HAL open science

Maintaining a relevant dataset for data-driven MPC

Alexandre Faye-Bédrin, Stanislav Aranovskiy, Paul Chauchat, Romain Bourdais

To cite this version:

Alexandre Faye-Bédrin, Stanislav Aranovskiy, Paul Chauchat, Romain Bourdais. Maintaining a relevant dataset for data-driven MPC. Premier congrès annuel de SAGIP(Société d'Automatique, de Génie Industriel et de Productique), SAGIP, Jun 2023, Marseille, France. hal-04073049

HAL Id: hal-04073049

https://centralesupelec.hal.science/hal-04073049v1

Submitted on 18 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1er Congrès Annuel
Marseille 7-9 Juin 2023

Maintaining a relevant dataset for data-driven MPC

Alexandre Faye-Bédrin, Stanislav Aranovskiy, Paul Chauchat, Romain Bourdais

IETR, CentraleSupélec, France

alexandre.faye-bedrin@centralesupelec.fr

Mots-clés : *MPC, Data-Driven, Persistency of excitation*

1 Introduction

Model Predictive Control (MPC) is a mature research area and one of the most industrially accepted advanced control solutions. For MPC, the accuracy of the process model is crucial, which motivates and gives rise to research on Adaptive MPC capable of online adjustments for model uncertainties and variations. In this work, we are interested in the Data-Driven MPC (DD MPC). This approach does not include an explicit parametric model; in contrast, it uses a dataset of past measurements to formulate an optimization problem.

One possible approach in DD MPC is based on Willems *et al*'s fundamental lemma [4]. It says that any trajectory of a discrete-time linear time-invariant system can be derived from a sufficiently large dataset of past trajectories if these trajectories satisfy the persistence of excitation condition. Then an optimization problem can be formulated for the future trajectory using this dataset and avoiding explicit dynamics formulation.

This idea is used in [3], where a DD MPC solution for a nonlinear system is developed. To this end, a record of recent input-output measurements is used as a linearization around the current position, and this linearization is then used for MPC. In [3], the crucial assumption is that the dataset remains persistently excited when operating in the closed loop. Such an assumption requires a dataset-maintaining strategy. The authors highlight that developments of efficient dataset update heuristics remain an open question. Specifically, they propose to collect measurements until the system arrives at the desired position, and then the collected dataset is frozen; no future measurements are used. While this approach works for constant reference tracking, it may yield an error if the reference changes. The dataset collected around a linearization point may be irrelevant when used for regulation around *another* linearization point (another reference). Thus, the dataset-maintaining problem remains open.

The contribution of this paper is as follows. We apply the DD MPC [3] to a nonlinear MIMO testbed representing a heating/ventilation system, where the temperature and the airflow set points (references) are piece-wise constants. Due to the changes in the set points, the heuristic proposed in [3] yields a steady-state tracking error. We propose a novel dataset-maintaining strategy that ensures the use of the most recent sufficiently excited dataset; the proposed strategy tracks the smallest singular value of the Hankel matrix introduced in Willems' lemma. Our experimental results show that the proposed solution tracks the set point variations.

2 Main Result

2.1 Data-driven MPC

Consider an nonlinear dynamic model of the form

$$x_{k+1} = f(x_k) + Bu_k; \qquad y_k = g(x_k) + Du_k \tag{1}$$

with $u_k \in \mathbb{R}^m$, $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^p$. Data-driven MPC is an MPC scheme, where the prediction model is obtained from past data. In the linear [4, 1] and affine [3] cases, it is based on the so-called Willem's fundamental lemma which we briefly recall. Let $u^d = (u_0^T, \dots, u_N^T)^T$, $y^d = (y_0^T, \dots, y_N^T)^T$ be a series of inputs-outputs of the system. Under some excitation conditions, *any* other trajectory of some length $L < N$ can be obtained from (u^d, y^d). In turn, this result is used to derive a data-driven MPC scheme to locally control nonlinear systems which have been linearized around an operating point [3].

This produces a system's model injected in a control scheme similar to the model-based one [2], with a quadratic cost function and terminal constraints for stability.

2.2 Nonlinear systems need updated datasets

Let (u^d, y^d) be the dataset used for DD MPC. Since it serves as a model of the system, it must satisfy two criteria: the prediction must be accurate, and it must ensure the controllability of the system by spanning all feasible control sequences.

An easy strategy would be to form u^d with the last N samples at a given time t: $u^d = (u_{t-N}^T, \dots, u_{t-1}^T)^T$. However, the risk is to reduce the controllability of the system if these samples do not carry enough excitation, especially in a steady state. This is why it was proposed to set up an initialization phase during which an exciting enough input sequence u^d is injected and uses the corresponding outputs y^d to form the dataset. For linear or affine systems, this dataset will ensure accuracy and controllability for an arbitrary long time [4, 1].

In the nonlinear case, a similar heuristic was proposed [3]: the dataset is frozen once a steady state is reached to ensure accuracy around this operating point. However, this only works if the system has a unique setpoint. No guarantees can be given for a reference which would change over time, hence the need for refined heuristics.

Proposed heuristics Let (u^d, y^d) be the current dataset, and $(\widetilde{u}^d, \widetilde{y}^d)$ be the one formed from the last N steps. The question is: when do we replace (u^d, y^d) with $(\widetilde{u}^d, \widetilde{y}^d)$? Our rationale is as follows: a dataset update will be needed if the reference changes, and this change should trigger enough excitation. Mathematically speaking, Willem's lemma is based on a Hankel matrix built from u^d, which must be full-rank. We propose to track this condition over time through the smallest singular value of this matrix $\sigma_{min}(u^d)$. Since in practice the presence of noise will almost always induce $\sigma_{min}(\widetilde{u}^d) > 0$, we propose the following heuristics: If $\sigma_{min}(\widetilde{u}^d) > \sigma$ for a given threshold σ, then $(\widetilde{u}^d, \widetilde{y}^d)$ replaces (u^d, y^d) as the new dataset.

The threshold σ has to be small enough to allow for frequent updates, but also large enough so that the dataset will ensure the controllability of the system.

3 Experimental results

We compare the data-driven MPC scheme from [3] and the one using the proposed update heuristics on a heat blower shown in Figure 1. The system is equipped with a blower fan and heating resistance. It is instrumented with a thermocouple and a flow meter.

In the experiment, the reference is set to change twice, one time for each output but at different times. The results obtained with both heuristics are shown in Figure 2.

The original heuristic [3] keeps updating the dataset with the last N samples until the system reaches a steady state at the first reference point (around $t = 200s$). When the reference

FIG. 1: Instrumented heat blower. The actuators (blower fan and heating resistance) are situated on the left of the picture, while the flow meter and thermocouple are on the right.

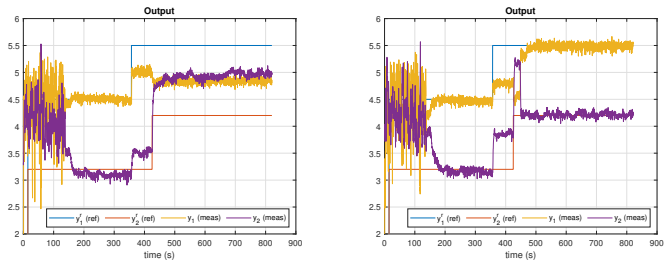

FIG. 2: Heat blower experiment results. The flow meter reference (resp. measurement) is in blue (resp. yellow) and the thermocouple reference (resp. measurement) in red (resp. purple). Left: Original heuristics [3]. Right: Proposed heuristics.

changes (at $t = 350s$ and $t = 415s$), strong static errors appear, which account for the loss of accuracy of the dataset. This means that the dataset made of inputs from before $t = 200s$ become irrelevant over time.

The proposed heuristic shows similar behavior for the first reference point and after the thermocouple reference changes ($t = 350s$), with a comparable static error. However, when the flow meter reference changes, we see that the controller manages to recover and ensure a null static error. Thus, the first change of reference does not induce an exciting enough signal to trigger the update, but the second one does. This validates that, under some conditions on the tracked reference, the proposed heuristic for DD MPC allows to keep an up-to-date dataset guaranteeing both the accuracy and the controllability of the system.

4 Conclusion

We highlight some challenges of the current DD MPC methods, regarding the quality of the data they use. A heuristic was proposed to overcome it while ensuring controllability and it was validated experimentally.

References

[1] Julian Berberich, Johannes Köhler, Matthias A Müller, and Frank Allgöwer. Data-driven model predictive control with stability and robustness guarantees. *IEEE Transactions on Automatic Control*, 66(4):1702–1717, 2020.

[2] Julian Berberich, Johannes Köhler, Matthias A Müller, and Frank Allgöwer. Linear tracking MPC for nonlinear systems—Part I: The model-based case. *IEEE Transactions on Automatic Control*, 67(9):4390–4405, 2022.

[3] Julian Berberich, Johannes Köhler, Matthias A Müller, and Frank Allgöwer. Linear tracking MPC for nonlinear systems—Part II: The data-driven case. *IEEE Transactions on Automatic Control*, 67(9):4406–4421, 2022.

[4] Jan C Willems, Paolo Rapisarda, Ivan Markovsky, and Bart LM De Moor. A note on persistency of excitation. *Systems & Control Letters*, 54(4):325–329, 2005.