



HAL
open science

Data Science and Model Predictive Control: A Survey of Recent Advances on Data-Driven MPC Algorithms

Marcelo Menezes Morato, Monica Spinola Felix

► **To cite this version:**

Marcelo Menezes Morato, Monica Spinola Felix. Data Science and Model Predictive Control: A Survey of Recent Advances on Data-Driven MPC Algorithms. 2024. hal-04747340

HAL Id: hal-04747340

<https://hal.science/hal-04747340v1>

Preprint submitted on 23 Oct 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Data Science and Model Predictive Control: A Survey of Recent Advances on Data-Driven MPC Algorithms

Marcelo M. Morato^{a,b}, Monica S. Felix^b

^a*Departamento de Automação e Sistemas, Universidade Federal de Santa Catarina, Brazil.*

^b*Univ. Grenoble Alpes, CNRS, Grenoble INP[†], GIPSA-lab, 38000 Grenoble, France.*

[†]*Institute of Engineering Univ. Grenoble Alpes.
(marcelo.menezes@ufsc.br)*

Abstract

Model Predictive Control (MPC) is an established control framework, based on the solution of an optimisation problem to determine the (optimal) control action at each discrete-time sample. Accordingly, major theoretical advances have been provided in the literature, such as closed-loop stability and recursive feasibility certificates, for the most diverse kinds of processes descriptions. Nevertheless, identifying *good*, trustworthy models for complex systems is a task heavily affected by uncertainties. As of this, developing MPC algorithms **directly from data** has recently received a considerable amount of attention over the last couple of years. In this work, we review the available **data-based MPC** formulations, which range from reinforcement **learning** schemes, **adaptive** controllers, and novel solutions based on behavioural theory and **trajectory** representations. In particular, we examine the recent research body on this topic, highlighting the main features and capabilities of available algorithms, while also discussing the fundamental connections among approaches and, comparatively, their advantages and limitations.

Keywords: Model Predictive Control, Data-driven control, Reinforcement Learning, Trajectory representation, Adaptive control.

1. Introduction

Dating from the original algorithms proposed from the process industry in the 80's, e.g. [1, 2], Model Predictive Control (MPC) has since become a widely used control technique for the regulation of constrained systems [3]. Over the last decades, a considerable amount of research has been devoted to the study of MPC algorithms, considering different system models and settings, e.g. [4, 5, 6, 7]. Corresponding theoretical certificates for closed-loop stability (and recursive feasibility of the (recurrent) optimisation problem) have been established since the seminal results provided by Mayne et al. [8] - which recently have been extended to broader settings by means of dissipativity theory in [9].

The MPC framework is set in discrete-time: at each sampling instant, the (optimal) control action is generated through the solution of a constrained optimisation problem, which embeds the performance objectives along a future horizon window, as well as the considered constraints. The general form of this optimisation, at each discrete time sample k , is given, generically, by:

$$\begin{aligned} \min_{U_k} \quad & \overbrace{\left(\sum_{j=0}^{N_p-1} \ell(x(k+j|k), u(k+j|k)) \right)}^{J(x(k), U_k)} + V(x(k+N_p|k)), \quad (1) \\ \text{subject to: } \quad & x(k+j+1|k) = f(x(k+j|k), u(k+j|k)), \quad \text{for all } j \in \mathbb{N}_{[0, N_p-1]}, \\ & x(k+j|k) \in \mathcal{X}, \quad \text{for all } j \in \mathbb{N}_{[1, N_p]}, \\ & u(k+j-1|k) \in \mathcal{U}, \quad \text{for all } j \in \mathbb{N}_{[1, N_p]}, \\ & x(k+N_p|k) \in \mathbf{X}_t, \end{aligned}$$

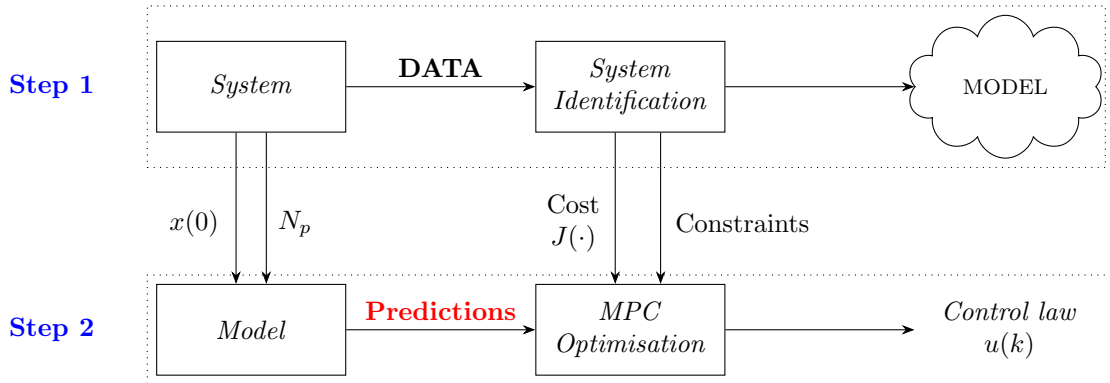


Figure 1: Traditional model-based MPC applications.

where x and u denote, respectively, the (predicted) process state and input variables, while

$$U_k := [u^T(k|k) \quad u^T(k+1|k) \quad \dots \quad u^T(k+N_p-2|k) \quad u^T(k+N_p-1|k)]^T$$

represents the optimisation¹ decision variable, i.e. the sequence of control actions along the prediction window N_p . From the optimal solution U_k^* , the first entry $u^*(k|k)$ is applied to the system.

It is indisputable that MPC has great theoretical and practical value - yet, it requires a reliable **process model** for a correct operation, as its name indicates. In the optimisation presented in Eq. (1), this model is marked by the generic function $f(\cdot, \cdot)$, which gives the state-space transition between samples. In the simpler setting, of systems with sufficient measurements and a reduced number of involved variables, the task of obtaining an accurate model (i.e. *system identification*) is standard and widely documented [10]. Nevertheless, in the case of system with complex dynamics and inaccessible variables, obtaining a trustworthy model is costly and ponderous task, affected by uncertainties and disturbances, as warns De Persis and Tesi [11], Bisoffi et al. [12], Steentjes et al. [13].

Over the last few years, developing control laws **directly from data**, measured from the controlled process, has received significant attention. We highlight that the notion of *data-driven* control - i.e. synthesis direct from data, skipping the system identification step - has indeed been under the spotlight: Figure 2 shows the *interest*² for the word “*data-driven*” in the Google online search engine along the last fifteen years - since 2017, the interest has been consistently growing.

In order to further elucidate the differences between traditional model-based MPC schemes and data-based formulations, we provide a schematic representation of the model-based schemes in Fig. 1: during an offline phase (Step 1), system data is collected and an identification procedure is applied in order to generate a model; then, during the online implementation (Step 2), the model is used to span process predictions along a horizon window of N_p steps, given an initial condition $x(0)$. By means of an optimisation, e.g. Eq. (1), the control law is generated. In data-driven MPC formulations, one seeks to eliminate Step 1, make it simpler or combine it together to the online phase (Step 2).

Notably, within the control community, the topic of data-driven control has been assessed specially in terms of adaptive methods, e.g. [14, 15], and reinforcement learning techniques, c.f. [16, 17]. More recently, several results have been presented using behavioural theory as a basis for data-driven control, c.f. the survey papers by Markovsky and colleagues [18, 19] and references therein.

¹In the MPC optimisation given in Eq. (1), \mathcal{X} and \mathcal{U} known sets used to represent the process constraints, while \mathbf{X}_t is a *terminal* set used for stability-related features. The optimisation cost $J(x(k), U_k)$ comprises a performance-related stage cost $\ell(\cdot, \cdot)$, summed along the future horizon, and a *terminal* cost $V(\cdot)$, related to the state prediction at the end of the horizon. We note that Eq. 1 is initialised with the current sampled state measurement $x(k)$; its solution is the minimiser U_k^* .

²The *interest* measure is related frequency of search queries online on the Google platform for a given term. In particular, an *interest* value of 100 represents the peak popularity of a given term, over a considered time window.

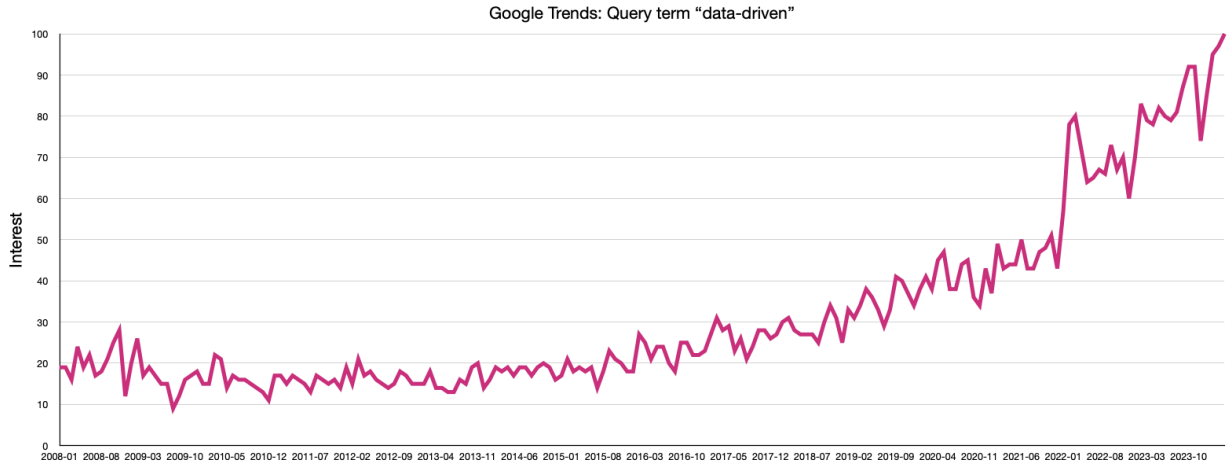


Figure 2: *Google Trends*: Online searches for *data-driven* in the Google engine along the last fifteen years.

Based on the aforementioned context, we focus henceforth on debating and detailing the available *data-*
 40 *based* MPC techniques³. In particular, by *data science MPC applications*, we specifically refer to:

- (i) MPC techniques purely based on measured process data, that *do no require* any kind of model description in the formulation, i.e. [20, 22, 23] - that is, data-driven MPC algorithms for which the optimisation is based on future predictions made directly from process data measured during the implementation; and
- 45 (ii) MPC schemes based on an initial process description (baseline model) that is iteratively refined, during the online implementation, in an adaptive fashion, based on the available process measurements collected over samples [24, 25, 26]. Here, we also consider schemes for which the process description is fixed, but the MPC optimisation cost function is adapted online, e.g. [15, 27, 28].

Despite the growing enthusiasm on such data-driven MPC schemes (as indicates Fig. 2), no study has
 50 formally **discussed** and compared the advantages and deficiencies of the available approaches, up to our best knowledge. Accordingly, the main contribution of this work is to survey and review the current of body of research on the topic, indicating fundamental connections between the available algorithms. In particular, we examine algorithms that are categorised along the three following branches, as indicates Fig. 3:

- 55 1. Adaptive schemes (Sec. 2), based on the online adaptation of the MPC problem, such as the techniques by Lorenzen et al. [15] and Tanaskovic et al. [27];
2. Methods based on reinforcement learning results (Sec. 3), as the formulations by Berkenkamp et al. [29] and Zanon and Gros [25];
3. Techniques based on behavioural theory and trajectory features (Sec. 4), e.g. [20, 30, 21], which exploit the persistency of excitation condition and Willem's Fundamental Lemma, i.e. [31].

60 **Remark 1.** With regard to scope of this survey, we emphasise that we consider herein, primarily, works which include theoretical *certificates* [32], that is, data-driven MPC schemes that are formally demonstrated

³In some references, MPC schemes based on data are referred to as *data-enabled* predictive control (DeePC, e.g. [20]) or *DPC* (used for data-based predictive control, c.f. [21]). For notation consistency, we use either the term *data-driven* MPC or *data-based* MPC, throughout this article.

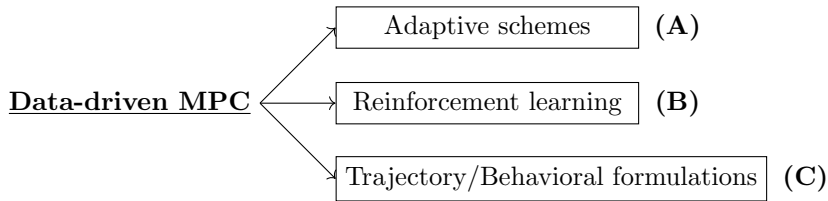


Figure 3: Surveyed categories of algorithms.

to stabilise the controlled process in closed-loop, e.g. [33, 26]. In many cases, these certificates also relate to the optimisation itself, which should be shown to be recursively feasible for a coherent implementation of the control strategy. **Moreover, we stress that, although** some authors refer to MPC schemes based on neural networks, and variations, e.g. [34, 35, 36], as data-driven MPC methods, we do not classify them as so. We opt not to evaluate these algorithms in this survey due to the fact that an offline *training* phase is forcefully required (just as an identification step when using a traditional model-based design). Nevertheless, we do not exclude reinforcement learning techniques that use Neural Networks as the basis for *approximators* when estimating adaptive model parameters.

Before presenting further details with regard to these three categories, and taking into account the model-based MPC schematic in Fig. 1, we provide a comparative schematic to the data-driven case in Figure 4. **In the case of** categories (A) and (B), i.e. adaptive and learning-based formulations, a simple system identification procedure is performed in order to obtain a baseline model, whose predictions are adapted online with adaptation (learning) laws based on new process data. **Moreover, regarding** trajectory formulations, i.e. category (C), an initial data dictionary is used to construct, combined with new data, an online data-driven optimisation problem.

With regard to these three aforementioned categories, we include Figure 5 in order to show a corresponding co-relation network cluster⁴. In particular, the cluster network was generated in order to illustrate the relations between the papers that are surveyed along this work. Author keywords with at least two occurrences were selected to identify correlations among the bibliography under review⁵. As indicated by the cluster, category A (adaptive schemes) is tightly linked to optimal control and system identification theory, for which robust MPC algorithms are applied; category B (reinforcement learning algorithms) takes into account data models, online learning and adaptation, as well as stochastic inference tools; notwithstanding the prior, category C (trajectory/behavioural schemes) is related to linear system theory and corresponding data-driven system analysis, especially due to the intrinsic use of the persistency of excitation condition, as **detailed in the sequel**.

Remark 2. The Reader can find an extensive overview of data-driven approaches to classical control theory in [38]. Furthermore, a brief historical recap of data-based strategies and a discussion of the future landscape of (direct and indirect) data-driven control can be found in the recent opinion articles by F. Dörfler, e.g. [39, 40]. As indicated therein, we care to highlight that data and learning-based algorithms already have considerable scientific impact and continue to gain increasing attention within academia and the industry. Therefore, we indicate that the current survey comprises, as a highlight, qualitative analyses w.r.t. the

⁴This network visualisation was generated by VOSviewer - a widely used software tool for visualising and analysing networks, particularly in bibliometrics, often employed to examine citation networks, co-authorship networks, and other academic data structures.

⁵In the figure, items are depicted as circles with labels, for which the size of the circles is determined by the weight of the corresponding items. Items are color-coded according to the clusters to which they belong, and lines represent connections between them. The proximity of the items in the network indicates their level of relatedness; the closer two items are, the more closely related they are, often in terms of co-citation links. Lines between journals typically represent the strongest co-citation connections. **A cluster represents a group of related items, sometimes referred to as a “community” [37]. We care to highlight that certain keywords were excluded from this process, including “model predictive control”, “predictive control”, “data-driven control,” and “data-based control”,** because they were common across all the papers being analysed. Furthermore, terms like “linear parameter varying” and “model” were excluded due to their weak correlations with other keywords.

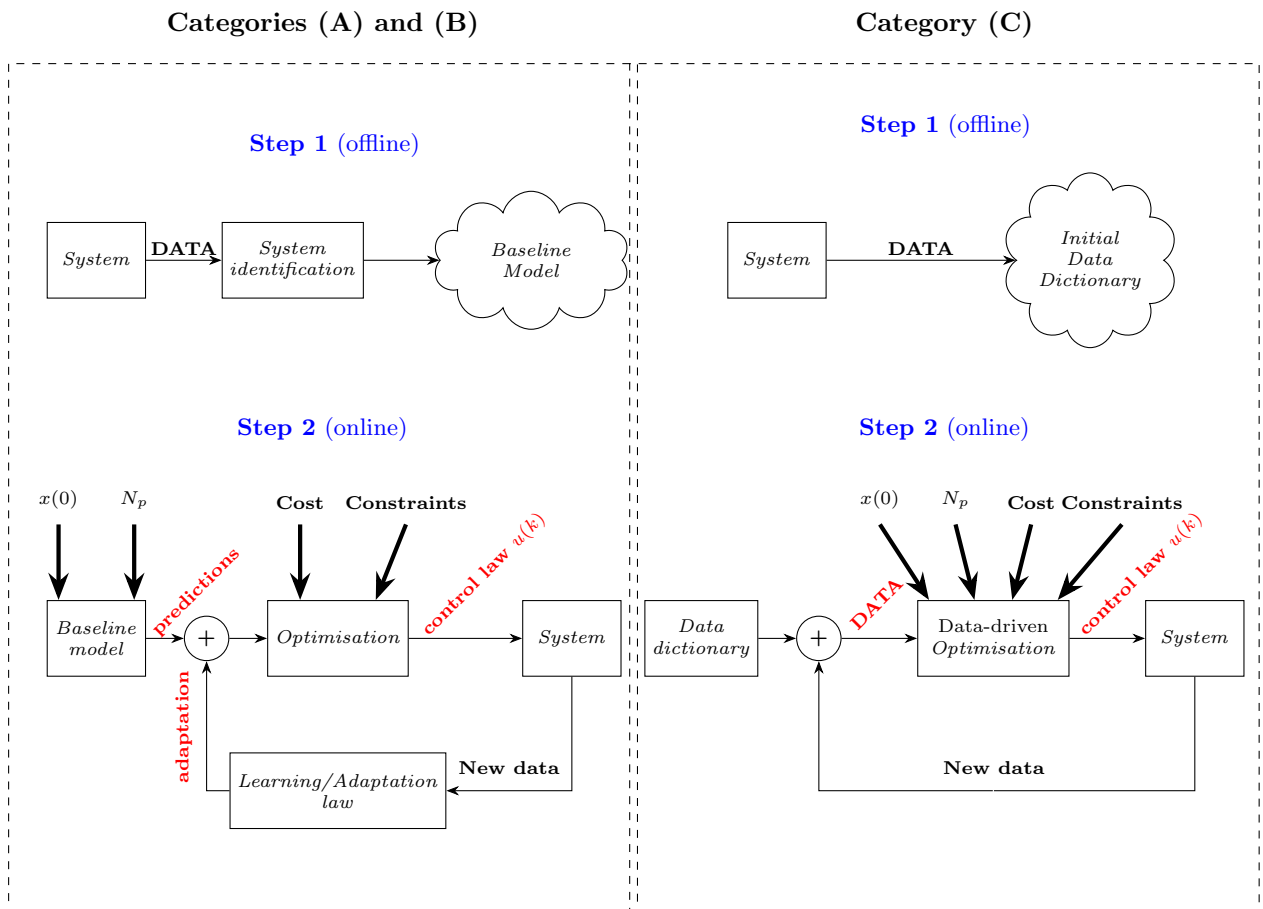


Figure 4: Data-driven MPC applications.

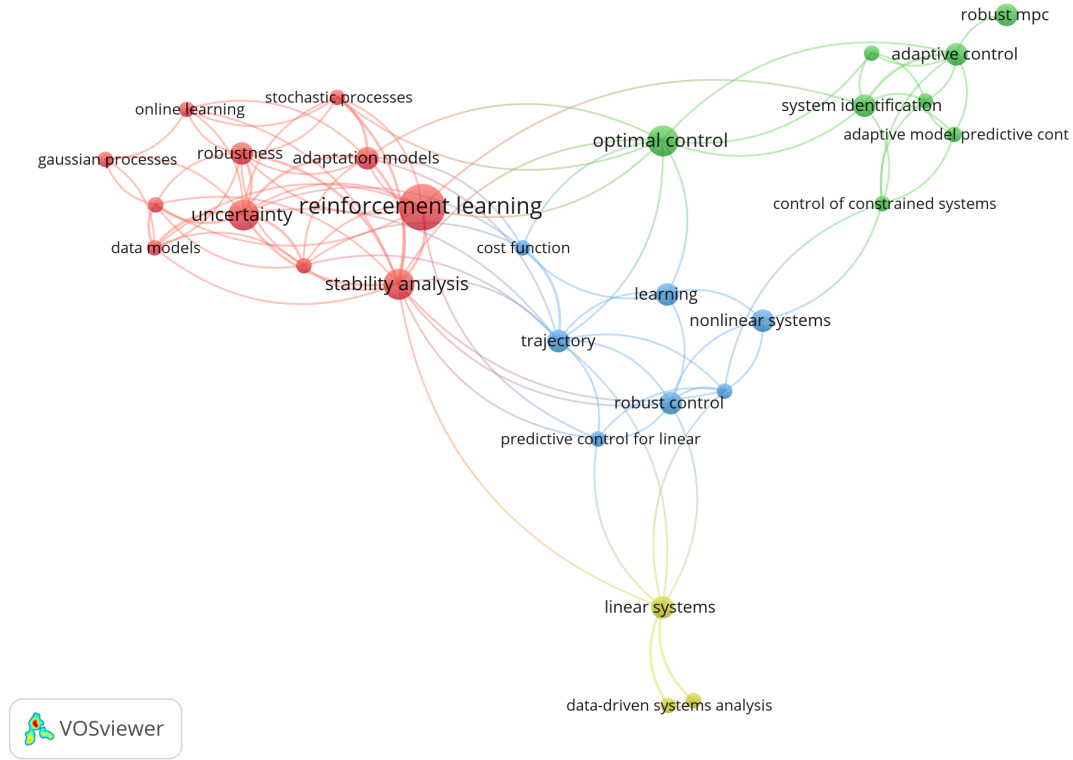


Figure 5: Network of co-citation of author keywords categorized by clusters.

data-driven MPC categories given in Fig.3; we also **discuss** advantages, limitations, and interconnections among the control approaches.

95 **Outline.** This paper is structured as follows: In Section 2, we discuss the available adaptive schemes, while machine learning solutions and trajectory-based algorithms are surveyed in Sections 3 and 4, respectively; we also elaborate on research gaps for further investigation on this field. General conclusions are drawn in Section 5.

Notation. The index set $\mathbb{N}_{[a,b]}$ represents $\{i \in \mathbb{N} | a \leq i \leq b\}$, with $0 \leq a \leq b$, and we use $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. The identity matrix of size j is denoted as I_j . The predicted value of a given variable $v(k)$ at time instant $k+i$, computed based on the information available at instant k , is denoted as $v(k+i|k)$. In particular, $v(k|k) = v(k)$. A class \mathcal{C}^1 function $f: \mathbb{R}^m \rightarrow \mathbb{R}$ is such that it is differentiable with continuous derivatives. $\|\cdot\|$ denotes the 2-norm. For a discrete-time signal $v: \mathbb{N} \rightarrow \mathbb{R}^{n_v}$, we denote $v(k) \in \mathbb{R}^{n_v}$ each of its entries and $\{v(k)\}_{k=0}^{N-1}$ the corresponding sequence of N data entries, or just v in short. We use
105 $\text{col}\{v\} := [v(0)^T \dots v(N-1)^T]^T$ to denote the column vectorisation, and $\text{diag}\{v\}$ as the block-diagonal matrix formed with $\text{col}\{v\}$. The Kronecker product is represented by \otimes ; the corresponding block-diagonal operator is denoted \circledast , implying that $(v \circledast I_\xi) = \text{diag}\{v(0) \otimes I_\xi \dots v(N-1) \otimes I_\xi\}$. For a sequence $\{v(k)\}_{k=0}^{N-1}$, we have the corresponding Hankel matrix, for a window of L entries, given by:

$$H_N(v) := \begin{bmatrix} v(0) & v(1) & \dots & v(N-L) \\ v(1) & v(2) & \dots & v(N-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ v(L-1) & v(L) & \dots & v(N-1) \end{bmatrix}.$$

Definition 1.1 (Persistent excitation [31]). A signal $\{u(k)\}_{k=0}^{L-1}$, with $u(k) \in \mathbb{R}^{n_u}, \forall k \geq 0$, is persistently exciting of order N if the rank of $H_N(u)$ greater or equal than $n_u N$.

2. Category A: Adaptive algorithms

The first set of methods that we consider as data-driven MPC schemes comprises the techniques based on online model adaptation (using “traditional” control and system identification arguments, as evidenced in the cluster provided in Fig. 5), such as the works [41], [42], and [43].

Synthetically, these algorithms are deployed from an initial model description - usually written in terms of time-varying parameters or an additive disturbance term, which is recursively re-identified, during the online implementation of the MPC. Typically, the optimisation is designed in order to impose certificates for the whole expected set of different models (which are implicitly mapped by the time-varying parameters or disturbances) and, thus, the implementation only requires the model to be updated at each sample.

In order to properly survey these adaptive data-driven MPC algorithms, we first give an overview of their intrinsic mechanism, with more details, then we review the available formulations (stochastic, scenario-based, and so forth), and, finally, we provide corresponding discussions and some perspectives. A corresponding [investigation](#) on the topic of adaptive (learning) data-driven MPC schemes is presented in [44], considering the problem of automated insulin delivery in diabetes.

2.1. Outline of the mechanism

For exposition clarity, we consider a parametric adaptive formulation⁶ as [discussed](#) by Fagiano et al. [48], and Lorenzen et al. [15]. Thus, consider the following parameter-varying state-space model:

$$x(k+1) = A(\theta(k))x(k) + B(\theta(k))u(k), \quad (2)$$

where $x \in \mathbb{R}^{n_x}$ are the system states, $u \in \mathbb{R}^{n_u}$ is the control input and $\theta \in \mathbb{R}^{n_\theta}$ are the unknown uncertainty parameter, for which an initial value θ_0 is available. In general, papers assume that the model matrices are affine on the adaptation parameter, that is: $(A(\theta), B(\theta)) := (A_0, B_0) + \sum_{i=1}^{n_\theta} (A_i, B_i) \theta_i$. Finally, we consider that the space of the adaptive parameter is known, that is $\theta(k) \in \Theta, \forall k \geq 0$, being Θ a known (generally convex) set.

Remark 3. In some formulations, e.g. [15], the parameter set Θ is also assumed time-varying and adapted online with respect to measured data. That is, a non-falsified smaller set is considered in the identification procedure; this set tends to shrink over time, as the system parameter identification becomes stationary.

During the online implementation, at each discrete-time instant $k \geq 0$, not only an MPC optimisation in similar form of Eq. (1) is solved, but also a complementary recursive data-driven model adaptation step. Accordingly, consider the collected state-input data set from previous samples: $\{x(i)\}_{i=0}^{k-1}$ and $\{u(i)\}_{i=0}^{k-1}$, and the corresponding filter matrix $D(x(k-1), u(k-1))$. Then, [Lorenzen et al. \[49\]](#) propose an adaptation parameter estimate based on the new available state measurement data $x(k)$ - [which can be generated by the means of a recursive law in the form of:](#)

$$\theta(k) = \theta(k-1) + q_\mu D(x(k), u(k-1))^T (x(k) - A(\theta(k-1))x(k-1) - B(\theta(k-1))u(k-1)), \quad (3)$$

being q_μ is a forgetting/update factor of appropriate dimension.

⁶Many alternatives formulations are used in adaptive MPC schemes, including: nonlinear state-space models subject to time-varying additive disturbances, which imply the adaptation, e.g. [24, 43]; step-response input-output models with time-varying transition matrices, c.f. [41, 45]; linear state-space models with output disturbances, e.g. [46]; and so forth. We highlight the discussion from [47]: in the case of parametric models, seen in most control engineering applications, the adaptation is set in terms of a finite number of parameters (that often have no physical interpretation), which are often linearly tied. In these approaches, uncertainty regarding the model structure [is not considered](#). In the case of non-parametric models, the available data is used to infer the current state and local data (i.e. smoothing). For these models, the uncertainty can be included with respect to local data density and model complexity, which can be helpful for sparsely sampled transient dynamics.

The adaptation parameter estimation rule in Eq. (3) can be synthesised using different metrics and syntheses, such as set-membership rules, c.f. [50, 51], least-square criteria, c.f. [52, 53], and so forth. These criteria relate both to the adaptive parameter set Θ and the structure of the considered system model (e.g. the affine dependency).

Based on the adaptation parameter estimate $\theta(k)$, the MPC problem is solved based on predictions spanned via Eq. (2). In order to provide performance and stability certificates, the MPC optimisation is forcefully designed to be robust with respect to the parameter variation over samples, mapped implicitly through Eq. (3), and also to parameter space Θ . Many of the available formulations rely on using robust terminal ingredients, as done in [41] and [54], or robust optimisation solutions, such as tube MPC, c.f. [55, 51], and min-max MPC, see [48, 56].

2.2. Existing formulations

Regarding the different formulations to these data-driven, adaptive MPC methods, we consider four distinct categories, as detailed in the sequel: (i) schemes based on robust synthesis; (ii) approaches based on stochasticity of the adaptive parameters; and (iii) schemes that rely on the adaptation of the MPC optimisation itself.

The first class of adaptive MPC schemes that we survey, herein considering the data-driven context, is those that are *robust* with respect to the adaptation parameter θ . That is, algorithms that are able to ensure the satisfaction of both state and input constraints, despite the variation of the process dynamics over time.

In the robust setting, the adaptive controller is synthesised to tolerate the dynamics imposed by a set of models, implicitly mapped by the (bounded) time-varying uncertainty. In these data-driven, adaptive MPC formulations with robust design, theoretical features from robust MPC methods (such as tubes, e.g. [14], min-max. arguments, e.g. [24, 56], results based on set-membership, e.g. [54], and so forth) are coupled with online system identification tools in order to guarantee the safe learning of the process model together with performance certificates.

In recent literature, one can find multiple references which propose the design of robust MPC schemes for adaptive models based on data. Next, we detail the main elements and some pertinent works.

First, we can observe several formulations (e.g. [41, 49, 57]) for which the control law is expressed in terms of a nominal control action supplemented by an additional robust stabilisation bias. The fundamental idea resides in tuning the first part of the control law (the nominal MPC action) to stabilise the nominal (initial) process model, while the remainder of the control law is preoccupied with the adaptation of the model, thus being a data-driven signal. We detail how these references relate to each other:

- Tanaskovic et al. [41] propose a real-time set membership identification law in the form of Eq. (3) in order to qualify the adaptation parameter online, based on data, and, from this, selects an instantaneous candidate model that describes the system; finally, the robust control law is based on the minimisation problem in the form of Eq. (1), whose terminal set is real-time refined in terms of θ . The main advantage of the adaptive MPC in [41] is that the sets within which the true system trajectories lie is identified with precision; however, as this set grows in volume, the resulting control becomes more conservative.
- Lorenzen et al. [49] provide a similar approach, for which a recursive identification law is proposed, also based on set-related arguments. In this approach, the control is robustified using a an invariant tube, whose size is scaled online in terms of the adaptation parameter. W.r.t. the prior approach, the key advantage is that the conservativeness is reduced, being implied more strongly for larger-horizon predictions (i.e. the tube - and conservativeness - grows along the predictions).
- Following the same rationale as the prior, the methods coined by Adetola and Guay [24] and Zhang and Shi [57] formulate the robust MPC as a min-max problem that is able to ensure constraint satisfaction and input-to-state practical stability of the closed-loop. The main advantage of these methods is that the performance can be improved online by shrinking the size of the uncertain parameter set at each time step.

Similarly to the previous two schemes, an adaptive scheme based on min-max MPC is proposed in [56] for the fault-tolerant control of autonomous vehicles, considering a formulation based on a Linear Parameter Varying (LPV) model. The main difference here is that the proposed control law is not written as the above-mentioned sum of a nominal and an adaptive law, but rather as parameter-dependent law, which varies according to the adaptation parameter measured from process itself. In particular, the fault-related parameters adapt the LPV model online, and the controller is robust with respect to possible variations of these scheduling variables. The main feature is that the MPC is robustified by using parameter-dependent terminal ingredients, which are shown to provide better performances than when using classical quadratic ones.

Also in scope of robust MPC formulations for adaptive data-driven control, Saffi et al. [58] use general control contraction metrics which are adapted online with respect to estimated uncertainties using set-membership, in order to parameterise a robust homothetic tube MPC formulation. The main difference of this work w.r.t. the prior tube MPCs is how the tubes are derived based on contraction metrics able to account for the state- and input-dependency of the time-varying model uncertainties, according to new data available online.

Finally, in the so-called dual adaptive schemes, e.g. [46, 53], strong duality theory is used to recast set-membership parameter estimate equations *exactly* within the MPC optimisation. By doing so, these works are able to guarantee robust constraint satisfaction and recursive feasibility of the optimisation. Essentially, the main drawback is that the dual effect on the initial model estimates may impose considerable performance degradation, with a corresponding large variation in closed-loop control costs.

Next, we analyse other relevant data-driven adaptive MPC schemes based on theoretical foundations that take into account the stochastic nature of the controlled processes, e.g. [48, 43, 59]. The main difference from the prior is that these MPCs are tuned by incorporating stochastic knowledge from the uncertainty of the controlled system, such as probability distribution. Their main advantage is that, by doing so, the corresponding controllers are able to ensure performances, stability, and constraints satisfaction for the given stochastic range of the uncertainty. Then, the adaptation feature is embedded by harnessing the uncertainty estimate information online. Typically, the incorporation of the stochastic knowledge - inferred from data-enhances the resulting control w.r.t. the aforementioned adaptive robust MPC schemes. Next, we detail these main formulations with regard to this scope.

In [43], a so-called *cautious* adaptive MPC schemes is formulated: the controller considers a nominal system subject to an additive nonlinear uncertainty modelled as a Gaussian process, as well as chance constraints written with respect to the residual uncertainties given by the difference between a Gaussian distribution model and online-generated uncertainty estimates. The main advantage in the cautious adaptive MPC formulation is that (performance and safety) improvements enabled over nominal (robust MPC) schemes - as shown in the several simulation (and comparison) examples in [43]. Yet, the drawback is the requirement of a Gaussian-type model for the related uncertainty, which is not applicable for all kinds of systems.

By relieving this Gaussian-process requirement, Bujarbaruah et al. [59] propose a comparable framework, also using chance constraints to account for stochastic characteristics of the considered uncertainty. In their work, as new data becomes available during the implementation, the feasible parameter set is redefined, considering a set-membership approach. The method is somewhat in between set-based robust adaptive MPC schemes and the cautious adaptive MPC formulation - the drawback is that, if the uncertainty stochasticity is too significant, the resulting controller can be quite conservative, as when using [49].

More recently, Aboudonia and Lygeros [60] proposed an alternative adaptive MPC scheme based on set membership identification. The main difference w.r.t. the method by Bujarbaruah et al. [59] is that a rigid tube-based robust trajectory is computed, instead of the chance constraints formulation. The major difference is that the MPC (terminal) ingredients - which ensure stability - are adapted online, during an adaptation phase that takes the learnt uncertainty set into account. Results indicate that the approach is promising, especially in the case of interconnected systems, for which the average computational time of the resulting MPC becomes significantly reduced. We highlight that tube-based formulations, such as [58] (nominal) and [60] (stochastic) are often less conservative than set-based counterparts, given that the tubes typically scale in volume w.r.t. predictions made along the horizon.

Finally, as another sub-category of adaptive data-based MPCs, we refer to approaches that do not adapt the constraints of the MPC optimisation online, but rather its form in terms of the cost elements that are minimised, c.f. [61, 50, 62, 63]. Accordingly, these schemes make use of new data in order to re-select the tuning weights of the performance costs, according to the estimated level of uncertainty. **We detail the main elements of these approaches and how they related to the prior:**

- In the formulations proposed by Salvador et al. [61] and Collet et al. [28], the MPC frameworks are set to optimise the weights of a linearly combined cost function, which minimises not only control performance but also the variance of the prediction error. In synthesis, the online operation requires a separate quadratic optimisation to determine the optimal weights, taking into account new process data. Then, these weights are used in the MPC to generate the adaptive control law. The main drawback is the increased computational burden of the scheme, due to the two optimisation procedures that are solved - which complicates real-time applications of the methods. In comparison to the prior adaptive data-driven methods, the key novelty is that the prediction error decreases over time, which means that, as the controlled system dynamics settle, the predictions become more accurate. For systems which exhibit slowly-varying model changes, the methods are an appropriate solution, given that after some samples the predictions typically converge to the true system dynamics. For processes with significant time-varying uncertainties, the resulting performances may be poorer than adaptive stochastic and well-tuned (tube) robust formulations.
- Differently from the prior, the algorithm proposed in [63] is centered in adapting the MPC prediction matrices online, based on data. In particular, an estimation prediction procedure is used to adapt these matrices, based on integral action and a rate-based formulation. The main advantage w.r.t. [61, 28] is that only one optimisation problem is solved online; nevertheless, formal comparisons and experimental validations are lacking to concretely conclude how these schemes compare performance-wise.
- More recently, the controllers proposed in [50, 62] use (data-based) *active learning* to account for uncertainties in an adaptive manner. That is, the MPC costs, c.f. Eq. (1), $\ell(\cdot, \cdot)$ (stage cost) and $V(\cdot)$ (terminal cost), as well as the terminal set \mathbf{X}_f are learnt online in terms of a baseline model description and a corresponding data-driven model update rule⁷. These works, in practice, also include a performance cost to excite the system in such a way that (system) identification is made possible altogether with the control action. For instance, the learning costs can include, e.g., least-square arguments, Gaussian rules and enforce persistence of excitation. The fundamental highlight related to these scheme is how they can incorporate stochastic information, system identification and learning arguments in order to adapt the MPC costs, thus combining features from the previous surveyed algorithms. Yet, their main drawback is that potentially conflicting objectives (of model identification and control) may result in inappropriate performance results, which may include even deterioration, oscillatory behaviours and permanent steady-state error.

2.3. Overall discussion: Summary of Strengths and Limitations

In Table 1, we present a broad overview of the surveyed adaptive MPC methods, detailing the main features and characteristics of each work. We also point out the kinds of systems considered in each surveyed article, as well as the type of validation available (experimental or numerical). **Next**, with regard to these schemes, we offer some general perspectives on their advantages and drawbacks.

Considering the emphasis on data-driven MPCs, we highlight that the fundamental relation among all these adaptive schemes is that they require, for a start, a baseline process model - which may be phenomenological or locally identified. Then, the model is adapted online according to new measurements, in terms of an uncertainty term. During their implementation, the methods depend on the data-driven estimation of the uncertainty parameter - using set-membership tools, Least-Square arguments and so forth.

⁷We note that these allow for both MPC optimisation with and without terminal ingredients.

Therefore, if the estimates are not accurate, the performance of the resulting adaptive controller may be below expectations, i.e. worst than when using nominal control laws.

²⁹⁰ **In some cases**, when the initial model is unreasonably biased, the uncertainty that is to be estimated online belongs to a set which is too large with respect to state feasibility set, which conversely implies in excessively conservative control - this feature can be seen both in robust and stochastic formulations. **Furthermore**, for some systems, the synthesis of the control contraction sets, terminal ingredients, and robust tubes related to the uncertainty may be infeasible or non-existing, that is the formulation offers ²⁹⁵ sufficient conditions (in general, LMIs) that lead to empty sets.

An interesting perspective (which surely deserves further investigation) is to design these adaptive schemes by means of LPV models, which are able to very naturally embed the adaptation feature to the controller. Moreover, the MPC can be robustified using parameter-dependent terminal ingredients, thus tying together both the MPC prediction model and the optimisation arguments to the adaptation parameter ³⁰⁰ (that is determined online).

Table 1: Overview of surveyed adaptive MPC schemes.

Reference	Certificates	Robustness tools	Design arguments	Application
[41]	Input and output constraints satisfaction	-	System identification: Set-membership	Numerical simulation: Mass, spring and damper system
[49]	Closed-loop stability and recursive feasibility	Min-max optimisation, constraint tightening (tube)	System identification: Set-membership	Numerical simulation: Random system
[57]	Constraint satisfaction, recursive feasibility, closed-loop input-to-state practical stability	Min-max optimisation, constraint tightening (tube)	System identification: Set-membership	Numerical simulation
[24]	Closed-loop robust stability	Min-max optimisation	System identification: Set-membership	Numerical simulation: Chemical reactor system
[56]	Closed-loop stability, recursive feasibility	LPV parameter-dependent terminal ingredients, min-max optimisation	-	Numerical simulation: Realistic autonomous vehicle system

Continued on next page

Table 1: Overview of surveyed adaptive MPC schemes. (Continued)

Reference	Certificates	Robustness tools	Design arguments	Application
[58]	Closed-loop stability, constraints satisfaction, recursive feasibility	Control contraction metrics, constraints tightening (homothetic tube)	System identification	Numerical simulation: Planar quad-copter system
[46]	-	Dual formulation	System identification: Set-membership, Stochastic framework: Gaussian process, probabilistic constraints, expected covariance	Numerical simulation: FIR system
[53]	Recursive feasibility, robust constraints satisfaction	Strong duality, constraint tightening (homothetic tube)	System identification: Set-membership	Numerical simulation: Polytopic system
[43]	-	Constraint tightening (tube)	System identification, Stochastic framework: Gaussian uncertainty description	Experimental validation: Autonomous racing of remote controlled race car
[59]	Robust constraint satisfaction, recursive feasibility, closed-loop stability	-	System identification: Set-membership, feasible parameter set, Stochastic framework: Stochastic (chance) constraints	Numerical simulation

Continued on next page

Table 1: Overview of surveyed adaptive MPC schemes. (Continued)

Reference	Certificates	Robustness tools	Design arguments	Application
[61]	-	-	Stochastic framework: Variance of the model estimation error	Numerical simulation: Quadruple tank system
[28]	-	-	Stochastic framework: Parameter variance estimate	Numerical simulation: Horizontal axis wind turbine
[50]	Closed-loop finite-gain \mathcal{L}_2 stability, robust recursive feasibility, constraint satisfaction	Constraints tightening (tube), robust terminal ingredients	System identification: Least mean square estimates	Numerical simulation
[63]	-	-	System identification: Least mean square estimates	Experimental validation: DC-DC Buck converter
[62]	Safe performance bounds, recursive feasibility	Terminal ingredients	-	Numerical simulation: Mass, spring and damper system

305 3. Category B: Machine learning schemes

In the previous Section, we revisited relevant results on adaptive MPC formulations that can be understood, in our perspective, as data-driven schemes, given that process data is iteratively used to refine an initial model and improve the control performances. Indeed, we can find diverse approaches in recent literature (often linked to the field of machine learning) that address the issue of adaptive data-driven predictive control. As surveyed in the previous section, several of these works improve an initial process model to the data measured during the implementation, e.g. [64, 51], while others opt to re-select the optimisation cost [28] or tighten the system constraints [62].

Albeit these interesting schemes, we recap, next, another branch of existing formulations (as evidenced in Figures 3 and 5) with a similar conceptual idea, which have attracted a significant amount of attention in recent literature: the so-called *learning*-based MPCs, e.g. [65, 16, 66, 29, 67, 68, 69, 25, 70, 26, 71]. **The fundamental difference among the adaptive schemes is summarised** in Sec. 2 (category A) and these learning MPCs resides in the methods that are used to infer on the sampled uncertainty, given new data. The key similarities, anyhow, reside in the different synthesis approach for the MPC optimisation itself (i.e. using tubes, set-membership arguments, etc).

320 Gros and Zanon [72] offer solid theoretical foundation on why to use reinforcement learning to improve
 adaptive MPC schemes. In particular, these approaches, often explicitly referred to as *model-free*, c.f.
 [73], reside in combining the MPC framework with machine learning techniques in order to reduce the
 dependency of the controller in the an initial process model (as in adaptive MPC formulations), while
 325 exploiting the growing amount of data available (during the implementation) by means of machine learning
 and reinforcement learning tools.

We highlight that reinforcement learning is indeed a powerful theoretical framework to learn optimal
 control laws from data. Yet, standard reinforcement learning schemes are based on *exploration* - i.e. apply-
 ing all possible control actions to the system to observe *what happens* -, which is obviously not a feasible
 approach for any real system. As argued by Berkenkamp et al. [29], several recent papers have provided
 330 learning algorithm that explicitly consider safety - defined in terms of stability guarantees and constraints
 satisfaction. In particular, we focus, in this survey, on the works that combine MPC and reinforcement learn-
 ing frameworks with *guarantees*, i.e. the theoretical properties that enable stability and imply in a control
 action that is *explainable* and *safe* [25, 26], w.r.t the reinforcement learning task. That is, techniques which
 account for online model parameter updates (delivered by the learning tools, during the implementation,
 335 based on available data) that maintain the safety and stability of the MPC in closed loop.

Remark 4. As indicated by Zanon and Gros [25], the combination of *learning* and predictive control methods
 has been proposed in several papers as, such as [16, 74, 29, 75, 76, 77, 78]. Nevertheless, we emphasise that
 the papers by Zanon et al. [78], Gros et al. [17], Zanon and Gros [25] and Amos et al. [77] are the first
 to actually use a nonlinear MPC formulation as an *approximator* in reinforcement learning (RL). That is,
 340 the optimal control policy generated by the RL layer is indirectly obtained by the minimisation of a given
 function - both the control policy and the associated value function are, in these references, written in terms
 of a nonlinear MPC formulation. In many other references, these are obtained by approximators based on
 Neural Networks, c.f. [79, 80].

3.1. Outline of the mechanism

345 In broad terms⁸, reinforcement learning is a method originally conceived for solving Markov decision
 processes [81]. Instead of using state transition probabilities as in regular Markov approaches, RL techniques
 are based on measured (state) samples and observed costs (often named *rewards*) associated to each possible
 state. Recently, RL schemes have been successfully applied to wide range of practical applications, including
 playing board-games and robots learning to walk without supervision, c.f. [82, 83].

When RL schemes are used for control applications, the implementation consists in computing an optimal
 control policy for each given state, based on the minimisation of a *action-value* function. According to data,
 this function is adapted over time according the observed closed-loop behaviour. Even tough RL methods
 have been employed for a wide range of systems, demonstrating great potential, many standard theoretical
 properties - that are typically expected from any simple controller -, such as closed-loop stability and
 355 satisfaction of constraints are tough to demonstrate. In synthesis, as argued extensively by Garcia and
 Fernández [84], the RL-based methods which are *safe* (i.e. exhibit the theoretical guarantees as in regular
 control approaches) tend to require a form of model of the underlying uncertainty of the controlled system.

For safe RL schemes, the controlled system is not modelled by a state transition law as in Eq. (2), but
 rather by a conditional probability density for a given state x to commute to another successor state x^+ ,
 360 due to a control action u ; this probability density is usually given in the form of $\mathbb{P}[x^+ | x, a]$. Then, the
 deterministic optimal control policy is obtained using $u(k) = \kappa(x(k))$, where $\kappa^* := \arg \min_{\kappa} J(\kappa)$, being
 this cost given by the expected value of an infinite-horizon sum of a stage cost, much similar to what is
 done in standard MPC, that is: $\mathbb{E} \left[\sum_{k=0}^{+\infty} \gamma^k \ell(x(k), \kappa(x(k))) \right]$. Typically, the differences lie in the form of
 the stage cost $\ell(\cdot, \cdot)$, in the non-unitary discount factor $\gamma \in]0, 1]$ and the infinite-horizon.

⁸Seeking brevity, we omit here the fundamental **aspects regarding the appropriate** selection of the expectancy value supports.
 A thorough **discussion** can be found in [26]. Moreover, in general, we can understand reinforcement learning as a sub-category
 of machine learning techniques. Yet, we use these two terms as equivalencies in this work, for simplicity.

In RL formulations, the optimal value function is given by the application of the optimal policy κ^* :

$$\mathcal{V}^*(x(k)) := \mathbb{E} \left[\sum_{k=0}^{+\infty} \gamma^k \ell(x(k), \kappa(x(k))) \mid x(k) \right], \quad (4)$$

while the associated Bellman equation gives the action-value function:

$$\mathcal{Q}^*(x, u) := \ell(x, u) + \gamma \mathbb{E} [\mathcal{V}^*(x^+) \mid x, u], \quad (5)$$

where the expectation is computed according to the considered conditional probability density related to the states commuting from x to the successor x^+ .

The distinct RL algorithms apply different approximations for the optimal maps \mathcal{V}^* , \mathcal{Q}^* and κ^* . In synthesis, these are parametrically expressed in terms of an optimal *adaptation* parameter θ^* (much in the likes of the uncertainty parameter that appears in adaptive MPC formulations, as gives Eqs. (2)-(3)). This optimal parameter is used to solve the optimal RL problem either by directly or approximately solving Eq. (5) in a sampled-data fashion.

As do Gros and Zanon [26], we compact this parameter estimation task by means of the the following optimisation problem $\theta^* := \min_{\theta} \sum_{k=0}^{n_{\theta}} \psi(x(k+1), x(k), u(k), \theta)$, given an estimation function candidate $\psi(\cdot, \cdot, \cdot, \cdot)$, which depends on the specific RL algorithm that is used. For example, in the so-called *Q-learning* schemes, e.g. [73, 78], this estimation function is given, with $n_{\theta} = 1$, by a quadratic difference between the current action-value and its θ -related counterpart, that is:

$$\psi(x(k+1), x(k), u(k), \theta) = (\ell(x(k), u(k)) + \gamma \mathcal{V}_{\theta}(x(k+1)) - \mathcal{Q}_{\theta}(x(k), u(k)))^2, \quad (6)$$

where \mathcal{V}_{θ} and \mathcal{Q}_{θ} stand, respectively, for the parametric approximations of \mathcal{V}^* and \mathcal{Q}^* .

In RL, the optimal adaptation parameter support θ is also found by taking into account a sample-based dispersion set parametric \hat{S}_+ . In practice, the state transitions are assumed to be confined, at each sample, within an estimated dispersion set \hat{S}_+ . Accordingly, the RL problem takes the following form:

$$\begin{aligned} \theta^* &:= \min_{\theta} \sum_{k=0}^{n_{\theta}} \psi(x(k+1), x(k), u(k), \theta), \\ \text{s.t. } &x(k+j+1|k) \in \hat{S}_+(x(k+j|k), u(k+j|k), \theta), \forall j \in \mathbb{N}_{[1, n_{\theta}]}. \end{aligned} \quad (7)$$

Based on the previous discussion, the operation of data-driven RL-based MPC schemes is, thus, implemented through a twofold:

1. At each discrete-time sample k , a standard (possibly robust) MPC in the form of Eq. (1) is solved, considering a baseline process model description that relies on θ ; the corresponding MPC input $u^*(k)$ is applied to the system and new process data is collected;
2. Then, based on the available data, the RL problem from Eq. (7) is solved⁹ and the parameter θ is updated.

⁹We highlight that in some of the formulations surveyed next, the RL mechanism operates at a lower sampling rate than the MPC. With respect to this matter, we stress that performing learning and control at the same time requires to regularly update both the model parameters and the corresponding control law that are implemented. Accordingly, when parameter updates are implemented *while* the system is being operated, as argued herein, it should be remarked that implementing a safe and stable control result at a given instant **does not** necessarily ensure that the overall system dynamics are stable (nor that constraints are continuously satisfied) in closed-loop - this property can only be ensured when the parameter updates occurs while the system trajectories are found within a specific set (i.e. the dispersion set estimate \hat{S}_+). In synthesis, thus, guaranteeing coherent closed-loop performance requires both a safe RL approach and a robust MPC formulation with respect to these updates.

390 *Remark 5.* In the so-called *safe* RL formulations, the dispersion set parametric estimate \hat{S}_+ not only confines
the state transitions but also the systems' constraints. Basically, \hat{S}_+ is taken as an outer approximation
of the real state transition dispersion set S_+ , thus also encompassing safety remarks. As in robust MPC
formulations, the ideal formulation would depend on a closed-loop dispersion set; yet, what is done is to
(recursively) select an open-loop profile reconstructed using an affine feedback law, and iterate this task at
395 each sample.

Remark 6. As discussed in [85], MPC algorithms are differentiable if, essentially, they are solvable. Thus,
Problem 7 can often be solved using sensitivity-based algorithms. For such, in the context of adaptive
schemes, one should verify that the function approximator is differentiable with respect to the adaptation
400 parameter θ - which implies in linear independence constraints qualification, strong second order sufficiency
and strict complementarity of the optimisation (accounting for variations in θ). Accordingly (even if the
later property is not satisfied), an RL step can be computed with appropriate estimates for the dispersion
set \hat{S}_+ and for the parametric terms \mathcal{V}_θ and \mathcal{Q}_θ .

3.2. Existing formulations

405 Taking into account the previous discussion on how data-driven MPC schemes with RL-based model
adaptation operate and are implemented, we provide, next, an overview of the available formulations. Indeed,
recent literature exhibits a vast amount of works that combine learning techniques with control application,
c.f. [16, 86, 87, 66, 88, 67, 76, 78, 25, 89, 90, 91]. In particular, in this survey, we review first the methods which
enable *safe* performances, then we discuss alternative formulations and, finally, methods with registered
410 experimental validation or benchmark comparisons.

With respect to the category of safe schemes, we stress that combining robust MPC with RL has been
extensively studied by Gros and Zanon [26] - including in several recent works by these authors, as cited
in this review. In synthesis, their formulations offers a direct path to seemingly merge an RL-based model
adaptation layer with a robust predictive controller, thereby providing performance certificates (which are
415 thoroughly theoretically validated). These safe RL MPC schemes do not have to resort to any kind of
offline certification of the resulting parameter estimates (i.e. *Monte Carlo* simulations - which is often
difficult to perform in practical situations), as done in *learning MPC schemes* based on Neural Networks,
e.g. [35, 34, 36]. *Next, we highlight the main elements of these safe RL MPC techniques.*

Even though there are currently some strategies available to induce safe control laws in the context of
420 RL, c.f. the survey [84] and references mentioned therein, none of them are able to strictly ensure constraints
satisfaction for all sampling instants - constraint violation is typically considered as a possible outcome and
strongly penalised in [72] and in some of the methods detailed in [84].

- While [92] is the first work to ensure the robust satisfaction of constraints (using an LQR control
425 formulation), the works [25, 17, 26] are the first concrete propositions where an RL formulation based
on (nonlinear) MPC that include the theoretical guarantees of safety. We indicate that a corresponding
formulation for systems with mixed-integer dynamics is available in [93];
- In [16], one can find an alternative formulation to the safe learning-based MPC problem, which also
offers deterministic guarantees on robustness. The main difference w.r.t. the algorithms in [25, 17, 26,
430 93] reside in decoupling the problem of safe exploration and performance, considering two models of
the controlled process: (i) one initial (approximate) model subject to a bounded uncertainty term; and
(ii) another statistical model related to the uncertainty. Accordingly, statistical (parametric or non-
parametric) identification tools are used within the learning formulation in order to address the problem
of estimating the later model, while the control formulation is used to ensures safety and robustness
435 by maintaining the predictions of the future dynamics (using the approximate model) stable. A key
feature of method is that it is shown to ensure closed loop stability; moreover, probabilistic arguments
are used to demonstrate that the resulting control performances converges to what would be obtained
using a deterministic MPC with the true model at hand.

- More recently, Lin et al. [71] propose an RL-based MPC scheme integrated by the means of policy iteration. That is, the proposed MPC is responsible for generating a given control policy which is, then, evaluated by a RL layer according to measured data. In particular, similarly to what is done in [25, 17, 26, 93], the obtained value function is taken as the terminal cost of MPC. The main benefit of using the policy iteration formulation is that it removes the necessity of using stabilising terminal ingredients in the MPC, does eliminating the need of the offline preparation of the optimisation. Safety, stability and performance are thus verified with regard to the RL layer only.

With respect to alternative of learning-based MPC schemes, which do not employ the safety-related arguments as in, e.g., [25], we highlight, **next**, some relevant formulations.

First, we refer to the scheme proposed by Esfahani et al. [90], which relate the use of the RL layer to a Moving Horizon Estimation problem (MHE). Not only the MHE is parametrised by RL techniques in order to estimate the adaptation parameter θ online, but it is also used to provide full state information of the real process. In particular, RL is used to jointly tune the parameters of the MHE and MPC optimisation problems, enclosing considerations on closed loop performances. The main drawback of the scheme is that the convergence of uncertainty inference from the MHE scheme is only demonstrated, to the best of our knowledge, under linearly independent constraints in the formulation - which restricts the systems for which the method can be applied.

The recent output-feedback algorithm by Maiworm et al. [94] combines a robust adaptive-kind MPC algorithm with machine learning and Gaussian-inference. In particular, the scheme comprises terminal ingredients and ensures guaranteed ISS with constraint satisfaction for general prediction models that are learned online and satisfy some Gaussian-related conditions. With regard to the adaptive MPC schemes surveyed in Sec. 2, the method is much alike the robust algorithm proposed by Tanaskovic et al. [41] - the fundamental difference is the use of the machine learning tools instead of set membership identification. Accordingly, similar discussions on resulting performances apply.

Bellegarda and Byl [89] combines RL MPC with the so-called trajectory optimisation (TO) problem. Specifically, TO is used in MPC schemes in order to analyse stability and robustness with regard to the generated motions and trajectories in closed-loop. Typically, TO is operated in a separate layer, coordinating the MPC scheme that lies below, using simplified (sometimes static) models of the process. The referred work, by coupling TO with RL, is able to decrease the RL sample complexity with knowledge of the real-time optimal control solution. Moreover, it allows to compute the MPC law by using the TO model for nominal predictions, while continuously improving it in the upper layer via RL.

The work [95] considers that the controlled system has a fixed model structure, with time-varying parameters. Accordingly, system identification techniques are mixed to the RL layer, which is, thus, responsible for two tasks: (i) to parametrise the MPC in order to optimise performance, in terms of an adaptation parameter; and (ii) to match the process data to a given model structure using system identification arguments (i.e. Least-Squares)

In [96], a cautious adaptive MPC is formulated using an RL layer using a Bayesian approximator. The method is much similar to the surveyed adaptive cautious MPC [43], yet the main innovation lies in using a constraint tightening mechanism in order to bound the expected number of unsafe episodes. A significant result of the paper is how the method is tested, in simulation, for different kinds of systems and settings, which indicate its competitiveness to other adaptive schemes and how an RL layer could be better than typical identification arguments (such as, e.g., set-membership and least-squares).

We care to highlight, finally, that most of the surveyed RL MPC method require gradient-based solvers to obtain a feasible solution, which actually restricts their application to larger datasets. Accordingly, Sawant et al. [97] propose a solution to this issue by formulating the learning procedure online, directly connected to the measured data. The proposed approach offers an MPC scheme that does not have to be solved over any collected data-set, thereby softening the resulting computational complexity and rendering it possible for big data problems.

We remark that, recently, there has been considerable research on data-driven MPC schemes, with theoretical features from reinforcement learning, with experimental validation (or applied in realistic, high-fidelity numerical simulations). We recap some of the main findings that are indicated in the available

literature:

- 490 • Considering a realistic autonomous surface vehicle (ASV) model, Cai et al. [98] propose an optimal control policy for freight missions which include the tasks of collision-free path following and autonomous docking. In particular, the proposed parametrised MPC is based on a deterministic gradient method to update the adaptive model parameters in the context of RL. The published results indicate how the method is able to improve closed-loop performances. Similarly, in [99], the same freight mission problems are addressed, whereas a scenario-tree robust MPC is used to deal with possible thruster faults. 495 In particular, the RL layer is tuned with Q -learning arguments and an economic cost is considered in the control formulation. Results also indicate good performances in practice;
- Rosolia et al. [69] provide a thorough survey of data-driven MPC schemes applied for autonomous systems. In this work, several learning MPC schemes are applied to the control often inverted pendulum with an assistive motor held in the palm of a hand. In particular, the learning task is thoroughly analysed and **discussed** in terms of how the scheme improves over time (as the RL converges and the MPC is deployed with a every-iteration-closer-to-reality model);
- In [68], we can find experimental validation results of an RL-based MPC for the control of an autonomous vehicle in aggressive driving situations. The provided results, yet empirical, demonstrate 500 how the scheme is able to provide good performances using only data collected from the system.

3.3. Overall discussion: Summary of Strengths and Limitations

There has been a consistent growing interest in applying learning schemes to control problems. Specifically, using RL techniques has been shown (in theory and practice - through simulation and experimental validation) to be quite interesting when coupled with MPC: an adaptive process model update layer based 510 on RL is an adequate solution, providing good estimation result for a wide variety of nonlinear models and settings. For the Reader's convenience, we provide, in Table 2, a broad overview of the surveyed methods, indicating the main features of each approach (used approximator, learning tools, etc) and detailing the kinds of systems considered, as well as the type of validation available (experimental or numerical).

The first schemes that combined RL and MPC were typically not shown to provide any strict certificate, 515 which complicates realistic validation and any kind of implementation for complex systems subject to hard constraints. Yet, very recently, tools have been presented for *safe exploration*, which, in practice, means the availability of closed loop performance guarantees, c.f. [26].

We emphasise that, for many of the surveyed techniques, the MPC optimisation is tuned similarly to what is done in adaptive formulations. The main difference resides in the approach used to infer on the 520 uncertainty from data. Nevertheless, even though the research community has given a great deal of attention to these learning schemes, Recht [100] extensively argues that these methods consistently require large data sets - typically much larger than data used for classical Least-Squares or set-membership approaches. This issue becomes a particularly difficult aspect in the case of complex systems with inaccessible variables, for which data sets are not so robust and can be quite difficult to generate.

Moreover, despite of the recent available safe formulations, the majority of RL estimation schemes tend 525 to lack formal guarantees on stability and performance of the resulting closed-loop. In synthesis, further coordinated research on the topic of RL MPC with guarantees can certainly be of interest in order to fustigate the establishment of theoretically-validated implementations.

Table 2: Overview of surveyed machine learning MPC schemes.

Reference	Certificates	Robustness arguments	Learning tools	Application
[16]	Closed-loop asymptotic stability , robust feasibility, robust constraint satisfaction	Probabilistic and deterministic robust invariant set, constraints tightening (tube)	Approximator: Parametric, stochastic, least-squares, Safe exploration	Numerical simulation: Moore-Greitzer compressor system, Experimental validation: Energy-efficient building automation system, quadcopter flight
[72, 25, 17, 26]	Robust recursive feasibility, closed-loop stability, constraint satisfaction	Robust invariant sets	Approximator: MPC costs and structure, Safe exploration	Numerical simulation (several)
[29]	Closed-loop stability	-	Approximator: Lyapunov candidate, Safe exploration	Numerical simulation: Inverted pendulum benchmark
[71]	Parameter estimation convergence, recursive feasibility, closed-loop stability	Robust terminal ingredients	Approximator: Policy iterated MPC costs and structure	Numerical simulation: Linear and Hammerstein-Wiener systems
[94]	ISS, recursive feasibility, constraint satisfaction	Robust terminal ingredients	Approximator: Gaussian process	Numerical simulation: CSTR system
[90]	-	-	Approximator: MHE-based structure and <i>Q</i> -learning	Numerical simulation: Mass, spring and damper system

Table 2: Overview of surveyed machine learning MPC schemes. (Continued)

Reference	Certificates	Robustness arguments	Learning tools	Application
[89]	-	-	Approximator: Proximal policy optimisation	Numerical simulation: Trajectory generation for autonomous cars
[95]	-	KKT arguments	MPC cost and structure, Approximator: Q-learning, system identification arguments	Numerical simulation
[96]	Yes	Constraints tightening (tube)	Approximator: Bayesian representation , Safe exploration	Numerical simulation: Thermal process and drone system
[97]	-	-	Approximator: Markov decision process, MPC cost and structure	Numerical simulation: Tracking problem, pendulum swing-up task, cart-pole swing-up and balancing issue
[98]	-	-	Approximator: MPC parametrisation, least squares temporal difference policy gradient	Numerical simulation: Freight mission of autonomous surface vehicles
[69]	Recursive feasibility, non-decreasing performance, closed-loop stability	Terminal ingredients	Approximator: Q-learning, Safe exploration	Numerical simulation: Dubins' car problem

Continued on next page

Table 2: Overview of surveyed machine learning MPC schemes. (Continued)

Reference	Certificates	Robustness arguments	Learning tools	Application
[68]	-	-	Approximator: Multi-layer neural networks	Numerical simulation: Cart-pole swing-up problem, quadcopter navigation task, Experimental validation: Aggressive driving task in autonomous vehicles

4. Category C: Behavioural/trajjectory-based formulations

Over the last few years, an alternative class of data-driven predictive control algorithms has been consistently developed, which we name herein as MPC algorithms based on *trajectory features*, sometimes referenced as methods based on *behavioural theory*, c.f. the discussion presented by Markovsky and Dörfler [18].

The first two categories of data-driven MPCs - i.e. adaptive schemes (Category A, Sec. 2) and learning-based algorithms (Category B, Sec. 3) - held considerable similarities between techniques. Moreover, the MPC optimisation was maintained in the same format. Now, this final surveyed category of work fundamentally differs from the prior: as indicated in Figure 4, the so-called behavioural formulations do not require a baseline prediction model, but rather only an initial data dictionary. Then, during the online implementation, the MPC optimisation itself is written in terms of measured data.

Specifically, these data-driven MPC algorithms derive from the unified framework originally celebrated¹⁰ in [11, 101]. This framework, in fact, exploits and remembers the well-known result from the 00’s by Willems et al. [31] (Willems’ Fundamental Lemma) in the context of data-based control synthesis: a persistency of excitation condition over the applied control input is used in order to characterise all possible trajectories of an unknown system, considering a single measured input–output dictionary of a sufficient length. This framework has been thoroughly discussed in the context of LTI systems, with links provided for stability properties and the analysis of dissipativity and passivity, c.f. [11, 102, 103, 104, 105, 106]. More recently, the thorough “handbook” paper by Verheijen et al. [107] details the fundamental details of these data-driven predictive control schemes, from theory to implementation aspects.

Next, we recall and outline the main arguments used to develop such trajectory-based representations, detailing how they can be applied in the context of MPC. We also present a broad overview of the existing formulations and **investigate** their capabilities and disadvantages.

Remark 7. In [18, 19], the Reader can find comprehensive overviews of the scope of behavioural theory in data-driven control. Furthermore, in [107, 108] application-oriented recaps of corresponding data-driven MPC schemes can be found.

¹⁰We note that, recently, an alternative synthesis approach using Peterson’s Lemma has been presented by Bisoffi et al. [12]. Since only few papers use this formulation, we focus henceforth only on the results that arise from the application of Willems’ Fundamental Lemma.

4.1. Outline of the mechanism

The condition of persistent excitation (see Def. 1.1) is widely used in system identification theory. Accordingly, as shows Berberich and Allgöwer [101], a persistently exciting input signal of sufficient order, applied to an LTI system, allows us to construct a vector space that comprises all possible trajectories of this system. Under the assumption that the system order n_x is known (or, at least, an upper bound estimate $\nu \geq n_x$), the following result demonstrates how to map system trajectories based on a single measured input-output sequence.

Proposition 4.1 (LTI Data-driven predictor [101]). *Suppose that $\{\tilde{u}(i), \tilde{y}(i)\}_{i=0}^{L-1}$ is a fixed (measured) input-output trajectory of an LTI system, being u a control input which is persistently exciting of order $N_p + \nu$. Let $\{u(i), y(i)\}_{i=k-\nu}^{k-1}$ be another input-output trajectory dictionary of this same system (of “initial conditions”, sampled at instant $k - 1$). Then, let $\{\hat{u}(i)\}_{i=k}^{k+N_p-1}$ be a sequence of future inputs that will be applied to this system. Then, as long as $\nu \geq n_x$, there exists a vector $\alpha \in \mathbb{R}^{L-N_p+1}$ for $\xi = L - N_p + \nu - 1$ such that:*

$$\begin{bmatrix} H_\nu(u^\perp) \\ H_\nu(y^\perp) \\ H_{N_p}(u) \end{bmatrix} \alpha = \begin{bmatrix} \text{col}\{\tilde{u}\} \\ \text{col}\{\tilde{y}\} \\ \text{col}\{\hat{u}\} \end{bmatrix},$$

where the compacted data trajectories u^\perp and y^\perp comprise $L+\nu+1$ entries, ordered as follows: $\{u^\perp(i)\}_{i=0}^{L-1} = \{\tilde{u}(i)\}_{i=0}^{L-1}$, i.e. fixed data, and $\{u^\perp(i)\}_{i=L}^{L+\nu} = \{u(i)\}_{i=k-\nu}^{k-1}$, i.e. sampled data (the same holds for the output data). Accordingly, the future output behaviour of this system is given by is

$$\text{col}\{\hat{y}\} = [\hat{y}^T(k) \quad \dots \quad \hat{y}^T(k + N_p - 1)]^T = H_{N_p}(y)\alpha.$$

We note that Proposition 4.1 represents a data-driven prediction rule for LTI systems. It consists in combining an initially measured data dictionary $\{\tilde{u}(i), \tilde{y}(i)\}_{i=0}^{L-1}$ (of L samples, fixed) with a sampled data dictionary of initial conditions $\{u(i), y(i)\}_{i=k-\nu}^{k-1}$ in order to generate predict the future outputs of the system, given a sequence of inputs. Naturally, it can be directly used as a basis to develop data-driven MPC algorithms, as detailed next.

Remark 8. Recently, Klädtke and Darup [109] detail how the data-driven trajectory representation from [101, 11], recapped in Proposition 4.1, implicitly characterises an input-output predictor just as in traditional model or subspace predictive control, even if it is not explicitly enforced as an equality constraint. That is, in some sense, the corresponding data-driven MPC schemes (as discussed in the sequel) implicit impose a certain kind of model-related representation recovered from data.

Remark 9. For brevity, and seeking alleviated notation, we indicate, herein, only the data-driven predictor based on trajectory features for the case of LTI systems. Nevertheless, we stress that there currently exists several corresponding counterparts for specific classes of nonlinear systems. In particular, in [101], the previous result is also extended for the case of nonlinear systems composed by a an LTI part an an input (or output) static nonlinearity, such as Lur’e, Hammerstein and Weiner dynamics. Also, extensions for second-order Volterra systems are available in [110], flat nonlinearities are addressed by Alsalti et al. [111], and Koopman operators are used in [112]. Also, Strässer et al. [113] show how such prediction rule can be extended to the case of nonlinear processes with rational or general non-polynomial dynamics. Furthermore, recent works [21, 106] demonstrate how these trajectory features can be extended to the case of LPV systems, which can be used to represent nonlinear dynamics under some assumption, given that the (current and future) scheduling parameter data is also known. The approach can be used to control generic nonlinear systems using the quasi-LPV embedding approach, considering that the function to map the scheduling parameters from the measured system variables is explicitly known (i.e. partial model information of the process is available, the same hypothesis brought forward by Berberich et al. [103]. A corresponding MPC application using Taylor-based linear expansions is also shown in [23].

4.2. Existing formulations

In terms of data-driven predictive control applications using the trajectory approach, we consider the following quadratic regulation performance cost, written in terms of the future input and output variables, from the viewpoint of sampling instant k : $J(\hat{y}, \hat{u}) := \sum_{i=k}^{k+N_p-1} (\|\hat{y}(i)\|_Q^2 + \|\hat{u}(i)\|_R^2)$, where Q and R are positive definite weighting matrices and N_p gives the size of the prediction horizon window. Note that, within this framework, input and output representations are used, and state trajectories, thus, are usually not considered. Then, given a fixed data dictionary $\{\tilde{u}(k), \tilde{y}(k)\}_{k=0}^{L-1}$ and a sampled data set of initial conditions $\{u(i), y(i)\}_{i=k-\nu}^{k-1}$, the original works [20, 114, 115] show how an MPC scheme can be synthesised based on the recurrent solution (at each discrete-time sampling instant k) of the following constrained quadratic programming problem:

$$\begin{aligned} \alpha^*(k) &= \min_{\alpha(k), \hat{\mathbf{y}}, \hat{\mathbf{u}}} J(\hat{\mathbf{y}}, \hat{\mathbf{u}}) + V(\hat{\xi}(k + N_p | k)), \\ \text{s.t.}: & \begin{bmatrix} \text{col}\{\tilde{u}\} \\ \text{col}\{\tilde{y}\} \\ \text{col}\{\hat{\mathbf{u}}\} \end{bmatrix} = \begin{bmatrix} H_\nu(u^\perp) \\ H_\nu(y^\perp) \\ H_{N_p}(u) \end{bmatrix} \alpha(k), \\ & \hat{\mathbf{y}} = H_{N_p}(y)\alpha(k), \\ & \hat{y}(k + j - 1 | k) \in \mathcal{Y}, \forall j \in \mathbb{N}_{[1, N_p]}, \\ & \hat{u}(k + j - 1 | k) \in \mathcal{U}, \forall j \in \mathbb{N}_{[1, N_p]}, \\ & \xi(k + N_p) \in \mathbf{X}_t. \end{aligned} \tag{8}$$

where the bold font emphasis is given to the optimisation decision variables. In this optimisation procedure, the system operational constraints are represented by the corresponding (compact and convex) input and output sets \mathcal{U} and \mathcal{Y} , respectively.

Let $J^*(\hat{y}, \hat{u})$ be the optimal solution of data-driven MPC optimisation, from which $\alpha(k)^*$ is the optimal minimiser. Then, the MPC application is implemented by applying data-driven control input $u^*(k)$, which is the first entry of the predicted optimal input sequence $\hat{u}^* = H_{N_p}(u)\alpha^*(k)$.

Some variations to the data-driven optimisation in Eq. (8) are possible, such as the inclusion of (ridge) regularisation and slack variables, e.g. [116], compare also to the experimentally-validated result presented in [117]. We note nevertheless, that the core structure of the method remains the same: a fixed sequence of input-output data is used to setup a prediction matrix, which includes the currently sampled initial conditions. In contrast to a model-based predictive control synthesis, as in [118], Eq. (8) does not have the sequence of inputs as its minimiser, but rather the (rectangular) predictor gain $\alpha(k)$, which, in turn, affects both the future sequence of inputs and outputs. In Algorithm 4.2, we summarise how these methods are implemented.

In Eq. (8), we purposely include the presence of the so-called terminal ingredients: the terminal set \mathbf{X}_t , the terminal cost $V(\hat{\xi}(k + N_p | k))$, and a fictive terminal variable $\hat{\xi}$. In the recent work by Berberich et al. [33], formal arguments are provided on how to generate these elements, thus ensuring closed-loop stability guarantees¹¹. We stress that their synthesis does not require any model knowledge and are enabled by fractional transformations of data-dependent matrices based on the trajectory representation discussed in the prequel¹². Stability and robustness properties of these algorithms are also investigated in [116, 11, 119, 23], with experimental validation demonstrated in [22]. We also stress that adaptations of the previous optimisation for the tracking problem (ensuring that the output follows a given reference signal) has also been addressed in the literature, with special focus to the use of artificial tracking arguments, e.g. [30, 33, 115].

¹¹The use of terminal equality constraints can also be considered as an alternative to the terminal ingredients, as shown Berberich et al. [119]. Yet, although being easier to implement (no offline synthesis of the terminal ingredients is required), the equality constraints often render MPC schemes with poorer robustness properties and smaller corresponding regions of attraction.

¹²Due to scope of this paper, we opt not to discuss the synthesis of these ingredients herein.

Algorithm 4.2 (Data-driven trajectory-based MPC).

- *Offline procedure:*
 1. Determine the system order upper-bound $\nu \geq n_x$;
 2. Collect the initial fixed input-output data set $\{\tilde{u}(i), \tilde{y}(i)\}_{i=0}^{L-1}$, considering the application of a persistently exciting input;
 3. Define the control objective, choosing weights to the performance objective $J(\hat{y}, \hat{u})$.
 - *Online implementation, for each sampling instant k :*
 1. Collect the input-output initial conditions $\{u(i), y(i)\}_{i=k-\nu}^{k-1}$;
 2. Solve the data-driven optimisation from Eq. (8), thus obtaining $\alpha^*(k)$;
 3. Apply the first input of the corresponding control input sequence $u^* = H_{N_p}(u)\alpha^*(k)$.
-

630 *Remark 10.* The surveyed trajectory-based algorithms indeed do not require any kind of model knowledge, in the sense of representation structure (matrices, input-state channels with direct transfers, direct output-state maps, etc) or process information itself (time response, frequency range, etc). Nevertheless, we emphasise that the methods do require the knowledge of the system order n_x (or, at least, an upper bound estimate $\nu \geq n_x$).

635 4.3. Overall discussion: Summary of Strengths and Limitations

As done in the previous sections, we provide, in Table 3, a broad overview of the surveyed trajectory-based MPC methods, detailing the main features of each scheme and which kind of validation (experimental, numerical) is currently available.

640 Even though these data-driven MPC schemes based on trajectory features have been very extensively discussed over the last few years, we offer, **next**, some general perspectives on these methods, taking into account the previous discussions.

645 **As detailed throughout** this work, MPC schemes generate (predictive) control policy by means of optimisation programs - which rely on the *prediction* of the system trajectories. In both model-based and (trajectory) data-driven formulations, the discrepancy between the prediction rule within the optimisation and the true system trajectories plays a significant role w.r.t. robustness and optimality and of the resulting control action.

In terms of robustness, typical uncertainty propagation and constraint tightening methods (tubes) can also be used in the data-driven context, c.f. [105, 120]. **Yet**, as argues Anand et al. [121], the latter property can be analysed in terms of the *self-consistency*¹³ of the MPC algorithm - which cannot be explicitly 650 verified for the available behavioural data-driven formulations. This lack of optimality can be critical in many practical situations.

Furthermore, these behavioural MPC schemes require the knowledge of a sufficiently large upper bound ν with respect to the system dimension n_x . Even though a rough approximation can be easily obtain, it may lead to an MPC algorithm which is excessively costly in the computational sense, which may be an 655 impediment for real-time applications:

¹³As proposed by Anand et al. [121], self-consistency is implied if the predictive control value function satisfies a Bellman relation. In synthesis, a self-consistent MPC algorithm can be re-stated as a time-invariant sequential decision process. In the model-based context, MPC is self-consistent by construction; in the data-driven formulations, the prediction rule does not derive from an open-loop simulation of the controlled system, which means that, in general, self-consistency cannot be inferred.

- Since the optimisation decision is made in terms of the rectangular matrix $\alpha(k)$, which is of dimension $L - N_p + 1$, the resulting problem may be computationally more complex than a model-based scheme, which (usually) scales numerically in terms of $n_x N_p$;
- Another issue is related to the size of the initial fixed data window L , which must be sufficiently large in order to bound $L \geq (n_u + 1)(N_p + \nu) - 1 \geq (n_u + 1)(N_p + n_x) - 1$. As a real, an increase arises on the computational load required to solve the optimisation (due to the size of the resulting prediction matrix);
- Despite the increased numerical complexity of the resulting optimisation, (very) recent advances have provided relevant relaxations (such as dimension reduction for the decision variables) for the DeePC algorithm, making its numerical toughness comparable to model-based MPCs, c.f. [122, 123]. These advances can be exploited for other data-driven MPC schemes based on trajectory representations, which indicates that the numerical aspect is not a limitation that tends to persist;

We stress that one of the major issues of the current literature on this topic is that the closed-loop stability certificates essentially rely on the assumption that the controlled system truly displays LTI dynamics, which is often a false premise. Moreover, if the system is indeed LTI, it seems unreasonable to use these formulations in practice, since LTI identification is very well established and mature.

In terms of nonlinear schemes using the LPV approach from [21], we emphasise that the methods require the condition of a persistently exciting control input, which must be maintained over time both for the original signal and the LPV one, which is generated by the means of a Kronecker product between the scheduling data and the input data. This requirement, in practice, can only be guaranteed if a small noisy signal is added to the applied control signal, in such a way that the sum of signals remains persistently exciting in closed-loop. This condition was not necessary in the original LTI formulations from Berberich et al. [30], which maintained the same data set over time. However, in the nonlinear scenario using qLPV descriptions, the trajectory-based predictions are only valid due to a local dependence on the outputs mapped to the parameters, and, therefore, they should be reassessed at each sample. Discussions on this subject were also presented in [103].

The inclusion of a bias signal over the applied control input, in order to maintain closed-loop persistency of excitation, may also degrade the control results in real experimental tests, since the magnitude of the noise signal, which guarantees that the rank of $H_{N_p}(u) \geq n_u N_p$, can damage actuators due to excessive vibrations, for example.

The issue of noisy data also affects the obtained performances by these controllers. In synthesis, the effects measurement noise can be alleviated by the means of ridge regularisation, for instance, e.g. [116].

Finally, we also emphasise that recent results provided by Dinkla et al. [124] reveal, through simulation and theoretical assessments, that the direct application of these trajectory-based data-driven MPC schemes may lead to performances which are impacted by an identification bias that arises from correlation between control inputs and noise in closed-loop.

Since the majority of processes is not LTI, we note that the interest of using these trajectory-based schemes remain rather diluted while concrete formal certificates for nonlinear systems are available. Further nonlinear representation schemes with corresponding MPC stability analyses would help leverage the interest of data-driven framework for systems with complex dynamics.

In general, the available literature indicates the the current state-of-the-art on *behavioural* data-driven MPC schemes can achieve comparable performances to model-based counter parts [107]. This feature is particularly consistent for the case of LTI systems.

When more complex systems are considered, traditional model-based predictive control typically provides better performances than data-driven schemes (i.e. less conservative control, with a larger region of attraction). However, recent advances on LPV embeddings for data-driven control, c.f. [106, 21], are promising for the more general nonlinear, time-varying setting. We care to emphasise that, in the case of really complex systems, identifying a trustworthy model is very difficult (or even impossible), which means that traditional MPC schemes will mostly rely on *biased* prediction models. Overall, such data-driven MPC

705 formulations based on behavioural theory, especially in the nonlinear and LPV setting, certainly deserve further scholastic attention.

710 With respect to the aforementioned **context**, we highlight the main message: so far, the model-based predictive control design remains, in most cases, a superior solution (with more accurate predictions and therefore allowing for greater robustness), as long as there are enough measurements so that it is possible to perform a reliable identification procedure of the controlled process. The data-driven approach emerges as a plausible alternative primarily for processes with few measurements available, for which the identification of a model is not only a costly step, but also not truly viable.

715 Finally, we highlight that, even though better performances have not yet been enabled by these schemes (in comparison to model-based counterparts), trajectory formulations can certainly be integrated to adaptive MPC schemes, in such a way that local linear models (or LPV representations) based on data are used to enhance a baseline process description. The topic is definitely open for further investigation.

Table 3: Overview of surveyed trajectory-based MPC schemes.

Reference	Certificates	Robustness arguments	Nonlinearity issues	Application
[114, 101, 30, 22]	Recursive feasibility, closed-loop stability (LTI case), performance guarantees (LTI case)	Robust terminal ingredients, noise effect minimisation	Possibility to consider Hammerstein-Weiner dynamics	Numerical simulation (several), Experimental validation (simple systems)
[11, 20, 116, 117, 122, 123]	Recursive feasibility (LTI case), closed-loop asymptotic stability (LTI case)	Noise effect minimisation, (ridge) regularisation	-	Numerical simulation (several), Experimental validation : Nano quadcopter system
[102, 104]	-	Dissipativity arguments noise effect minimisation	-	Numerical simulation (several)
[106, 21]	Recursive feasibility, stability (hard assumptions on the LPV scheduling signal)	-	LPV embeddings	Numerical simulation

Continued on next page

Table 3: Overview of surveyed trajectory-based MPC schemes. (Continued)

Reference	Certificates	Robustness arguments	Nonlinearity issues	Application
[23]	Recursive feasibility, closed-loop stability (within a neighbourhood of the measured data)	Noise effect minimisation	qLPV-embeddings under scheduling proxy knowledge	Numerical simulation: rotational pendulum, Experimental validation: Electro-mechanical positioning system
[111]	-	-	Flatness-based nonlinearities	Numerical simulation
[110]	-	-	Volterra systems	Numerical simulation
[113]	-	-	Koopman-based nonlinearities	Numerical simulation
[124]	Recursive feasibility (LTI case), closed-loop stability (LTI case)	Noise minimisation, identification bias reduction	-	Numerical simulation

720 **5. Conclusions**

In this work, we reviewed the current relevant literature available regarding data-driven MPC algorithms. In particular, we surveyed *safe* methodologies - i.e. those with theoretical *certificates* that allow for closed-loop stability (and performance and constraints satisfaction) and recursive feasibility of the MPC optimisation.

725 Our aim in this work was to provide a better comprehension of the state-of-the art on the analysed topic, drawing conclusions on the capabilities and deficiencies of the possible design alternatives. Specifically, we discussed how MPC schemes can be formulated under a data-driven perspective by means of adaptive algorithms, solutions based on reinforcement learning, or derived from the more recent design alternative based on input-output trajectories (data dictionaries). **Our main findings are recalled next.**

730 **First, we highlight that** the available adaptive algorithms often require an initial model that is a sufficiently fair description of the process, in such a way that stability can be ensured despite the online adaptation procedure. That is, most methods rely on a given structure with known initial parameters or assume some form of bounded (or a given stochastic distribution of) model mismatches over time. Several of the alternatives used for online adaptive estimation schemes are correlated with system identification tools, such as recursive least-squares inferences. There are also formulations which reside on using a single baseline model and, during the implementation, vary the form and structure of the MPC itself. Many results indicate interesting performances of these adaptive schemes, yet formal performance certificates frequently depend on worst-case demonstrations (example: showing that stability is enable for all models considering a given set of bounded parameter values). We emphasise, nevertheless, that the use of LPV models is a viable

740 path to synthesise adaptive MPC schemes with certificates with potentially less conservative results - given that the robustness can be verified only with respect to the variation rate of the scheduling parameters over samples, which are typically small.

We also stress that several recent works have shown how techniques based on reinforcement learning can be provably robust and safe for a variety of situations. In synthesis, further research can still be devoted in 745 the topic of safe RL MPC schemes, considering that the majority of experimentally-validated schemes exhibit only empirically-verified performance guarantees. Moreover, the majority of these RL MPC techniques often require large amounts of data and stochastic assumptions regarding the process, which can be a difficult issue to solve for complex systems with inaccessible measurements.

750 Finally, we discussed how works based on trajectory features assume an underlying persistency of excitation condition and a baseline LTI behaviour, which means it is possibly not adequate for hardly nonlinear dynamics and that, in such cases, the closed-loop implemented signal must be sufficient rich, which may deteriorate actuators.

Acknowledgements

M. M. Morato acknowledges the financial support from FAPESC (2023TR001506) and CNPq (304032/2019–0, 403949/2021–1, 406477/2022–1. This work was partially supported by CAPES under grant 88881.878833/2023–01 (OPCOMPLEX, SticAmSud) and partially funded by the National R+D+I Plan, Spain Project PID2021–126889O B-I00, funded by MCIN/AEI/10.13039/501100011033 and by “ERDF A way of making Europe”.



Funded by
the European Union

This work has been funded by European Union’s Horizon Europe research and innovation programme under the *Marie Skłodowska-Curie Action* grant agreement No 101149263 (*REALOPT4CONTROL* project).

760 The Authors thank for the support from the Department of Automation and Systems at UFSC (Brazil), and from GIPSA-Lab, at Université Grenoble-Alpes (France).

Author contributions

M. M. Morato: Writing and Revision.

M. S. Felix: Writing and Revision.

765 References

- [1] C. R. Cutler, B. L. Ramaker, Dynamic matrix control: a computer control algorithm, in: Joint Automatic Control Conference, 17, p. 72.
- [2] D. W. Clarke, C. Mohtadi, P. Tuffs, Generalized predictive control—part i. the basic algorithm, *Automatica* 23 (1987) 137–148.
- 770 [3] E. F. Camacho, C. Bordons, Model predictive control, Springer Science & Business Media, 2013.
- [4] M. Alamir, A framework for real-time implementation of low-dimensional parameterized NMPC, *Automatica* 48 (2012) 198–204.
- [5] F. Allgöwer, A. Zheng, Nonlinear model predictive control, volume 26, Birkhäuser, 2012.
- 775 [6] D. Limon, A. Ferramosca, I. Alvarado, T. Alamo, Nonlinear mpc for tracking piece-wise constant reference signals, *IEEE Transactions on Automatic Control* 63 (2018) 3735–3750.
- [7] M. M. Morato, J. E. Normey-Rico, O. Sename, Model predictive control design for linear parameter varying systems: A survey, *Annual Reviews in Control* 49 (2020) 64–80.
- [8] D. Q. Mayne, J. B. Rawlings, C. V. Rao, P. O. Scokaert, Constrained model predictive control: Stability and optimality, *Automatica* 36 (2000) 789–814.
- 780 [9] M. M. Morato, T. Holicki, C. Scherer, Stabilizing model predictive control synthesis using integral quadratic constraints and full-block multipliers, *International Journal of Robust and Nonlinear Control* (2023).
- [10] L. Ljung, Perspectives on system identification, *Annual Reviews in Control* 34 (2010) 1–12.
- [11] C. De Persis, P. Tesi, Formulas for data-driven control: Stabilization, optimality, and robustness, *IEEE Trans. on Automatic Control* 65 (2019) 909–924.
- 785 [12] A. Bisoffi, C. De Persis, P. Tesi, Data-driven control via petersen’s lemma, *Automatica* 145 (2022) 110537.

- [13] T. R. Steentjes, M. Lazar, P. M. Van den Hof, On data-driven control: informativity of noisy input-output data with cross-covariance bounds, *IEEE Control Systems Letters* 6 (2021) 2192–2197.
- [14] D. Muñoz-Carpintero, M. Cannon, B. Kouvaritakis, Robust MPC strategy with optimized polytopic dynamics for linear systems with additive and multiplicative uncertainty, *Systems & Control Letters* 81 (2015) 34–41.
- 790 [15] M. Lorenzen, F. Allgöwer, M. Cannon, Adaptive model predictive control with robust constraint satisfaction, *IFAC-PapersOnLine* 50 (2017) 3313–3318.
- [16] A. Aswani, H. Gonzalez, S. S. Sastry, C. Tomlin, Provably safe and robust learning-based model predictive control, *Automatica* 49 (2013) 1216–1226.
- [17] S. Gros, M. Zanon, A. Bemporad, Safe reinforcement learning via projection on a safe set: How to achieve optimality?, *IFAC-PapersOnLine* 53 (2020) 8076–8081.
- 795 [18] I. Markovsky, F. Dörfler, Behavioral systems theory in data-driven analysis, signal processing, and control, *Annual Reviews in Control* 52 (2021) 42–64.
- [19] I. Markovsky, L. Huang, F. Dörfler, Data-driven control based on the behavioral approach: From theory to applications in power systems, *IEEE Control Systems Magazine* 43 (2023) 28–68.
- 800 [20] J. Coulson, J. Lygeros, F. Dörfler, Data-enabled predictive control: In the shallows of the DeePC, in: *Proceedings of the 18th European Control Conference (ECC)*, IEEE, pp. 307–312.
- [21] C. Verhoek, H. S. Abbas, R. Tóth, S. Haesaert, Data-driven predictive control for linear parameter-varying systems, *IFAC-PapersOnLine* 54 (2021) 101–108.
- [22] J. Berberich, J. Köhler, M. A. Müller, F. Allgöwer, Data-driven model predictive control: closed-loop guarantees and experimental results, *at-Automatisierungstechnik* 69 (2021) 608–618.
- 805 [23] M. M. Morato, J. E. Normey-Rico, O. Sename, Nonlinear data-driven control Part II: qLPV predictive control with parameter extrapolation, *Journal of Control, Automation and Electrical Systems* (2024) 1–13.
- [24] V. Adetola, M. Guay, Robust adaptive MPC for constrained uncertain nonlinear systems, *International Journal of Adaptive Control and Signal Processing* 25 (2011) 155–167.
- 810 [25] M. Zanon, S. Gros, Safe reinforcement learning using robust MPC, *IEEE Transactions on Automatic Control* 66 (2020) 3638–3652.
- [26] S. Gros, M. Zanon, Learning for MPC with stability & safety guarantees, *Automatica* 146 (2022) 110598.
- [27] M. Tanaskovic, L. Fagiano, V. Gligorovski, Adaptive model predictive control for linear time varying MIMO systems, *Automatica* 105 (2019) 237–245.
- 815 [28] D. Collet, M. Alamir, D. Di Domenico, G. Sabiron, Data-driven fatigue-oriented MPC applied to wind turbines individual pitch control, *Renewable Energy* 170 (2021) 1008–1019.
- [29] F. Berkenkamp, M. Turchetta, A. Schoellig, A. Krause, Safe model-based reinforcement learning with stability guarantees, *Advances in neural information processing systems* 30 (2017).
- [30] J. Berberich, J. Köhler, M. A. Müller, F. Allgöwer, Data-driven tracking MPC for changing setpoints, *IFAC-PapersOnLine* 53 (2020) 6923–6930.
- 820 [31] J. C. Willems, P. Rapisarda, I. Markovsky, B. L. De Moor, A note on persistency of excitation, *Systems & Control Letters* 54 (2005) 325–329.
- [32] M. Alamir, From certification of algorithms to certified MPC: The missing links, *IFAC-PapersOnLine* 48 (2015) 65–72.
- 825 [33] J. Berberich, J. Köhler, M. A. Müller, F. Allgöwer, On the design of terminal ingredients for data-driven MPC, *IFAC-PapersOnLine* 54 (2021) 257–263.
- [34] S. Piche, B. Sayyar-Rodsari, D. Johnson, M. Gerules, Nonlinear model predictive control using neural networks, *IEEE Control Systems Magazine* 20 (2000) 53–62.
- [35] B. M. Åkesson, H. T. Toivonen, A neural network model predictive controller, *Journal of Process Control* 16 (2006) 937–946.
- 830 [36] P. Kittisupakorn, P. Thitayasook, M. A. Hussain, W. Daosud, Neural network based model predictive control for a steel pickling process, *Journal of process control* 19 (2009) 579–590.
- [37] N. J. Van Eck, L. Waltman, Vosviewer manual, *Manual for VOSviewer version 1* (2023).
- [38] A. S. Bazanella, L. Campestrini, D. Eckhard, The data-driven approach to classical control theory, *Annual Reviews in Control* 56 (2023) 100906.
- 835 [39] F. Dörfler, Data-driven control: Part one of two: A special issue sampling from a vast and dynamic landscape, *IEEE Control Systems Magazine* 43 (2023) 24–27.
- [40] F. Dörfler, Data-driven control: Part two of two: Hot take: Why not go with models?, *IEEE Control Systems Magazine* 43 (2023) 27–31.
- [41] M. Tanaskovic, L. Fagiano, R. Smith, P. Goulart, M. Morari, Adaptive model predictive control for constrained linear systems, in: *2013 European Control Conference (ECC)*, IEEE, pp. 382–387.
- 840 [42] M. Tanaskovic, D. Sturzenegger, R. Smith, M. Morari, Robust adaptive model predictive building climate control, *Ifac-Papersonline* 50 (2017) 1871–1876.
- [43] L. Hewing, J. Kabzan, M. N. Zeilinger, Cautious model predictive control using Gaussian process regression, *IEEE Transactions on Control Systems Technology* 28 (2019) 2736–2743.
- 845 [44] M. R. Askari, I. Hajizadeh, M. Rashid, N. Hobbs, V. M. Zavala, A. Cinar, Adaptive-learning model predictive control for complex physiological systems: Automated insulin delivery in diabetes, *Annual Reviews in Control* 50 (2020) 1–12.
- [45] M. Tanaskovic, L. Fagiano, R. Smith, M. Morari, Adaptive receding horizon control for constrained MIMO systems, *Automatica* 50 (2014) 3019–3029.
- [46] T. A. N. Heirung, B. E. Ydstie, B. Foss, Dual adaptive model predictive control, *Automatica* 80 (2017) 340–348.
- 850 [47] R. Murray-Smith, D. Sbarbaro, C. E. Rasmussen, A. Girard, Adaptive, cautious, predictive control with gaussian process

priors, IFAC Proceedings Volumes 36 (2003) 1155–1160.

- [48] L. Fagiano, G. Schildbach, M. Tanaskovic, M. Morari, Scenario and adaptive model predictive control of uncertain systems, IFAC-PapersOnLine 48 (2015) 352–359.
- [49] M. Lorenzen, M. Cannon, F. Allgöwer, Robust MPC with recursive model update, Automatica 103 (2019) 461–471.
- 855 [50] J. Köhler, E. Andina, R. Soloperto, M. A. Müller, F. Allgöwer, Linear robust adaptive model predictive control: Computational complexity and conservatism, in: IEEE 58th Conference on Decision and Control, IEEE, pp. 1383–1388.
- [51] J. Köhler, P. Kötting, R. Soloperto, F. Allgöwer, M. A. Müller, A robust adaptive model predictive control framework for nonlinear uncertain systems, International Journal of Robust and Nonlinear Control 31 (2021) 8725–8749.
- [52] M. M. Morato, J. E. Normey-Rico, O. Sename, Novel qLPV MPC design with least-squares scheduling prediction, IFAC-PapersOnLine 52 (2019) 158–163.
- 860 [53] A. Parsi, D. Liu, A. Iannelli, R. S. Smith, Dual adaptive MPC using an exact set-membership reformulation, 2022. [Pre-Print, Submitted].
- [54] S. Zhang, L. Dai, Y. Xia, Adaptive MPC for constrained systems with parameter uncertainty and additive disturbance, IET Control Theory & Applications 13 (2019) 2500–2506.
- 865 [55] J. Köehler, M. A. Müller, F. Allgöwer, A nonlinear model predictive control framework using reference generic terminal ingredients, IEEE Trans. on Automatic Control (2019).
- [56] M. M. Morato, M. Jungers, J. E. Normey-Rico, O. Sename, A predictive fault tolerant control method for qLPV systems subject to input faults and constraints, Journal of the Franklin Institute 359 (2022) 9129–9167.
- [57] K. Zhang, Y. Shi, Adaptive model predictive control for a class of constrained linear systems with parametric uncertainties, Automatica 117 (2020) 108974.
- 870 [58] A. Sasfi, M. N. Zeilinger, J. Köhler, Robust adaptive MPC using control contraction metrics, 2022. [Pre-Print, Submitted].
- [59] M. Bujarbaruah, X. Zhang, M. Tanaskovic, F. Borrelli, Adaptive stochastic MPC under time-varying uncertainty, IEEE Transactions on Automatic Control 66 (2020) 2840–2845.
- [60] A. Aboudonia, J. Lygeros, Adaptive learning-based model predictive control for uncertain interconnected systems: A set membership identification approach, 2024. [arXiv:2404.16514v1].
- 875 [61] J. R. Salvador, D. M. de la Peña, T. Alamo, A. Bemporad, Data-based predictive control via direct weight optimization, IFAC-PapersOnLine 51 (2018) 356–361.
- [62] R. Soloperto, J. Köhler, F. Allgöwer, Augmenting MPC schemes with active learning: Intuitive tuning and guaranteed performance, IEEE Control Systems Letters 4 (2020) 713–718.
- 880 [63] P. C. N. Verheijen, G. R. G. da Silva, M. Lazar, Data-driven rate-based integral predictive control with estimated prediction matrices, in: 25th International Conference on System Theory, Control and Computing, IEEE, pp. 630–636.
- [64] L. Hewing, K. P. Wabersich, M. Menner, M. N. Zeilinger, Learning-based model predictive control: Toward safe learning in control, Annual Review of Control, Robotics, and Autonomous Systems 3 (2020) 269–296.
- [65] D. Ernst, M. Glavic, F. Capitanescu, L. Wehenkel, Reinforcement learning versus model predictive control: a comparison on a power system problem, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 39 (2008) 517–529.
- 885 [66] U. Rosolia, F. Borrelli, Learning model predictive control for iterative tasks. a data-driven control framework, IEEE Transactions on Automatic Control 63 (2017) 1883–1896.
- [67] A. Jain, F. Smarra, R. Mangharam, Data predictive control using regression trees and ensemble learning, in: 2017 IEEE 56th annual Conference on Decision and Control (CDC), IEEE, pp. 4446–4451.
- 890 [68] G. Williams, N. Wagener, B. Goldfain, P. Drews, J. M. Rehg, B. Boots, E. A. Theodorou, Information theoretic MPC for model-based reinforcement learning, in: 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp. 1714–1721.
- [69] U. Rosolia, X. Zhang, F. Borrelli, Data-driven predictive control for autonomous systems, Annual Review of Control, Robotics, and Autonomous Systems 1 (2018) 259–286.
- 895 [70] G. Ceusters, R. C. Rodríguez, A. B. García, R. Franke, G. Deconinck, L. Helsen, A. Nowé, M. Messagie, L. R. Camargo, Model-predictive control and reinforcement learning in multi-energy system case studies, Applied Energy 303 (2021) 117634.
- [71] M. Lin, Z. Sun, Y. Xia, J. Zhang, Reinforcement learning-based model predictive control for discrete-time systems, IEEE Transactions on Neural Networks and Learning Systems (2023).
- 900 [72] S. Gros, M. Zanon, Data-driven economic NMPC using reinforcement learning, IEEE Transactions on Automatic Control 65 (2019) 636–648.
- [73] H. Shah, M. Gopal, Model-free predictive control of nonlinear processes based on reinforcement learning, IFAC-PapersOnLine 49 (2016) 89–94.
- 905 [74] C. J. Ostafew, A. P. Schoellig, T. D. Barfoot, Robust constrained learning-based NMPC enabling reliable mobile robot path tracking, The International Journal of Robotics Research 35 (2016) 1547–1563.
- [75] T. Koller, F. Berkenkamp, M. Turchetta, A. Krause, Learning-based model predictive control for safe exploration, in: IEEE 57th Conference on Decision and Control, IEEE, pp. 6059–6066.
- [76] R. Murray, M. Palladino, A model for system uncertainty in reinforcement learning, Systems & Control Letters 122 (2018) 24–31.
- 910 [77] B. Amos, I. Jimenez, J. Sacks, B. Boots, J. Z. Kolter, Differentiable MPC for end-to-end planning and control, Advances in neural information processing systems 31 (2018).
- [78] M. Zanon, S. Gros, A. Bemporad, Practical reinforcement learning of stabilizing economic MPC, in: 2019 18th European Control Conference (ECC), IEEE, pp. 2258–2263.
- 915 [79] G. Dalal, K. Dvijotham, M. Vecerik, T. Hester, C. Paduraru, Y. Tassa, Safe exploration in continuous action spaces,

2018. [Pre-Print].
- [80] T.-H. Pham, G. De Magistris, R. Tachibana, Optlayer-practical constrained optimization for deep reinforcement learning in the real world, in: International Conference on Robotics and Automation (ICRA), IEEE, pp. 6236–6243.
- [81] L. P. Kaelbling, M. L. Littman, A. W. Moore, Reinforcement learning: A survey, *Journal of artificial intelligence research* 4 (1996) 237–285.
- [82] T. Haarnoja, S. Ha, A. Zhou, J. Tan, G. Tucker, S. Levine, Learning to walk via deep reinforcement learning, 2018. [Pre-Print].
- [83] K. Xenou, G. Chalkiadakis, S. Afantenos, Deep reinforcement learning in strategic board game environments, in: Multi-Agent Systems: 16th European Conference, Revised Selected Papers, Springer, pp. 233–248.
- [84] J. Garcia, F. Fernández, A comprehensive survey on safe reinforcement learning, *Journal of Machine Learning Research* 16 (2015) 1437–1480.
- [85] J. Drgoňa, K. Kiš, A. Tuor, D. Vrabie, M. Klaučo, Differentiable predictive control: Deep learning alternative to explicit model predictive control for unknown nonlinear systems, *Journal of Process Control* 116 (2022) 80–92.
- [86] F. L. Lewis, D. Vrabie, Reinforcement learning and adaptive dynamic programming for feedback control, *IEEE circuits and systems magazine* 9 (2009) 32–50.
- [87] F. L. Lewis, D. Vrabie, K. G. Vamvoudakis, Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers, *IEEE Control Systems Magazine* 32 (2012) 76–105.
- [88] M. Brunner, U. Rosolia, J. Gonzales, F. Borrelli, Repetitive learning model predictive control: An autonomous racing example, in: 2017 IEEE 56th annual conference on decision and control (CDC), IEEE, pp. 2545–2550.
- [89] G. Bellegarda, K. Byl, An online training method for augmenting MPC with deep reinforcement learning, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, pp. 5453–5459.
- [90] H. N. Esfahani, A. B. Kordabad, S. Gros, Reinforcement learning based on MPC/MHE for unmodeled and partially observable dynamics, in: 2021 American Control Conference (ACC), IEEE, pp. 2121–2126.
- [91] U. Rosolia, A. D. Ames, Iterative model predictive control for piecewise systems, *IEEE Control Systems Letters* 6 (2021) 842–847.
- [92] S. Dean, S. Tu, N. Matni, B. Recht, Safely learning to control the constrained linear quadratic regulator, in: 2019 American Control Conference (ACC), IEEE, pp. 5582–5588.
- [93] S. Gros, M. Zanon, Reinforcement learning for Mixed-Integer problems based on MPC, *IFAC-PapersOnLine* 53 (2020) 5219–5224.
- [94] M. Maiworm, D. Limon, R. Findeisen, Online learning-based model predictive control with Gaussian process models and stability guarantees, *International Journal of Robust and Nonlinear Control* 31 (2021) 8785–8812.
- [95] A. B. Martinsen, A. M. Lekkas, S. Gros, Combining system identification with reinforcement learning-based mpc, *IFAC-PapersOnLine* 53 (2020) 8130–8135.
- [96] K. P. Wabersich, M. N. Zeilinger, Cautious Bayesian MPC: Regret analysis and bounds on the number of unsafe learning episodes, *IEEE Transactions on Automatic Control* (2022).
- [97] S. Sawant, A. S. Anand, D. Reinhardt, S. Gros, Learning-based MPC from big data using reinforcement learning, 2023. [Pre-Print].
- [98] W. Cai, A. B. Kordabad, H. N. Esfahani, A. M. Lekkas, S. Gros, MPC-based reinforcement learning for a simplified freight mission of autonomous surface vehicles, in: 2021 60th IEEE Conference on Decision and Control (CDC), IEEE, pp. 2990–2995.
- [99] A. B. Kordabad, H. N. Esfahani, A. M. Lekkas, S. Gros, Reinforcement learning based on scenario-tree MPC for ASVs, in: 2021 American Control Conference (ACC), IEEE, pp. 1985–1990.
- [100] B. Recht, A tour of reinforcement learning: The view from continuous control, *Annual Review of Control, Robotics, and Autonomous Systems* 2 (2019) 253–279.
- [101] J. Berberich, F. Allgöwer, A trajectory-based framework for data-driven system analysis and control, in: 2020 European Control Conference (ECC), IEEE, pp. 1365–1370.
- [102] A. Romer, J. Berberich, J. Köhler, F. Allgöwer, One-shot verification of dissipativity properties from input–output data, *IEEE Control Systems Letters* 3 (2019) 709–714.
- [103] J. Berberich, A. Koch, C. W. Scherer, F. Allgöwer, Robust data-driven state-feedback design, in: 2020 American Control Conference (ACC), IEEE, pp. 1532–1538.
- [104] A. Koch, J. Berberich, J. Köhler, F. Allgöwer, Determining optimal input–output properties: A data-driven approach, *Automatica* 134 (2021) 109906.
- [105] J. Bongard, J. Berberich, J. Köhler, F. Allgöwer, Robust stability analysis of a simple data-driven model predictive control approach, *IEEE Transactions on Automatic Control* (2022).
- [106] C. Verhoek, R. Tóth, S. Haesaert, A. Koch, Fundamental lemma for data-driven analysis of linear parameter-varying systems, in: 60th IEEE Conference on Decision and Control (CDC 2021), pp. 5033–5039.
- [107] P. C. N. Verheijen, V. Breschi, M. Lazar, Handbook of linear data-driven predictive control: Theory, implementation and design, *Annual Reviews in Control* 56 (2023) 100914.
- [108] M. M. Morato, J. E. Normey-Rico, O. Sename, Nonlinear data-driven control Part I: An overview of trajectory representations, *Journal of Control, Automation and Electrical Systems* (2024) 1–19.
- [109] M. Klädtke, M. S. Darup, Implicit predictors in regularized data-driven predictive control, *IEEE Control Systems Letters* (2023).
- [110] J. G. Rueda-Escobedo, J. Schiffer, Data-driven internal model control of second-order discrete Volterra systems, in: 59th IEEE Conference on Decision and Control, IEEE, pp. 4572–4579.
- [111] M. Alsalti, J. Berberich, V. G. Lopez, F. Allgöwer, M. A. Müller, Data-based system analysis and control of flat nonlinear

- systems, in: 60th IEEE Conference on Decision and Control, IEEE, pp. 1484–1489.
- [112] Y. Lian, R. Wang, C. N. Jones, Koopman based data-driven predictive control, 2021. [Pre-Print, arXiv: 2102.05122].
- [113] R. Strässer, J. Berberich, F. Allgöwer, Data-driven control of nonlinear systems: Beyond polynomial dynamics, in: 2021 60th IEEE Conference on Decision and Control (CDC), IEEE, pp. 4344–4351.
- 985 [114] J. Berberich, J. Köhler, M. A. Müller, F. Allgöwer, Data-driven model predictive control with stability and robustness guarantees, *IEEE Transactions on Automatic Control* 66 (2020) 1702–1717.
- [115] J. Berberich, J. Köhler, M. A. Müller, F. Allgöwer, Linear tracking MPC for nonlinear systems part ii: The data-driven case, *IEEE Transactions on Automatic Control* (2022).
- [116] J. Coulson, J. Lygeros, F. Dörfler, Regularized and distributionally robust data-enabled predictive control, in: 2019 990 IEEE 58th Conference on Decision and Control (CDC), IEEE, pp. 2696–2701.
- [117] E. Elokda, J. Coulson, P. N. Beuchat, J. Lygeros, F. Dörfler, Data-enabled predictive control for quadcopters, *International Journal of Robust and Nonlinear Control* 31 (2021) 8916–8936.
- [118] M. M. Morato, V. M. Cunha, T. L. M. Santos, J. E. Normey-Rico, O. Sename, A robust nonlinear tracking MPC using qLPV embedding and zonotopic uncertainty propagation, 2023. [Pre-Print, Submitted].
- 995 [119] J. Berberich, J. Köhler, M. A. Müller, F. Allgöwer, Stability in data-driven MPC: an inherent robustness perspective, 2022. [Pre-Print, arXiv: 2205.11859].
- [120] J. Berberich, J. Köhler, M. A. Müller, F. Allgöwer, Robust constraint satisfaction in data-driven MPC, in: 2020 59th IEEE Conference on Decision and Control (CDC), IEEE, pp. 1260–1267.
- [121] A. S. Anand, S. Sawant, D. Reinhardt, S. Gros, Data-driven predictive control and MPC: Do we achieve optimality?, 1000 2024. [Pre-Print, Available at arXiv: 2405.17892].
- [122] S. Baros, C.-Y. Chang, G. E. Colon-Reyes, A. Bernstein, Online data-enabled predictive control, *Automatica* 138 (2022) 109926.
- [123] K. Zhang, Y. Zheng, C. Shang, Z. Li, Dimension reduction for efficient data-enabled predictive control, *IEEE Control Systems Letters* (2023).
- 1005 [124] R. Dinkla, S. P. Mulders, J.-W. van Wingerden, T. Oomen, Closed-loop aspects of Data-enabled Predictive Control, *IFAC-PapersOnLine* 56 (2023) 1388–1393.