Abstract

The performance of model predictive control (MPC) largely depends on the accuracy of the prediction model and of the constraints the system is subject to. In several applications, obtaining an accurate knowledge of these elements might be expensive in terms of money and resources, if at all possible. Therefore, by starting from an initial guess, we investigate how their accuracy can be enhanced by combining MPC with active learning approaches. Even though learning-based MPC approaches are appealing due to the potential performance improvement, the inclusion of online learning can also lead to instability, constraint violations, arbitrary performance deterioration, as well as lack of recursive feasibility of the considered MPC scheme. In this thesis, we develop novel learning-based MPC frameworks that actively incentivize learning of the underlying system dynamics or of the constraints, while ensuring recursive feasibility, constraint satisfaction, and performance bounds for the closed loop.

In the first part of the thesis, we focus on the case of inaccurate models, and analyze learning-based MPC schemes that include, in addition to the primary cost, a learning cost that aims at generating informative data by inducing excitation in the system. First, we propose a nonlinear MPC framework that ensures desired performance bounds for the resulting closed loop, which can be intuitively tuned compared to existing MPC designs. The scheme is recursively feasible, can deal with any learning cost function, and is applicable to a general class of nonlinear MPC design procedures. In the second approach, we focus on the more specific setup of linear systems subject to uncertain parameters and noisy output measurements. The considered learning cost function is based on the reformulation of the stochastic dual control problem, and the proposed scheme is recursively feasible, ensures closed-loop constraint satisfaction, and input-to-state stability.

However, even though the previously mentioned schemes can incorporate a learning cost, due to mismatch between the predicted trajectory and the closed-loop one, there are no guarantees that the desired learning phase occurs in closed-loop operations. Therefore, we propose an MPC framework that is able to guarantee closed-loop learning of the controlled system, while ensuring constraint satisfaction and suitable performance bounds.

Finally, motivated by practical applications, in the last part of the thesis, we investigate the scenario where the system is known but evolves in a partially unknown environment. In such a setup, exploration of unknown areas can lead to a potential performance improvement, and therefore it is incentivized. In order to safely explore the surrounding environment with the goal to learn the constraints, we propose a learning-based MPC scheme that incentivizes safe exploration if and only if this might yield to a performance improvement. The scheme ensures constraint satisfaction and convergence to the optimal steady-state, even in case of local minima appearing due to potentially non-convex constraints.

The applicability and practical benefits of all the proposed MPC approaches are demonstrated by means of numerical examples.

Deutsche Kurzzusammenfassung

Die Leistung der modellprädiktiven Regelung (MPC) hängt weitgehend von der Genauigkeit des Vorhersagemodells und den Systembeschränkungen ab. Bei vielen Anwendungen kann eine genaue Bestimmung dieser, wenn überhaupt möglich, sehr aufwendig und kostspielig sein. Daher untersuchen wir, wie die Genauigkeit ausgehend von einer anfänglichen Schätzung, durch die Kombination von MPC mit aktiven Lernansätzen verbessert werden kann. Obwohl lernbasierte MPC-Ansätze aufgrund der potenziellen Leistungsverbesserung attraktiv sind, kann die Einbeziehung des Online-Lernens auch zu Instabilität, Verletzungen der Beschränkungen, Leistungsverschlechterung sowie Verlust der rekursiven Lösbarkeit des MPC-Problems führen. In dieser Arbeit entwickeln wir neuartige lernbasierte MPC-Verfahren, die aktive Anreize für das Lernen der zugrunde liegenden Systemdynamik oder der Beschränkungen schaffen und gleichzeitig die rekursive Lösbarkeit, die Erfüllung der Beschränkungen und die Leistungsvorgaben für den geschlossenen Regelkreis sicherstellen.

Im ersten Teil der Arbeit konzentrieren wir uns auf den Fall eines ungenauen Modells und analysieren lernbasierte MPC-Konzepte, die zusätzlich zu den primären Kosten auch Lernkosten beinhalten, welche darauf abzielen, informative Daten zu generieren, indem sie eine Anregung im System induzieren. Zunächst schlagen wir ein nichtlineares MPC-Verfahren vor, das die gewünschten Leistungsvorgaben für den resultierenden Regelkreis sicherstellt, die im Vergleich zu bestehenden MPC-Designs intuitiv eingestellt werden können. Dieses Verfahren ist rekursiv lösbar, kann mit beliebigen Lernkostenfunktionen umgehen und ist auf eine allgemeine Klasse von nichtlinearen MPC-Designverfahren anwendbar. In einem zweiten Ansatz konzentrieren wir uns auf den spezifischeren Fall linearer Systeme mit unsicheren Parametern und verrauschten Ausgangsmessungen. Die betrachtete Lernkostenfunktion basiert auf der Neuformulierung des stochastischen dualen Steuerungsproblems, und das vorgeschlagene Schema ist rekursiv lösbar, gewährleistet die Erfüllung der Beschränkungen im geschlossenen Regelkreis und die Stabilität zwischen Eingang und Zustand.

Auch wenn die zuvor erwähnten Verfahren bereits Lernkosten berücksichtigen, gibt es aufgrund der Diskrepanz zwischen der vorhergesagten Trajektorie und derjenigen im geschlossenen Regelkreis keine Garantie dafür, dass die gewünschte Lernphase im geschlossenen Regelkreis stattfindet. Daher schlagen wir ein MPC-Verfahren vor, das das Lernen des geregelten Systems im geschlossenen Regelkreis garantiert und dabei zugleich die Erfüllung von Beschränkungen und geeigneter Leistungsvorgaben sicherstellt.

Schließlich, motiviert durch praktische Anwendungen, untersuchen wir im letzten Teil der Arbeit den Fall, dass das System bekannt ist, sich aber in einer teilweise unbekannten Umgebung befindet. In einem solchen Szenario kann die Erkundung unbekannter Bereiche zu einer potenziellen Leistungsverbesserung führen und wird deshalb belohnt. Um Beschränkungen zu lernen, muss die Umgebung sicher erkundet werden, dazu schlagen wir ein lernbasiertes MPC-Verfahren vor, das eine sichere Erkundung belohnt, genau dann, wenn diese zu einer Leistungsverbesserung führen kann. Das Verfahren gewährleistet die Erfüllung der Beschränkungen und Konvergenz zum optimalen stationären Zustand trotz der durch die potenziell nicht-konvexen Randbedingungen verursachten lokalen Minima.

Die Anwendbarkeit und die praktischen Vorteile aller vorgeschlagenen MPC-Ansätze werden anhand von numerischen Beispielen demonstriert.

Chapter 1.

Introduction

1.1. Motivation

In several applications, the desired goal can be described by the solution of an optimal control problem (OCP) with infinite horizon. However, except for special cases, solving an infinite horizon control problem is not computationally tractable. Among the different approaches proposed in the literature to circumvent such a shortcoming, a highly successful method is model predictive control (MPC), also referred to as receding horizon control. MPC approximates the solution of the computationally intractable infinite horizon OCP by repeatedly solving an appropriately designed finite horizon optimal control problem. The popularity of MPC is largely due to its applicability to general nonlinear systems with multiple-input multiple-output, while taking into account state and input constraints.

A detailed description of MPC can be found in Grüne and Pannek 2017; Rawlings, Mayne, et al. 2017a. Here, we describe its basic working principle: at each time instant the current state of the system is measured, and the related finite-horizon OCP over the finite horizon is solved. The OCP optimizes a user-defined cost function, and its solution are the optimal open loop state and input trajectories over the chosen prediction horizon, which are based on the given model of the system. The first predicted optimal control input is then applied to the system for the current time

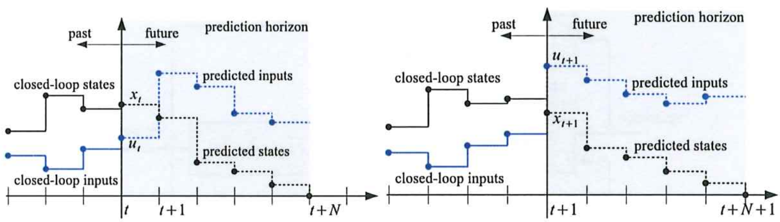

Figure 1.1.: Basic concept of MPC

instant, and the optimization is repeated at the next time instant based one the new measured state. Even though the optimization problem only predicts the open-loop evolution of the system, a feedback-loop is indirectly introduced by re-evaluating the state of the system and solving the optimization problem at each time instant. The steps described above are also summarized in Figure 1.1.

Since its first appearance in academia in the works Dreyfus 1965; Propoi 1963, MPC received continuous and constant attention not only on its theoretical aspects, but also from researcher with an application-oriented point of view. Most of the MPC literature focuses on the case where the control goal is to stabilize the origin of the system. Such approaches are therefore named as "stabilizing MPC", and their theoretical properties are closed-loop stability for the case of nominal MPC, i.e., when the system is exactly known, and robustness against external disturbances and model uncertainty for robust MPC approaches, which instead deal with perturbed systems.

Even though the presence of disturbances and model uncertainty is the key motivation of feedback control, in MPC they are only treated indirectly, i.e., by repeatedly solving the optimal control problem based on a new state measurement. During the last few decades, several approaches able to incorporate disturbance and uncertainty into the MPC design received a significant attention due to their ability to bound the evolution of the real system based on the open loop of the nominal model. Particularly, the case of bounded disturbances, considered in the field of robust MPC, has been deeply investigate for the case of linear system, providing not only accurate but also computationally tractable results, Chisci et al. 2001; Mayne, Seron, et al. 2005, while existing solutions for the nonlinear case aim at finding a trade-off between conservatism and computational tractability, Houska and Villanueva 2019; Köhler, Soloperto, et al. 2020; Limon, Alvarado, et al. 2010. On the contrary, despite the intense ongoing research, the case of nonlinear systems subject to unbounded disturbances, modeled by random variables, is less developed due to the its theoretically challenging aspects and computationally expensive nature, Farina et al. 2016; Mesbah 2016. Even though closed loop constraint satisfaction cannot be achieved due to the presence of unbounded disturbances, stochastic MPC schemes are nevertheless able to provide formal guarantees for the simpler case of linear systems subject to additive Gaussian disturbances, Arcari, Iannelli, et al. 2022; Cannon et al. 2009; Hewing and

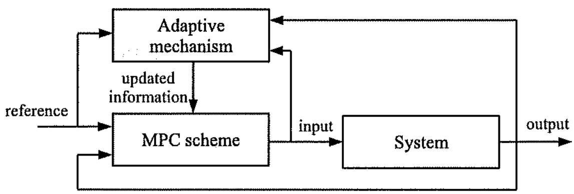

Figure 1.2.: Block scheme of adaptive MPC

Zeilinger 2018; Lorenzen, Dabbene, et al. 2016. However, any more elaborated setup, e.g., nonlinear systems subject to Gaussian disturbances, is still an open point. The main objective of robust and stochastic MPC schemes is to ensure safety by enforcing, up to a certain degree, closed-loop constraint satisfaction. Such a condition can be achieved robustly, i.e., for all the possible disturbance realization for the case of robust MPC, or only with a certain probability in stochastic MPC.

The performance of MPC depends largely on whether the chosen stage cost, prediction model, and constraints reflect the actual goal we want to achieve. In several applications, obtaining an accurate model of the controlled system might be expensive in terms of money and resources, if at all possible. Thus, starting from an initial model guess, online model refinement in MPC, e.g., using machine learning approaches, or (robust) adaptive methods is an active research topic, see Figure 1.2. The online updated model can then be used for prediction as an exact representation of the system, in a so-called certainty equivalence (CE) fashion. However, such adaptive control schemes have the only goal to steer the system towards a user defined setpoint, while estimation of the underlying system is only done as a side effect of the control action itself. The disadvantages of such *passive learning* approaches is the lack of informative data, which is for example especially relevant when the system operates at steady states, and can reflect in slow learning Mesbah 2018. On the other hand, *active learning* approaches explicitly aim at improving the accuracy of the model while simultaneously optimizing a user-defined cost function, e.g., steering the system towards the desired setpoint. Active learning in MPC can be achieved by introducing a learning cost function that has the goal to induce excitation in the system, and, when possible, by explicitly predicting how the nominal model might improve over the prediction horizon. Due to their dual objective, i.e., learning and stabilizing, such approaches are referred with dual adaptive MPC schemes.

In addition to a model adaption, the performance of an MPC scheme can be further enhanced by improving the accuracy of the constraint set, which might be time-varying, or partially unknown. For example, in applications related to autonomous driving, the constraint set might be online adapted due to the presence of moving obstacles, as well as the changing environment resulting from the fact that the vehicle moves. In order to improve the performance, exploration of unknown areas might be necessary and, therefore, actively incentivized by the MPC scheme.

The two discussed problems setups are contained in the broad field of learning-based MPC. Even though combining learning with MPC can be beneficial, it can also lead the system to instability, constraint violations, as well as lack of recursive feasibility of the considered MPC scheme. Therefore, the goal of learning-based MPC schemes is to enhance the performance of the system by improving the accuracy of the designed elements, while possibly providing theoretical guarantees.