On the significance of the noise model for the performance of a linear MPC in closed-loop operation

Hagdrup, Morten; Boiroux, Dimitri; Mahmoudi, Zeinab; Madsen, Henrik; Poulsen, Niels Kjølstad; Poulsen, Bjarne; Jørgensen, John Bagterp

Published in:
IFAC-PapersOnLine

Link to article, DOI:
10.1016/j.ifacol.2016.07.241

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
On the significance of the noise model for the performance of a linear MPC in closed-loop operation

Morten Hagdrup *, Dimitri Boiroux ***, Zeinab Mahmoudi ***, Henrik Madsen *, Niels Kjølstad Poulsen *, Bjarne Poulsen *, John Bagterp Jørgensen *

* DTU Compute, Technical University of Denmark, Kgs. Lyngby, Denmark (e-mail: jbjo@dtu.dk)
** Danish Diabetes Academy, Odense, Denmark

Abstract: This paper discusses the significance of the noise model for the performance of a Model Predictive Controller when operating in closed-loop. The process model is parametrized as a continuous-time (CT) model and the relevant sampled-data filtering and control algorithms are developed. Using CT models typically means less parameters to identify. Systematic tuning of such controllers is discussed. Simulation studies are conducted for linear time-invariant systems showing that choosing a noise model of low order is beneficial for closed-loop performance.

Keywords: Closed-loop control, predictive control, model-based control, Kalman Filters, state estimation

1. INTRODUCTION

Model Predictive Control (MPC) is a control methodology that uses a model of the system to be controlled to predict its output over a future horizon. At each time instance a control sequence is calculated online as the solution to an open-loop control problem based on the model, the current state and specified reference trajectory. Only the first element of the control sequence is applied to the system and feedback is obtained by repeating this procedure when the next measurements are received. A notable advantage of MPC is the way constraints are handled directly when solving the optimization problem resulting in the control sequence. The performance of the controller thus hinges on the quality of the system model, but not only on that. Also noise and possible disturbances must be catered for, see Pannocchia and Rawlings (2003), Gopaluni et al. (2004), Gevers (2005), Shah and Engell (2010), Huusom et al. (2012) and references therein. One should therefore consider a system model comprising a deterministic as well as a stochastic or noise part. Selecting a noise model involves a trade-off between conflicting requirements namely those of low variance set-point tracking, disturbance rejection and fast response to unmeasured disturbances.

Boiroux et al. (2015) provided a comparative study of the effects of choosing different deterministic model parts in MPC-based Artificial Pancreas technology keeping the stochastic part fixed. The goal of the present paper is to study the role played by the stochastic part of the model. This term is intended to absorb not only the presence of unmeasured disturbances but also more generally unmodelled system dynamics. The ultimate test of the suitability of a given noise model is therefore whether the system performs adequately in closed-loop (CL). The message of this paper is that closed-loop performance may benefit from selecting a suitable low-order noise model.

CL performance is evaluated for model structures corresponding to different filter orders. For each model structure the noise term is identified using the Maximum Likelihood (ML) criterion from measurements collected before closing the loop, see Jørgensen and Jørgensen (2007b). An efficient MPC implementation is developed based on continuous-time transfer functions keeping the deterministic and stochastic model parts separate. The stochastic part will determine the Kalman Filter and Predictor while the deterministic model part, set-point and filtered state estimates will determine the optimal control problem to be solved.

The paper is structured as follows. Section 2 recalls the basic theory of realization and discretization of linear time-invariant (LTI) systems given in terms of transfer functions. The following section continues by focusing on the case of LTI systems with continuous-time white noise input. Section 4 develops the Kalman Filter and Predictor for the resulting discrete-time state space model and is followed by a section developing the Model Predictive Controller. We round off with a section discussing the outcome of a concrete closed-loop control simulation for noise models of different orders.

2. REALIZATION OF LINEAR SYSTEMS

We consider a linear system described in continuous time in terms of transfer functions $G(s)$ and $H(s)$ and with discrete measurements $y(t_k)$ at times $t = t_k$:

$$Z(s) = G(s) U(s) + H(s) W(s)$$

$$y(t_k) = z(t_k) + v_k, \quad k = 0, 1, 2, ...$$

(1)
Here, $U$ denotes the input to the deterministic part of the model and $W$ the white noise input to the stochastic part of the model. We shall assume that $G$ and $H$ are strictly proper. Finally, $\{v_k\} \sim N_{iid}(0, r^2)$ is a sequence of independent and identically distributed Gaussian random variables representing the measurement noise. With a view to using (1) as a model of a system to be subjected to Model Predictive Control we now turn to realizing it as a discrete-time state-space model. In doing so we consider the deterministic and stochastic parts separately and assume that the Zero-Order-Hold (ZOH) condition applies to the deterministic part. The stochastic part on the other hand involves sampling a certain Stochastic Differential Equation (SDE). We deal with this in section 3. We rely on the following lemma:

**Lemma 1.** Let a continuous-time system $S$ be described by $Z(s) = G(s)U(s)$ where $G(s)$ is assumed to be a proper transfer function. When $S$ is subjected to ZOH-input then there exist matrices $A$, $B$, $C$ and $D$ such that the state space model

$$
x_{k+1} = Ax_k + Bu_k
z_k = Cx_k + Du_k
$$

provides a realization of $S$ in discrete-time, when equidistantly sampled. If $G(s)$ is strictly proper we have that $D = 0$.

The deterministic part of the system description

$$Z_d(s) = G(s)U(s)$$

may be realized as a state space model by

$$Z_d(s) \sim \begin{cases} x_{k+1}^d = A_dx_k + B_du_k \\ z_k^d = C_dx_k \end{cases}$$

and the stochastic part

$$Z_s(s) = H(s)W(s)$$

as

$$Z_s(s) \sim \begin{cases} x_{k+1}^s = A_sx_k + B_sw_k \\ z_k^s = C_sx_k \end{cases}$$

Using that

$$Z(s) = Z_d(s) + Z_s(s)$$

we find that there exist matrices $A, B, C, G$ expressible in terms of the system matrices of (4) and (6) such that

$$x_{k+1} = Ax_k + Bu_k + Gw_k
z_k = Cx_k + v_k$$

provides a state space realization of (1) with an added equation accounting for measurement noise $v_k$. In fact (8) results by taking

$$x_k = \begin{bmatrix} x_k^d \\ x_k^s \end{bmatrix}
A = \begin{bmatrix} A_d & 0 \\ 0 & A_s \end{bmatrix}
B = \begin{bmatrix} B_d \\ 0 \end{bmatrix}
G = \begin{bmatrix} 0 \\ B_s \end{bmatrix}
C = \begin{bmatrix} C_d & C_s \end{bmatrix}$$

The process noise, $\{w_k\}$, and the measurement noise, $\{v_k\}$ are assumed to be sequences of Gaussian random variables with the joint distribution of $(w_k, v_k)$ given by

$$\{w_k\} \sim N_{iid}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\right)$$

3. REALIZATION OF STOCHASTIC TRANSFER FUNCTIONS

We now consider a continuous-time LTI system with transfer function $H(s)$ subjected to continuous-time white noise input given in the Laplace domain by $W(s)$:

$$Z(s) = H(s)W(s)$$

Assuming equidistant sampling with sampling time $T_s$ the transfer function and the associated measurement equation may be realized in the form of an SDE in the sense of Itô

$$dx(t) = A_s^*x(t)dt + B_s^*d\omega(t)
\begin{align*}
y(t_k) &= C_s^*x(t_k) + v(t_k)
\end{align*}$$

where $\omega$ denotes standard Brownian Motion and

$$x(t_0) \sim N(x_0, P_0)
\begin{align*}
d\omega(t) &= N_{iid}(0, I)dt
v(t_k) &= N_{iid}(0, R)
\end{align*}$$

with $R = r^2$.

We now discretize assuming equidistant sampling at integer multiples of $T_s$ and obtain a discrete-time state space model (6) by taking

$$A_s = e^{A_sT_s}
B_s = I
C_s = C_s^*$$

and

$$w_k \sim N_{iid}(0, Q)$$

with

$$Q = \int_0^{T_s} e^{A_s^*\sigma}B_s^*(B_s^*)^*e^{(A_s^*)^*}\sigma d\sigma$$

The reader is referred to Åström (1970) for a proof of these discretization results. Since $(A_s^*, B_s^*)$ is controllable it follows from Zhou et al. (1995) that $Q$ is positive definite. According to Van Loan (1978)

$$Q = \Phi_{12}'\Phi_{12}$$

where

$$\exp\left(\begin{bmatrix} -A_s^* & B_s^*(B_s^*)^* \\ 0 & (A_s^*)^* \end{bmatrix}T_s\right) = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ 0 & \Phi_{22} \end{bmatrix}$$

The formula (17) may be employed to calculate $A_s$ and $Q$ numerically by means of Padé approximation, but it is in fact possible to calculate exact analytical expressions for those matrices in the case of the simple transfer functions we consider here. The expressions, however, quickly become rather unwieldy with increasing order. We introduce the notation

$$\beta = \frac{T_s}{\tau}$$

and list values of $A_s$ and $Q$ in Table 1 for examples of the kind of transfer functions considered in this paper. We note that with $H(s) = \frac{k}{(\tau s + 1)^n}$ the continuous-time noise output power for white noise input derived from standard Brownian Motion becomes

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{k^2}{(1 + (\tau \omega)^2)^n} d\omega = \frac{k^2}{2\pi} I_n$$

where $I_n = \int_{-\infty}^{+\infty} \frac{1}{(1 + (\tau \omega)^2)^n} dt$ satisfies $I_1 = \pi$ and the recursion $I_{n+1} = \frac{2n+1}{2n} I_n$ holds for $n \geq 1$. For fixed filter order $n$ the continuous-time noise output power depends only on the ratio $\frac{k^2}{\tau}$ but the distribution of this power over the spectrum is depends on $\tau$, spreading out more as
Table 1. Observer canonical realization of stochastic transfer functions.

<table>
<thead>
<tr>
<th>$H(s)$</th>
<th>$A_s$</th>
<th>$Q/\tau^2$</th>
<th>$T_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{k}{\tau s + 1}$</td>
<td>$e^{-\beta}$</td>
<td>$\frac{1}{\tau^2} \left[ 1 - e^{-2\beta} \right]$</td>
<td>$T_s$</td>
</tr>
<tr>
<td>$\frac{k}{\tau s + 1}$</td>
<td>$\begin{bmatrix} e^{-\beta} \tau (1 - e^{-\beta}) &amp; 0 \ 0 &amp; 1 \end{bmatrix}$</td>
<td>$\tau \left( \frac{1 - e^{-\beta}}{\tau^2} \right)$</td>
<td>$\beta - 1 + e^{-\beta}$</td>
</tr>
</tbody>
</table>

$\tau$ decreases. Considering in particular the case $n = 1$ it is clear from the expression $Q = \frac{k^2}{\tau^2} (1 - \exp(-2T_s/\tau))$ that even when keeping $\frac{k^2}{\tau^2}$ fixed, the value of $Q$ depends on the relative size of $T_s$ w.r.t. $\tau$. In fact $\tau$ becomes a measure of the coherence time of the state $x$ in the sense that for $T_s \ll \tau$ we have $Q \approx 0$ and for $T_s \gg \tau$ we have $Q \approx \frac{k^2}{\tau^2}$.

4. FILTERING AND PREDICTION

One could implement MPC for the model (1) based on the prediction formulae

$$
\hat{x}_{k+j+1|k} = A_s \hat{x}_{k+j|k} + B u_{k+j}, \quad j = 0, ..., N - 1
$$

and

$$
\tilde{y}_{k+j|k} = C \hat{x}_{k+j|k}, \quad j = 1, ..., N
$$

for the realization (8) with noise specification (9). However, it will be advantageous to exploit the splitting of the system into a deterministic and a stochastic part as one only needs to run a Kalman Filter for the stochastic part. The idea is to consider the modified output $\tilde{y}_k = y_k - z^d_k$ where the deterministic part $z^d_k$ is obtained from (4) with initial condition $z^d_0 = 0$. Assuming that stationary conditions have been reached and that at time $t = t_k$ we have just received the measurement $y_k = y(t_k)$ we may perform the online filtering

$$
\tilde{y}_{k-1} = C_s \tilde{z}_{k-1}
$$

and

$$
\tilde{x}_{k-1} = \tilde{x}_{k-1} + K_f \tilde{x}_{k-1}
$$

using the coefficients

$$
R_e = C_s P C_s^T + R
$$

and

$$
K_f = P C_s R_e^{-1}
$$

calculated offline. Here $P$ denotes the stationary one-step-ahead state error covariance matrix obtained from the Discrete-time Algebraic Riccati Equation (DARE):

$$
P = A_s P A_s^T - A_s P C_s (C_s P A_s^T + R)^{-1} C_s P A_s^T + Q
$$

From equations (22) and (23) we observe that the $K_f$ obtained for given values of $(Q, R)$ also results by replacing the pair of matrices by $(\alpha Q, \alpha R)$ for any positive scalar $\alpha$. Hence we deduce that for fixed $T_s$ and $\tau$ the Kalman Gain $K_f$ depends on $k$ and $r$ solely through the ratio $\frac{k}{\tau^2}$. In Jørgensen and Jørgensen (2007a) the example

$$
H(s) = \frac{k}{\tau s + 1}
$$

is considered and it is observed from numerical simulations that it is impossible to estimate the quantities $k$ and $\tau$ by the non-linear LS method. This we may now explain as follows. Since $K_f$ depends on $k$ and $r$ only through the ratio $\frac{k}{\tau^2}$ and $A_s, C_s$ do not depend on these quantities it is clear that the non-linear LS cost function to be minimized for a given set of observations depends only on $k$ and $r$ through the ratio $\frac{k}{\tau^2}$. Hence one can not identify $k$ and $r$ by the non-linear LS method.

5. MODEL PREDICTIVE CONTROL

In this section we shall use the predictions presented in the previous sections to develop receding horizon optimal regulators. We define the output penalty function $\rho$ by

$$
\rho(z, \bar{z}, \xi, \eta) = \|z - \bar{z}\|^2 + \kappa \|\xi\|^2 + \gamma \|\eta\|^2
$$

and the objective function by

$$
\phi = \frac{1}{2} \sum_{j=1}^{N} \rho(\hat{z}_{k+j|k}, \hat{z}_{k+j|k}, \hat{\xi}_{k+j|k}, \hat{\eta}_{k+j|k}) + \lambda \|\Delta u_{k+j|k}\|^2
$$

In (26) the first term of $\rho$ penalizes deviations of the predicted outputs, $\{\hat{z}_{k+j|k}\}_{j=1}^{N}$, from the anticipated setpoints, $\{\hat{z}_{k+j|k}\}_{j=1}^{N}$. The second and third terms of $\rho$ represent the penalty associated with the soft constraints on the output variable:

$$
\hat{z}_{k+j|k} \geq \min_{k+j} \hat{z}_{k+j|k} - \hat{\xi}_{k+j|k}
$$

and

$$
\hat{\xi}_{k+j|k} \geq 0
$$

for $j = 1, ..., N$. The final term of $\phi$ is a regularization term that penalizes change, $\Delta u_k = u_k - u_{k-1}$, in the manipulated variable in order to ensure smoothness of the solution

Introducing the (hard) input constraints

$$
u_{\min} \leq u \leq u_{\max}
$$

we may now state the finite horizon optimal control problem with the objective function (26), predictions (20), input constraints (28) and soft output constraints (27) as

$$
\min \\phi
s.t. (20), (27), (28)
$$

We substitute into the expression to be minimized the equations

$$
\hat{z}_{k+j|k} = z^d_{k+j|k} + \hat{z}_{k+j|k}
$$

(30)
resulting from the splitting into deterministic and stochastic parts and obtain the modified objective \( \tilde{\phi} \)
\[
\tilde{\phi} = \frac{1}{2} \sum_{j=1}^{N} \rho (\bar{z}_{k+j}^d, \bar{z}_{k+j}^s, \hat{z}_{k+j}|k, \hat{\eta}_{k+j}|k) + \lambda \| u_{k+j}|k \|^2
\]
with the modified reference trajectory
\[
\bar{z}_{k+j}|k = \bar{z}_{k+j}|k - \bar{z}_{k+j}|k
\]
(31)
Adjusting the expressions for the soft constraints (27) to
\[
\bar{z}_{k+j}|k \geq z_{\text{min},k+j}|k - \hat{\xi}_{k+j}|k
\]
(33a)
\[
\bar{z}_{k+j}|k \leq z_{\text{max},k+j}|k + \hat{\eta}_{k+j}|k
\]
(33b)
(33c)
we see that solving the problem (29) is equivalent to solving the constrained optimization problem
\[
\min_{\{u_j: j = 1, \ldots, N\}} \tilde{\phi}
\]
\[\text{s.t.} \quad (4), (28), (33)\]
(35a)
(35b)
Thus effectively we have to solve an MPC problem for the deterministic part where at each step we modify the reference trajectory and the soft constraints based on the latest estimates of the state of the stochastic part of the model. Algorithm 1 sums up in compact form the calculations and instructions performed by the MPC each time a new measurement \( y_k \) has been received.

**Algorithm 1 MPC implementation, step \( k \)**

**Require:** \( y_k; \{\bar{z}_{k+j}|k\}_{j=1}^{N}, \bar{z}_{k+k-1}^s, x_k, u_{k-1} \)

**Filter:**
\[
\begin{align*}
\bar{z}_{k|k} &= C_d \bar{x}_{k|k}^d \\
y_k^s &= y_k - \bar{z}_{k|k} \\
\bar{z}_{k|k-1} &= C_a \bar{x}_{k|k-1} \\
\hat{e}_k &= y_k - \bar{z}_{k|k-1} \\
\hat{e}_{k+k-1} &= \hat{e}_{k+k-1} + K_f \hat{e}_k
\end{align*}
\]
**Stochastic Prediction:**
For \( j = 1, 2, \ldots, N \):
\[
\begin{align*}
\bar{z}_{k+j|k} &= A_s \bar{z}_{k+j-1|k} \\
\bar{z}_{k+j|k}^s &= C_s \bar{x}_{k+j|k}^s \\
z_{\text{min},k+j|k} &= z_{\text{min},k+j}|k - \bar{z}_{k+j|k} \\
z_{\text{max},k+j|k} &= z_{\text{max},k+j}|k + \hat{\eta}_{k+j}|k \\
z_{\text{bounds},k+j|k} &= (z_{\text{min},k+j}|k, z_{\text{max},k+j}|k) \\
\bar{z}_{k+j|k}^* &= \bar{z}_{k+j|k} - \bar{z}_{k+j|k} \\
\bar{z}_{k+j|k} &= \bar{z}_{k+j|k} - \bar{z}_{k+j|k}
\end{align*}
\]
**Regulator:**
\[
\begin{align*}
u_k &= \mu(x_k; \{\bar{z}_{k+j}|k\}_{j=1}^{N}, \{z_{\text{bounds},k+j}|k\}_{j=1}^{N}) \\
\text{One-step predictor:} & \quad x_{k+1} = A d \bar{x}_{k} + B_d u_k \\
\text{Return:} & \quad u_k, x_{k+1}^d, \hat{e}_{k+1}
\end{align*}
\]

6. CLOSED-LOOP SIMULATION STUDY

This section presents closed-loop simulation studies based on the MPC scheme developed above. We consider a plant given by (1) with \( G = G_p \) and \( H = H_p \) where
\[
G_p(s) = \frac{-1}{(1 + \tau_I s)^2} \quad H_p(s) = \frac{k_s}{(1 + \tau_I s)^5}
\]
(36)
for \( n = 1, 2, 3, 5 \) leaving \( \tau \) and \( k \) to be estimated in each case. We compare with the performance achieved for the nominal case where we assume full knowledge of the 5th order noise model. Since the parameters of the process noise model are to be estimated the ratio of process noise to measurement noise will play a role for the quality of the parameter estimates and for the ensuing controller. The noise model parameters are estimated using the Maximum Likelihood (ML) method on a batch of 1000 samples before the loop is closed. Different seeds are used for closed-loop simulations and for generating the data for parameter estimation.

An important step in the tuning of a Model Predictive Controller is that of settling on a suitable value in (26) for the parameter \( \lambda \) which implicitly determines the bandwidth of the controller. A good starting point for the tuning of \( \lambda \) is to assume that possible constraints, hard as well as soft, are inactive. We let \( \lambda \) sweep through a wide range of values and plot the resulting variances of control input \( u \) and process output \( y \) in a Pareto plot for the case of the set-point identically equal to 0. The number of samples based on which each variance is calculated is equal to 900. We step through the range of \( \lambda \)’s such that \( \log_{10} \lambda \) is changed by 0.25 at each step going from -9 and up to +6. For further details on tuning of Model Predictive Controllers the reader is referred to Hovorka et al. (2012), Huusom et al. (2010), Olesen et al. (2013) and Shah and Engell (2010). For an example of MPC tuning in the field of Artificial Pancreas technology the reader may consult Hovorka et al. (2010).

Figures 1 and 2 show the Pareto plots obtained for \( r = 0.001 \) and \( r = 0.01 \) respectively. The parts of the curves of interest to us are those in the vicinity of the ‘knee’ of each curve, since there we have a good balance between low output variance and control effort spent. For both sets of plots we observe that the closest set-point tracking is obtained for the nominal case, that is when one in addition to the deterministic part also has full knowledge of the process noise model. However, if only the model structure is known but not the correct parameters then for both plots the results for the 5th order models are the worst of those for which estimated parameters are used. In both cases it seems that the 2nd order model is a good choice. In particular in the case with higher levels of measurement noise does the 2nd order model seem to be the most robust...
References


Fig. 3. Process output $y$ and reference $r$ plotted along with control signal $u$ for estimated 1st and 2nd order noise models and for the nominal case with perfectly known 5th order noise model. Those three plot are shown for their Pareto optimal $\lambda$-values which are $10^{-3.5}$, $10^{-3.5}$ and $10^{-1.5}$ respectively. For comparison we include the nominal case with relaxed penalty $\lambda = 10^{-3.5}$. Input constrained to $-4 \leq u \leq 4$.


