deviates from the one calculated in the algorithm.

This algorithm is based on a standard MPC formulation, where an optimal control problem is solved at each sampling instant using the current system state as the initial state. The solution to the control problem is the control sequence $\{u^{\mathrm{o}}(k\,|\,t)\}_{k=t}^{N-1}$, the first element of which, $u^{\mathrm{o}}(t\,|\,t)$, is implemented as the control input (Mayne et al., 2000). Our algorithm differs from standard MPC in that the system state is given by the hyper-state $\left(\hat{\theta}(t),\ \{u(k)\}_{k=t-1}^{t-n_b+1},\ R(t)\right)$. Furthermore, (25) is not a true open-loop problem since the uncertainty predictions implicitly anticipate a closed loop, which we discuss in Section 6. Our algorithm is also similar to a certainty-equivalence controller in that the true parameter values are not used, but rather their expected values; however, the expected values are not used with the assumption that they equal the true parameter values.

5. EXAMPLE

We now demonstrate our algorithm on a small example. The simulations go from $t_0 = 0$ to $t_f = 10$ with $y^* = 2$, $\theta = [1.00,\ 0.20]^\top$, $\hat{\theta}(0) = [1.50,\ 0.50]^\top$, $P(0) = 10^2 I$, $\lambda = 1$, $y(0) = -1.0$, $u(-1) = -1$, $N = 4$, $w_2 = 0$, $w_3 = 0.1$, $-y_{\min} = y_{\max} = 10$, $-u_{\min} = u_{\max} = 5$, $-\Delta u_{\min} = \Delta u_{\max} = \infty$.

The system was simulated using MATLAB and the QCQP was solved using the local NLP solver IPOPT (Wächter and Biegler, 2005) under GAMS (GAMS Development Corporation), which provides gradients obtained using automatic differentiation. The example is run on a standard laptop computer and the average solution time for the optimal control problems is 0.43 s with a standard deviation of 0.49 s. Figure 1 shows the results.

The first plot in Figure 1 shows the output $y(t)$ for the two values of r and the reference value y^*. The two corresponding input sequences are plotted directly below, together with the optimal steady-state input $u^* = y^*/(b_1 + b_2)$. The true values and the estimates of the system parameters b_1 and b_2 are shown in the third and fourth plots, respectively.

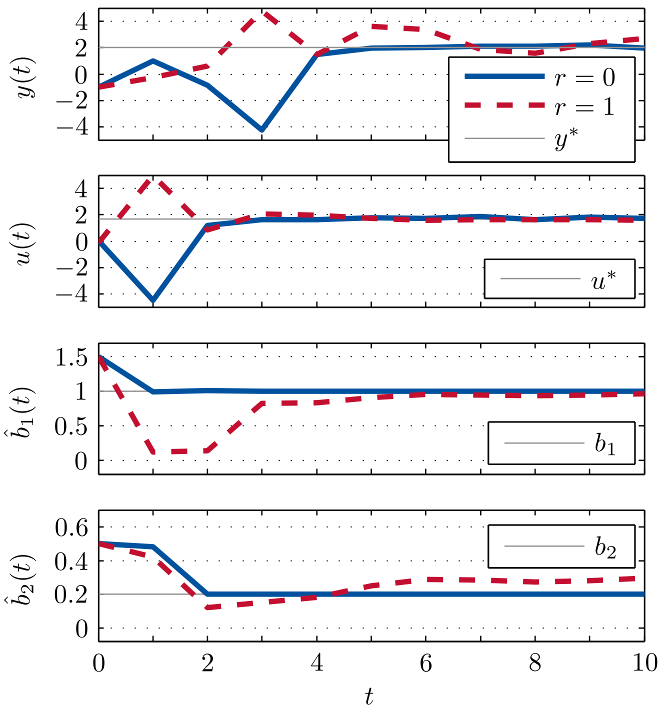

Fig. 1. Simulation example demonstrating the algorithm for both the noiseless (noise variance $r = 0$) and noise-corrupted (noise variance $r = 1$) case.

Both parameters are correctly identified after two time steps when $r = 0$. With the low signal-to-noise ratio ($r = 1$) it takes the parameters longer to approach their true values, but both are close after three time steps.

In both cases the first control $u(0)$ is small but nonzero, due to the large values in $P(0)$ which rewards caution in the control. From $t = 1$ there is some initial excitation (or probing) motivated by the future reduction of uncertainty. The probing is followed by the inputs quickly approaching u^*. The vanishing excitation coincides with the uncertainty going to zero (all elements of $P(2)$ are very close to zero for both cases), which illustrates how the objective $J_N(t)$ converges to the certainty-equivalence objective (10) as $P(t) \to 0$.

The output approaches the reference value after four time steps in both cases, which is consistent with the system parameters converging after two or three time steps and the output depending on the two previous inputs.

6. DISCUSSION

The dual control formulation we present here includes no notion of a trade-off or balance between the control and excitation parts of the objective. One can construct controllers that balances a control effort, a certainty-equivalence output objective like (10), and excitation, but this approach is heuristic and leads to suboptimal control. We here analyze the expected value with respect to future decisions; our reformulation anticipates future information and predicts future uncertainty and as a result provides the optimal combination with no notion of a trade-off or freedom to trade one for the other.

One consequence of the suggested approach is that the learning is endogenized in the controller, meaning there is no separation of control and excitation. The controller is not only able to actively excite the system, it is also aware of the mechanisms for learning through the constraints that describe the uncertainty propagation. The uncertainty awareness also affects the cautiousness of the controls.

Since our DMPC anticipates future decisions in the prediction of uncertainty it is in some sense aware of the closed control loop. The DMPC is therefore a type of closed-loop feedback controller, in contrast to a standard MPC formulation that makes open loop predictions without any built-in knowledge of the control loop being closed.

The complexity of the optimal control problem (25) increases moderately with the number of unknown model parameters n_p and the length of the prediction horizon N. The objective function (24) contains n_pN bilinear terms of the form $\varphi_j(k)z_j(k)$ (j denotes vector element), whereas the uncertainty-propagation constraints (25c) and (25d) contain n_p^2N bilinear terms each. The symmetric nature of the quadratic equality constraints can be exploited in implementation to reduce this number. The quadratic growth cannot be avoided, however.

It is important to note that the algorithm shares with all finite-horizon MPC controllers that endpoint constraints or cost are needed to prove stability. These issues are under current investigation.

7. CONCLUSIONS AND FUTURE WORK

The exact reformulation of the stochastic optimal-control problem for FIR systems that we derive in this article provides a clear illustration of caution, probing, certainty-equivalence, and the dual effect in stochastic control. Formulating the control problem in terms of an information matrix that is the inverse of the covariance allows us to cast the problem as a QCQP. This problem formulation enables the use of standard algorithms for solving dual control problems for FIR systems with a receding horizon and reduces the complexity of the optimal control problem so that ϵ-global optimal solutions can be obtained efficiently. The performance of the DMPC is illustrated with an example that demonstrates caution, probing, the convergence to certainty equivalence, and the dual effect.

Future work includes studying the effect of running the DMPC algorithm with global solutions to the QCQPs instead of the local solutions we use here, as well as deriving a quadratic formulation for probabilistic output constraints. The practicality of solving nonconvex optimal control problems to global optimality online will be investigated. Deriving tight variable bounds greatly facilitates faster convergence to global optimality and will be the topic of a future paper.

REFERENCES

K. J. Åström. Theory and applications of adaptive control—a survey. *Automatica*, 19(5):471–486, 1983.

K. J. Åström and A. Helmersson. Dual control of an integrator with unknown gain. *Computers and Mathematics with Applications*, 12(6):653–662, 1986.

K. J. Åström and B. Wittenmark. *Adaptive Control*. Dover, second edition, 1995.

B. J. Allison, J. E. Ciarniello, P. J. C. Tessier, and G. A. Dumont. Dual adaptive control of chip refiner motor load. *Automatica*, 31(8):1169–1184, August 1995.

Y. Bar-Shalom. Stochastic dynamic programming: Caution and probing. *IEEE Transactions on Automatic Control*, 26(5):1184–1195, October 1981.

D. S. Bayard and A. Schumitzky. Implicit dual control based on particle filtering and forward dynamic programming. *International Journal of Adaptive Control and Signal Processing*, 24(3):155–177, 2010.

J. C. Doyle, B. A. Francis, and A. Tannenbaum. *Feedback Control Theory*. Macmillan, 1990.

A. A. Feldbaum. Dual-control theory. I. *Automation and Remote Control*, 21(9):874–880, 1961.

GAMS Development Corporation. The General Algebraic Modeling System. URL `http://www.gams.com/`.

M. Gevers, X. Bombois, R. Hildebrand, and G. Solari. Optimal experiment design for open and closed-loop system identification. *Communications in Information and Systems*, 11(3):197–224, 2011.

T. A. N. Heirung, B. E. Ydstie, and B. Foss. An MPC approach to dual control. In *IFAC International Symposium on Dynamics and Control of Process Systems*, pages 69–74, Mumbai, India, 2013.

T. A. N. Heirung, B. Foss, and B. E. Ydstie. MPC-based dual control with online experiment design. *Journal of Process Control*, conditionally accepted, 2015.

J. M. Lee and J. H. Lee. An approximate dynamic programming based approach to dual adaptive control. *Journal of Process Control*, 19(5):859–864, 2009.

L. Ljung. *System Identification. Theory for the User*. Prentice-Hall, second edition, 1999.

D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

R. Misener and C. A. Floudas. GloMIQO: Global mixed-integer quadratic optimizer. *Journal of Global Optimization*, 57(1):3–50, 2012.

M. Tawarmalani and N. V. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming. Theory, Algorithms, Software, and Applications*. Kluwer, 2002.

A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, 2005.

