Proceedings of the 14th International Symposium on Process Systems Engineering – PSE 2021+
June 19–23, 2022, Kyoto, Japan

Data-driven online scenario selection for multistage NMPC

Zawadi Mdoe, Mandar Thombre, Johannes Jäschke*

Department of Chemical Engineering, Norwegian University of Science & Technology, Sem Sælands vei 4, 7491 Trondheim, Norway.

johannes.jaschke@ntnu.no

Abstract

This paper aims at reducing the conservativeness of the robust and computationally efficient sensitivity assisted multistage nonlinear model predictive controller. The approach uses a hyperbox over-approximation for the parametric uncertainty set that often results into conservativeness. We propose the use of principal component analysis (PCA) on available process data to extract more information to tighten the approximation of the parametric uncertainty set. It is approximated by a polytope whose vertices lie on the principal components. Then we define the multistage nonlinear problem with a linear transformation of the uncertain parameters. This transformation ensures consistency with the required conditions for sensitivity assisted multistage MPC algorithm used for scenario tree pruning. Finally, the method was implemented on a case study of a system of four tanks and the controller exhibited reduced conservativeness and fast computational performance.

Keywords: Robust MPC, Dynamic optimization, Parametric uncertainty, Data-driven

1. Introduction

Model predictive control (MPC) is a model based control strategy that reoptimizes a nonlinear process system with respect to a control objective subject to constraints at each sampling time. MPC includes constraints for online decision making, and has good control performance even when the system is disturbed away from the desired reference trajectory (Rawlings & Mayne, 2009). Although MPC has inherent robustness against uncertainty, the property may break when there are significant disturbances, causing infeasibilities. As a result, robust MPC approaches have been developed. One of them was proposed by Lucia et al. (2013) and is known as the multistage MPC.

1.1. Multistage MPC

Multistage MPC explicitly considers a selection of possible future scenarios along a prediction horizon to formulate its optimization problem. The scenarios are determined by propagating from the current state to the end of the prediction horizon, a finite number of uncertain parameter realizations using a scenario tree. When the prediction horizon is long the number of scenarios in the scenario tree increases exponentially resulting into an intractable problem. Lucia et al. (2013) proposed a robust horizon where the scenario tree branching is stopped before the end of the horizon, and the uncertain parameters are kept constant until the end of the prediction horizon. The robust horizon makes the problem practically feasible to solve but can still be expensive, especially for nonlinear problems, leading to a significant computational delay. In order to reduce the computational cost and computational delay of the multistage MPC, Thombre et al. (2020) proposed the sensitivity assisted multistage MPC. It has an algorithm to prune irrelevant scenarios from the scenario tree using NLP sensitivities in order to speed up computations. The sensitivity assisted multistage MPC is discussed further in Section 2.

1.2. Motivation

Even though multistage MPC is robust against constraint violations, it is rather conservative, resulting into performance loss. The conservativeness is highly dependent on how uncertainty set is represented. So far, its implementation has mainly been done using a hyperbox over-approximation of the uncertainty set. The over-approximation is often very poor if the true uncertainty set is ellipsoidal. Although the computational delay of multistage MPC can be reduced by the sensitivity assisted algorithms (Thombre et al., 2020), it has been implemented with an over-approximation of the uncertainty set leading to conservative control performance. However, in combination with statistical data analysis methods used for uncertainty identification, one can significantly reduce the conservativeness. Krishnamoorthy et al. (2018) suggested that detailed information on process uncertainty could be extracted via statistical data analysis to obtain more representative scenarios. Moreover, Shang & You (2019) rigourously present on calibration of approximate uncertainty sets for a scenario-based stochastic MPC in linear systems using support vector clustering with stability guarantees based on some mild assumptions. The contribution of this paper is to demonstrate how principal component analysis can be specifically applied to the sensitivity assisted multistage MPC framework in order to reduce conservativeness and retain its computational efficiency.

1.3. Notation

We assume a nonlinear system model $z_{i+1} = f(z_i, \nu_i, d_i)$ that predicts the evolution of the states z_i from time t_{k+i} with control actions ν_i and uncertain parameters d_i. Let us define the notation used in this manuscript. The time index $k \geq 0$ corresponds to sampling time t_k. A perfect state measurement is always assumed, and the state at time t_k is denoted by x_k. The time index of a model prediction is denoted by $i \in \mathbb{Z}_+$ which corresponds to sample time t_{k+i}. The nominal parameters are denoted as d_i^0 such that the nominal model becomes $z_{i+1}^0 = f(z_i^0, \nu_i^0, d_i^0)$. For a nonlinear system we obtain a nonlinear optimization problem (NLP) resuling inot a class of MPC known as nonlinear MPC (NMPC).

2. Sensitivity assisted multistage NMPC

The algorithm for the sensitivity assisted multistage NMPC (samNMPC) that performs online critical scenario selection based on NLP sensitivities was first developed by Holtorf et al. (2019). This selection is done by solving the NMPC problem for the nominal scenario together with a lower level optimization problem (LLP) that gives the parametric realizations that maximize the inequality constraints. This gives the constraints that are most likely violated. However, when the inequalities are interval bounds there exists a trivial solution to the LLP that lies on the vertices of the uncertainty hyperbox. Assume that the constraints are monotonically increasing or decreasing in the uncertain parameter space. The multistage MPC problem is parametric in the disturbances thus the online critical scenarios selection is based on parametric NLP sensitivities from the nominal scenario. This algorithm determines the realization most likely to violate a constraint using the sign of the parametric sensitivity. It formulates a pruned scenario tree with only the critical scenarios and the nominal, leading to a smaller NMPC problem that is cheaper to solve. The stability and recursive feasibility properties of the samNMPC were established by Thombre et al. (2020). A sensitivity assisted multistage NMPC problem at time t_k is written as follows:

$$
\begin{aligned}
V_N^{sam}(x_k) = \min_{\substack{z_i^c,\nu_i^c \\ c\in\widehat{\mathbb{C}}\cup\{0\}}} \sum_{c\in\widehat{\mathbb{C}}\cup\{0\}} \omega_c\Big(\psi(z_N^c, d_{N-1}^c) + \sum_{i=0}^{N-1} \ell(z_i^c,\nu_i^c,d_i^c)\Big) + \\
\sum_{c\in\bar{\mathbb{C}}} \omega_c\Big(\psi(z_N^0+\Delta z_N^c, d_{N-1}^c) + \sum_{i=0}^{N-1} \ell(z_i^0+\Delta z_i^c, \nu_i^0+\Delta\nu_i^c, d_i^c)\Big)
\end{aligned} \tag{1a}
$$

$$
\text{s.t. } z_{i+1}^c = f(z_i^c,\nu_i^c,d_i^c), \quad i = 0,\dots,N-1 \tag{1b}
$$

$$
z_0^c = x_k,\ z_N^c \in \mathbb{X}_f, \tag{1c}
$$

$$
\nu_i^c = \nu_i^{c'}, \quad \{(c,c') \mid z_i^c = z_i^{c'}\} \tag{1d}
$$

$$
d_{i-1}^c = d_i^c, \quad i = N_R,\dots,N-1 \tag{1e}
$$

$$
z_i^c \in \mathbb{X},\ \nu_i^c \in \mathbb{U},\ d_i^c \in \mathbb{D}, \forall\, c,c' \in \widehat{\mathbb{C}}\cup\{0\} \tag{1f}
$$

where the sets $\widehat{\mathbb{C}}$ and $\bar{\mathbb{C}}$ are the critical and noncritical scenario index sets, respectively and $\{0\}$ repesents the nominal scenario. $\mathbb{D} \in \mathbb{R}^{n_d}$ is the uncertain parameter set containing a finite number of realizations, $\mathbb{X} \in \mathbb{R}^{n_x}$, $\mathbb{U} \in \mathbb{R}^{n_u}$ are the feasible sets for states and inputs, respectively and $\mathbb{X}_f$ represents the terminal set. N is the prediction horizon length and N_R is the robust horizon. z_i^c and ν_i^c are the predicted state and control variable vectors for scenario c at time t_{k+i}, respectively. The stage cost function is given by ℓ, terminal cost is denoted by ψ, and ω_c represents the weights on scenario c to the objective function. The variables and constraints in problem (1) are only those associated with critical constraints, thus making the problem smaller than that of the ideal multistage NMPC with a robust horizon.

3. Data driven sensitivity assisted multistage NMPC

This section presents the main idea which is to integrate principal component analysis (PCA) and samNMPC in order to reduce its conservativeness, hence enhancing its performance. The goal is to achieve that while retaining the computational speed of samNMPC.

3.1. Principal component analysis

Principal component analysis (PCA) is a multivariate data analysis tool that reveals hidden information from data. This method evaluates the variability in the data set and identifies principal components (PC) which are the unit directions that explain the total variation in the data. As a result, PCA fits a hyperellipsoid to the data with the principal components corresponding to the ellipsoids axes. The principal components are listed in order of decreasing component variance.

Assume we have a data set with n_s samples for each uncertain parameter and the data set is a represented by a matrix $\mathbf{D} \in \mathbb{R}^{n_s \times n_d}$. Before decomposition, the data set must be mean centered and scaled because PCA is sensitive to scale differences. Let the scaled and mean centered data corresponding to $\mathbf{D}$ be denoted as $\mathbf{D_0} \in \mathbb{R}^{n_s \times n_d}$. PCA on $\mathbf{D_0}$ results in the linear model $\mathbf{D_0} = \mathbf{\Lambda}\mathbf{C}^\top$ where $\mathbf{\Lambda} \in \mathbb{R}^{n_s \times n_p}$ is a matrix with the scores corresponding to each data sample. The scores are a projection of the data points onto the principal components directions. The matrix $\mathbf{C} \in \mathbb{R}^{n_p \times n_p}$ is made up of the weights on the original samples required to obtain the component score.

3.2. Algorithm for scenario selection using both data and NLP sensitivities

This algorithm combines PCA that determines the maximum and minimum scores in the principal component directions with the samNMPC algorithm presented by Thombre et al. (2020). In order to use the samNMPC algorithm with data, we make a linear transformation of the uncertain parameters in the optimization problem using the PCA matrix. The algorithm has the following steps

(a) Scale or normalize and mean-center the data set $\mathbf{D}$ to obtain $\mathbf{D_0}$.
(b) Perform PCA on $\mathbf{D_0}$ to determine the principal component scores $\mathbf{\Lambda}$ and the corresponding principal component matrix $\mathbf{C}$.
(c) Transform the uncertain parameter vectors d_i^c into the new orthogonal space using the matrix $\mathbf{C}$, such that, $d_i^c = \mathbf{C}\overline{d}_i^c + d_i^0$ where $\overline{d}_i^c$ are the transformed parameters.
(d) Substitute the transformation from step (c) above in problem (1) to obtain an NLP in terms of the transformed parameters.
(e) At the current time t_k, determine critical scenarios $\hat{\mathbb{C}}$ and non-critical scenarios $\bar{\mathbb{C}}$ with respect to the transformed parameters using the samNMPC algorithm.
(f) Generate a pruned scenario tree with only the critical scenarios and the nominal scenario and then solve the transformed problem (1).

4. Case study

Consider the quadtank problem with a four tank configuration from Raff et al. (2006). The levels of water in the four tanks are described by the following set of differential equations:

$$\dot{x_1} = -\frac{a_1}{A_1}\sqrt{2gx_1} + \frac{a_3}{A_1}\sqrt{2gx_3} + \frac{\gamma_1}{A_1}u_1 \qquad \dot{x_3} = -\frac{a_3}{A_3}\sqrt{2gx_3} + \frac{1-\gamma_2}{A_3}u_2$$
$$\dot{x_2} = -\frac{a_2}{A_2}\sqrt{2gx_2} + \frac{a_4}{A_2}\sqrt{2gx_4} + \frac{\gamma_2}{A_2}u_2 \qquad \dot{x_4} = -\frac{a_4}{A_4}\sqrt{2gx_4} + \frac{1-\gamma_1}{A_4}u_1$$

where the states x_i are the tank levels, the inputs u_i are pump flow rates, and the uncertain parameters are the valve coefficients γ_1 and γ_2. The controller tracks setpoint levels x_1 and x_2 with minimum input usage such that the objective is $\ell = (x_1 - x_1^*)^2 + (x_2 - x_2^*)^2 + r_1 u_1^2 + r_2 u_2^2$. There are constraints on x_3 and x_4 and the system experiences predefined pulses in x_1 as described by Thombre et al. (2020).

4.1. Data analysis

The uncertain parameters have a process data cloud shown in the left plot of Figure 1. PCA on the data gives $\mathbf{C} = [0.6571,\ -0.7538;\ 0.7538,\ 0.6571]$. The red circled points

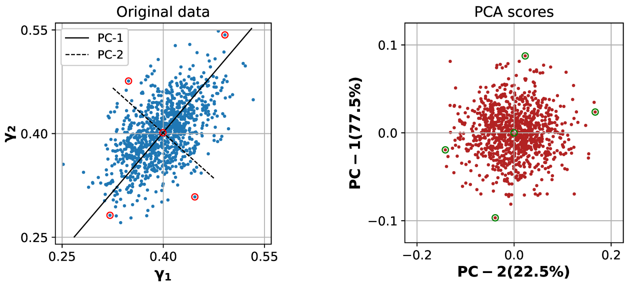

Figure 1: PCA on process data. Left shows original data, right shows the PCA scores.

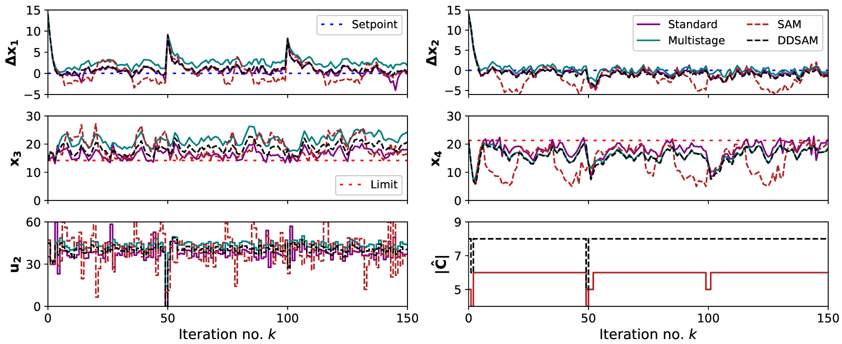

Figure 2: Comparing the control performance of the data-driven samNMPC with standard (nominal) NMPC, multistage, sensitivity-assisted multistage NMPC in the quadtank problem.

are the data points corresponding to the extreme scores on each principal component. The scores in the principal components are shown in plot on the right of Figure 1.

4.2. Results

The uncertain parameters γ_1 and γ_2 are random values generated from the multivariate distribution of the process. Then simulations were performed for both standard NMPC, multistage NMPC, samNMPC and the data-driven samNMPC. It was done for 150 iterations and the results for robust horizon $N_R = 2$ are shown in Figure 2. The tracking performance of the samNMPC is improved by the data transformation. Data-driven samNMPC tracks closer to the set point hence it is less conservative than original samNMPC and multistage NMPC. It is also robust against constraint violations for x_3 and x_4. To show the improvement of the tracking performance, we computed the accumulated cost in the simulation as shown in the bar chart on the right of Figure 3. For robust horizons lengths 1 to 3, data-driven samNMPC shows a slightly better setpoint tracking performance than the standard NMPC. It also shows a significant improvement from the original samNMPC tracking performance. In terms of computational efficiency, Figure 3 shows that the data-driven samNMPC is as fast as the original samNMPC.

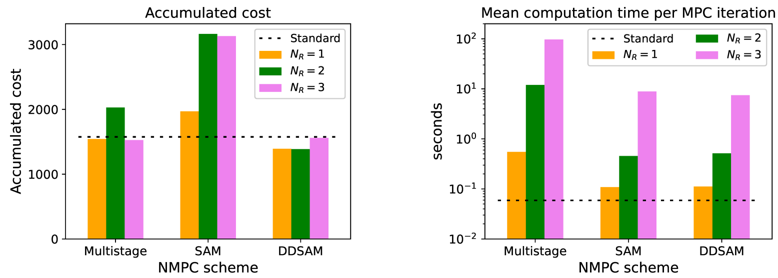

Figure 3: Comparing tracking costs (left - absolute scale) and the average computation time (right - logarithmic scale) for different robust horizons (N_R).

5. Conclusions

We have demonstrated how analysis on process data can extract more information on the uncertainty set used to formulate the sensitivity assisted multistage MPC problem. The integration of data with samNMPC requires transforming the uncertain parameters into new variables corresponding to the principal components. As a result, the samNMPC becomes less conservative while still being computationally efficient. A caveat to the method is that the uncertainty set representation is an under-approximation using a polytope whose vertices are the maximum and minimum PCA scores. There is still a chance that the process may be outside the polytope especially when a dominant principal component does not exist. However, we expect such cases to be rare and we include soft constraints on the state bounds with penalties to avoid the infeasibilities. Future work would be investigating how scaling up to a higher-order system with more uncertain parameters affects the method's performance.

References

F. Holtorf, A. Mitsos, & L. T. Biegler (2019). Multistage NMPC with on-line generated scenario trees: Application to a semi-batch polymerization process. Journ. Proc. Control.

D. Krishnamoorthy, M. Thombre, S. Skogestad, & J. Jäschke (2018). Data-driven scenario selection for multistage robust model predictive control. IFACPapersOnLine.

S. Lucia, T. Finkler, S. Engell (2013). Multi-stage nonlinear model predictive control applied to a semi-batch polymerization reactor under uncertainty, Journ. Proc. Control.

T. Raff, S. Huber, Z. K. Nagy, & F. Allgower (2006). Nonlinear model predictive control of a four tank system: An experimental stability study. Int. Symp. on Intelligent Control. IEEE.

J. B. Rawlings, & D. Q. Mayne (2009). Model Predictive Control: Theory and Design.

C. Shang, & F. You (2019). A data-driven robust optimization approach to scenario-based stochastic model predictive control, Journ. Proc. Control.

M. Thombre, Y. Zhou, J. Jäschke, & L. T. Biegler (2020). Sensitivity-assisted multistage nonlinear model predictive control: Robustness, stability and computational efficiency, Comp. & Chem. Eng.