2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Adaptive Disturbance Estimation for Offset-Free SISO Model Predictive Control

Jakob Kjøbsted Huusom, Niels Kjølstad Poulsen, Sten Bay Jørgensen and John Bagterp Jørgensen

Abstract— **Offset free tracking in Model Predictive Control requires estimation of unmeasured disturbances or the inclusion of an integrator. An algorithm for estimation of an unknown disturbance based on adaptive estimation with time varying forgetting is introduced and benchmarked against the classical disturbance modelling approach, where the system description is augmented by a disturbance state. The time varying forgetting renders the new approach less sensitive to the nature of the disturbance. By simulation we demonstrate that this algorithm is advantageous in case of infrequent step disturbances of any magnitude.**

I. INTRODUCTION

Model Predictive Control (MPC) is a state of the art control technology which utilizes a model of the system to predict the process output over some future horizon and solve a quadratic optimization problem with the control signal as decision variables. Inequality constraints can be formulated for both manipulated variables and the process outputs. The first of the controls are implemented. After retrieving the next process output the problem is solved again for the next control etc. Early achievements and industrial implementations in Model Prediction Control include IDCOM and Dynamic Matrix Control [1], [2]. These early algorithms were based on step or impulse response models. More general linear input-output models structure were used in Generalized Predictive Control [3], but an interest in MPC implementations based on state space models were created by the seminal paper [4]. The state space approach provides a unified framework for discussion of the various predictive control algorithms and is well suited for stability analysis [5].

In this paper the plant is represented by the linear, discrete time, single input/single output ARX model (1). This model class is selected based on a system identification argument. This class is linear in the parameters and the parameter estimation problem is convex.

$$A(q^{-1})y(t) = B(q^{-1})u(t) + d + \varepsilon(t) \tag{1a}$$

where A and B are polynomials of order n in the backward shift operator q^{-1} and $\varepsilon(t) \in \mathcal{N}_{iid}(0, \sigma^2)$.

$$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_n q^{-n} \tag{1b}$$
$$B(q^{-1}) = b_1 q^{-1} + b_2 q^{-2} + \cdots + b_n q^{-n} \tag{1c}$$

J. K. Huusom and S. B. Jørgensen are with the Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 227. DK - 2800 Lyngby, Denmark jkh@kt.dtu.dk, sbj@kt.dtu.dk

N. K. Poulsen and J. B. Jørgensen are with the Department of Informatics and Mathematical Modelling, Technical University of Denmark, Building 321. DK - 2800 Lyngby, Denmark nkp@imm.dtu.dk, jbj@imm.dtu.dk

The time invariant input d is an unmeasured disturbance which is assumed constant over some horizon. Since the system is exposed to such a constant disturbance a MPC implementation will not provide offset-free tracking unless an integrator is implemented or the disturbance is estimated as part of the controller [6]. Performance of offset-free ARX-model based MPC with either an integrator or by the disturbance modelling approach is analyzed [7]. In this paper, we present an adaptive technique to estimate a time varying disturbance. A similar approach has successfully been applied to estimate unmeasured disturbances for a fatty acid distillation column MPC control problem [8]. Offset-free tracking is a property which must be required for any MPC implementation since unmeasured step disturbances in the input is common in the process industries when e.g. a feed source changes. Examples are refineries and cement industries where the composition of the crude oil or raw minerals may change significantly when feed is changed from one source to another. In biochemical production the problem may appear in continues downstream processing where the feed comes from batch processes. The inclusion of disturbance states may also be necessary for rendering offset-free tracking in presence of a model/plant mismatch which will be the case for all industrial implementations.

The main contribution of this paper is the application of adaptive disturbance estimation with time varying forgetting in an MPC to eliminate offset. This scheme is advantageous for system exhibiting infrequent step disturbances since the estimation algorithm acts fast. Increasing the sensitivity when large prediction errors are observed and decrease the sensitivity towards the random process noise when the disturbance remains at a constraint level. This feature is not possible in MPC implementations where the system is augmented with disturbance states. The paper is organized as follows. The basic MPC algorithm is introduced in Sec. II. Then the classical disturbance modelling approach and the novel adaptive disturbance estimation algorithm is presented in the two following sections. An illustrative simulation study is performed in Section V. Conclusions are drawn in Sec. VI.

II. MODEL PREDICTIVE CONTROL BASED ON ARX MODELS

The ARX model (1), excluding the constant disturbance, is realized as a stationary state space model in innovation form

$$x_{k+1} = Ax_k + Bu_k + K\varepsilon_k \tag{2a}$$
$$y_k = Cx_k + \varepsilon_k \tag{2b}$$

978-1-4577-0081-1/11/$26.00 ©2011 AACC

with the matrices (A,B,K,C) in observer canonical form

$$A = \left[\begin{array}{c|cccc} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -a_{n-1} & 0 & 0 & \cdots & 1 \\ \hline -a_n & 0 & 0 & \cdots & 0 \end{array}\right] B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} K = \begin{bmatrix} -a_1 \\ \vdots \\ -a_n \end{bmatrix}$$

$$C = \left[\begin{array}{cccc} 1 & 0 & \cdots & 0 \end{array}\right]$$

The optimal predictions in the stationary state space model in innovation form (2) is based on computation of the innovations

$$\varepsilon_k = y_k - \hat{y}_{k|k-1} \tag{3}$$

using the measurement y_k at time k and the one-step-ahead prediction, $\hat{y}_{k|k-1} = C\hat{x}_{k|k-1}$. The one-step-ahead prediction of the states and outputs are

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_{k|k} + K\varepsilon_k \tag{4a}$$
$$\hat{y}_{k+1|k} = C\hat{x}_{k+1|k} \tag{4b}$$

and similarly the $(j+1)$-step-ahead ($j \geq 1$) predictions are

$$\hat{x}_{k+1+j|k} = A\hat{x}_{k+j|k} + Bu_{k+j|k} \quad j = 1, \ldots, N-1 \tag{5a}$$
$$\hat{y}_{k+1+j|k} = C\hat{x}_{k+1+j|k} \quad j = 1, \ldots, N-1 \tag{5b}$$

The ℓ_2-based constrained predictive controller uses an objective function of the form

$$\phi = \frac{1}{2}\sum_{j=0}^{N-1} \|\hat{y}_{k+1+j|k} - r_{k+1+j|k}\|_2^2 + \rho\|\Delta u_{k+j|k}\|_2^2 \tag{6}$$

where ρ denotes the relative penalty on the control move compared to the tracking error. This objective function obviously depends on the control variables, hence the optimal control problem is

$$\min_{\{u_{k+j|k}\}_{j=0}^{N-1}} \quad \phi = \phi(\{u_{k+j|k}\}_{j=0}^{N-1}) \tag{7a}$$
$$s.t. \quad (4), (5) \tag{7b}$$
$$u_{\min} \leq u_{k+j|k} \leq u_{\max} \quad j \in \mathcal{N} \tag{7c}$$
$$\Delta u_{\min} \leq \Delta u_{k+j|k} \leq \Delta u_{\max} \quad j \in \mathcal{N} \tag{7d}$$

with $\Delta u_{k+j|k} = u_{k+j|k} - u_{k+j-1|k}$ ($j \in \mathcal{N}$), $u_{k-1|k} = \hat{u}_{k-1|k-1}$, and $\mathcal{N} = \{0, 1, \ldots, N-1\}$. The optimal solution is denoted $\{\hat{u}_{k+j|k}\}_{j=0}^{N-1}$. Only the first part of the solution, $\hat{u}_{k|k}$, is implemented on the process and the computations are repeated as new measurements arrive. The constrained optimal control problem (7) can be converted into a standard convex quadratic program [7].

III. DISTURBANCE MODELLING

In presence of unmeasured disturbances the classical approach to achieve offset free tracking performance for a model predictive control implementation, is to include disturbance states in the process model. This method was originally presented in [9] and a thorough presentation of disturbance models for linear model predictive control is given in [10] and [11] with conditions for detectability of the augmented systems.

Given a general system description on state space form

$$x_{k+1} = Ax_k + Bu_k + B_d d_k + Gw_k \tag{8a}$$
$$y_k = Cx_k + C_d d_k + v_k \tag{8b}$$

It is assumed that the disturbance evolves as

$$d_{k+1} = d_k + \xi_k \tag{9}$$

where the noise in the system is given by the following Gaussian distribution

$$\begin{bmatrix} w_k \\ \xi_k \\ v_k \end{bmatrix} = \mathcal{N}_{iid}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 & S \\ 0 & Q_\xi & 0 \\ S^T & 0 & R \end{bmatrix}\right) \tag{10}$$

The augmented system description becomes

$$\begin{bmatrix} x_{k+1} \\ d_{k+1} \end{bmatrix} = \begin{bmatrix} A & B_d \\ 0 & I \end{bmatrix}\begin{bmatrix} x_k \\ d_k \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_k + \begin{bmatrix} G & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} w_k \\ \xi_k \end{bmatrix} \tag{11a}$$

$$y_k = \begin{bmatrix} C & C_d \end{bmatrix}\begin{bmatrix} x_k \\ d_k \end{bmatrix} + v_k \tag{11b}$$

The general idea is to use a state estimator for the augmented system in the model predictive controller. The prediction equations become

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_{k|k} + B_d\hat{d}_{k|k} \tag{12a}$$
$$\hat{d}_{k+1|k} = \hat{d}_{k|k} \tag{12b}$$

and the stationary Kalman filter with the gains, $\{L_x, L_d\}$, calculated from the solution to the Riccati equation, are

$$\begin{bmatrix} \hat{x}_{k|k} \\ \hat{d}_{k|k} \end{bmatrix} = \begin{bmatrix} \hat{x}_{k|k-1} \\ \hat{d}_{k|k-1} \end{bmatrix} + \begin{bmatrix} L_x \\ L_d \end{bmatrix}(y_k - C\hat{x}_{k|k-1} - C_d\hat{d}_{k|k-1}) \tag{13}$$

This state estimator can estimate the unmeasured disturbance and render the controller capable of offset free tracking. Q_ξ only affect the gain L_d in the filter and for $Q_\xi \to \infty$, $L_d \to 1$ which would correspond to designing the MPC based on a ARIX model. For the ARX model structure in (1) the disturbance model is given by $\{B_d = G = K, C_d = 1, Q = R = S = \sigma^2\}$. [12] show that any choice of disturbance model can give the same closed loop performance despite the origin of the disturbance. The requirement is that the disturbance covariance, used in calculation of the estimator gain, is estimated from the autocovariance of plant data.

IV. ADAPTIVE DISTURBANCE ESTIMATION

If a disturbance is known, a MPC implementation using this knowledge in the predictions can be implemented. If we at time k have an estimate of $\hat{d}_k$, for future predictions it will be assumed that this level $\hat{d}$ of the disturbance will remain constant. The mapping from the constant disturbance to the states in the state space realization is the matrix K just as for the random disturbance. Hence the one step ahead predictions will be

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k-1} + B\hat{u}_{k|k} + K(\hat{d} + e_k) \\ \hat{y}_{k+1|k} &= C\hat{x}_{k|k-1} \end{aligned} \tag{14}$$

e_{k+j} is assumed zero in predictions more than one step ahead while $\hat{d}$ is constant. Note that $\hat{y}$ is only dependent on the disturbance through the state vector while y has a direct dependence cf. (1). The reason being that $\hat{u}$ is calculated by the MPC using knowledge of $\hat{d}$ in the predictions. Since the disturbance is unknown, we estimate it adaptively online. We use a recursive least squares algorithm with discounted measurements, minimize the following least squares criterion [13]

$$V_N = \sum_{i=1}^{N} \lambda^{N-i} e_i^2 \tag{15}$$

or written in recursive from

$$V_k = \lambda V_{k-1} + \frac{1}{2} e_k^2 \tag{16}$$

where $e_k = y_k - C\hat{x}_{k|k-1}$ is the prediction error at time instant k, N is the length of the current horizon and the constant $\lambda \in [0; 1]$ is the forgetting factor which state the relative importance of old data compared to new in the loss function. For a constant forgetting factor less than 1, the weighting of old data deceases exponentially in time, typical values of λ would be in the range from 0.9 to 0.99. An estimate of the disturbance such that the loss function (15) is minimized is obtained by the recursive algorithm

$$\hat{d}_k = \hat{d}_{k-1} + \kappa_k e_k \tag{17a}$$

$$\kappa_k = \frac{P_{k-1}}{\lambda + P_{k-1}} \tag{17b}$$

$$P_k = (1 - \kappa_k) P_{k-1} \lambda^{-1} \tag{17c}$$

In this recursion P_k is the estimated variance of the disturbance estimate $\hat{d}_k$. κ_k is a gain from the prediction error to the change in $\hat{d}$ from time $k-1$ to k. Since both the disturbance and the variance of $\hat{d}_k$ is unknown this recursion can be initiated by setting $\hat{d}_0 = 0$ and P_0 equal to some small value. Typically the covariance P is initiated with a large value since the objective is to achieve a fast estimation of an unknown parameter. Here however the unknown parameter, d, is initially assumed to be zero so a large value of P, which would give a large gain, is not expected to be necessary.

The future predictions needed in the MPC will include the disturbance estimate $\hat{d}$ and assume it to be constant over the prediction horizon. For this implementation of MPC with disturbance estimation there are two free tuning parameters which affect the closed loop performance. ρ in the cost function (6) and the forgetting factor λ in the estimation of the disturbance in (17). In a situation where the system exhibits an abrupt change in the disturbance, as e.g. a step, a value of λ close to one will give a slow convergence of $\hat{d}_k$ to the new level. For a smaller value of λ the estimation is more sensitive to the new value of the output prediction error e_k, hence it will react faster but also be more influenced by noise. The choice of λ in this method plays a similar role as the choice of Q_ξ in the disturbance modelling method. A good value for the forgetting factor depends on the system. A reasonable choice of the forgetting factor is given in term of the equivalent horizon N_∞. The equivalent horizon is defined as the horizon over which the parameters can be expected to be constant.

$$\lambda = 1 - \frac{1}{N_\infty} \tag{18}$$

A. Time Varying Forgetting

In recursive estimation of time varying systems, the criterion which is minimized, when using exponential forgetting, is (16). An alternative method was proposed by [14] which keeps the value of the criterion constant between samples.

$$\begin{aligned} V_k &= \lambda_k V_{k-1} + \frac{1}{2} e_k^2 \\ s.t. \quad V_k &= V_{k-1} = V_0 \end{aligned} \tag{19}$$

In this extension to the algorithm (17) the forgetting factor becomes time varying according to the size of the prediction error. The time varying forgetting factor is approximated by:

$$\lambda_k \approx 1 - \frac{e_k^2}{N_\infty \sigma^2 (1 + P_{k-1})} \tag{20}$$

where σ^2 is the variance of process noise. The recursive estimator for the forgetting factor is based on the following distribution for the prediction error

$$e_k \in \mathcal{N}_{iid}(0, \sigma^2(1 + P_k)) \tag{21}$$

This method achieves an estimation which gives fast convergence when a change is observed (small values of λ_k) and it will reduce the effect of noise otherwise ($\lambda_k \approx 1 - {}^{1}\!/_{N_\infty}$). A change in the system is observed when the ratio between the prediction error and its variance becomes large. I.e. the method will adapt itself according to the observer's ability to predict the output one step ahead. This method is equivalent to both disturbance modelling and the disturbance estimation with exponential forgetting reported earlier, with respect to design variables. Only the equivalent horizon N_∞ is left as a free tuning parameter. In this method N_∞ should be chosen according to which value of λ_k is preferred under steady conditions. I.e. conditions where the estimate of the unknown disturbance is close to the true value and small prediction errors are achieved. On the other hand N_∞ should not be so large that the method is insensitive to an abrupt change in the prediction error.

When implementing a time varying forgetting factor by (20) it is necessary to include a lower limit λ_{min} since negative values will not be meaningful. A reasonable value for the lower limit would base the estimation on at least 2 to 4 data points. Here $\lambda_{min} = 1 - {}^{1}\!/_{3} \approx 0.67$ is used. Furthermore it is desired to lower the sensitivity to high frequency noise in the recursion for the forgetting factor (20) [15]. The squared prediction error normalized by $(1 + P_{k-1})$ will be replaced with a low pass filtered signal of the process noise variance ν_k with $\nu_0 = 0$ since the expected value of the prediction error is zero. Given these extensions the full algorithm for disturbance estimation with Fortescue's method is given by

$$\hat{d}_k = \hat{d}_{k-1} + \kappa_k e_k \tag{22a}$$

$$\kappa_k = \frac{P_{k-1}}{\lambda_k + P_{k-1}} \tag{22b}$$

$$\lambda_k = \max\left\{1 - \frac{\nu_k}{N_\infty \sigma^2}, \lambda_{min}\right\} \tag{22c}$$

$$P_k = (1-\kappa_k)P_{k-1}\lambda_k^{-1} \tag{22d}$$

where the filter equations for estimation of the prediction error variance are

$$T_k = \tau_T T_{k-1} + 1 \tag{22e}$$

$$\nu_k = \nu_{k-1} + \frac{1}{T_k}\left(\frac{e_k^2}{1+P_{k-1}} - \nu_{k-1}\right) \tag{22f}$$

It is seen that the recursion (22) is just an extended version of (17) with a time varying forgetting factor and a additional recursion with a low pass filter for ν. We choose the filter constant $\tau_T = 1/2$ and initialize $T_0 = 1$. I.e. the filter will be based on approximately two data points, which gives some smoothing without removing a change which may occur.

It is seen that (22) requires the knowledge of the variance of the process noise, σ^2. This requirement is in correspondence with the disturbance modelling approach where this variance is used in the design of the Kalman filter. We will in this paper assume this variance known. In case it is not, a process noise variance estimate ς can be estimated recursively and initialized by $\varsigma_0 = e_0^2$ [15].

$$S_k = \tau_S S_{k-1} + 1 \tag{23a}$$

$$\varsigma_k = \varsigma_{k-1} + \frac{1}{S_k}\left(\frac{e_{k-N_d}^2}{1+P_{k-1-N_d}} - \varsigma_{k-1}\right) \tag{23b}$$

The filter constant τ_S should be chosen close to one such that the estimate is based on a long horizon and less sensitive to jumps in the prediction error, e.g. 0.99. S_0 is also initialized equal to one. It is seen that the recursion of ς uses values of the prediction error and the variance P which are delayed by N_d [15]. The reason for this delay is that, when a step disturbance occurs and the prediction error gets large, it will increase both ν_{k+1} and ς_{k+1} simultaneously if $N_d = 0$. This simultaneous increase will render the recursion of the forgetting factor λ less sensitive to abrupt changes in the prediction error. If N_d is an integer larger than zero, ν_{k+1} and λ_{k+1} will react to an abrupt change while ς_{k+1} is not affected immediately. This situation better reflects the intention in (20) when σ^2 is known.

V. SIMULATION EXAMPLE

In the following a series of closed loop simulations with the two different MPC control implementations will be performed and their performance compared on a numerical example. The example will use the same ARX model for simulation of the true system and for predictions in the MPC. The model is

$$A(q^{-1})y(t) = B(q^{-1})u(t) + d(t) + \varepsilon(t) \tag{24a}$$

where $\varepsilon(t) \in \mathcal{N}_{iid}(0, 0.1^2)$ and

$$A(q^{-1}) = 1 - 2.4q^{-1} + 2.05q^{-2} - 0.63q^{-3} \tag{24b}$$

$$B(q^{-1}) = 0.5q^{-1} \tag{24c}$$

The closed loop performance will be quantified by the function (25) which reflects the MPC performance cost evaluated over the entire simulation horizon of 1000 samples.

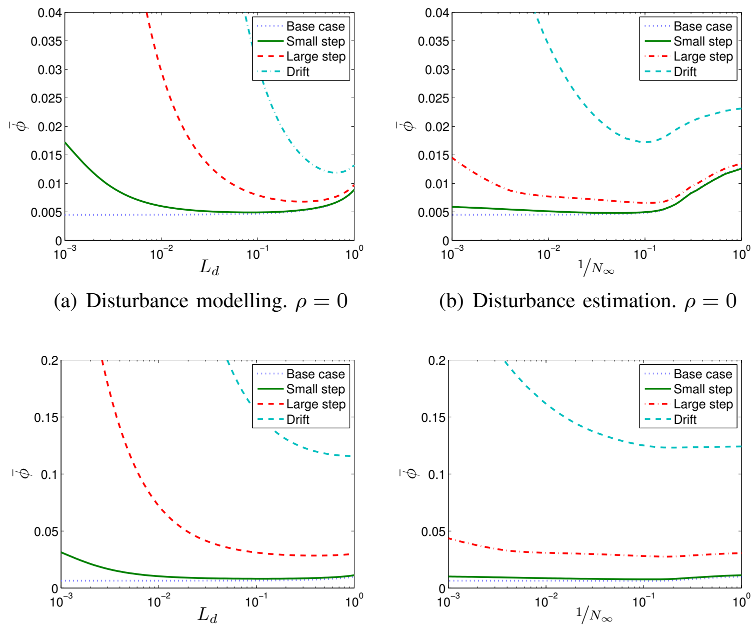
$$\bar{\phi} = \frac{1}{2(t_f - t_0)} \sum_{t=t_0}^{t_f} y_t^2 + \rho(\Delta u_t)^2 \tag{25}$$

In all simulations the same seed of random noise is applied to the process.

At first the methods will be tested on four different cases of the disturbance profile in addition to the process noise and for a wide span of their respective tuning parameters.

Base case	$d_k = 0 \,\forall k \in 1,..,N$
Small step	$d_k = 1/4 H(k-50) \,\forall\, k \in 1,..,N$
Large step	$d_k = 1H(k-50) \,\forall\, k \in 1,..,N$
Drift	$d_k = \frac{1}{1-q^{-1}} w_k$, where $w_k \in \mathcal{N}_{iid}(0, 0.1^2)$

where the function $H(k-k_0)$ is the Heaviside step function. Note that the drift disturbance is integrated noise of the same variance as the process noise. The disturbance state variance is $Q_\xi \in [10^{-9}, 10^3]$ for the disturbance model approach and the equivalent horizon $N_\infty \in [1; 1000]$ for the disturbance estimation approach with time varying forgetting. The closed loop performance is shown in Figure 1. In this figure, the closed loop performance is not plotted against the actual tuning parameters. The performance is plotted against the gain L_d in the disturbance state estimator in the disturbance modelling case and against $1/N_\infty$ for the disturbance estimation approach with time varying forgetting. This conversion is performed to achieve a comparable scaling and interpretation of the abscissa in the figures. Both axis are now scaled between 0 and 1. A value close to 1 indicates a high sensitivity of the prediction error to the update of the disturbance while 0 indicates that no update will occur. L_d in the disturbance modelling is fixed when Q_ξ is chosen for a specific system. The gain κ_k in the disturbance update for the disturbance estimation algorithm, change in time cf. (22) when changes in the prediction error occur. Under stationary conditions and no disturbance the quantity $1/N_\infty = 1 - \lambda_k = \kappa_k$, ignoring the lower bound λ_{min}. We therefore have $1/N_\infty$ as an indication of the lower bound on κ_k since a change in the disturbance will provide a change on λ_k and increase the gain until the disturbance is estimated after which it will decrease to the lower bound again. Figure 1 shows that the two methods do differ in closed loop performance and sensitivity to the nature of the disturbance. It is seen that the disturbance modelling approach needs a value of L_d larger than 0.1 if any disturbance enters the system. The best performance, for these cases, is achieved for different values of the gain dependent on the disturbance and the value of ρ in performance cost function. It is further more seen that if the disturbance drifts, the disturbance modelling achieves the best performance for a value of L_d which corresponds to



(a) Disturbance modelling. $\rho = 0$

(b) Disturbance estimation. $\rho = 0$

(c) Disturbance modelling. $\rho = 0.01$

(d) Disturbance estimation. $\rho = 0.01$

Fig. 1. Closed loop performance for the four simulation scenarios using both MPC implementations with $\rho \in \{0, 0.01\}$. The performance is plotted as function of the disturbance state estimator gain L_d for the disturbance model approach and the inverse of the equivalent horizon N_∞ for the disturbance estimation approach with time varying forgetting.

$Q_\xi = 0.1^2$ as expected. The disturbance estimation approach does in general achieve good performance for small values of $1/N_\infty$ when the disturbance does not have character of a drift. $N_\infty \approx 10$ seems to perform well in all cases. Good performance is however observed if N_∞ does not exceed 500 when an infrequent step of any magnitude enters the system. This performance is not sensitive to the value of ρ, except for when the methods are tuned very hard for fast disturbance rejection. The performance have also been tested outside the range of ρ presented here. In general both methods have comparable performance when choosing the best tuning for each case. The disturbance estimation is clearly superior in robustness when the disturbance occurs as infrequent steps while the disturbance modelling is superior when the disturbance is a drift with a known variance. It is however doubtful whether this variance can be obtained.

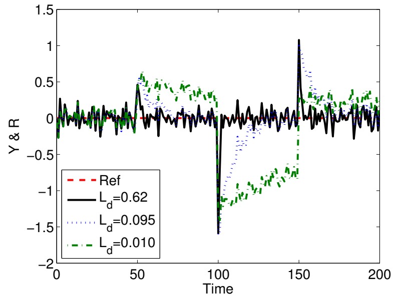
Secondly both methods are applied to a new simulation case over 200 samples. Here the disturbance is initially equal zero. At time $t = 50$ the disturbance steps to $d = 0.5$, at $t = 100$ a large step to $d = -1$ and finally at $t = 150$ the disturbance steps back to 0. For $Q_\xi = \{10^{-2}, 10^{-4}, 10^{-6}\}$ and for $N_\infty = \{3, 10, 100\}$ the controllers are tested and the loop response is show in Fig. 2 for ρ equal to 0 and 0.01. The values of the performance cost are given in Table I. It is seen from this study that for the disturbance estimation strategy, good disturbance rejection if achieved for the entire range of N_∞ at reasonably low values of the performance cost. For the disturbance modelling only the very hardly tuning controller with $Q_\xi = 10^{-2}$ is able to give comparable, and even faster, disturbance rejection.

TABLE I

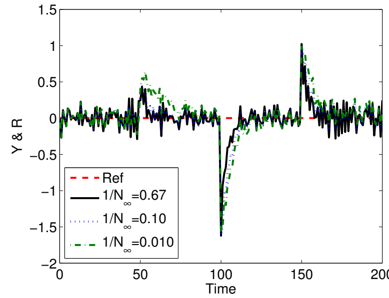
THE TUNING PARAMETERS AND THE CLOSED LOOP PERFORMANCE GIVEN BOTH MPC IMPLEMENTATIONS FOR THE SIMULATION CASE WITH A SERIES OF SMALL AND LARGE STEPS IN THE DISTURBANCE.

Q_{xi}	L_d	$\bar{\phi}(\rho = 0)$	$\bar{\phi}(\rho = 0.01)$
10^{-2}	0.62	0.0177	0.0341
10^{-4}	0.095	0.0544	0.106
10^{-6}	0.010	0.150	0.283
N_∞	$1/N_\infty$	$\bar{\phi}(\rho = 0)$	$\bar{\phi}(\rho = 0.01)$
3	0.67	0.0261	0.0485
10	0.10	0.0350	0.0701
100	0.010	0.0501	0.0991

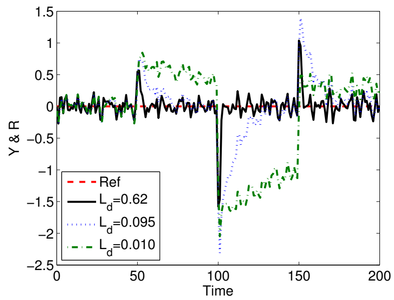
For the less aggressive tuning the speed of the disturbance rejection is significantly decreased which is seen in the values for the performance cost. Unfortunately this type of hard tuning is cannot be recommended in case the steps in the disturbance is infrequent since a high value of L_d results in a higher input and output variance for the loop. Therefore the strategy with adaptive disturbance estimation with an algorithm which is sensitive to the magnitude of the prediction error is better suited to insure offset free tracking for model predictive control implementations. This algorithm is not in the same way dependent on a tuning which balance disturbance rejection versus noise sensitivity since the method tune the gain κ itself based on the observations. Hence the tuning of the disturbance estimation method is related to the noise sensitivity during steady state simulation, and it is therefore much less sensitive to the nature and how



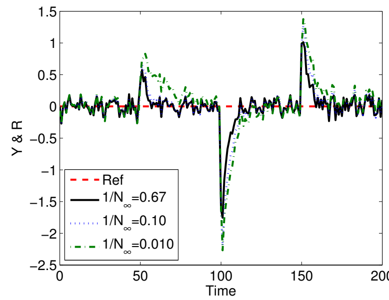
(a) Disturbance modelling. $\rho = 0$



(b) Disturbance estimation. $\rho = 0$



(c) Disturbance modelling. $\rho = 0.01$



(d) Disturbance estimation. $\rho = 0.01$

Fig. 2. Closed loop performance for a series of small and large steps in the disturbance using both MPC implementations with $\rho \in \{0, 0.01\}$.

frequent disturbances occur.

VI. CONCLUSION

A novel method for achieving offset free tracking in MPC applications is presented. It is based on adaptive estimation of the unknown disturbance with time varying forgetting. This method is compared to a classical approach where the system is augmented with a disturbance state which is estimated as part of the state estimation. Both methods have one free tuning parameter which expresses the sensitivity to large prediction errors in the equation for the disturbance update. It is shown that while the disturbance modelling approach is superior when the disturbance is drifting with a known variance, the adaptive disturbance estimation with a time varying forgetting is more robust to the nature of the disturbance and performs very well when a system is subjected to infrequent step disturbances of any size.

VII. ACKNOWLEDGMENTS

The first author gratefully acknowledge the Danish Council for Independent Research, Technology and Production Sciences (FTP) for funding through grand no. 274-08-0059

REFERENCES

[1] J. Richalet, A. Rault, J. L. Testud, and J. Papon, "Model predictive heuristic control: Application to industrial processes," *Automatica*, vol. 14, no. 5, pp. 413 – 428, 1978.

[2] C. Cutler and B. Ramaker, "Dynamic matrix control - a computer control algorithm," in *Proceedings of the Joint Automatic Control Conference*, 1980.

[3] D. Clarke, C. Mohtadi, and P. S. Tuffs, "Generalized predictive control - part i. the basic algorithm," *Automatica*, vol. 23, no. 2, pp. 137 – 148, 1987.

[4] K. R. Muske and J. B. Rawlings, "Model predictive control with linear models," *AIChE Journal*, vol. 39, no. 2, pp. 262 – 287, 1993.

[5] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789 – 814, 2000.

[6] U. Maeder, F. Borrelli, and M. Morari, "Linear offset-free model predictive control," *Automatica*, vol. 45, no. 10, pp. 2214 – 2222, 2009.

[7] J. K. Huusom, N. K. Poulsen, S. B. Jørgensen, and J. B. Jørgensen, "Tuning of methods for offset free mpc based on arx model representations," in *Proceedings of the American Control Conference - ACC*, 2010, pp. 2355 – 2360.

[8] M. Ohshima, H. Ohno, I. Hashimoto, M. Sasajima, M. Maejima, K. Tesuto, and T. Ogawa, "Model predictive control with adaptive disturbace prediction and its application to fatty acid distillation column control," *Jounal of Process Control*, vol. 5, no. 1, pp. 41 – 48, 1995.

[9] E. J. Davison and H. W. Smith, "Pole assignment in linear time-invariant multivariable systems with constant disturbances," *Automatica*, vol. 7, no. 4, pp. 489 – 498, 1971.

[10] K. R. Muske and T. A. Badgwell, "Disturbance modeling for offset-free linear model predictive control," *Journal of Process Control*, vol. 12, no. 5, pp. 617 – 632, 2002.

[11] G. Pannocchia and J. B. Rawlings, "Disturbance models for offset-free model-predictive control," *AIChE journal*, vol. 49, no. 2, pp. 426 – 437, 2003.

[12] M. R. Rajamani, J. B. Rawlings, and S. J. Qin, "Achiving state estimation equivalence for misassigned disturbances in offset-free model predictive control," *AIChE Journal*, vol. 55, no. 2, pp. 396 – 407, 2009.

[13] L. Ljung and T. Söderstöm, *Theory and Practice of Recursive Identification*, A. S. Willsky, Ed. MIT Press, 1983.

[14] T. R. Fortescue, L. S. Kershenbaum, and B. E. Ydstie, "Implementation of self-tuning regulators with variable forgetting factors," *Automatica*, vol. 17, no. 6, pp. 831 – 835, 1981.

[15] N. K. Poulsen, "Robust self tuning controllers," Ph.D. dissertation, Institute of Mathematical Statistics and Operational Research (IMSOR), Technical University of Denmark, 1985.