Adaptive Robust Model Predictive Control with Matched and Unmatched Uncertainty

Rohan Sinha, James Harrison, Spencer M. Richards, and Marco Pavone

Abstract— **We propose a learning-based robust predictive control algorithm that can handle large uncertainty in the dynamics for a class of discrete-time systems that are nominally linear with an additive nonlinear dynamics component. Such systems commonly model the nonlinear effects of an unknown environment on a nominal system. Motivated by an inability of existing learning-based predictive control algorithms to achieve safety guarantees in the presence of uncertainties of large magnitude in this setting, we achieve significant performance improvements by optimizing over a novel class of nonlinear feedback policies inspired by certainty equivalent "estimate-and-cancel" control laws pioneered in classical adaptive control. In contrast with previous work in robust adaptive MPC, this allows us to take advantage of the structure in the a priori unknown dynamics that are learned online through function approximation. Our approach also extends typical nonlinear adaptive control methods to systems with state and input constraints even when an additive uncertain function cannot directly be canceled from the dynamics. Moreover, our approach allows us to apply contemporary statistical estimation techniques to certify the safety of the system through persistent constraint satisfaction with high probability. We show that our method allows us to consider larger unknown terms in the dynamics than existing methods through simulated examples.**

I. Introduction

Learning-based control offers promising methods to enable the deployment of autonomous systems in diverse, dynamic environments. Such methods do so by learning from data to improve closed-loop performance over time. Upon deployment, these methods should provide safety guarantees and quickly adapt in the face of uncertainty; to this end, estimates of system uncertainties must be maintained and updated as new data becomes available. However, many recently proposed learning-based control algorithms rely on uncertainty estimation methods that are either too conservative (e.g., yielding limited performance to remain safe) or too fragile (e.g., in the face of large uncertainties).

Related Work. We briefly review two significant paradigms for the control of uncertain systems, namely *adaptive control* and *robust* control. We then discuss works that combine ideas from both paradigms, before introducing our contributions.

Adaptive control concerns the joint design of a parametric feedback controller and a parameter adaptation law to improve closed-loop performance over time when the dynamics are (partially) unknown [1], [2]. Such design for nonlinear systems commonly relies on the ability to express unknown dynamics terms as linear combinations of *known basis functions*, i.e., *features*. The adaptation law updates the weights online, and the controller applies part of the control input to cancel the corresponding estimate of the unknown dynamics [1], [3], [4]. This can achieve tracking convergence up to an error threshold that depends on the representation capacity of the features relative to the true dynamics [3], [4]. Thus, recent works have sought to blend high-capacity parametric or nonparametric models from machine learning with classical adaptive control designs. This includes the use of deep neural networks via online back-propagation [5], Gaussian processes [6], and Bayesian neural networks [7], [8] via online Bayesian updates and meta-learned features [9], [10]. However, these approaches are fundamentally limited by common assumptions in classical adaptive control, namely that uncertain dynamics terms can be stably cancelled by the control input, i.e., that these terms are so-called *matched uncertainties* [1]–[4]. Moreover, most of these works do not explicitly consider state and input constraints, which are essential to safe control in practice.

Robust control generally seeks to achieve consistent performance despite uncertainty in the dynamics. In particular, robust Model Predictive Control (MPC) algorithms consider the control of a system subject to bounded uncertain dynamics terms, i.e., *disturbances*, and explicit state and input constraints as an optimization program. Some methods optimize the worst-case performance of the controller [11], while others tighten the constraints to accommodate the trajectory "tubes" induced by the disturbances and optimize the nominal predicted trajectory instead [12], [13]. To account for future information gain and reduce conservatism, these methods either fix a disturbance recourse policy [12] or optimize over state feedback policies [14].

Adaptive robust MPC incorporates the online improvement of adaptive control methods into robust MPC to explicitly handle constraints and process noise during learning [15]–[21]. A straightforward approach is to maintain an estimate of the bounds on an unknown dynamics component over the state space and use those bounds in any chosen robust MPC scheme [17], [18], [20], [21]. However, robust MPC algorithms are only able to tolerate a comparatively mild amount of bounded uncertainty. Therefore, most existing adaptive MPC algorithms cannot be used if the true support of the unknown dynamics is too large because they do not explicitly take advantage of the structure in the learned dynamics terms. In addition, these algorithms often require specific assumptions on the estimation procedure to guarantee the bounds become less conservative over time, even though those assumptions are not satisfied by common statistical

The authors are with Stanford University, Stanford, CA. {rhnsinha, jharrison, spenrich, pavone}@stanford.edu

estimators like least-squares [15]–[17], [20]. For example, some approaches try to identify constant or slowly varying offsets with set-membership estimation [15], [16], even though this can be more fragile than statistical methods to outliers or model misspecification. Overall, existing adaptive robust MPC methods use only an approximate outer envelope of unknown nonlinear dynamics terms, thereby rendering them overly conservative.

Contributions. In this work, we present an adaptive robust MPC method for systems with unknown nonlinear dynamics terms with *unknown support*, subject to state and input constraints. Rather than construct an outer envelope for such terms, we develop theoretical guarantees for a broad class of *function approximators*, including set membership and least-squares methods for certain noise models. Our insight is to decompose uncertain dynamics terms into a *matched* component, which lies in a subspace that can be stably cancelled by the control input, and an *unmatched* component, which lies in an orthogonal complement to this subspace. We use certainty equivalent adaptive control to stably cancel the *matched* component and then apply robust MPC by treating the unmatched component as a bounded disturbance. We prove our method is both recursively feasible and input-to-state stable. Moreover, we demonstrate on a variety of simulated systems that our method reduces the conservatism and increases the feasible domain of the resulting robust MPC problem in comparison to existing robust MPC methods.

II. Problem Formulation

We consider nonlinear, discrete-time systems of the form

$$x_{t+1} = Ax_t + Bu_t + f(x_t) + v_t, \tag{1}$$

where $x_t \in \mathbb{R}^n$ is the system state, $u_t \in \mathbb{R}^m$ is the control input, $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are known matrices, $v_t \in \mathcal{V}$ is a disturbance in some known, bounded *convex* set $\mathcal{V} \subset \mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}^n$ is an *unknown* nonlinear dynamics term. We assume that B has full column rank to guarantee the existence of the Moore-Penrose pseudo-inverse $B^\dagger = (B^\top B)^{-1}B^\top$. This assumption is not restrictive, as it simply ensures there are no redundant actuators. We want to control the system state from some initial condition $x_0 \in \mathcal{X}$ to zero according to the robust optimal control problem

$$\begin{aligned} J(x_0) = \min_{x_t,u_t} \max_{v_t} & \sum_{t=0}^{\infty} h(x_t,u_t) \\ \text{s.t. } & x_{t+1} = Ax_t + Bu_t + f(x_t) + v_t \\ & u_t \in \mathcal{U},\ x_t \in \mathcal{X},\ v_t \in \mathcal{V},\ \forall t \in \mathbb{N}_+ \end{aligned}, \tag{2}$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{U} \subseteq \mathbb{R}^m$ are *convex* constraint sets, and $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ is a jointly continuous, positive-definite, convex stage cost function. The infinite horizon and the non-convex dynamics constraints caused by the nonlinearity of f make (2) intractable to solve in its current form. As is standard in MPC, in §III we will detail feedback policies that solve a finite horizon approximation of (2). Since a worst-case objective as (2) is intractable for general disturbance bounds, we consider policies that optimize the nominal objective instead. In addition, we restrict the unknown dynamics f to a class of parametric nonlinear functions that will allow us to construct a provably safe, learning-based control algorithm in §III.

Function Approximation of Uncertain Dynamics. We assume f is linearly parameterizable, i.e.,

$$f(x) = W^*\phi(x), \tag{3}$$

where $\phi : \mathbb{R}^n \to \mathbb{R}^d$ is a *known* nonlinear feature map, and $W^* \in \mathbb{R}^{n\times d}$ is a matrix of *unknown* weights. We also assume that at $t = 0$ we have an initial estimate $\widehat{W}(0)$ of W^*, and a confidence interval $\mathcal{W}(0) = \mathcal{W}_1(0)\times\cdots\times\mathcal{W}_n(0)$ such that $\hat{w}_i - w_i^* \in \mathcal{W}_i(0)$ for each $i \in \{1,\dots,n\}$, where $\hat{w}_i(0), w_i^* \in \mathbb{R}^d$ are rows of $\widehat{W}$ and W^*, respectively.

We assume the features $\phi(x)$ are bounded for all $x \in \mathcal{X}$, which still allows for function approximators such as neural networks with sigmoid output activation units. More specifically, we make the following assumption:

Assumption 1: The feature map $\phi : \mathbb{R}^n \to \mathbb{R}^d$ is *unity bounded* on $\mathcal{X}$, i.e., $\|\phi(x)\|_2 \le 1$ for all $x \in \mathcal{X}$.

Matched and Unmatched Uncertainty. Distinguishing how much of $f(x)$ can be stably cancelled by the control input u will be critical to our approach in §III. Therefore, we classify the dynamic uncertainty $f(x)$ as follows:

Definition 1: The term $f(x)$ in (1) is a *matched uncertainty* if $f(x) \in \operatorname{Range}(B)$ for all $x \in \mathcal{X}$. If instead $f(x) \notin \operatorname{Range}(B)$ for all $x \in \mathcal{X}$, then $f(x)$ is an *unmatched uncertainty*.

If $f(x)$ is a matched uncertainty, then there exists a unique function $g : \mathbb{R}^n \to \mathbb{R}^m$ such that $f(x) = Bg(x)$ for all $x \in \mathcal{X}$ since we assume B has full column rank. Critically, in this case we could cancel $f(x)$ from (1) if we knew $g(x)$ by setting $u = -g(x)$; in §III, we will consider both this setting and the case where $f(x)$ must be decomposed into matched and unmatched parts.

ISS Stability. The disturbances $\{v_t\}_{t=0}^{\infty}$ in (1) make it impossible to reason about the asymptotic stability of the closed-loop system. Therefore, we briefly review relevant results of Input-to-State Stability (ISS) theory, which is often used to analyze robust control algorithms [13], [14], [17].

Definition 2: [22, Def. 1] The system $x_{t+1} = g(t, x_t, v_t)$ with external disturbance v_t is globally *input-to-state stable (ISS)* if there exists a class-$\mathcal{KL}$ function $\beta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ and a class-$\mathcal{K}$ function $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\|x_t\| \le \beta(\|x_0\|, t) + \gamma(\|v_t\|_{\mathcal{L}_\infty}),\ \ \forall x_0 \in \mathbb{R}^n\ t \ge 0, \tag{4}$$

where $\|v_t\|_{\mathcal{L}_\infty}$ is the *signal norm*

$$\|v_t\|_{\mathcal{L}_\infty} = \sup_{k\in\{0,\dots,t\}} \|v_k\|. \tag{5}$$

In essence, ISS requires that the nominal system is asymptotically stable and that the influence of the disturbances is bounded. This makes it a convenient framework to analyze the stability of systems subject to random disturbances. Similarly to regular nonlinear stability analysis, we can show a system is ISS if there exists an ISS-Lyapunov function. Since our algorithms are adaptive, the closed-loop system is time-varying. To keep consistency with standard ISS notation [23], we use a more restricted version of the results in [22].

Definition 3: [22] The function $V : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is an *ISS-Lyapunov function* for the system $x_{t+1} = g(t, x_t, v_t)$ if it is continuous at the origin for all $t \geq 0$ and there exist three class-$\mathcal{K}_\infty$ functions $\alpha_1, \alpha_2, \alpha_3$ and a class-$\mathcal{K}$ function σ such that $\forall x_t \in \mathbb{R}^n$

$$\begin{aligned} \alpha_1(\|x_t\|) \leq V(t, x_t) \leq \alpha_2(\|x_t\|), \\ V(t+1, g(t, x_t, v_t)) - V(t, x_t) \leq -\alpha_3(\|x_t\|) + \sigma(\|v_t\|). \end{aligned} \tag{6}$$

Theorem 1: [22, Thm. 1] A system is globally ISS if and only if the system admits an ISS-Lyapunov function.

The above definitions naturally extend to local ISS stability; for a detailed discussion, we refer readers to [13], [22], [23].

III. Proposed Approach

We present our approach in two stages. First, we introduce the idea behind our algorithm on a restricted class of *matched* unknown functions in the dynamics. We then generalize this approach and construct a robust adaptive MPC.

A. The Fully Matched Uncertainty Case

Assuming *matched uncertainty* is common in adaptive control for a simple reason: it allows us to estimate $g(x_t)$ and cancel it from the dynamics [3], [4]. We can extend standard adaptive control approaches for systems with matched uncertainty to a constrained discrete-time predictive control setting by introducing a simple nonlinear recourse law we refer to as a certainty equivalent controller.

Definition 4: The set of *certainty equivalent* controllers for a *matched system* is the function class whose elements $\pi : \mathbb{R}^n \to \mathbb{R}^m$ are of the form

$$\pi(x_t) = u^*(x_t) - \hat{g}(x_t), \tag{7}$$

where $\hat{g}(x_t) = \widehat{W}\phi(x_t)$ with $\widehat{W} \in \mathbb{R}^{m \times d}$ is an estimate of the true dynamics $g(x_t)$ and $u^*(x_t)$ is a robust control input that we will define in (20) as the solution of a robust MPC problem. Under a certainty equivalent feedback law (7), the closed loop dynamics of the *matched system* become

$$\begin{aligned} x_{t+1} &= Ax_t + B(\pi(x_t) + g(x_t)) + v_t \\ &= Ax_t + Bu^*(x_t) - B(\hat{g}(x_t) - g(x_t)) + v_t \\ &= Ax_t + Bu^*(x_t) + d_t. \end{aligned} \tag{8}$$

We use the *compound disturbance* d_t to represent the effect of both the estimation error and the regular disturbance. The dynamics (8) under the *certainty equivalent* control law (7) make it explicit that the control performance no longer depends on the value of $g(x_t)$, but rather on the quality of our approximation $\hat{g}(x_t)$. Moreover, if we refine the estimate online, yielding time-varying policies $\pi_t(\cdot) = u_t^*(\cdot) - \hat{g}_t(\cdot)$, the resulting adaptive controller will improve its performance over time. This will allow us to achieve safe control performance subject to much larger disturbances than feasible under a naive robust MPC approach.

To guarantee safety and performance under the *certainty equivalent* control law (7), we need to:

1) Guarantee the closed loop dynamics (8) do not result in constraint violations through a standard robust MPC approach.
2) Ensure that the *certainty equivalent* cancellation of $g(x_t)$ does not result in input constraint violations.

We start by assuming that we have an online function estimation (or adaptation) strategy that guarantees that bounds on the quality of the parameter estimate $\widehat{W}(t)$ can only improve over time. This will guide the control design and we will discuss two commonplace function estimators that satisfy these properties in §IV.

Assumption 2: We have an *online parameter estimator* that takes the initial estimate $\widehat{W}(0)$, confidence interval $\mathcal{W}(0)$, and trajectory history $\mathcal{H}(t) = \{x_0, u_0, \ldots, x_t, u_t\}$ and returns an online parameter estimate $\widehat{W}(t)$ and confidence interval $\mathcal{W}(t)$ such that $\hat{w}_i(t) - w_i^* \in \mathcal{W}_i(t) \subseteq \mathcal{W}_i(t-1)$.

Assumption 2 does not require the point estimate $\widehat{W}$ to remain constant in time, and therefore the range of the estimated function $\hat{f}_t(x) = \widehat{W}(t)\phi(x)$ can change over time to include values that previously were not. This differs from approaches such as [15], [20] where the support of the unknown dynamics function is approximated as shrinking. We will see that Assumption 2 is more reflective of standard function estimation schemes in §IV.

B. Certainty Equivalent Adaptation for General Systems

We now generalize the certainty equivalent control laws (7) for matched systems to the unmatched dynamics considered in (1) and use them to construct an effective MPC with safety guarantees for system (1). We propose a class of nonlinear feedback policies that cancel as much of the nonlinear dynamics as possible for the general system with unmatched uncertainty (1).

Definition 5: We define the set of *matching certainty equivalent policies* as the time-varying function class whose elements $\pi_t : \mathcal{X} \to \mathcal{U}$ at time t are of the form

$$\pi_t(x_t) = u_t^*(x_t) - K_m \hat{f}_t(x_t) \tag{9}$$

where $K_m \in \mathbb{R}^{m \times n}$ is the fixed *matching* feedback gain.

In this work, we take inspiration from adaptive control and take K_m as the projection of $\hat{f}_t(x_t)$ onto $\mathrm{Range}(B)$, cancelling out as much of the disturbance as we can in the 2-norm sense. Therefore, we take $K_m = B^\dagger$, such that

$$K_m \hat{f}_t(x_t) = \arg\min_z \|Bz - \hat{f}_t(x_t)\|.$$

Remark 1: If the unknown dynamics are matched, then $K_m \hat{f}_t(x) = B^\dagger B \hat{g}_t(x) = \hat{g}_t(x)$. Therefore the matching certainty equivalent policy (9) generalizes approaches for systems with matched uncertainty to those with unmatched uncertainty.

To construct a robust adaptive MPC, we first identify a convex approximation of the range of values $\hat{f}_t(x_t)$ can take. We take a simple approach and make a polytopic approximation that is guaranteed to improve over time.

Lemma 1: If we approximate $f(x_t)$ online with features that satisfy Assumption 1 and an estimator that satisfies Assumption 2, then for any $x \in \mathcal{X}$ from time t onwards

$$\hat{f}_{t+k}(x), f(x) \in \mathcal{F}_t \quad \forall k \geq 0 \tag{10}$$

where the set $\mathcal{F}_t$ is defined as

$$\mathcal{F}_t = \{z \in \mathbb{R}^n : |z_i| \leq \|\hat{w}_i(t)\| + 2 \max_{\tilde{w}_i \in \mathcal{W}_i(t)} \|\tilde{w}_i\|\} \tag{11}$$

Proof: We show that $\mathcal{F}_t$ over-approximates the range of values $z = \hat{f}_{t+k}(x)$ can take for any $x \in \mathcal{X}$ with $t+k \geq t$. Note that

$$\begin{aligned} |z_i| &= |\hat{w}_i(t+k)^\top \phi(x)| \\ &\leq \|\hat{w}_i(t+k)\| \\ &\leq \|\hat{w}_i(t)\| + \|\hat{w}_i(t+k) - \hat{w}_i(t)\| \\ &\leq \|\hat{w}_i(t)\| + \|\hat{w}_i(t+k) - w_i^*\| + \|\hat{w}_i(t) - w_i^*\| \end{aligned} \tag{12}$$

Assumption 2 gives $\mathcal{W}_i(t+k) \subseteq \mathcal{W}_i(t)$ for $k \geq 0$, therefore we have that

$$|z_i| \leq \|\hat{w}_i(t)\| + 2 \max_{\tilde{w}_i \in \mathcal{W}_i(t)} \|\tilde{w}_i\|, \tag{13}$$

which proves that $\hat{f}_{t+k}(x) \in \mathcal{F}_t$ for all $k \geq 0$. In addition, let $y = f(x)$. Then,

$$\begin{aligned} |y_i| &= |{w_i^*}^\top \phi(x)| \\ &\leq \|w_i^*\| \\ &\leq \|\hat{w}_i(t)\| + \|\hat{w}_i(t) - w_i^*\| \\ &\leq \|\hat{w}_i(t)\| + \max_{\tilde{w}_i \in \mathcal{W}_i(t)} \|\tilde{w}_i\|. \end{aligned} \tag{14}$$

Hence $f(x) \in \mathcal{F}_t$. ■

It is not possible to create a tighter approximation (i.e., eliminate the factor of 2) without additional assumptions. To see this, consider a unit norm ball confidence interval constant in time. In the worst case, the true parameter lies on the boundary of the ball around the current estimate. This means all future estimates may lie a Euclidean distance of 2 units away from the current estimate, yielding the bound in Lemma 1.

Remark 2: Solving for the set $\mathcal{F}_t$ defined in Lemma 1 requires finding the max-norm element of a convex set. For many convex confidence intervals, such as ellipsoids, this is easy to compute. For polytopes, this requires vertex enumeration, incurring exponential complexity in the dimension of the state space n.

Moreover, Lemma 1 does not require that $\mathcal{F}_{t+1} \subseteq \mathcal{F}_t$, so we need to do more work to create an approximation that is non-increasing in size. We have the following corollary.

Corollary 1: At time t, $\mathcal{F}_i$ are known for $i = 0, \ldots, t$. Therefore, Lemma 1 implies for all $x \in \mathcal{X}$ that we must have

$$\hat{f}_{t+k}(x), f(x) \in \bigcap_{i=0}^{t} \mathcal{F}_i =: \widehat{\mathcal{F}}_t \tag{15}$$

for all $k \geq 0$.

$\widehat{\mathcal{F}}_t$ is easily defined as

$$\widehat{\mathcal{F}}_t = \{z : |z_i| \leq \min_{j \in \{0,\ldots,t\}} [\|\hat{w}_i(j)\| + 2 \max_{\tilde{w}_i \in \mathcal{W}_i(j)} \|\tilde{w}_i\|]\} \tag{16}$$

and can be computed recursively in time.

The matching certainty equivalent (CE) law (9) results in the following closed-loop dynamics:

$$\begin{aligned} x_{t+1} &= Ax_t + B\pi(x_t)) + f(x_t) + v_t \\ &= Ax_t + Bu^*(x_t) + v_t \\ &\quad + (I - BB^\dagger)f(x_t) + BB^\dagger(f(x_t) - \hat{f}_t(x_t)) \\ &= Ax_t + Bu^*(x_t) + d_t, \end{aligned} \tag{17}$$

where we use d_t as shorthand to consolidate disturbances driven by the process disturbance v_t, the function estimation error $\hat{f}_t(\cdot) - f(\cdot)$, and the imperfect matching using $K_m = B^\dagger$. (17) makes it clear that under the CE law (9) the closed loop dynamics are driven by the estimation error and the component of the unknown function that cannot be cancelled from the dynamics. We now identify an approximation of the support of the compound disturbance d_t.

Lemma 2: Assume the function estimator satisfies Assumption 2 with features that satisfy the boundedness Assumption 1. If we control the system (1) using the certainty equivalent control law (9), then for all $k \geq 0$, the compound disturbance d_{t+k} in the dynamics (17) is contained in the set $\widehat{\mathcal{D}}_t \subseteq \widehat{\mathcal{D}}_{t-1}$, defined as

$$\widehat{\mathcal{D}}_t = (I - BB^\dagger)\widehat{\mathcal{F}}_t \oplus BB^\dagger \mathcal{D}_t \oplus \mathcal{V}. \tag{18}$$

Here $\oplus$ indicates the Minkowski sum operator, a matrix-set multiplication indicates the linear transformation of the set's elements, and the approximation error support $\mathcal{D}_t$ is given as

$$\mathcal{D}_t = \{z \in \mathbb{R}^n : |z_i| \leq \max_{\tilde{w}_i \in \mathcal{W}_i} \|\tilde{w}_i\|, \forall i \in \{1, \ldots, n\}\}. \tag{19}$$

Proof: Let the approximation error at a state $x \in \mathcal{X}$ at time $t+k$ be $z = \hat{f}_{t+k}(x) - f(x) = (\widehat{W}(t+k) - W^*)\phi(x) = \widetilde{W}\phi(x)$ for some $\widetilde{W} \in \mathcal{W}(t+k) \subseteq \mathcal{W}(t)$. Then

$$|z_i| = |\tilde{w}_i^\top \phi(x_t)| \leq \|\tilde{w}_i\| \leq \max_{\tilde{w}_i \in \mathcal{W}_i(t)} \|\tilde{w}_i\|.$$

Hence $\hat{f}_{t+k}(x) - f(x) \in \mathcal{D}_t$ for all $k \geq 0$. Since by Assumption 2 we have $\mathcal{W}_i(t) \subseteq \mathcal{W}_i(t-1)$, this immediately implies $\mathcal{D}_t \subseteq \mathcal{D}_{t-1}$. Then, by Corollary 1, the nonlinear dynamics $f(x) \in \widehat{\mathcal{F}}_t$, which implies the compound disturbance d_t in the closed loop dynamics (17) is contained in $\widehat{\mathcal{D}}_t$. In addition, since both $\widehat{\mathcal{F}}_t \subseteq \widehat{\mathcal{F}}_{t-1}$ and $\mathcal{D}_t \subseteq \mathcal{D}_{t-1}$, we have that the support of the compound disturbance is nested over time: $\widehat{\mathcal{D}}_t \subseteq \widehat{\mathcal{D}}_{t-1}$. ■

Remark 3: To interpret (18), consider the case when the dynamics have *matched uncertainty*. We can then write $\widehat{\mathcal{F}}_t = B\widehat{\mathcal{G}}_t$ and $\mathcal{D}_t = B\bar{\mathcal{D}}_t$ for two analogous approximations of the support and approximation error $\widehat{\mathcal{G}}_t, \bar{\mathcal{D}}_t \subseteq \mathbb{R}^m$ for $f(\cdot) = Bg(\cdot)$. In this case, $\widehat{\mathcal{D}}_t$ reduces to $B\bar{\mathcal{D}}_t \oplus \mathcal{V}$, showing us the compound disturbance only depends on the quality of the estimate $\hat{g}(\cdot)$ as in (8) rather than its magnitude.

Remark 4: If the estimation error is small, (18) shows that the compound disturbances d_t will only depend on the process noise and the component of the unknown function f that cannot be cancelled using the CE control law (9). Therefore, if the estimation error is small, our method will tolerate nonlinear disturbances with significantly larger magnitude than a conventional robust MPC scheme.

We can now define the constraint-tightened robust MPC problem we solve online as the following constrained finite time optimal control problem:

$$\begin{aligned} J_N^*(t,x_t) = & \min_{u_{t:t+N-1|t}(\cdot)} \sum_{k=0}^{N-1} h(\bar{x}_{t+k|t},\bar{u}_{t+k|t}) + V_N(\bar{x}_{t+N|t}) \\ \text{s.t.} \quad & \bar{x}_{t+k+1|t} = A\bar{x}_{t+k|t} + B\bar{u}_{t+k|t}, \\ & x_{t+k+1|t} = Ax_{t+k|t} + Bu_{t+k|t}(\cdot) + d_{t+k|t}, \\ & u_{t+k|t}(x_{t+k|t}) \in \mathcal{U} \ominus B^\dagger \widehat{\mathcal{F}}_t, \\ & \bar{u}_{t+k|t} = u_{t+k|t}(\bar{x}_{t+k|t}), \\ & x_{t+k|t} \in \mathcal{X}, \\ & x_{t+N|t} \in \mathcal{O}_t, \\ & x_{t|t} = \bar{x}_{t|t} = x_t, \\ & \forall d_{t+k|t} \in \widehat{\mathcal{D}}_t, \quad \forall k \in \{0,1,\ldots,N-1\} \end{aligned} \tag{20}$$

Problem (20) optimizes a time-varying feedback policy with a cost on the nominal trajectory $(\bar{x},\bar{u})$ subject to state and input constraints on the realized trajectory. We take a standard tube-MPC approach and assume access to a robust control invariant $\mathcal{O}_t$ and terminal cost $V_N(\cdot)$ to guarantee stability.

Assumption 3: There exists a policy $\pi_N : \mathcal{X} \to \mathcal{U}$ associated with the robust positive invariant set $\mathcal{O}_t \subseteq \mathcal{X}$ such that for all $d_t \in \widehat{\mathcal{D}}_t$ we have that

$$x \in \mathcal{O}_t \implies Ax+B\pi_N(x)+d_t \in \mathcal{O}_t,\ \pi_N(x) \in \mathcal{U}\ominus B^\dagger\widehat{\mathcal{F}}_t. \tag{21}$$

Assumption 4: We assume that the terminal cost $V_N : \mathbb{R}^n \to \mathbb{R}_+$ is a continuous Lyapunov function for the policy $\pi_N(x)$ associated with the terminal set $\mathcal{O}_t$ in Assumption 3, i.e. that there exists a class-$\mathcal{K}_\infty$ function α_N such that

$$V(Ax + B\pi_N(x)) - V(x) \le -\alpha_N(\|x\|). \tag{22}$$

Assumption 3 and Assumption 4 are standard and easily satisfied by taking $\pi_N(x) = -Kx$ as the solution to an LQR problem with h as the stage cost. Then, $\mathcal{O}_t$ can be computed efficiently using the standard algorithms in [24]. Note that Lemmas 1 and 2 imply that $\mathcal{O}_{t-1} \subseteq \mathcal{O}_t$ since $\widehat{\mathcal{D}}_t \subseteq \widehat{\mathcal{D}}_{t-1}$. Therefore, the terminal constraint in problem (20) will become less conservative over time. In this work, we optimize problem (20) over *feedback policies* that maintain convexity of the constraints to plan for future information gain. For example, we could optimize over *causal affine recourse* policies of the form

$$u_{t+k|t}(\cdot) = \sum_{j=0}^{k-1} K_j d_{t+j|t} + \bar{u}_{t+k|t} \tag{23}$$

It is well known that we can solve problem (20) with policies of the form in (23) by solving a convex program [14]. We take a nominal tube MPC approach in favor of a min-max method [11] to guarantee an efficient solution exists.

Let $[u^*_{t|t}(\cdot),\ldots,u^*_{t+N-1|t}(\cdot)]$ denote the solution of the MPC problem (20) at time t, then we choose the robust control term of the *certainty equivalent* control policy (7) as

$$u_t^*(x_t) = u^*_{t|t}(x_t) \tag{24}$$

C. Properties

We prove the stability of our algorithm through a standard recursive feasibility and input-to-state stability argument.

Theorem 2: Consider a system of the form in (1), a parameter estimator that satisfies Assumption 2 with features that satisfy Assumption 1 in closed loop feedback with the *matching certainty equivalent* control law (9)-(24). If the tube MPC problem (20) is feasible at $t = 0$, then for all $t \ge 0$ we have that (20) is feasible and the closed loop system (1)-(9) satisfies $x_t \in \mathcal{X}$, and $\pi(x_t) \in \mathcal{U}$.

Proof: Suppose the optimal control problem (20) is feasible at time t, with solution $[u^*_{t|t}(\cdot),\ldots,u^*_{t+N-1|t}(\cdot)]$. By Lemma 1 and Corollary 1 we have that $\hat{f}_{t+k} \in \widehat{\mathcal{F}}_t \ \forall t+k \ge t$. Therefore, the time varying CE control law

$$\pi^*_{t+k|t}(\cdot) = u^*_{t+k|t}(\cdot) - B^\dagger \hat{f}_{t+k}(x_t) \tag{25}$$

satisfies the input constraints for $t \in [t, t+N-1]$, since (20) then implies $u^*_{t+k|t}(\cdot) \in \mathcal{U} \ominus B^\dagger\widehat{\mathcal{F}}_t$. Moreover, by Lemma 2 the disturbance support shrinks in time: $\widehat{\mathcal{D}}_{t+1} \subseteq \widehat{\mathcal{D}}_t$. Therefore, we have that under policy (25) the closed-loop trajectory formed by (1)-(25) satisfies $x_{t+k} \in \mathcal{X}$ for all $k \in [0,N]$ and that $x_{t+N} \in \mathcal{O}_t$. Hence, if we apply the CE policy (9)-(24) at time t, then $x_{t+1} \in \mathcal{X}$ and $u_t \in \mathcal{X}$.

By Assumption 3, for any $x \in \mathcal{O}_t \subseteq \mathcal{O}_{t+1}$, applying the policy $\pi_N(x) \in \mathcal{U}\ominus B^\dagger\widehat{\mathcal{F}}_t$ implies that $Ax+B\pi_N(x)+d \in \mathcal{O}_{t+1}$ for any $d \in \widehat{\mathcal{D}}_t$. Therefore, the policy sequence $[u^*_{t+1|t}(\cdot),\ldots u^*_{t+N-1|t}(\cdot),\pi_N(\cdot)]$ is feasible for the tube MPC problem (20) at time $t+1$, since its feasible set can only increase over time. Therefore, if the MPC program (20) is feasible at time $t = 0$, it is also feasible for all $t \ge 0$ and the closed-loop system formed by the matching CE law (1), (9)-(24) must robustly satisfy state and input constraints by induction.

■

Remark 5: Theorem 2 and Assumption 3 imply that the robust positive invariant set $\mathcal{O}_t$ or disturbance sets $\widehat{\mathcal{D}}_t, \widehat{\mathcal{F}}_t$ need not be updated at every timestep to guarantee recursive feasibility (nor stability). For high-dimensional systems operating at fast control rates, this allows us to reduce online computation and update the constraints in batches. Moreover, it also enables us to apply this algorithm in an episodic or iterative setting and update the models only between episodes.

Theorem 3: Consider a system of the form in (1), a parameter estimator that satisfies Assumption 2 with features that satisfy Assumption 1 in closed loop feedback with the *certainty equivalent* control law (9)-(24). Let $\mathcal{X}_N \subseteq \mathcal{X}$ denote the set of states for which the tube MPC problem (20) is feasible. Then the closed loop system is locally input-to-state stable with region of attraction $\mathcal{X}_N$.

Proof: Our proof closely resembles [16, Thm. 2]. We argue that the nominal system is stable by a standard

MPC argument, and that the closed-loop system is ISS since the disturbances are bounded. Since we assume the stage cost is positive definite and continuous, there exist two class-$\mathcal{K}_\infty$ functions α_1, α_2 such that for all $t \geq 0$, $\alpha_1(\|x\|) \leq J_N^*(t,x) \leq \alpha_2(\|x\|)$ and that $J^*(t,0) = 0$ (see [13, Prop. 1], [16, Thm. 2]), and a class-$\mathcal{K}_\infty$ function α_3 such that $h(x,u) \geq \alpha_3(\|x\|)$. Let $J_N^*(t,x_t)$ be the solution of (20) associated with the nominal prediction $[\bar{x}^*_{t|t},\dots,\bar{x}^*_{t+N|t}]$, nominal inputs $[\bar{u}^*_{t|t},\dots,\bar{u}^*_{t+N-1|t}]$, and feedback policies $[u^*_{t|t}(\cdot),\dots,u^*_{t+N-1|t}(\cdot)]$. By Theorem 2, we have that if we apply the CE control law (9)-(24) at time t, then the policies $[u^*_{t+1|t},\dots,u^*_{t+N-1|t},\pi_N]$ are a feasible solution for (20) at time $t+1$. Let $\bar{J}_N(t,x)$ be the cost associated with forward simulating the nominal system using the policies $[u^*_{t+1|t},\dots,u^*_{t+N-1|t},\pi_N]$ with x as initial condition. i.e. set $u^*_{t+N|t} = \pi_N(\cdot)$ and $\bar{u}^*_{t+N|t} = \pi_N(\bar{x}^*_{t+N|t})$ and let $\bar{x}_{t+1|t+1} = x$, $\bar{x}_{k+1|t+1} = A\bar{x}_{k|t+1} + Bu^*_{k|t}(\bar{x}_{k|t+1})$ for $k \in \{t+1,\dots,t+N\}$ so

$$\bar{J}_N(t,x) = \sum_{k=t+1}^{t+N} h(\bar{x}_{k|t+1}, u^*_{k|t}(\bar{x}_{k|t+1})) + V_N(\bar{x}_{t+N+1|t+1}).$$

Note that $J_N^*(t+1,x_{t+1}) \leq \bar{J}_N(t+1,x_{t+1})$ and that by continuity of the stage cost and Assumption 4, $\bar{J}(t+1,\cdot)$ is uniformly continuous for all $t \geq 0$ on the state space. It follows that there exists a $\mathcal{K}_\infty$ function α_J such that for all $t \geq 0$, $|\bar{J}_N(t,x_1) - \bar{J}_N(t,x_2)| \leq \alpha_J(\|x_1 - x_2\|)$ (see [13, Lem. 1]). Therefore,

$$\begin{aligned}
&J_N^*(t+1,x_{t+1}) - J_N^*(t,x_t)\\
&\leq \bar{J}(t+1,x_{t+1}) - J_N^*(t,x_t)\\
&= \bar{J}(t+1,x_{t+1}) - \bar{J}(t+1,\bar{x}^*_{t+1|t})\\
&\quad + \bar{J}(t+1,\bar{x}^*_{t+1|t}) - J_N^*(t,x_t)\\
&\leq |\bar{J}(t+1,x_{t+1}) - \bar{J}(t+1,\bar{x}^*_{t+1|t})| - h(x_t,\bar{u}^*_{t|t}(x_t))\\
&\leq \alpha_J(\|d_t\|) - \alpha_3(\|x_t\|)
\end{aligned}$$

Finally, since $d_t \in \widehat{\mathcal{D}}_t \subseteq \widehat{\mathcal{D}}_{t-1}$ is contained in a closed and bounded set, there must exist a class-$\mathcal{K}$ function σ such that $\alpha_J(\|d_t\|) \leq \sigma(\|d_t\|)$. Hence

$$J_N^*(t+1,x_{t+1}) - J_N^*(t,x_t) \leq -\alpha_3(\|x_t\|) + \sigma(\|d_t\|)$$

Therefore, the system must be ISS stable by Theorem 1. ∎

Remark 6: We remark that the ISS result in Theorem 3 does not explicitly show that improvements in the confidence of the model lead to better performance of the controller, since we only assume the model confidence is non-decreasing in Assumption 2.

However, if we can guarantee that the confidence in the estimator increases over time (i.e. that $\mathcal{W}(t) \to \{0\}$ as $t \to \infty$), then our approach converges to a smaller neighborhood of the origin than robust MPC methods without the certainty equivalent matching term in equation (9). This is a consequence of the asymptotic gain interpretation of ISS [13] and the fact that $(I - BB^\dagger)\widehat{\mathcal{F}}_\infty \subset \widehat{\mathcal{F}}_\infty$. In adaptive control, such guarantees are typically made under *persistence of excitation* assumptions [1].

IV. Adaptation Laws & Learning Algorithms

In this section we highlight two common online function approximation schemes that satisfy the decaying confidence interval of Assumption 2 that we used to construct the MPC algorithm in the previous section.

Set Membership Estimation. A common approach in the adaptive MPC literature is to estimate constant, or slowly changing, disturbances through *set membership* estimation [15], [20]. These estimators maintain a feasible parameter set that is refined as more data becomes available. The feasible parameter set contains all credible model parameters that explain previous observations, which means that the feasible parameter sets are nested over time. We propose learning the parameters of a nonlinear uncertainty model of the form in (3) directly using set-membership estimation. Assuming access to an initial range of feasible parameters $\Theta = \widehat{W}_0 \oplus \mathcal{W}$, the feasible parameter set at time t is obtained as

$$\begin{aligned}\Theta(t) = \{W \in \Theta : x_{k+1} - Ax_k - Bu_k - W\phi(x_k)\\ \in \mathcal{V} \quad \forall\, k \in [0, t-1]\}.\end{aligned} \tag{26}$$

This estimator maintains independent feasible sets for each row of W^* and can be updated recursively in time with simple polytopical set intersections by rewriting (26) in terms of the row-wise vectorization of W. Clearly, $\Theta_t \subseteq \Theta_{t-1}$. As is common practice in the literature [17], we propose generating a point estimate of the parameters as the *Chebyshev center* of the feasible parameter set:

$$\widehat{W}_t = \arg\min_{\widehat{W}} \max_{W^* \in \Theta_t} \|\widehat{W} - W^*\|_F^2 \tag{27}$$

By definition, this approach minimizes the worst-case error of the point estimates and is efficient for convex sets. If we denote the Chebyshev radius for the feasible parameter set associated with the i'th row of W^* as $r_i(t) = \min_{w_i} \max_{w_i^* \in \Theta_i(t)} \|w_i - w_i^*\|_2$. Then we can take the confidence interval on $\hat{w}_i - w^*$ as $\mathcal{W}_i(t) = \{\tilde{w}_i : \|\tilde{w}_i\|_2 \leq r_i(t)\}$.

By definition, since $\Theta_t \subseteq \Theta_{t-1}$, the Chebyshev radii must be decreasing over time: $r_i(t) \leq r_i(t-1)$. Therefore, a set-membership estimator with point estimates as the Chebyshev center can be used with the MPC framework we proposed in §III. In fact, since a set membership approach ensures the true parameter lies in a shrinking, known set, this estimator has stronger properties than we assume in Assumption 2. However, set-membership estimation is not robust to outliers or model misspecification. Therefore, statistical estimation is favored in practice.

Recursive Least Squares. In general, it is not easy to guarantee that the feasible parameter sets (26) are nested over time (i.e. $\Theta(t) \subseteq \Theta(t-1)$) when we use a statistical estimation technique without making restrictive assumptions. In the case of least squares estimation, we can generate confidence intervals directly from the distribution over parameters if we know the disturbance distribution. We outline this approach under a simple, standard assumption.

Assumption 5: We assume that each entry of the process noise is bounded $v_t = [v_t^1,\dots,v_t^n]^\top \in \mathcal{V} = \{v : \|v\|_\infty \leq \sigma\}$, that the disturbances are sampled i.i.d. from some distribution

$v_t \sim P$ with zero-mean, and that each entry v_t^i is independent of the others. Hence, v_t^i is sub-Gaussian with variance proxy σ^2.

Under Assumption 5, we can essentially treat the noise as both normally distributed for convenient analysis and provide safety guarantees for the algorithm proposed in §III. The least squares estimate is given as

$$\widehat{W}_t = \arg\min_W \sum_{i=0}^{t} \|W\phi(x_i) - y_i\|^2 \tag{28}$$

where $y_t = x_{t+1} - Ax_t - Bu_t$. Let the prediction at time t be $\hat{y}_t = \widehat{W}(t)\phi(x_t)$. The estimator $\widehat{W}(t+1)$ can then be updated in constant time using the recursive update equations

$$\widehat{W}(t+1) = \widehat{W}(t) - \frac{(\hat{y}_t - y_t)\phi(x_t)^\top \Lambda_t^{-1}}{1 + \phi(x_t)^\top \Lambda_t^{-1}\phi(x_t)} \tag{29}$$

$$\Lambda_{t+1}^{-1} = \Lambda_t^{-1} - \frac{\Lambda_t^{-1}\phi(x_t)\phi(x_t)^\top \Lambda_t^{-1}}{1 + \phi(x_t)^\top \Lambda_t^{-1}\phi(x_t)} \tag{30}$$

where, in a Bayesian formulation, we have placed a subjective prior over the rows of W of the form $w_i^* \sim \mathcal{N}(\hat{w}_i(0), \Lambda_0\sigma^2)$. We can recover the frequentist ordinary least-squares estimator if we assume a flat prior, which requires the availability of some amount of prior data to yield the initial values $\widehat{W}_0$, Λ_0.

If we take a risk tolerance of $\delta \in (0,1)$, we could naively define the confidence interval for the i-th row of $\widehat{W}(t)$ as

$$\mathcal{W}_i^{\text{naive}}(t) = \{\tilde{w}_i \in \mathbb{R}^n : \tilde{w}_i^\top \Lambda_t \tilde{w}_i \le \sigma^2 \chi_n^2(1 - \frac{\delta}{n})\} \tag{31}$$

where $\chi_n^2(1-\delta)$ is the $1-\delta$ quantile of the chi-square distribution with n degrees of freedom. However, the confidence interval in (31) does not capture the fact that we want to certify the safety of the policy for all time with high probability. We cannot achieve this with a single confidence interval of a point estimate at time t, as it ignores the correlations between the model estimates over time. For robust control, we instead desire confidence intervals such that with probability at least $1-\delta$,

$$\widehat{W}(t) - W^* \in \mathcal{W}(t) \quad \forall t \ge 0. \tag{32}$$

Recent work applied a Martingale argument originating from the Bandits literature to generate such confidence intervals [25]. The authors scaled the naive confidence intervals by a time-varying parameter $\beta_t(\frac{\delta}{n})$ such that the sets $\mathcal{W}_i(t)$ as defined below ensure that the confidence interval $\mathcal{W}(t) = \mathcal{W}_1(t) \times \cdots \times \mathcal{W}_n(t)$ contains the model mismatch[1].

Theorem 4: [25, Thm 1] For the Bayesian recursive least-squares filter (30), with probability at least $1-\delta$, we have that for all $t \ge 0$ the estimation error $\hat{w}_i(t) - w_i^* \in \mathcal{W}_i(t)$ where

$$\mathcal{W}_i(t) = \{\tilde{w}_i : (\tilde{w}_i^\top \Lambda_t \tilde{w}_i)^{\frac{1}{2}} \le \sigma\beta_t(\delta/n)\} \tag{33}$$

$$\beta_t(\delta) = \sqrt{2\log\left(\frac{\det(\Lambda_t)^{1/2}}{\delta\det(\Lambda_0)^{1/2}}\right)} + \sqrt{\frac{\lambda_{\max}(\Lambda_0)}{\lambda_{\min}(\Lambda_t)}\chi_n^2(1-\delta)} \tag{34}$$

The confidence intervals resulting from Theorem 4 unfortunately do not immediately satisfy Assumption 2 without a *persistence of excitation* or *active exploration* assumption as is made in [26], so we cannot directly apply them to our control design. A simple workaround is to update the estimate (28) fed into to the controller only when the associated confidence intervals (33) have shrunk, effectively disregarding new data until the system has been excited sufficiently. Although this approach was shown to perform well in practice [27], future work should explore strategies to guarantee confidence intervals constructed using Theorem 4 (or other equivalent results) satisfy Assumption 2 more naturally.

V. Experiments

Simple Matched System. We illustrate the properties of the adaptive MPC algorithm (7)-(24) that we developed in section §III on the following double-integrator system:

$$x_{t+1} = \begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (u_t + w_1 \tanh([0,1]x_t)) + v_t \tag{35}$$

The true parameter $w_1 = .5$ is unknown, and must be estimated online to improve the controller performance. We take the disturbance as an isotropic Gaussian with $\sigma^2 = 5 \times 10^{-3}$ clipped at its 95% confidence intervals.

In addition, the system is subject to the state and input constraints $(-4,-3) \preceq x \preceq (4,3)$ and $-2 \le u \le 2$, respectively.

The goal of the control task is to regulate the system to 0 subject to the state and input constraints from the initial condition $x_0 = [2.5, 2]^\top$ while minimizing a quadratic cost function $h(x,u) = x^\top Qx + u^\top Ru$ over a horizon of length $N = 3$. We take $Q = I_2$, $R = 1$. We compare the adaptive MPC from §III-A with an analogous tube MPC that considers the estimated support of the nonlinearity as a disturbance and does not cancel the uncertainty or adapt the model. We use the least-squares estimator with a flat prior and collect $k = 50$ data points of the system evolution in a unit box near the origin to form an initial estimate of the model parameters.

The closed-loop system evolution is shown in Fig. 1 (left, middle). The adaptive certainty-equivalent MPC algorithm is able to effectively control the system. We also plot the reachable sets associated with the first and last predicted trajectory of the system in Fig. 1. This makes it clear that the adaptive MPC resolves uncertainty in the system, since these sets shrink over time. Online learning has little consequence for the naive tube MPC, as increasing the confidence in the model does not significantly reduce the estimated range of values that the nonlinear function takes. Moreover, the naive tube MPC is overly conservative, such that the tube MPC problem is infeasible from the given initial condition.

[1] Subject to assumptions on the calibration of the prior, for which we refer the reader to [25, Assumption 3]. This assumption is trivially satisfied for flat priors, and we assume this assumption is satisfied for subjective priors.

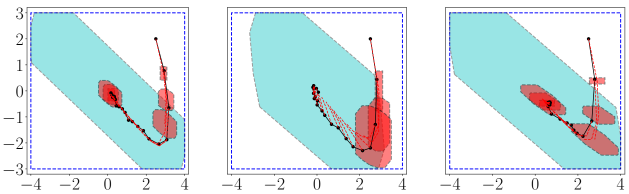

Fig. 1. Closed-loop trajectories of different systems. The true system evolution is in black, the predicted nominal trajectories as well as predicted reachable sets are red. The dark blue line indicates the state constraints. The terminal invariant $\mathcal{O}_t$ is shown in cyan. Left: Adaptive CE MPC on the system with *matched* uncertainty (35). Middle: Standard tube MPC on the *matched* system (35). Right: Adaptive CE MPC on the system with *unmatched* uncertainty (36).

In addition, we compare the asymptotic performance of our algorithm with the naive tube MPC as a function of w_1. This allows us to control the magnitude of the nonlinearity. To do this, we set the number of data points used to generate the initial model estimate to $k = 10^3$ and collect trajectory rollouts for various values of w_1 from the fixed initial condition. As shown in Fig. 2 (left), our adaptive MPC is feasible for disturbances more than twice the magnitude of those the tube MPC can tolerate for the given initial condition. In addition, the realized control cost does not differ significantly from regular tube MPC.

Finally, we illustrate the ability of our algorithm to tolerate larger dynamic uncertainty by comparing the size of the feasible envelope (i.e. the set of initial conditions for which the MPC problem is feasible) as a function of w_1. We set the number of datapoints to inform our prior to a modest $k = 50$, grid the state-space, and take the feasible region as the convex hull of the initial conditions for which the MPC problem is feasible. We estimate the percentage of states $x_0 \in \mathcal{X}$ in the feasible envelope as the ratio of volumes between the feasible envelope and the state space $\mathcal{X}$, illustrated with solid lines in Fig. 2 (right). Our adaptive MPC algorithm can tolerate much larger disturbances than a tube MPC that does not consider the structure of the uncertainty. In these experiments, the feasible envelope of the naive tube MPC becomes empty when the maximal robust invariant is null ($\mathcal{O}_t = \emptyset$), indicating that there is no subset of $\mathcal{X}$ in which the LQR policy associated with h results in provably safe behavior [24].

Simple Unmatched System. We now extend the simulations of the simple matched system to the unmatched case to understand the effect of additive nonlinear terms that cannot be cancelled from the dynamics. We keep the nominal dynamics identical to (35) and take the nonlinear dynamics

$$f(x) = [w_1 \sin(4x_1), w_2 \tanh(x_2)]^\top, \tag{36}$$

where w_1 and w_2 are unknown parameters. Similar to the previous experiments, we initialize the model with $k = 100$ data-points sampled around the origin. In this example, the certainty equivalent policy (9) can only compensate for the second component of the nonlinear dynamics (36). Therefore, a disturbance of magnitude $|w_1|$ will enter the closed loop system. We set $w_1 = .2$ and $w_2 = .3$. As is apparent from Fig. 1 (left, right), the size of the reachable sets increases if we simulate the system with the unmatched dynamics (36). Still, our method clearly outperforms a standard tube MPC on the same task. Next, we fix $w_2 = .5$ and vary w_1 to understand the impact of an estimated, unmatched dynamics components on our method. Fig. 2 (dashed, right) shows that in our experiment, matching as much of the nonlinear dynamics as possible allows us to handle unmatched dynamics of about twice the magnitude as the benchmark tube MPC scheme. We conclude that our method is a more effective strategy even if large components of the dynamics are not matched. Naturally, Fig. 2 (right) also shows that the benefit of our method diminishes as the component of the nonlinear dynamics $f(x_t)$ in $\mathrm{Range}(B)$ becomes smaller.

Planar Quadrotor. Finally, we simulate a simplified example of a quadrotor in a windy environment. The force field induced by the wind varies spatially, modelling real-world scenarios such as down-wash from another quadrotor. We consider a planar version of the quadrotor for simplicity [28]. We model the wind disturbance as incident at a fixed angle with a velocity that drops of according to an inverse square exponential normal to the direction of incidence. We linearize the dynamics around $x = 0$, $u = \frac{mg}{2}[1,1]^\top$ and discretize the simulation using Euler's method. We only set constraints on the pose of the drone (x, y, θ). Its linear and angular velocities are unconstrained. The quadrotor is an underactuated system, and therefore the discretized simulation has unmatched dynamics terms. Still, a drone controller can always match disturbance forces along the y axis. We take a Bayesian approach, and model the unknown wind disturbance as a function of $d = 20$ normalized random Fourier features $\phi_i(x) = \frac{1}{\sqrt{d}} \cos(\alpha_i^\top x + \beta_i)$ where $\alpha \sim \mathcal{N}(0, I_n)$ and $\beta_i \sim \mathrm{U}[0, 2\pi]$. To calibrate the Bayesian prior, we first select a zero-mean prior with a variance chosen such that the confidence interval reflects a conservative bound on the wind disturbance. We then calibrate the Bayesian prior using $k = 100$ historical data-points of the wind disturbance.

We compare our approach with a naive tube MPC that does not account for the wind disturbance, since the support of the estimated wind disturbance was too large for a benchmark tube MPC to be feasible in our experiments. As shown in

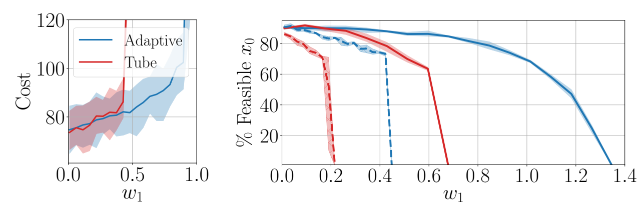

Fig. 2. Left: Closed-loop realized trajectory cost for the matched system (35) as a function of w_1. Exploding cost indicates infeasibility. The error bars indicate 2σ bounds. Right: Solid lines indicate size of feasible envelopes as a function of w_1 for a matched system. Dashed lines indicate the size of the feasible envelopes for the unmatched system (36) as a function of the magnitude of the unmatched dynamics w_1 with $w_2 = .5$.

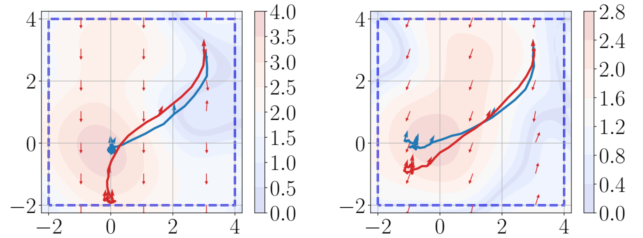

Fig. 3. Learned trajectories of the adaptive MPC (blue) compared to a naive, unsafe tube MPC (red) in the xy-plane for a simulated planar quadrotor. The quivers show the orientation of the quadrotors over time. The contours show the *learned* wind disturbance, the colorbar indicates the learned wind speed in (m/s). Left: Wind comes straight from above. Right: Wind comes at $\theta_w = 22.5°$.

Fig. 3 (left), if the wind disturbance is axis aligned, the adaptive MPC learns to match the wind forces and reaches the origin quickly. In contrast, the naive tube MPC misses the origin and drifts significantly. In addition, if we set the angle of incidence of the wind as $\theta_w = 22.5°$, Fig. 3 (right) shows that our approach still achieves decent control performance. The certainty equivalent controller (9) cancels the y-component of the disturbance and converges to a small steady-state offset in the x direction. In contrast, a regular tube MPC could not guarantee safety for any of the drone tasks, and a naive unsafe tube MPC that does not consider the wind disturbance at all performs poorly.

VI. Conclusions and Future Work

The simulated experiments in §V show that our method achieves substantial performance improvements compared to existing approaches, even when significant components of the nonlinear dynamics are unmatched and cannot be cancelled by a certainty equivalent control policy. We conclude that by extending certainty equivalent control laws from classical adaptive control, we can reduce the conservatism of robust MPC approaches. As a result, our method can tolerate much larger nonlinear terms in the dynamics.

Since our control algorithm allows for adaptation laws based on statistical estimation techniques that are more robust to outliers, future work should extend our simulations to hardware experiments. In addition, our method requires that the features are known a priori, an assumption that future work could relax. For example, this could be done by applying meta-learning algorithms [8] and reasoning about the misspecification of the model.

Acknowledgements

We thank Monimoy Bajarbaruah for his thoughtful comments on an early manuscript. This research was supported in part by the National Science Foundation (NSF) via Cyber-Physical Systems (CPS) award #1931815, and the National Aeronautics and Space Administration (NASA) via University Leadership Initiative grant #80NSSC20M0163 and via an Early Stage Innovations grant. Spencer M. Richards and James Harrison were also supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). This article solely reflects our own opinions and conclusions, and not those of any NSF, NASA, or NSERC entity.

References

[1] J.-J. E. Slotine and Weiping Li. *Applied nonlinear control*. 1991.
[2] P. Ioannou and J. Sun. *Robust Adaptive Control*. 2012.
[3] P. Ioannou and B. Fidan. *Adaptive Control Tutorial*. 2006.
[4] K. Wise E. Lavretsky. *Robust and Adaptive Control with Aerospace Applications*. 2012.
[5] G. Joshi, J. Virdi, and G. Chowdhary. Asynchronous Deep Model Reference Adaptive Control. *arXiv:2011.02920*, 2020.
[6] G. Chowdhary, H. A. Kingravi, J. P. How, and P. A. Vela. Bayesian Nonparametric Adaptive Control Using Gaussian Processes. *TNNLS*, 2015.
[7] D. D. Fan, J. Nguyen, R. Thakker, N. Alatur, A.-a Agha-mohammadi, and E. A. Theodorou. Bayesian Learning-Based Adaptive Control for Safety Critical Systems. In *ICRA*, 2020.
[8] J. Harrison, A. Sharma, and M. Pavone. Meta-learning Priors for Efficient Online Bayesian Regression. In *WAFR*, 2018.
[9] J. Harrison, A. Sharma, R. Calandra, and M Pavone. Control adaptation via meta-learning dynamics. In *NeurIPS Workshop on Meta-Learning*, 2018.
[10] S. M. Richards, N. Azizan, J.-J. E. Slotine, and M. Pavone. Adaptive-control-oriented meta-learning for nonlinear systems. *arxiv:2103:04490*, 2021.
[11] P. O. M. Scokaert and D. Q. Mayne. Min-max feedback model predictive control for constrained linear systems. *TAC*, 1998.
[12] D. Q. Mayne, M. M. Seron, and S. V. Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 2005.
[13] D. Limon, T. Alamo, D. M. Raimondo, D. M. De la Peña, J. M. Bravo, A. Ferramosca, and E. F. Camacho. Input-to-state stability: A unifying framework for robust model predictive control. In *Nonlinear Model Predictive Control*. 2009.
[14] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 2006.
[15] M. Bujarbaruah, X. Zhang, U. Rosolia, and F. Borrelli. Adaptive MPC for Iterative Tasks. In *CDC*, 2018.
[16] M. Bujarbaruah, X. Zhang, M. Tanaskovic, and F. Borrelli. Adaptive stochastic MPC under time varying uncertainty. *TAC*, 2021.
[17] M. Bujarbaruah, S. H. Nair, and F. Borrelli. A Semi-Definite Programming Approach to Robust Adaptive MPC under State Dependent Uncertainty. In *ECC*, 2020.
[18] R. Soloperto, M. A. Müller, S. Trimpe, and F. Allgöwer. Learning-Based Robust Model Predictive Control with State-Dependent Uncertainty. *IFAC NMPC*, 2018.
[19] S. Di Cairano. Indirect adaptive model predictive control for linear systems with polytopic uncertainty. In *ACC*, 2016.
[20] J. Köhler, E. Andina, R. Soloperto, M. Müller, and F. Allgöwer. Linear robust adaptive model predictive control: Computational complexity and conservatism. In *CDC*, 2019.
[21] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin. Provably safe and robust learning-based model predictive control. *Automatica*, 2013.
[22] H. Li, A. Liu, and L. Zhang. Input-to-state stability of time-varying nonlinear discrete-time systems via indefinite difference Lyapunov functions. *ISA Transactions*, 77:71–76, June 2018. Publisher: Elsevier.
[23] Z.-P. Jiang and Y. Wang. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 2001.
[24] F. Borrelli, A. Bemporad, and M. Morari. *Predictive control for linear and hybrid systems*. 2017.
[25] T. Lew, A. Sharma, J. Harrison, and M. Pavone. Safe Model-Based Meta-Reinforcement Learning: A Sequential Exploration-Exploitation Framework. *arXiv:2008.11700*, 2020.
[26] H. Mania, M.I. Jordan, and B Recht. Active learning for nonlinear system identification with guarantees, 2020.
[27] T. Koller, F. Berkenkamp, M. Turchetta, and A. Krause. Learning-Based Model Predictive Control for Safe Exploration. In *CDC*, 2018.
[28] R. Tedrake. *Underactuated Robotics: Algorithms for Walking, Running, Swimming, Flying, and Manipulation (Course Notes for MIT 6.832)*.