Delft University of Technology

Fac. of Information Technology and Systems **Control Systems Engineering**

Technical report CSE02-016a

Complexity reduction in MPC for stochastic max-plus-linear discrete event systems by variability expansion: Extended report*

T.J.J. van den Boom, B. Heidergott, and B. De Schutter

December 2006

A short version of this report has been published in *Automatica*, vol. 43, no. 6, pp. 1058–1063, June 2007.

Control Systems Engineering
Faculty of Information Technology and Systems
Delft University of Technology
Delft, The Netherlands
Current URL: https://www.dcsc.tudelft.nl

* This report can also be downloaded via https://pub.bartdeschutter.org/abs/02_016a.html

Complexity reduction in MPC for stochastic max-plus-linear discrete event systems by variability expansion: Extended report*

T.J.J. van den Boom‡, B. Heidergott§, and B. De Schutter‡

‡ Delft Center for Systems and Control, Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands
email: {t.j.j.vandenboom,b.deschutter}@dcsc.tudelft.nl

§ Department of Econometrics, Vrije Universiteit
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
email: bheidergott@feweb.vu.nl

Abstract

Model predictive control (MPC) is a popular controller design technique in the process industry. Recently, MPC has been extended to a class of discrete event systems that can be described by a model that is "linear" in the max-plus algebra. In this context both the perturbations-free case and for the case with noise and/or modeling errors in a bounded or stochastic setting have been considered. In each of these cases an optimization problem has to be solved on-line at each event step in order to determine the MPC input. This paper considers a method to reduce the computational complexity of this optimization problem, based on variability expansion. In particular, it is shown that the computational load is reduced if one decreases the level of "randomness" in the system.

1 Introduction

Model predictive control (MPC) [5, 11] is a well-established technology for the control of multivariable systems in the presence of input, output and state constraints. Usually, MPC uses (non)linear discrete-time models. However, the attractive features mentioned above have led us to extend MPC to discrete event systems (DES). The class of DES essentially consists of man-made systems that contain a finite number of resources (such as machines, communications channels, or processors) that are shared by several users (such as product types, information packets, or jobs) all of which contribute to the achievement of some common goal (the assembly of products, the end-to-end transmission of a set of information packets, or a parallel computation) [1]. In this paper we focus on the class of DES with synchronization but no concurrency. Such DES can be described by a model that is "linear" in the max-plus algebra [1, 3, 8], and therefore they are called max-plus-linear (MPL) DES.

In [4, 18] MPC has been extended to MPL DES, and a comparison was made to other control method for MPL DES [2, 10, 13]. For conventional linear systems noise and disturbances are usually modeled by including an extra term in the system equations (i.e., the noise is considered to be

*This paper is an extended version of the paper "Complexity reduction in MPC for stochastic max-plus-linear discrete event systems by variability expansion," by T.J.J. van den Boom, B. Heidergott, and B. De Schutter, *Automatica*, vol. 43, no. 6, pp. 1058–1063, June 2007.

additive). However, in MPL DES the influence of noise and disturbances is usually not max-plus-additive, but max-plus-multiplicative (cf. [1] or Section 6). As regards modeling errors, uncertainty in the modeling or identification phase leads to errors in the (estimates of the) system matrices. Since both noise/disturbances and modeling errors perturb the system by introducing uncertainty in the system matrices, both features can be treated in one single framework, as was already shown in [18, 19]. The characterization of the perturbation will then determine whether it describes model mismatch or disturbance. In [19] an MPC controller has been developed for the uncertainty setting described above and there it was also shown that under quite general conditions the resulting MPC optimization problem is a convex optimization problem. However, for many practical situations, the computational complexity will increase significantly as the prediction horizon and the system order increase.

In this paper, we will present a novel approach to the approximate calculation of stochastic integrals, called variability expansion. Since variability expansion is an analytical method and does not resort to simulation, it is, in principle, possible to compute higher moments of performance characteristics of stochastic systems. We combine this general method with max-plus systems, which enables us to solve the MPC optimization problem for MPL DES very efficiently. The results on variability expansion in this paper are an extension of a conference paper [6] where an example from queuing theory has been studied. The present paper provides a thorough technical analysis of variability expansion and includes the proofs, which were lacking in [6].

The paper is organized as follows. In Section 2 we introduce max-plus algebra and stochastic MPL DES. In Section 3 we give a short overview of the MPC algorithm for MPL DES. Section 4 introduces the method of variability expansion and describes how the complexity of the MPC optimization problem for MPL DES can be reduced significantly by using this method. The technical analysis of this method is postponed to the Appendix of the paper. Section 5 gives a more detailed analysis of the computational complexity of the developed algorithm. Finally, Section 6 gives a worked example and a comparison of computational performance for the new method and previous methods.

2 Stochastic max-plus-linear systems

In this paper we consider MPL DES that include stochastic uncertainty (see also [19]). Define $\varepsilon = -\infty$ and $\mathbb{R}_\varepsilon = \mathbb{R} \cup \{\varepsilon\}$, and let the system matrices of such a system be given by $A(k) \in \mathbb{R}_\varepsilon^{n_x \times n_x}$, $B(k) \in \mathbb{R}_\varepsilon^{n_x \times n_u}$, $C(k) \in \mathbb{R}_\varepsilon^{n_y \times n_x}$; then the system is described by a state space model of the form

$$x_i(k) = \max\Big(\max_{j=1,\ldots,n_x} (A_{ij}(k) + x_j(k-1)) \,,\, \max_{j=1,\ldots,n_u} (B_{ij}(k) + u_j(k)) \Big) \qquad i = 1,\ldots,n_x, \quad (1)$$

$$y_\ell(k) = \max_{j=1,\ldots,n_x} (C_{\ell j}(k) + x_j(k)) \qquad \ell = 1,\ldots,n_y\,. \quad (2)$$

The index k in (1)–(2) is called the event counter. The state $x(k)$ typically contains the time instants at which the internal events occur for the kth time, the input $u(k)$ contains the time instants at which the input events occur for the kth time, and the output $y(k)$ contains the time instants at which the output events occur for the kth time[1].

Remark 1 *Recurrence relations (1) and (2) can be written in a concise way using max-plus-algebra [1, 3, 8]. To see this, let* $x \oplus y = \max(x,y)$ *and* $x \otimes y = x + y$ *for* $x, y \in \mathbb{R}_\varepsilon$. *For matrices* $A \in \mathbb{R}_\varepsilon^{n \times m}$

[1] More specifically, for a manufacturing system, $x(k)$ contains the time instants at which the processing units start working for the kth time, $u(k)$ the time instants at which the kth batch of raw material is fed to the system, and $y(k)$ the time instants at which the kth batch of finished product leaves the system.

and $B \in \mathbb{R}_\varepsilon^{m\times l}$, their $\otimes$-product is defined by

$$[A \otimes B]_{ij} \;=\; \bigoplus_{k=1}^{m} A_{ik} \otimes B_{kj} = \max_{k=1,\ldots,m} (A_{ik} + B_{kj}).$$

In the same vein, $\oplus$-addition of matrices $A \in \mathbb{R}_\varepsilon^{n\times m}$ and $B \in \mathbb{R}_\varepsilon^{n\times m}$ is defined by

$$[A \oplus B]_{ij} \;=\; A_{ij} \oplus B_{ij} = \max(B_{ij}, A_{ij}).$$

With these definitions, the system equations (1) and (2) become

$$\begin{aligned} x(k) &= A(k) \otimes x(k-1) \oplus B(k) \otimes u(k) \\ y(k) &= C(k) \otimes x(k)\ . \end{aligned}$$

The system equations become thus linear in the max-plus algebra, and therefore the system is called a max-plus linear system.

The entries of system matrices $A(k)$, $B(k)$ and $C(k)$ are uncertain due to modeling errors or disturbances. Usually fast changes in the system matrices will be considered as noise and disturbances, whereas slow changes or permanent errors are considered as model mismatch. In this paper both features will be treated within one single framework.

The uncertainty caused by disturbances and errors in the estimation of physical variables, can be gathered in the uncertainty vector $e(k)$. In this paper we assume that the uncertainty has stochastic properties. Hence, $e(k)$ is a stochastic variable. We assume that $e(k)$ captures the complete event-varying aspect of the system.

Now we will describe how the entries of $e(k)$ enter the system. Let $\mathscr{S}_{\text{mpns}}$ be the set of max-plus-nonnegative-scaling functions, i.e., functions f of the form

$$f(z) = \max_{i=1,\ldots,m} (\mu_i + \nu_{i,1} z_1 + \ldots + \nu_{i,n} z_n)$$

with variable $z \in \mathbb{R}_\varepsilon^n$ and constants $\nu_{i,j} \in \mathbb{R}^+$ and $\mu_i \in \mathbb{R}$, where $\mathbb{R}^+$ is the set of nonnegative real numbers. If we want to stress that f is a function of z we will denote this by $f \in \mathscr{S}_{\text{mpns}}(z)$.

Note that the system matrices of an MPL model usually consist of sums or maximizations of internal process times, transportation times, etc. (see, e.g., [1] or Section 6). Since the entries of $e(k)$ directly correspond to the uncertainties in the duration times, and using the fact that the set $\mathscr{S}_{\text{mpns}}$ is closed under the operations max, $+$, and scalar multiplication by a nonnegative scalar [18], we know that the entries of the uncertain system matrices belong to $\mathscr{S}_{\text{mpns}}$:

$$A(k) \in \mathscr{S}_{\text{mpns}}^{n_x\times n_x}(e(k)),\quad B(k) \in \mathscr{S}_{\text{mpns}}^{n_x\times n_u}(e(k)),\quad C(k) \in \mathscr{S}_{\text{mpns}}^{n_y\times n_x}(e(k))\ . \tag{3}$$

System (1)–(2) with system matrices (3) will be called a stochastic MPL DES. Some results for the analysis of stochastic MPL DES can be found in [15, 16].

3 Model predictive control for stochastic MPL systems

In [4, 18, 19] the MPC framework has been extended to MPL models (1)–(2) as follows. Just as in conventional MPC [5, 11] we define at each event step k a cost criterion $J(k)$ that reflects the output and input cost functions ($J_{\text{out}}(k)$ and $J_{\text{in}}(k)$) in the event period $[k, k+N_{\text{p}}-1]$:

$$J(k) = J_{\text{out}}(k) + \lambda J_{\text{in}}(k) \tag{4}$$

where N_p is the prediction horizon, λ is a weighting parameter, and[2], e.g.,

$$J_{\mathrm{out}}(k) = \sum_{j=0}^{N_\mathrm{p}-1} \sum_{i=1}^{n_y} I\!E[\eta_i(k+j)] \tag{5}$$

$$J_{\mathrm{in}}(k) = -\sum_{j=0}^{N_\mathrm{p}-1} \sum_{\ell=1}^{n_u} u_\ell(k+j) \tag{6}$$

where $I\!E[\eta_i(k)]$ denotes the expected value of the ith "tardiness" $\eta_i(k)$. This tardiness is given by

$$\eta_i(k) = \max(\, y_i(k) - r_i(k)\ ,\ 0\ )\ , \tag{7}$$

where $r(k)$ is the due date for output signal $y(k)$. Note that this choice of $J_{\mathrm{out}}(k)$ favors on-time delivery and penalizes late delivery.

Define the vectors

$$\tilde{u}(k) = \begin{bmatrix} u(k) \\ \vdots \\ u(k+N_\mathrm{p}-1) \end{bmatrix}, \tilde{r}(k) = \begin{bmatrix} r(k) \\ \vdots \\ r(k+N_\mathrm{p}-1) \end{bmatrix}, \tilde{y}(k) = \begin{bmatrix} y(k) \\ \vdots \\ y(k+N_\mathrm{p}-1) \end{bmatrix}, \tilde{e}(k) = \begin{bmatrix} e(k) \\ \vdots \\ e(k+N_\mathrm{p}-1) \end{bmatrix}.$$

The aim is now to compute an optimal input sequence $u(k), \ldots, u(k+N_\mathrm{p}-1)$ that minimizes $J(k)$ subject to some linear constraints on the inputs and outputs (e.g., minimal and maximal input or output rates[3], hard due dates) of the form [4]

$$A_{\mathrm{constr}}(k)\tilde{u}(k) + B_{\mathrm{constr}}(k) I\!E[\tilde{y}(k)] \leq c_{\mathrm{constr}}(k)\ . \tag{8}$$

where the matrices A_{constr} and B_{constr} and vector c_{constr} model the constraints of the overall control problem, have the appropriate dimensions and do not depend on $\tilde{u}(k)$. As the $u(k)$'s correspond to consecutive event occurrence times, we have to add the condition

$$\Delta u(k+j) = u(k+j) - u(k+j-1) \geq 0 \qquad \text{for } j = 0, \ldots, N_\mathrm{p}-1. \tag{9}$$

Furthermore, in order to reduce the number of decision variables and the corresponding computational complexity we introduce a control horizon N_c ($\leq N_\mathrm{p}$) and we impose the additional condition that the input rate should be constant from event step $k+N_\mathrm{c}-1$ on:

$$\Delta u(k+j) = \Delta u(k+N_\mathrm{c}-1) \qquad \text{for } j = N_\mathrm{c}, \ldots, N_\mathrm{p}-1. \tag{10}$$

MPC uses a receding horizon principle. This means that after computation of the optimal control sequence $u(k), \ldots, u(k+N_\mathrm{c}-1)$, only the first control sample $u(k)$ will be implemented, subsequently the horizon is shifted one event step, the state and/or model is updated with new information of the measurements, and the optimization is restarted.

The MPL-MPC problem for event step k can be defined as:

$$\min_{\tilde{u}(k)} J_{\mathrm{out}}(k) + \lambda J_{\mathrm{in}}(k) \quad \text{subject to (1), (2), (8), (9) and (10).}$$

[2]Other choices for $J_{\mathrm{out}}(k)$ and $J_{\mathrm{in}}(k)$ are given in [4].

[3]For a manufacturing system the input (output) rate corresponds to the rate at which raw material/external resources (finished products) are fed to (leave) the system.

In order to compute the optimal MPC input signal, we need the expected value of the signals $\eta_i(k+j)$ and $y_i(k+j)$. We will now consider the computation of $\mathbb{E}[\eta_i(k+j)]$ and $\mathbb{E}[y_i(k+j)]$. In [19] it is shown that $\eta_i(k+j)$ and $y_i(k+j)$ are max-plus-nonnegative-scaling functions of the variable $w(k) = \begin{bmatrix} -\tilde{r}^T(k) & x^T(k-1) & \tilde{u}^T(k) \end{bmatrix}^T$.

The following proposition is proved in [19]:

Proposition 2 *Consider a signal $v(k)$ that is a max-plus-nonnegative-scaling function of $w(k)$ and $\tilde{e}(k)$:*

$$v(k) = \max_{j=1,\ldots,n_v} \left(\alpha_j + \beta_j^T w(k) + \gamma_j^T \tilde{e}(k) \right) , \tag{11}$$

where $\alpha_j \in \mathbb{R}_\varepsilon$, $\beta_j \in (\mathbb{R}^+)^{n_w}$, $\gamma_j \in (\mathbb{R}^+)^{n_{\tilde{e}}}$, and $\tilde{e}(k) \in \mathbb{R}^{n_{\tilde{e}}}$ is a stochastic variable with probability density function p. If we define the sets $\Phi_j(w(k))$, $j = 1,\ldots,n_v$ such that[4]

$$\forall \tilde{e}(k) \in \Phi_j(w(k)) : v(k) = \alpha_j + \beta_j^T w(k) + \gamma_j^T \tilde{e}(k)$$

and $\bigcup_{j=1}^{n_v} \Phi_j(w(k)) = \mathbb{R}^{n_{\tilde{e}}}$, then the expected value of $v(k)$ is given by

$$\mathbb{E}[v(k)] = \sum_{j=1}^{n_v} \int \ldots \int_{\tilde{e} \in \Phi_j(w)} \left(\alpha_j + \beta_j^T w(k) + \gamma_j^T \tilde{e} \right) p(\tilde{e})\, d\tilde{e}$$

where $d\tilde{e} = d\tilde{e}_1\, d\tilde{e}_2 \ldots d\tilde{e}_{n_{\tilde{e}}}$. Furthermore, the function $\mathbb{E}[v(k)]$ is convex in $w(k)$ and a subgradient $g_v(w(k))$ of $\mathbb{E}[v(k)]$ is given by

$$g_v(w(k)) = \sum_{\ell=1}^{n_v} \beta_\ell^T \int_{\tilde{e} \in \Phi_\ell(w(k))} \cdots \int p(\tilde{e})\, d\tilde{e} \, .$$

Now consider the MPL-MPC problem for event step k. First note that $\eta_i(k+j)$ and $y_i(k+j)$ depend on $\tilde{e}(k)$ and can both be written as a function $v(\tilde{e}(k))$ of the form (11), and that, because of Proposition 2, $\mathbb{E}[\eta_i(k+j)]$ and $\mathbb{E}[y(k+j)]$ are convex in $w(k)$. This means that $J_{\text{out}}(k)$ and $J(k)$ are convex in $\tilde{u}(k)$. Hence, it is easy to verify that the following lemma holds.

Lemma 3 *If the linear constraints are monotonically nondecreasing as a function of $\mathbb{E}[\tilde{y}(k)]$ (in other words, if $[B_c]_{ij} \geq 0$ for all i, j), constraint (8) becomes convex in $\tilde{u}(k)$.*

So, if the linear constraints are monotonically nondecreasing, the MPL-MPC problem turns out to be a convex problem in $\tilde{u}(k)$, and both a subgradient of the constraints and a subgradient of the cost criterion can easily be derived using Proposition 2. Note that convex optimization problems can be solved using reliable and efficient optimization algorithms, based on, e.g., interior point methods [14, 20].

4 Variability expansion

The algorithm described in the previous section has a complexity that is growing fast with an increasing number of stochastic variables $n_{\tilde{e}}$ due to the numerical integration that is required when computing the expected values of $\eta_i(k+j)$ and $y_i(k+j)$.

In this section we will approximate the expected value of $v(\tilde{e}(k))$ using the method of variability expansion. To this end, we assume that the entries of $\tilde{e}(k)$ are independent and identically distributed

[4] I.e., for $\tilde{e}(k) \in \Phi_{j'}(w(k))$ the maximum in (11) is reached for the index $j = j'$.

(i.i.d) and we introduce an artificial parameter θ. We replace with probability $1-\theta$ the ith entry of random vector $\tilde{e}(k)$ by its mean. The result is denoted by $\tilde{e}_\theta(k)$. The parameter θ allows controlling the level of randomness in the system, and letting θ go from 0 to 1 increases the level of stochasticity in the system.

The main idea of variability expansion is the following. Considering $I\!E[v(\tilde{e}_\theta(k))]$ as a function in θ, it can be developed into a Taylor series in θ that converges to the true function on $\mathbb{R}$; for a proof we refer to the Appendix. Note that only $\theta \in [0,1]$ has an interpretation in terms of our model. In particular, if we denote the value of $d^m/d\theta^m I\!E[v(\tilde{e}_\theta(k))]$ for $\theta = 0$ by $d^m/d\theta^m I\!E[v(\tilde{e}_0(k))]$, then $I\!E[v(\tilde{e}_1(k))]$, the "true" expected value of $v(\tilde{e}(k))$, is given by

$$I\!E[v(\tilde{e}(k))] = I\!E[v(\tilde{e}_1(k))] = \sum_{m=0}^{M} \frac{1}{m!} \frac{d^m}{d\theta^m} I\!E[v(\tilde{e}_0(k))] + R_M(k)\,,$$

where, for $M < n_{\tilde{e}}$,

$$R_M \leq \frac{1}{(M+1)!} \sup_{\theta\in[0,1]} \left| \frac{d^{M+1}}{d\theta^{M+1}} I\!E[v(\tilde{e}_0(k))] \right|$$

and $R_M = 0$ otherwise.

A closed-form expression for the mth order derivative $d^m/d\theta^m I\!E[v(\tilde{e}_0(k))]$ can be obtained as follows. Set for $0 \leq m \leq n_{\tilde{e}}$ and $i_1 < i_2 < \ldots < i_m$:

$$V(i_1, i_2, \ldots, i_m) = I\!E[v(\tilde{e}_\theta(k, i_1, i_2, \ldots, i_m))]$$

where $[\tilde{e}_\theta(k, i_1, i_2, \ldots, i_m)]_j$ equals the mean value of the jth element of $\tilde{e}_0$ for $j \notin \{i_1, i_2, \ldots, i_m\}$ and $[\tilde{e}(k)]_j$ for $j \in \{i_1, i_2, \ldots, i_m\}$, and where $V(0) = v(\tilde{e}_0(k))$. This means that $V(i_1, i_2, \ldots, i_m)$ is the estimation of v in the case where only the elements $[\tilde{e}(k)]_j$ for $j \in \{i_1, i_2, \ldots, i_m\}$ are stochastic, and the elements $[\tilde{e}(k)]_j$ for $j \notin \{i_1, i_2, \ldots, i_m\}$ are fixed to their mean.

In example 1, presented in the subsequent section, we will work with a Taylor series of degree 4 and we need the derivatives $d^m/d\theta^m I\!E[v(\tilde{e}_0(k))]$ for $m = 1, 2, 3, 4$. They are given by

$$\begin{aligned}
\frac{d}{d\theta} I\!E[v(\tilde{e}_0(k))] &= \sum_{i=1}^{n_{\tilde{e}}} (V(i) - V(0)) \\
\frac{d^2}{d\theta^2} I\!E[v(\tilde{e}_0(k))] &= 2 \sum_{i_1=1}^{n_{\tilde{e}}-1} \sum_{i_2=i_1+1}^{n_{\tilde{e}}} \Big(V(i_1, i_2) + V(0) - V(i_1) - V(i_2) \Big) \\
\frac{d^3}{d\theta^3} I\!E[v(\tilde{e}_0(k))] &= 6 \sum_{i_1=1}^{n_{\tilde{e}}-2} \sum_{i_2=i_1-1}^{n_{\tilde{e}}-1} \sum_{i_3=i_2+1}^{n_{\tilde{e}}} \Big(V(i_1, i_2, i_3) + V(i_1) + V(i_2) + V(i_3) \\
&\qquad - V(i_1, i_2) - V(i_1, i_3) - V(i_2, i_3) - V(0) \Big) \\
\frac{d^4}{d\theta^4} I\!E[v(\tilde{e}_0(k))] &= 24 \sum_{i_1=1}^{n_{\tilde{e}}-3} \sum_{i_2=i_1+1}^{n_{\tilde{e}}-2} \sum_{i_3=i_2+1}^{n_{\tilde{e}}-1} \sum_{i_4=i_3+1}^{n_{\tilde{e}}} \Big(V(i_1, i_2, i_3, i_4) + V(i_1, i_2) + V(i_1, i_3) + V(i_1, i_4) \\
&\qquad + V(i_2, i_3) + V(i_2, i_4) + V(i_3, i_4) + V(0) - V(i_1, i_2, i_3) - V(i_1, i_2, i_4) \\
&\qquad - V(i_1, i_3, i_4) - V(i_2, i_3, i_4) - V(i_1) - V(i_2) - V(i_3) - V(i_4) \Big).
\end{aligned}$$

There is the following simple building rule for the derivatives: The factor in front of the summation is $m!$ when the mth order derivative is evaluated. The outer summation ranges over all possible combinations of marking m out of $n_{\tilde{e}}$ random variables. The inner sum ranges over all possible ordered

combinations of letting the m marked variables be either stochastic or not. The sign of an element in the inner sum is given by -1 to the power of the number of deterministic substitutions among the m marked variables.

In example 1, to be presented in Section 6 below, we approximate $\mathbb{E}[v(\tilde{e}(k))]$ by a fourth-order Taylor expansion by ignoring the error term R_4:

$$\begin{aligned}\mathbb{E}[v(\tilde{e}(k))] \;\approx\; & v(\tilde{e}_0(k)) + \frac{d}{d\theta}\mathbb{E}[v(\tilde{e}_0(k))] + \frac{1}{2}\frac{d^2}{d\theta^2}\mathbb{E}[v(\tilde{e}_0(k))] + \frac{1}{6}\frac{d^3}{d\theta^3}\mathbb{E}[v(\tilde{e}_0(k))] \\ & + \frac{1}{24}\frac{d^4}{d\theta^4}\mathbb{E}[v(\tilde{e}_0(k))]\,.\end{aligned}$$

The above expression contains redundant terms and a simplified version can be obtained as follows. For $m \leq n_{\tilde{e}}$, set

$$\mathbf{V}(m) = \sum_{i_1=1}^{n_{\tilde{e}}-m} \sum_{i_2=i_1+1}^{n_{\tilde{e}}-m+1} \cdots \sum_{i_m=i_{m-1}+1}^{n_{\tilde{e}}} V(i_1, i_2, \ldots, i_m)\,.$$

The term $\mathbf{V}(m)$ yields the total effect of making m out of $n_{\tilde{e}}$ variables stochastic. For the nth derivative we mark in total n variables out of which m are stochastic. Hence, there are

$$\left(\begin{array}{c} n_{\tilde{e}} - m \\ n - m \end{array} \right)$$

possibilities of reaching at $n_{\tilde{e}} - m$ deterministic substitutions provided that there m stochastic ones, and, in accordance with our building rule for higher-order derivatives, we arrive at the following result, which will be proved in the Appendix.

Lemma 4 *Provided that $\tilde{e}(k)$ has a bounded support, the nth order derivative of $\mathbb{E}[v(\tilde{e}_\theta(k))]$ with respect to θ is for any $\theta \in \mathbb{R}$ given by*

$$\frac{d^n}{d\theta^n}\mathbb{E}[v(\tilde{e}(k))] = n! \sum_{l=0}^{n} \left(\begin{array}{c} n_{\tilde{e}} - l \\ n - l \end{array} \right) (-1)^{n-l} \mathbf{V}(l)\,,$$

for $n \leq n_{\tilde{e}}$, and zero otherwise.

By Lemma 4, $\mathbb{E}[v(\tilde{e}_\theta(k))]$ is infinitely many times differentiable with respect to θ. Moreover, the derivatives vanish for sufficiently high order, which implies that $\mathbb{E}[v(\tilde{e}_\theta(k))]$ as a function of θ can be represented on $\mathbb{R}$ by its Taylor series developed at $\theta = 0$. This train of thoughts leads to the following approximation for $\mathbb{E}[v(\tilde{e}(k))]$.

$$\begin{aligned}\mathbb{E}[v(\tilde{e}(k))] &= \sum_{n=0}^{M} \sum_{l=0}^{n} \left(\begin{array}{c} n_{\tilde{e}} - l \\ n - l \end{array} \right) (-1)^{n-l} \mathbf{V}(l) + R_{M+1} \\ &= \sum_{l=0}^{M} \left(\sum_{n=l}^{M} \left(\begin{array}{c} n_{\tilde{e}} - l \\ n - l \end{array} \right) (-1)^{n-l} \right) \mathbf{V}(l) + R_{M+1}\,.\end{aligned}$$

Note that

$$\frac{1}{(M-l)!} \prod_{j=l+1}^{M} (j - n_{\tilde{e}}) = \sum_{m=l}^{M} \left(\begin{array}{c} n_{\tilde{e}} - l \\ m - l \end{array} \right) (-1)^{m-l},$$

C_l^M	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$
$M=0$	1	0	0	0	0
$M=1$	$1-n_{\tilde{e}}$	1	0	0	0
$M=2$	$(1-n_{\tilde{e}})(2-n_{\tilde{e}})/2$	$(2-n_{\tilde{e}})$	1	0	0
$M=3$	$(1-n_{\tilde{e}})(2-n_{\tilde{e}})(3-n_{\tilde{e}})/6$	$(2-n_{\tilde{e}})(3-n_{\tilde{e}})/2$	$(3-n_{\tilde{e}})$	1	0
$M=4$	$(1-n_{\tilde{e}})(2-n_{\tilde{e}})(3-n_{\tilde{e}})(4-n_{\tilde{e}})/24$	$(2-n_{\tilde{e}})(3-n_{\tilde{e}})(4-n_{\tilde{e}})/6$	$(3-n_{\tilde{e}})(4-n_{\tilde{e}})/2$	$(4-n_{\tilde{e}})$	1

Table 1: Coefficients c_l^M of the Taylor series for $M,l=0,\ldots,4$.

see, for example, [9], p. 57 formula (18). Let

$$c_l^M = \frac{1}{(M-l)!}\prod_{j=l+1}^{M}(j-n_{\tilde{e}}),$$

where we set $c_l^M = 1$ for $l = M$. In Table 1 the coefficients c_l^M are computed for $M = 0,\ldots,4$. We summarize our analysis in the following theorem.

Theorem 5 *Let $\tilde{e}(k)$ has bounded support. For $M \in \mathbb{N}$, the Taylor polynomial for $I\!E[v(\tilde{e}(k))]$ of degree M is given by*

$$I\!E[v(\tilde{e}(k))] = \sum_{l=0}^{M} c_l^M \mathbf{V}(l) + R_M(k),$$

with $R_M(k) = 0$ for $M \geq \tilde{n}_{\tilde{e}}$.

By Theorem 5 it holds that

$$I\!E[v(\tilde{e}(k))] \approx \sum_{l=0}^{M} c_l^M \mathbf{V}(l). \tag{12}$$

The subgradient $\nabla_{\tilde{u}} I\!E[v(\tilde{e}(k))]$ can be computed using the same weighted summation. For example, the approximate subgradient for $M \leq n_{\tilde{e}}$ becomes:

$$\nabla_{\tilde{u}} I\!E[v(\tilde{e}(k))] \approx \sum_{l=0}^{h} c_l^M \nabla_{\tilde{u}} \mathbf{V}(l).$$

The values of $\mathbf{V}$ and $\nabla_{\tilde{u}}\mathbf{V}$ can be computed using Proposition 2. Because of the dramatic reduction in number of stochastic variables, these values are computed much faster than a full estimation of $v(\tilde{e}(k))$ and $\nabla_{\tilde{u}} I\!E[v(\tilde{e}(k))]$.

Note that because of the approximations, full convexity might be lost. However, if the approximations are close to the original functions, we still have a smooth optimization problem.

5 Computational complexity

Consider equation (11) where n_v is the number of max-terms, and $n_{\tilde{e}}$ is the number of stochastic variables. If we use variability expansion, we usually have a reduced number of stochastic variables, denoted by n_r with $n_r \leq n_{\tilde{e}}$. Let us assume that $e(k)$ is uniformly distributed (see the examples in Section 6)[5]. In the whole procedure, the computation of the sets Φ_j, $j = 1,\ldots,n_v$, is the most time-demanding step. Or to be more precise, we have to compute the vertices of n_v polytopes in $\mathbb{R}^{n_r}$. Every

[5]For other piecewise affine distributions we can make a similar analysis [19]. For more general distributions it may be much harder.

polytope is described by $2n_r+n_v-1$ inequality constraints. If we denote $\mathscr{C}_{\text{vert}}(\ell,n)$ as the complexity to compute all the vertices of a polytope defined by n inequality constraints in an ℓ dimensional space, then the complexity of computing the vertices of n_v polytopes in $\mathbb{R}^{n_r}$ is $\mathscr{C}_{\text{sets}} \sim n_v \mathscr{C}_{\text{vert}}(n_r, 2n_r+n_v-1)$, where $\mathscr{C}_{\text{vert}}(\ell,n)$ is given in [12] as follows:

$$\mathscr{C}_{\text{vert}}(\ell,n) = \left(\begin{array}{c} n-\left\lfloor \frac{\ell+1}{2} \right\rfloor \\ n-\ell \end{array} \right) + \left(\begin{array}{c} n-\left\lfloor \frac{\ell+2}{2} \right\rfloor \\ n-\ell \end{array} \right)$$

where $\lfloor * \rfloor$ denotes the greatest integer function and $\left(\begin{array}{c} * \\ * \end{array} \right)$ denotes the binomial coefficient. In the case of variability expansion of order M we solve $\left(\begin{array}{c} n_{\tilde{e}} \\ m \end{array} \right)$ problems of complexity $n_v \mathscr{C}_{\text{vert}}(m, 2m+n_v-1)$ for $m = 1,\ldots,M$ and so the total complexity becomes

$$\mathscr{C}_{\text{tot}} \sim \sum_{m=1}^{M} \left(\begin{array}{c} n_{\tilde{e}} \\ m \end{array} \right) n_v \mathscr{C}_{\text{vert}}(m, 2m+n_v-1).$$

The dominant factor is for $m = M$ and so the overall complexity will be

$$\mathscr{C}_{\text{tot}} \sim \left(\begin{array}{c} n_{\tilde{e}} \\ M \end{array} \right) n_v \mathscr{C}_{\text{vert}}(M, 2M+n_v-1).$$

For cost criterion (5), the due date error $\eta_i(k+j)$ will only depend on the elements of $\tilde{e}$ that correspond to $e(k+l)$, $l \leq j$ because of causality. This means that the complexity mainly depends on the complexity of computing $\eta_i(k+N_{\text{p}})$, $i = 1,\ldots,n_y$. The number of max-terms for computing $\eta_i(k+N_{\text{p}})$ is $n_v = n_x + N_{\text{p}} n_u$, and usually $n_{\tilde{e}} \approx N_{\text{p}} n_e$, and so the complexity to compute J_{out} is in the order of

$$\mathscr{C}_{\text{tot}} \sim \left(\begin{array}{c} N_{\text{p}} n_e \\ M \end{array} \right) (n_y n_x + n_y N_{\text{p}} n_u)$$

$$\times \left\{ \left(\begin{array}{c} 2M+n_x+N_{\text{p}} n_u-1-\left\lfloor \frac{M+1}{2} \right\rfloor \\ M+n_x+N_{\text{p}} n_u-1 \end{array} \right) + \left(\begin{array}{c} 2M+n_x+N_{\text{p}} n_u-1-\left\lfloor \frac{M+2}{2} \right\rfloor \\ M+n_x+N_{\text{p}} n_u-1 \end{array} \right) \right\}.$$

We see that the complexity will grow rapidly with increasing M. The optimal choice for M will depend on a trade-off between accuracy (larger M) and computation speed (small M). Note that the complexity for the original problem is easily recovered by substitution of $M = n_{\tilde{e}} = N_{\text{p}} n_e$ (i.e. no reduction).

6 Examples

Example 1: A production system

Consider the production system in Figure 1. This system consists of two machines M_1 and M_2 and operates in batches. The raw material is fed to machine M_1 where preprocessing is done. Afterwards

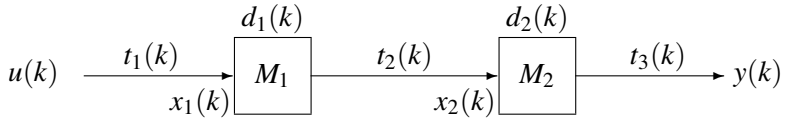

Figure 1: A production system.

the intermediate product is fed to machine M_2 and finally leaves the system. We assume that each machine starts working as soon as possible on each batch, i.e., as soon as the raw material or the required intermediate product is available, and as soon as the machine is idle (i.e., the previous batch of products has been processed and has left the machine).

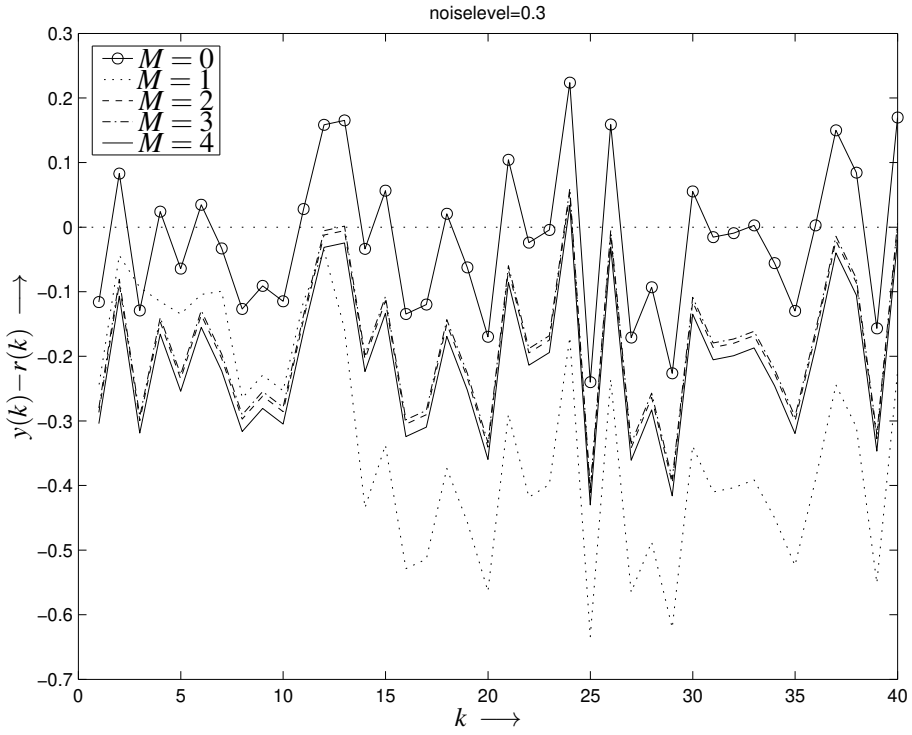

Figure 2: The due date error $y(k)-r(k)$ for MPC with an Mth order approximation, $M \in \{0,1,2,3,4\}$ and a noise level $\alpha = 0.3$.

Define $u(k)$ as the time instant at which the system is fed for the kth time, $y(k)$ as the time instant at which the kth product leaves the system, $x_i(k)$ as the time instant at which machine i starts for the kth time, $t_j(k)$ as the transportation time on link j for the kth batch and $d_i(k)$ as the processing time on machine i for the kth batch. The system equations are given by

$$\begin{aligned}
x_1(k) &= \max(x_1(k-1)+d_1(k-1), u(k)+t_1(k)),\\
x_2(k) &= \max(x_1(k)+d_1(k)+t_2(k), x_2(k-1)+d_2(k-1)),\\
&= \max(x_1(k-1)+d_1(k-1)+d_1(k)+t_2(k), u(k)+d_1(k)+t_1(k)+t_2(k),\\
&\qquad x_2(k-1)+d_2(k-1)),\\
y(k) &= x_2(k)+d_2(k)+t_3(k).
\end{aligned}$$

In matrix notation we obtain (1)-(2) where the system matrices A, B and C are given by

$$A(k) = \begin{bmatrix} d_1(k-1) & \varepsilon \\ d_1(k-1)+d_1(k)+t_2(k) & d_2(k-1) \end{bmatrix},$$

$$B(k) = \begin{bmatrix} t_1(k) \\ d_1(k)+t_1(k)+t_2(k) \end{bmatrix}, \quad C(k) = \begin{bmatrix} \varepsilon & d_2(k)+t_3(k) \end{bmatrix}.$$

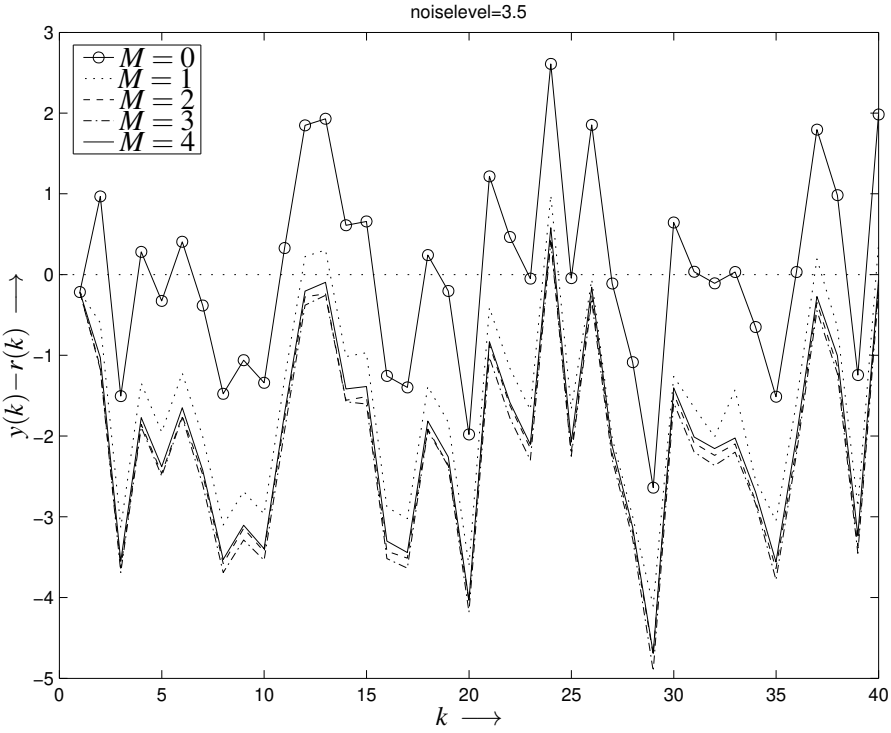

Figure 3: The due date error $y(k)-r(k)$ for MPC with an Mth order approximation, $M \in \{0,1,2,3,4\}$ and a noise level $\alpha = 3.5$.

Let us now solve the stochastic MPC problem for this perturbed MPL system. Assume that two of the transportation times are constant: $t_1(k) = 0$, $t_3(k) = 0$, and that transportation time $t_2(k)$ and the production times $d_1(k)$ and $d_2(k)$ are corrupted by noise:

$$\begin{aligned} d_1(k) &= 5 + \alpha\, 0.2\, e_1(k), \\ d_2(k) &= 1 + \alpha\, 0.5\, e_2(k), \\ t_2(k) &= 1 + \alpha\, 0.6\, e_3(k), \end{aligned}$$

where α is a nonnegative constant and $e(k) = \begin{bmatrix} e_1(k) & e_2(k) & e_3(k) \end{bmatrix}^T$ is a random signal with probability density function

$$p(e) = \begin{cases} 1/8 & \text{if} \quad \max\limits_{i=1,2,3} (|e_i|) \leq 1, \\ 0 & \text{if} \quad \max\limits_{i=1,2,3} (|e_i|) > 1. \end{cases} \tag{13}$$

Assume that the initial state is equal to $x(0) = [0\ 6]^T$, the due date signal is given by $r(k) = 4 + 6 \cdot k$ and the cost criterion (4) is optimized for $N_\mathrm{p} = 3$, $N_\mathrm{c} = 2$ and $\lambda = 0.1$. With the choice of the cost criterion (5)-(6), we can rewrite the stochastic MPC problem into a convex optimization problem. For the computation of the cost criterion we use an Mth order Taylor approximation with $M = 0,1,2,3,4$.

Next we apply MPC for the Mth order approximation for $M = 0,1,2,3,4$. The optimal input sequence is computed for $k = 1,\ldots,40$, and for each k, the first element $u(k)$ of the sequence $\tilde{u}(k)$ is

applied to the perturbed system (due to the receding horizon strategy). We perform two experiments with different noise levels $\alpha = 0.3$ and $\alpha = 3.5$. In the experiments, the true system is simulated for a random sequence $e(k)$, $k = 1,\ldots,40$, satisfying the probability density function (13). The due date error $y(k)-r(k)$ for MPC is given in Figure 2 for a noise level $\alpha = 0.3$, and in Figure 3 for a noise level $\alpha = 3.5$. The 0th order approximation is in fact equal to the case where no disturbance is taken into account. We see that for $M = 0$ the scheme leads to a frequent violation of the due dates (i.e. the difference signal $y(k) - r(k)$ is frequently positive). We see that for increasing approximation order M the due date error decreases and $y(k) - r(k)$ is below zero most of the time (which means that our product is delivered in time). Furthermore, the approximation seems to converge for increasing M.

	$M=0$	$M=1$	$M=2$	$M=3$	$M=4$
CPU time	1	16.5	470	3810	34900

Table 2: (Scaled) CPU times for different levels in approximation

In Table 2 the (scaled) CPU times are given for the computation of the cost criterion and its subgradient for $M \in \{0,1,2,3,4\}$. From Table 2 we see that computation time grows dramatically with increasing M. Depending on the application and the available computation interval, we can choose the level of approximation. In general, the above trade-off will give us the best possible approximation of the optimal solution, given the constraints in computation time. For this system $M = 2$ or $M = 3$ is probably sufficient for practical use.

Example 2: A batch process

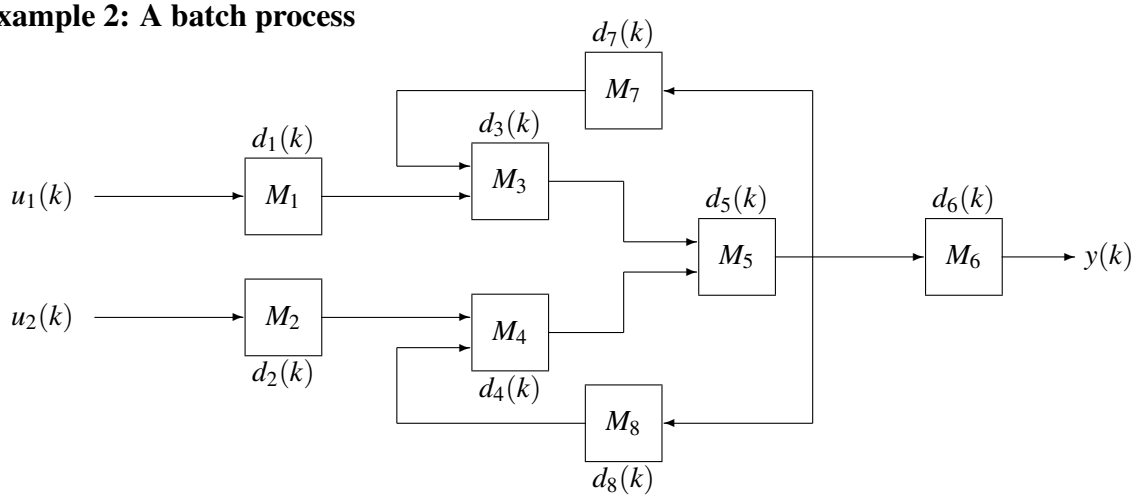

Figure 4: A batch process.

Figure 4 gives the schematic configuration of a batch process [17]. This system consists of six machines M_1 to M_6 operating in batches and two machines M_7 and M_8 working continuously. Two substances are fed into machines M_1 and M_2 where they are heated. In the stirred tank reactors M_3 and M_4 the substances are mixed with a solvent and a reaction takes place. The result flows into tank M_5. Then the solvent is separated from the product and stored into the tanks M_7 and M_8. The product is finalized in machine M_6.

We assume that each machine starts working as soon as possible. We define $u_i(k)$ as the time instant at which the subsystem M_i, $i = 1,2$, is fed for the kth time, $y(k)$ is the time instant at which the

kth product leaves the system, $x_i(k)$ is the time instant at which machine i starts for the kth time, and $d_i(k)$ is the processing time on machine i for the kth batch.

The system equations are given in matrix notation by (1)-(2) with system matrices

$$A(k) = \begin{bmatrix} d_1(k-1) & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & d_2(k-1) & \varepsilon & \varepsilon \\ d_1(k-1)+d_1(k) & \varepsilon & d_3(k-1) & \varepsilon \\ \varepsilon & d_2(k-1)+d_2(k) & \varepsilon & d_4(k-1) \\ d_1(k-1)+d_1(k)+d_3(k) & d_2(k-1)+d_2(k)+d_4(k) & d_3(k-1)+d_3(k) & d_4(k-1)+d_4(k) \\ d_1(k-1)+d_1(k)+d_3(k) & d_2(k-1)+d_2(k)+d_4(k) & d_3(k-1)+d_3(k) & d_4(k-1)+d_4(k) \end{bmatrix}$$

$$\begin{bmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ d_5(k-1)+d_7 & \varepsilon \\ d_3(k)+d_5(k-1)+d_7 & \varepsilon \\ \max(d_3(k)+d_5(k-1)+d_7, d_4(k)+d_5(k-1)+d_8) & \varepsilon \\ \max(d_3(k)+d_5(k-1)+d_7, d_4(k)+d_5(k-1)+d_8) & d_6(k-1) \end{bmatrix},$$

$$B(k) = \begin{bmatrix} 0 & \varepsilon \\ \varepsilon & 0 \\ d_1(k) & \varepsilon \\ \varepsilon & d_2(k) \\ d_1(k)+d_3(k) & d_2(k)+d_4(k) \\ d_1(k)+d_3(k)+d_5(k) & d_2(k)+d_4(k)+d_5(k) \end{bmatrix},$$

$$C(k) = \begin{bmatrix} d_1(k-1)+d_1(k)+d_3(k)+d_6(k) & d_2(k-1)+d_2(k)+d_4(k)+d_6(k) & d_3(k-1)+d_3(k)+d_6(k) \\ d_4(k-1)+d_4(k)+d_6(k) & \max(d_3(k)+d_5(k-1)+d_7+d_6(k), d_4(k)+d_5(k-1)+d_8+d_6(k)) & d_6(k-1)+d_6(k) \end{bmatrix}.$$

Let us now solve the stochastic MPC problem for this perturbed MPL system. Assume that two of the production times are constant: $d_7(k) = d_8(k) = 2$, and that the production times $d_1(k), \ldots, d_6(k)$ are corrupted by noise:

$$d_i(k) = d_{i,0} + \alpha_i e_i(k), \quad i = 1, \ldots, 6,$$

where $d_{1,0} = d_{2,0} = 1$, $d_{3,0} = d_{4,0} = 3$, $d_{5,0} = 4$, $d_{6,0} = 3$, $\alpha_1 = \alpha_2 = 0.2$, $\alpha_3 = \alpha_4 = 0.1$, $\alpha_5 = 0.3$, $\alpha_6 = 0.1$ and $e(k) = \begin{bmatrix} e_1(k) & \ldots & e_6(k) \end{bmatrix}^T$ is a random signal with probability density function

$$p(e) = \begin{cases} 1/64 & \text{if} \quad \max\limits_{i=1,\ldots,6}(|e_i|) \leq 1, \\ 0 & \text{if} \quad \max\limits_{i=1,\ldots,6}(|e_i|) > 1. \end{cases} \tag{14}$$

Assume that the due date signal is given by $r(k) = 6 + 10 \cdot k$ and the cost criterion (4) is optimized for $N_\mathrm{p} = 3$, $N_\mathrm{c} = 2$ and $\lambda = 0.001$. With the choice of the cost criterion (5)-(6), we can rewrite the stochastic MPC problem into a convex optimization problem. For the computation of the cost criterion we use a Taylor approximation with $M = 0, 1, 2$.

We apply MPC for the Mth order approximation for $M = 0, 1, 2$. The simulation with the MPC controller is done for $k = 1, \ldots, 40$. The due date error $y(k) - r(k)$ for MPC is given in Figure 5. The 0th order approximation is in fact equal to the case where no disturbance is taken into account.

We see that for $M = 0$ (no disturbance is taken into account) and $M = 1$ the scheme leads to a frequent violation of the due dates (i.e. the difference signal $y(k) - r(k)$ is frequently positive). We see that for $M = 2$ the due date error is small enough for a proper functioning of the process.

7 Discussion

We have discussed complexity reduction in MPC for max-plus linear discrete event systems with stochastic uncertainties. From the MPC framework, a convex optimization problem results if the constraints are a nondecreasing function of the output. With an increasing number of stochastic variables,

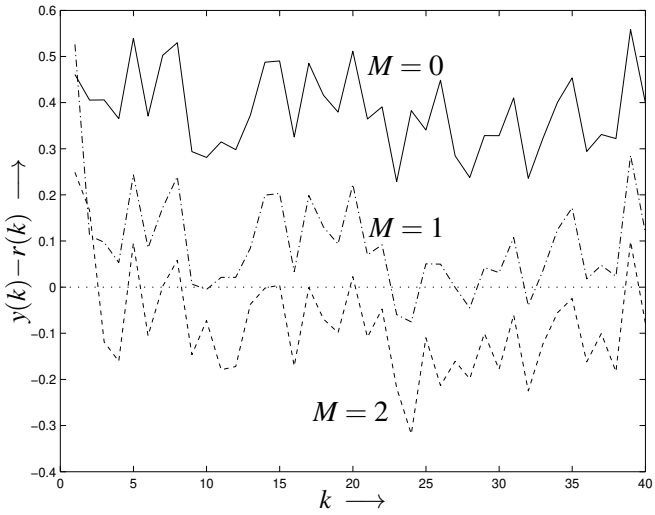

Figure 5: The due date error $y(k)-r(k)$ for MPC on a batch process with an Mth order approximation, $M \in \{0,1,2\}$.

the computational complexity of the optimization problem increases dramatically due to the numerical integrations required to evaluate the objective function. To tackle this increase of complexity, we use the method of variability expansion. The key idea of this method is to introduce a parameter θ that controls the level of stochasticity in the system. In this paper we have derived explicit expressions for the coefficients in the expansion (and we have provided the proofs that were lacking in [6]). Based on a Taylor expansion in the parameter θ, good approximations for the expectations of the cost criterion and the constraints can be computed, which leads to a significant reduction of the computational complexity of our approach. We have analyzed the computational complexity of the overall algorithm, and illustrated the theory with two worked examples. From the examples it becomes clear that if we do not take the stochastic perturbation into account (the case that the approximation order is $M = 0$), the due-date error will often be positive, which means for a production system that the products are finished too late. Even for small M the due-date error is reduced dramatically, and the system can deliver products in time.

Acknowledgments

Research partially funded by the Dutch Technology Foundation STW project "Model predictive control for hybrid systems" (DMR.5675), the European 6th Framework Network of Excellence "HYbrid CONtrol: Taming Heterogeneity and Complexity of Networked Embedded Systems (HYCON)" and by the European IST project "Modelling, Simulation and Control of Nonsmooth Dynamical Systems (SICONOS)" (IST-2001-37172).

References

[1] F. Baccelli, G. Cohen, G.J. Olsder, and J.P. Quadrat. *Synchronization and Linearity*. John Wiley & Sons, (out of print, accessible via the web at www.maxplus.org), New York, 1992.

[2] J.L. Boimond and J.L. Ferrier. Internal model control and max-algebra: Controller design. *IEEE Trans. on Aut. Control*, 41(3):457–461, March 1996.

[3] R.A. Cuninghame-Green. *Minimax Algebra*, volume 166 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, 1979.

[4] B. De Schutter and T. van den Boom. Model predictive control for max-plus-linear discrete event systems. *Automatica*, 37(7):1049–1056, July 2001.

[5] C.E. García, D.M. Prett, and M. Morari. Model predictive control: Theory and practice — A survey. *Automatica*, 25(3):335–348, May 1989.

[6] B. Heidergott. Variablity expansion for performance characteristic of (max,plus)-linear systems. In *Proc. of the Workshop on Discrete Event Systems (WODES'02)*, Zaragoza, Spain, 2002.

[7] B. Heidergott. *Max-Plus Linear Stochastic Systems and Perturbation Analysis*. Springer, New York, 2006.

[8] B. Heidergott, G.J. Olsder, and J. van der Woude. *Max Plus at Work: Modeling and Analysis of Synchronized Systems*. Princeton University Press, Princeton, 2006.

[9] D. Knuth. *The Art of Computing, Vol. I.* Addison-Wesley, Massachusetts, 1997.

[10] L. Libeaut and J.J. Loiseau. Admissible initial conditions and control of timed event graphs. In *Proc. of the 34th IEEE Conf. on Decision and Control*, pages 2011–2016, New Orleans, Louisiana, December 1995.

[11] J.M. Maciejowski. *Predictive control with constraints*. Prentice Hall, Pearson Education Limited, Harlow, UK, 2002.

[12] T.H. Mattheiss and D.S. Rubin. A survey and comparison of methods for finding all vertices of convex polyhedral sets. *Mathematics of Operations Research*, 5(2):167–185, May 1980.

[13] E. Menguy, J.L. Boimond, and L. Hardouin. Adaptive control for linear systems in max-algebra. In *Proc. of the International Workshop on Discrete Event Systems (WODES'98)*, pages 481–488, Cagliari, Italy, August 1998.

[14] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, Pennsylvania, 1994.

[15] G.J. Olsder, J.A.C. Resing, R.E. de Vries, M.S. Keane, and G. Hooghiemstra. Discrete event systems with stochastic processing times. *IEEE Transactions on Automatic Control*, 35(3):299–302, March 1990.

[16] J.A.C. Resing, R.E. de Vries, G. Hooghiemstra, M.S. Keane, and G.J. Olsder. Asymptotic behavior of random discrete event systems. *Stochastic Processes and their Applications*, 36:195–216, 1990.

[17] G. Schullerus and V. Krebs. A method for estimating the holding times in timed event graphs. In *Proc. of the Workshop on Discrete Event Systems (WODES'02)*, Zaragoza, Spain, October 2002.

[18] T.J.J. van den Boom and B. De Schutter. MPC for perturbed max-plus-linear systems. *Systems and Control Letters*, 45(1):21–33, January 2002.

[19] T.J.J. van den Boom and B. De Schutter. Model predictive control for perturbed max-plus-linear systems: A stochastic approach. *International Journal of Control*, 77(3):302–309, February 2004.

[20] S.J. Wright. *Primal-Dual Interior Point Methods*. SIAM, Philadephia, Pennsylvania, 1997.

Appendix

A1. Variability expansion (for more details see [7])

We show that $I\!E[v(\tilde{e}_\theta(k))]$ is infinitely often differentiable with respect to θ, and more particular for $\theta \in [0,1]$, which is the interval of interest for us due to the interpretation of θ as the level of randomness. Note that $\tilde{e}(k)$ can be written as $f(X_1,\ldots,X_{n_{\tilde{e}}})$ for some measurable mapping f and i.i.d. random noise variables X_i. Let μ denote the distribution of X_i and a its finite expected value, i.e., $I\!E[X_i] = \int_{\mathbb{R}} x\mu(dx) = a$, for $1 \le i \le n_{\tilde{e}}$. We formalize variability expansion as follows. Choose $l \in \{0,1\}^{n_{\tilde{e}}}$, let X_i have distribution μ (the "true" distribution) if $l_i = 1$ and let $X_i = a$ (with probability one) if $l_i = 0$. This is easily achieved be replacing those X_i in f for which $l_i = 1$ by a . The thus modified mapping f is denoted by f_l. Next, let the elements of l be independently distributed with $\mathbb{P}(l_i = 1) = 1-\theta$ and $\mathbb{P}(l_i = 0) = \theta$, where $\mathbb{P}(\cdot)$ denotes probability. Furthermore, let $X_i(\theta)$ have distribution μ with probability θ and let it be deterministic and equal to a with probability $1-\theta$. It then holds that

$$I\!E[v(\tilde{e}_\theta(k))] = I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))] = \sum_{l\in\{0,1\}^{n_{\tilde{e}}}} I\!E[f_l(X_1,\ldots,X_{n_{\tilde{e}}})]\,(1-\theta)^{\sum_{i=1}^{n_{\tilde{e}}} l_i}\,\theta^{n_{\tilde{e}}-\sum_{i=1}^{n_{\tilde{e}}} l_i}.$$

Note that the sum on the right-hand side of the above equation is finite and we may interchange the order of higher-order differentiation and summation. Since the distribution of l is a polynomial in θ of order $n_{\tilde{e}}$, it is infinitely differentiable and its derivatives of order $n_{\tilde{e}}+1$ and higher vanish.

A2. Proof of Lemma 4 (for more details see [7])

Let $\mu_\theta = \theta\mu + (1-\theta)\delta_a$, where δ_a denotes the Dirac measure in a (which carries only mass on point a). With this notation,

$$I\!E[v(\tilde{e}_\theta(k))] = I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))] = \int_{\mathbb{R}^{n_{\tilde{e}}}} f(x_1,\ldots,x_{n_{\tilde{e}}})\,\mu_\theta^{n_{\tilde{e}}}(dx_1)\cdots\mu_\theta^{n_{\tilde{e}}}(dx_{n_{\tilde{e}}}),$$

where

$$f(x_1,\ldots,x_{n_{\tilde{e}}}) = I\!E\left[v(\tilde{e}_\theta(k)) \,\middle|\, [\tilde{e}_\theta(k)]_j = x_j, 1 \le j \le n_{\tilde{e}}\right].$$

Provided that the elements of $\tilde{e}_\theta(k)$ have bounded support it follows from our basic model that $f(x_1,\ldots,x_{n_{\tilde{e}}})$ is bounded as a function in $x_1,\ldots,x_{n_{\tilde{e}}}$. Interchanging integration and differentiation is thus justified and we obtain

$$\frac{d^n}{d\theta^n} I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))] = \int_{\mathbb{R}^{n_{\tilde{e}}}} f(x_1,\ldots,x_{n_{\tilde{e}}})\,\frac{d^n}{d\theta^n}\mu_\theta(dx_1)\cdots\mu_\theta(dx_{n_{\tilde{e}}}).$$

Note that $\mu_\theta(dx) = \theta\mu(dx) + (1-\theta)\delta_a(dx)$, which implies $\frac{d}{d\theta}\mu_\theta(dx) = \mu(dx) - \delta_a(dx)$. After some computation one arrives at

$$\frac{d^n}{d\theta^n}\mu_\theta(dx_1),\ldots,\mu_\theta(dx_{n_{\tilde{e}}}) = n! \sum_{l\in\mathscr{L}(n_{\tilde{e}};n)} \left(\prod_{k=1}^{n_{\tilde{e}}} (\mu_\theta)^{(l_k)}(dx_k) - \prod_{k=1}^{n_{\tilde{e}}} (\mu_\theta)^{(l_k^-)}(dx_k)\right), \tag{15}$$

where the following definitions are in force. The set $\mathscr{L}(n_{\tilde{e}};n)$ is given by

$$\mathscr{L}(n_{\tilde{e}};n) \quad = \quad \left\{ l \in \{0,-1,1\}^{n_{\tilde{e}}} \;\middle|\; \sum_{k=1}^{n_{\tilde{e}}} l_k = n \text{ and } \prod_{\substack{l_1,\ldots,l_{n_{\tilde{e}}} \\ l_k \neq 0}} l_k = 1 \right\} .$$

For $l \in \mathscr{L}(m;n)$, the vector l^- is generated out of l by changing the sign of the highest non-zero element; more formally, let k^* be the highest position of a non-zero element in $l \in \mathscr{L}(m;n)$, that is, $l_k = 0$ for all $k > k^*$ and $l_{k^*} \in \{-1,+1\}$, and set

$$l^- = (l_1,\ldots,l_m)^- = (l_1,\ldots l_{k^*-1}, -l_{k^*}, l_{k^*+1},\ldots,l_m) .$$

Finally, we have set

$$\mu_\theta^{(0)} = \mu_\theta , \quad \mu_\theta^{(1)} = \mu \quad \text{ and } \mu_\theta^{(-1)} = \delta_a .$$

With (15), we obtain

$$\begin{aligned} &\frac{d^n}{d\theta^n} I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))] \\ &\qquad = n! \sum_{l \in \mathscr{L}(n_{\tilde{e}};n)} \int \cdots \int f(x_1,\ldots,x_{n_{\tilde{e}}}) \Big(\prod_{k=1}^{n_{\tilde{e}}} (\mu_\theta)^{(l_k)}(dx_k) - \prod_{k=1}^{n_{\tilde{e}}} (\mu_\theta)^{(l_k^-)}(dx_k) \Big). \end{aligned}$$

In order to interpret the integral expression on the right-hand side of the above equation as an expected value with respect to appropriate random variables, we introduce $X_i(r)(\theta)$ such that $X_i(r)(\theta)$ has distribution $\mu_\theta(r)$ for $r \in \{0,1,-1\}$. This yields

$$\begin{aligned} &\frac{d^n}{d\theta^n} I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))] \\ &\qquad = n! \sum_{l \in \mathscr{L}(n_{\tilde{e}};n)} \left(I\!E\left[f\left(X_1^{(l_1)}(\theta),\ldots,X_{n_{\tilde{e}}}^{(l_{n_{\tilde{e}}})}(\theta) \right) \right] - I\!E\left[f\left(X_1^{(l_1^-)}(\theta),\ldots,X_m^{(l_{n_{\tilde{e}}}^-)}(\theta) \right) \right] \right) . \end{aligned}$$

The key observation it that only those elements $X_i^{(l_i)}(\theta)$ with $l_i = 0$ (resp. $X_i^{(l_i^-)}(\theta)$ with $l_i^- = 0$) actually do depend on θ. Moreover, as θ tends to zero, those $X_i^{(l_i)}(\theta)$ (resp. $X_i^{(l_i^-)}(\theta)$) with $l_i = 0$ converge weakly to a, whereas all other noise variables remain unaffected. Hence, the limit of the higher order derivatives of $I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))]$ as θ tends to 0 exists and is given by

$$\begin{aligned} &\lim_{\theta \downarrow 0} \frac{d^n}{d\theta^n} I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))] \\ &= n! \sum_{l \in \mathscr{L}(n_{\tilde{e}};n)} \Big(I\!E\Big[f\big(1_{l_1=1} X_1 + 1_{l_1 \in \{0,-1\}} a, \ldots, 1_{l_{n_{\tilde{e}}}=1} X_{n_{\tilde{e}}} + 1_{l_{n_{\tilde{e}}} \in \{0,-1\}} a \big) \Big] \\ &\qquad\qquad - I\!E\left[f\left(1_{l_1^-=1} X_1 + 1_{l_1^- \in \{0,-1\}} a, \ldots, 1_{l_{n_{\tilde{e}}}^-=1} X_{n_{\tilde{e}}} + 1_{l_{n_{\tilde{e}}}^- \in \{0,-1\}} a \right) \right] \Big) , \end{aligned}$$

for $n_{\tilde{e}} \geq n$ and zero otherwise; where $1_A = 1$ if expression A holds and zero otherwise.

For $n = 1$, the set $\mathscr{L}(n_{\tilde{e}};n)$ contains vectors with all elements equal to zero except for one entry. Hence, the first order derivative at $\theta = 0$ can be written as follows

$$\lim_{\theta \downarrow 0} \frac{d}{d\theta} I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))]$$

$$= \sum_{l=1}^{n_{\tilde{e}}} \left(I\!E \left[f\left(1_{l=1} X_1 + 1_{l\neq 1}\, a, \ldots, 1_{l=n_{\tilde{e}}} X_{n_{\tilde{e}}} + 1_{l\neq n_{\tilde{e}}}\, a\right)\right] - I\!E\left[f\left(a,\ldots,a\right)\right]\right),$$

which in the notation of the paper reads

$$\lim_{\theta\downarrow 0} \frac{d}{d\theta} I\!E[f(X_1(\theta),\ldots,X_{n_{\tilde{e}}}(\theta))] = \sum_{l=1}^{n_{\tilde{e}}} \left(V(l) - V(0)\right).$$

Following this train of thought for higher-order derivatives proves Lemma 4.