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):
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 \( t = t_k \):

\[
Z(s) = G(s) U(s) + H(s) W(s)
\]

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

$$
\begin{align*}
    x_{k+1} &= Ax_k + Bu_k \\
    z_k &= Cx_k + Du_k 
\end{align*}
$$

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

$$
Z_d(s) \sim \begin{cases}
    x^d_{k+1} = A_dx^d_k + B_d u_k \\
    z^d_k = C_d x^d_k
\end{cases}
$$

and the stochastic part

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

as

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

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

$$
\begin{align*}
    x_{k+1} &= Ax_k + Bu_k + Gw_k \\
    z_k &= Cx_k \\
    y_k &= z_k + v_k
\end{align*}
$$

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

$$
\begin{align*}
    x_k &= \begin{bmatrix} x^d_k \\ x^s_k \end{bmatrix} \\
    A &= \begin{bmatrix} A_d & 0 \\ 0 & A_s \end{bmatrix} \\
    B &= \begin{bmatrix} B_d \\ 0 \end{bmatrix} \\
    G &= \begin{bmatrix} 0 \\ B_s \end{bmatrix} \\
    C &= \begin{bmatrix} C_d & C_s \end{bmatrix}
\end{align*}
$$

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

$$
\begin{align*}
    \{w_k\} &\sim N_{iid}(0, Q) \\
    \{v_k\} &\sim N_{iid}(0, R)
\end{align*}
$$

### 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

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

where $\omega$ denotes standard Brownian Motion and

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

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

$$
\begin{align*}
    A_s &= e^{A_s T_s} \\
    B_s &= I \\
    C_s &= C_s^c
\end{align*}
$$

and

$$
\begin{align*}
    w_k &\sim N_{iid}(0, Q) \\
    Q &= \Phi_{22}^s \Phi_{12}^s
\end{align*}
$$

where

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

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

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

where $I_n = \int_{-\infty}^{+\infty} \frac{k^2}{(1 + (\tau \omega)^2)^n} dt$ satisfies $I_1 = \pi$ and the recursion $I_{n+1} = \frac{2n}{2n+1} I_n$ 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...
\[ \begin{array}{|c|c|c|}
\hline
s & A_s & Q \\
\hline
\frac{k}{\tau s + 1} & e^{-\beta} & \frac{1}{\tau} [1 - e^{-2\beta}] \\
\hline
\frac{k}{s + (\tau s + 1)^2} & \begin{bmatrix} e^{-\beta} \tau (1 - e^{-\beta}) \\ 0 \end{bmatrix} & \frac{1}{2\tau} \begin{bmatrix} \beta - 1 + e^{-\beta} + e^{-2\beta} \\ \beta + 1 + e^{-\beta} \end{bmatrix} \\
\hline
\end{array} \]

\[ \text{Table 1. Observer canonical realization of stochastic transfer functions.} \]

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

\[ K_{fx} = PC^T \]

\[ u_{\text{min}} \leq u \leq u_{\text{max}} \]

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

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

\[ \text{4. FILTERING AND PREDICTION} \]

\[ \dot{x}_{k+j+1}|k = A_s \hat{x}_{k+j}|k + Bu_{k+j}, \quad j = 0, ..., N - 1 \]

\[ \hat{y}_{k+j}|k = C_s \hat{x}_{k+j}|k, \quad j = 1, ..., N \]

\[ \text{for the realization (8) with noise specification (9). However, it will be advantageous to exploit the splitting of the system into a deterministic and a stochastic part as one only needs to run a Kalman Filter for the stochastic part.} \]

\[ \text{The idea is to consider the modified output } \hat{y}_k = y_k - z^d_k \text{ where the deterministic part } z^d \text{ is obtained from (4) with initial condition } z^d_0 = 0. \text{ Assuming that stationary conditions have been reached and that at time } t = t_k \text{ we have just received the measurement } y_k = y(t_k) \text{ we may perform the online filtering} \]

\[ \hat{y}_{k+1} = C_s \hat{x}_{k+1}|k - 1 \]

\[ e_k = y_k - \hat{y}_{k+1} \]

\[ \text{using the coefficients} \]

\[ R_c = C_s P C_s^T + R \]

\[ K_{fx} = P C_s^T R_c^{-1} \]

\[ \text{calculated offline. Here } P \text{ denotes the stationary one-step-ahead error covariance matrix obtained from the Discrete-time Algebraic Riccati Equation (DARE):} \]

\[ P = A_s P A_s^T - A_s P C_s^T (C_s P C_s^T + R)^{-1} C_s P A_s^T + Q \]

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

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

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

\[ \text{we substitute into the expression to be minimized the equations} \]

\[ \hat{z}_{k+j}|k = z^d_{k+j}|k + \hat{z}^r_{k+j}|k \]
resulting from the splitting into deterministic and stochastic parts and obtain the modified objective $\bar{\phi}$

$$\bar{\phi} = \frac{1}{2} \sum_{j=1}^{N} \rho(\tilde{z}_{k+j|k}^d, \tilde{z}_{k+j|k}^s, \hat{\xi}_{k+j|k}, \hat{\eta}_{k+j|k}) + \lambda \|\Delta u_{k+j|k}\|^2$$ (31)

with the modified reference trajectory

$$\tilde{z}_{k+j|k}^e = \tilde{z}_{k+j|k} - \tilde{z}_{k+j|k}^s$$ (32)

Adjusting the expressions for the soft constraints (27) to

$$\tilde{z}_{k+j|k}^d \geq z_{\text{min},k+j|k}^s - \hat{\xi}_{k+j|k}$$ (33a)

$$\tilde{z}_{k+j|k}^d \leq z_{\text{max},k+j|k}^s + \hat{\eta}_{k+j|k}$$ (33b)

$$\hat{\xi}_{k+j|k} \geq 0$$ (33c)

$$\hat{\eta}_{k+j|k} \geq 0$$ (33d)

by introducing

$$z_{\text{min},k+j|k}^s = z_{\text{min},k+j|k} - \tilde{z}_{k+j|k}^s$$ (34a)

$$z_{\text{max},k+j|k}^s = z_{\text{max},k+j|k} - \tilde{z}_{k+j|k}^s$$ (34b)

we see that solving the problem (29) is equivalent to solving the constrained optimization problem

$$\min_{\bar{\phi}} \quad s.t. \quad (4), (28), (33)$$ (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:** $y_k; \{\tilde{z}_{k+j|k}^e\}_{j=1}^{N}, \tilde{x}_{k|k-1}^e, x_{d|k}, u_{k-1}$

**Filter:**

$$z_k^d = C_d x_k^d$$

$$\tilde{y}_k^s = y_k - \tilde{z}_k$$

$$\tilde{z}_{k|k-1}^e = C_k \tilde{x}_{k|k-1}^e$$

$$e_k = \tilde{y}_k - \tilde{z}_{k|k-1}$$

$$\hat{x}_{k|k} = \tilde{x}_{k|k-1}^e + K_{fx} e_k$$

**Stochastic Prediction:**

For $j = 1, 2, \ldots, N$:

$$\tilde{x}_{k+j|k}^s = A_{k+j|k} \tilde{x}_{k|k}^s$$

$$\tilde{z}_{k+j|k} = z_{\text{min},k+j|k} - \tilde{z}_{k+j|k}^s$$

$$\tilde{z}_{k+j|k}^s = C_s \tilde{x}_{k+j|k}^s$$

$$\tilde{z}_{\text{max},k+j|k}^s = z_{\text{max},k+j|k} - \tilde{z}_{k+j|k}^s$$

$$\tilde{z}_{\text{bounds},k+j|k} = (z_{\text{min},k+j|k}^s, z_{\text{max},k+j|k}^s)$$

**Regulator:**

$$u_k = \mu(x_k, \{\tilde{x}_{k+j|k}^s\}_{j=1}^{N}, \{\tilde{z}_{\text{bounds},k+j|k}\}_{j=1}^{N})$$

**One-step predictor:**

$$x_{k+1} = A_{d|k} x_k + B_d u_k$$

**Return:** $u_k, x_{d|k+1}^e, \tilde{x}_{d|k+1}$

---

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

$$G_p(s) = \frac{-1}{(1 + \tau_I s)^2}$$

$$H_p(s) = \frac{k_s}{(1 + \tau_s s)^5}$$ (36)

with $\tau_I = 5$, $\tau_S = 3$, $k_S = 1$ and $\{v_k\} \sim N_{\text{ind}}(0, r^2)$ for two different values of $r$. We investigate how the choice of noise-model impacts on the performance of the resulting controller in closed-loop. Sampling time $T_s = 1$ and we choose a prediction horizon $N = 120$ which is large relative to the dynamics of the deterministic part of the system so as to approximate an infinite horizon controller. Also note that we sample fast enough to capture virtually all of the deterministic dynamics. We assume that the deterministic dynamics are known but the stochastic part is unknown and hence must be estimated. We try out candidate noise models of the form

$$H(s) = \frac{k}{(1 + \tau s)^n}$$ (37)

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 Huusom 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 Hvorka et al. (2010).
This section presents closed-loop simulation studies based on solving the constrained optimization problem by introducing new measurement estimates of the state of the stochastic part of the reference trajectory and the soft constraints based on the given by (1) with $G$.

Stochastic Prediction:

Filter:

$$
\hat{z}_k = \hat{z}_{k-1} + \hat{z}_k
$$

$$
\hat{z}_k = \hat{z}_{k-1} + \hat{z}_k
$$

$$
\hat{z}_k = \hat{z}_{k-1} + \hat{z}_k
$$

Solving the constrained optimization problem will ensure that the controller based on the 2nd order model fares better than the nominal one.

Our study of closed-loop performance for MPC for linear systems has shown that performance benefits from choosing a low-order noise model. In the study we have considered disturbance rejection performance but not aimed for offset-free control as such. There are fields such as zone-MPC where this aspect is less important. A case in point is Artificial Pancreas technology for Type 1 Diabetes cf. Boiroux et al. (2010) and Gondhalekar et al. (2013).

Should offset-free control be a priority then integral action may be ensured by including a factor $s$ in the denominator of $H(s)$. Of course integral action comes at a price. The resulting variance in a set-point tracking scenario would increase. The appropriate tuning for such an MPC is a subject for future study.

7. CONCLUSION

This paper has introduced a framework for MPC for linear systems specified in terms of continuous-time models with deterministic and stochastic parts. It has been used to investigate the influence of the noise model on the resulting controller performance in closed-loop. The result is that CL-performance benefits from a low-order noise model even for higher-order process noise descriptions. A possible direction of further research would be to investigate whether the conclusions of this paper would hold also for control of nonlinear systems, e.g. the non-linear differential equations governing the insulin-glucose dynamics.

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 $-4 \leq u \leq 4$.


