Proceedings of the EuroGNC 2013, 2nd CEAS Specialist Conference on
Guidance, Navigation & Control, Delft University of Technology, Delft,
The Netherlands, April 10-12, 2013

Concurrent Learning Adaptive Model Predictive Control

Girish Chowdhary*, Maximilian Mühlegg *, Jonathan P. How, and
Florian Holzapfel

Abstract A concurrent learning adaptive-optimal control architecture for aerospace systems with fast dynamics is presented. Exponential convergence properties of concurrent learning adaptive controllers are leveraged to guarantee a verifiable learning rate while guaranteeing stability in presence of significant modeling uncertainty. The architecture switches to online-learned model based Model Predictive Control after an online automatic switch gauges the confidence in parameter estimates. Feedback linearization is used to reduce a nonlinear system to an idealized linear system for which an optimal feasible solution can be found online. It is shown that the states of the adaptively feedback linearized system stay bounded around those of the idealized linear system, and sufficient conditions for asymptotic convergence of the states are presented. Theoretical results and numerical simulations on a wing-rock problem with fast dynamics establish the effectiveness of the architecture.

1 Introduction

Model based optimal control of dynamical systems is a well studied topic. For example, one of the most commonly used techniques for linear and nonlinear systems with constraints is model predictive control (see e.g. [4, 30, 20]). While this technique has been heavily studied and implemented for slower industrial processes, only in the past decade enough computational power has become available to enable online optimization for fast system dynamics typical of aerospace applications

Girish Chowdhary and Jonathan P. How
Massachusetts Institute of Technology, e-mail: girishc@mit.edu, jhow@mit.edu

Maximilian Mühlegg and Florian Holzapfel
Technische Universität München, e-mail: maximilian.muehlegg@tum.de,
florian.holzapfel@tum.de

* equal contribution

(some relevant demonstrations are in [14, 15, 13, 32, 33, 5]). MPC depends on a dynamic predictive model of the system. However, unaccounted modeling errors and dynamic variations in any real world scenario often result in an *a-priori* generated model of a system becoming obsolete or inaccurate. In such cases, the stability and performance of an MPC approach cannot be guaranteed, especially if the underlying dynamics are nonlinear [31]. One way to deal with this is to estimate parameters of the dynamic model online, and then generate optimal controllers at each time step assuming that the estimated model at that time step is the correct one. This results in an indirect adaptive control approach that uses the principle of certainty equivalence (see e.g. [19, 2]). The benefit of this indirect-adaptive MPC approach is that it allows for a way to incorporate learning in the MPC framework. However, the main drawback of this method is that it is difficult to guarantee stability, especially during parameter estimation transient phases. This is one major challenge in synthesizing algorithms for online adaptive-optimal control [25].

Several authors have made key contributions to implementing such adaptive MPC architectures. Fukushima et al. used the comparison principle to develop adaptive MPC for linear systems [17]. Adetola et al. considered adaptive MPC of linearly parameterized nonlinear systems and showed that one way to guarantee stability is to ensure that the initial parameter errors are within certain bounds [1]. Aswani et al. explored and experimented in flight with the notion of safe-MPC by guaranteeing that control inputs are selected such that the system evolution is constrained to (approximations of) invariant reachable sets. Their work has clearly demonstrated that adaptive MPC can indeed result in improved flight performance through flight testing. However, they used an EKF for parameter estimation, which is known to not guarantee predictable and quantifiable learning rates under general operating conditions, and concentrate on linear dynamical systems [5, 3]. In general, while significant progress has been made in adaptive MPC, the results tend to be conservative, as the presence of learning transients prevent a general non-conservative solution to be formed.

On the other hand, adaptive control is one of the most well studied areas in control systems theory. In adaptive control algorithms and techniques are developed for dealing with modeling uncertainties and disturbances. Direct adaptive control methods directly modify the system input to account for modeling uncertainties. In a certain light, these techniques could be viewed as model-free, in the sense that they do not focus on learning the system model, but rather on suppressing the uncertainty pointwise-in-time to minimize the instantaneous tracking error. Direct adaptive controllers can guarantee stability, even during harsh transients, however, they do not offer any long-term improvement due to model learning unless the system states are persistently exciting (PE; see e.g. [6]). Furthermore, it is difficult to generate optimal solutions in presence of input and state constraints with direct adaptive control architectures.

Adaptive control literature also consists of hybrid-direct-indirect control architectures. For example, Duarte and Narendra, Lavretsky, and Chowdhary and Johnson have shown that modifying direct adaptive controllers such that they focus also on learning the uncertainty improves performance (see e.g. [12, 24, 9]). The power of these techniques is that they can handle harsh learning transients,guarantee learning of unknown model parameters subject to conditions on the system trajectories, and guarantee system stability during the learning. It is natural therefore, to hypothesize that adaptive-optimal control algorithms can be devised that use provable hybrid adaptive control techniques to guarantee stability in the learning phase and then switch automatically to model-based optimal control algorithms (e.g. MPC) after sufficient confidence in estimated parameters has been gauged online. One such architecture is proposed in this paper and displayed in Figure 1. The main challenges in developing such an architecture include guaranteeing a verifiable learning rate for the uncertainty estimation such that the uncertainty is approximated in finite time before the architecture switches to the online learned model-based optimal controller, guaranteeing stability before and after the switch, and guaranteeing that the architecture can switch back to the adaptive controller if ideal parameters of the system change.

In this paper, we present a Concurrent Learning based adaptive-optimal Model Predictive Controller (CL-MPC) to address these challenges. Our architecture leverages the CL algorithm of Chowdhary and Johnson [9, 8], which guarantees simultaneous system stability and exponential convergence to the ideal parameters without requiring persistency of excitation. This allows us to guarantee verifiable convergence rates. A online metric is developed to initiate a switch to MPC. Learning continues while the system is in MPC using a variant of the CL algorithm, and it is shown that exponential convergence of parameters can be guaranteed if the basis of the uncertainty is known. Furthermore, using a feedback linearization approach we show that a feedback linearizable nonlinear system can be transformed into a a linear system for which an optimal feasible MPC solution can be formulated in presence of constraints. This greatly helps in ensuring feasibility of obtaining an optimal solution for aerospace systems with fast dynamics, as one need only to solve the MPC problem for the ideal feedback linearized system. It is also shown that the actual feedback linearized system's solution is mean square exponentially bounded around the ideal system, and sufficient conditions are provided to guarantee asymptotic convergence to the ideal solution. The presented architecture is validated through simulation on a wing-rock dynamics system. The results show significant improvement over an adaptive-only approach in presence of significant modeling uncertainty.

2 Approximate Model Inversion based Model Reference Adaptive Control

Let $x(t) \in D_x \subset \mathbb{R}^n$, $\delta(t) \in D_\delta \subset \mathbb{R}^l$, and consider the following multiple-input nonlinear uncertain dynamical system

$$\dot{x}(t) = f(x(t), \delta(t)). \tag{1}$$

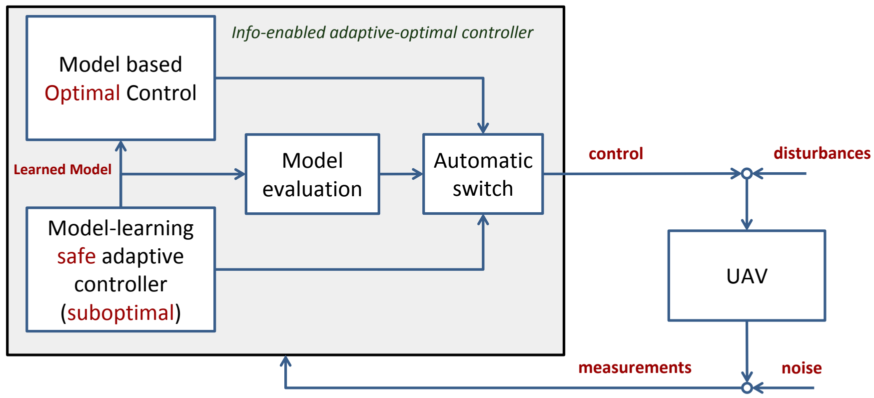

Fig. 1 An adaptive-optimal control architecture. A learning-focused adaptive controller guarantees stability while learning uncertain system parameters. Once sufficient confidence has been gauged online in the estimated parameters, the architecture switches to using an online model-based controller, such as MPC. The resulting switched adaptive-optimal controller is guaranteed to be stable without being conservative about initial parameter errors.

The unknown function $f(\cdot)$ is assumed to be globally Lipschitz and the control input δ is assumed to be bounded and piecewise continuous, so as to ensure the existence and uniqueness of the solution to (1). Furthermore, it is assumed that an admissible control input exists that drives the system from any initial condition in D_x to a neighborhood of any arbitrary point in D_x in finite time. It is further assumed that $l \leq n$ (while restrictive for overactuated systems, this assumption can be relaxed through the design of appropriate control assignment [16]).

The Approximate Model Inversion based MRAC approach used here feedback linearizes the system (1) by finding a pseudo-control input $\nu(t) \in \mathbb{R}^l$ that achieves a desired acceleration. If the exact plant model in equation (1) is known and invertible, the required control input to achieve the desired acceleration is computable by inverting the plant dynamics. However, since this usually is not the case, an approximate inversion model $\hat{f}(x,\delta)$ is employed. The inversion model is chosen to be invertible w.r.t. δ; the operator $\hat{f}^{-1} : \mathbb{R}^{n+l} \to \mathbb{R}^l$ is assumed to exist and assign for every unique element of $\mathbb{R}^{n+l}$ a unique element of $\mathbb{R}^l$.

The following assumption guarantees invertibility of $\hat{f}(x,\delta)$ w.r.t. δ [21].

Assumption 1. $\frac{\partial \hat{f}(x,\delta)}{\partial \delta}$ is continuous w.r.t δ and nonsingular over $D_x \times D_\delta$.

Given a desired pseudo-control input $\nu \in \mathbb{R}^l$ a control command δ can be found by approximate dynamic inversion:

$$\delta = \hat{f}^{-1}(x,\nu). \tag{2}$$

Let $z = (x,\delta)$ for brevity. The use of an approximate model results in a modeling error Δ for the system,

$$\Delta(z) = f(z) - \hat{f}(z). \tag{3}$$

It should be noted that if the control assignment function (the mapping between control inputs to states) is known and invertible with respect to δ, then an inversion model exists such that the modeling error is not dependent on the control input δ.

The modeling uncertainty can be assumed to be represented using a linear combination of basis functions. The basis functions can often be designed based on knowledge of the system dynamics (see e.g. [37, 9]). Alternatively, universally approximating bases, such as Gaussian radial basis functions, can be used ([23]). In either case, letting the basis be represented by $\phi(z) \in \mathbb{R}^m$, we assume the existence of an ideal weight matrix $W^* \in \mathbb{R}^{m \times l}$ such that

$$\Delta(z) = W^{*^T}\phi(z) + \eta(z), \tag{4}$$

where the representation error $\eta_{\text{sup}} = \sup_{z \in D_x} \|\bar{\eta}(z)\|$ is bounded over D_x.

A designer chosen reference model is used to characterize the desired response of the system

$$\dot{x}_{rm} = f_{rm}(x_{rm}, r), \tag{5}$$

where $f_{rm}(\cdot)$ denote the reference model dynamics, which are assumed to be continuously differentiable in x_{rm} for all $x_{rm} \in D_x \subset \mathbb{R}^n$. The reference command $r(t)$ is assumed to be bounded and piecewise continuous, furthermore, $f_{rm}(\cdot)$ is assumed to be such that x_{rm} is bounded for a bounded reference input.

Define the tracking error to be $e(t) = x_{rm}(t) - x(t)$, and the pseudo-control input ν to be

$$\nu = \nu_{rm} + \nu_{pd} - \nu_{ad}, \tag{6}$$

consisting of a linear feedback term $\nu_{pd} = Ke$ with $K \in \mathbb{R}^{l \times n}$; a linear feedforward term $\nu_{rm} = \dot{x}_{rm}$; and an adaptive term $\nu_{ad}(z)$. Since Δ is a function of ν_{ad} as per equation (3), and ν_{ad} needs to be designed to cancel Δ, the following assumption needs to be satisfied:

Assumption 2. The existence and uniqueness of a fixed-point solution to $\nu_{ad} = \Delta(\cdot, \nu_{ad})$ is assumed.

Sufficient conditions for satisfying this assumption are available in [40, 21]. Assumption 2 implicitly requires the sign of the control effectiveness matrix to be known ([21]).

Using equation (3) and the pseudo-control (6) the tracking error dynamics can be written as

$$\dot{e} = Ae + B[\nu_{ad}(z) - \Delta(z)], \tag{7}$$

where the state space model (A, B) is in canonical form with the eigenvalues of A assigned by ν_{pd}. The baseline full state feedback controller ν_{pd} is chosen to make A Hurwitz. Hence, for any positive definite matrix $Q \in \mathbb{R}^{n \times n}$, a positive definite solution $P \in \mathbb{R}^{n \times n}$ exists for the Lyapunov equation

$$0 = A^T P + PA + Q. \tag{8}$$

The adaptive controller framework described above guarantees that the tracking error is uniformly bounded if the following well known gradient based update laws that minimize a cost on the instantaneous tracking error $e^T e$ are used:

$$\dot{W}(t) = -\Gamma_W \phi(z(t)) e^T(t) PB. \tag{9}$$

However, this adaptive law guarantees that the adaptive parameters (W) stay bounded within a neighborhood of the ideal parameters (W^*) only if the regressor vector $\phi(z)$ is PE (see e.g. [35, 29, 28, 22]). Note that even when $\phi(z)$ is PE, $e(t) \nrightarrow 0$ uniformly only if $\sup_{z \in D_x} \|\bar{\eta}(z)\| = 0$. Therefore, this adaptive law cannot be used within the proposed framework as there is no guarantee that the weights will converge to their true values and an (approximate) representation of the system uncertainty will be learned. This has been a major reason why MRAC-MPC switching systems cannot be formulated easily.

Fortunately, it is possible to incorporate long term learning in the MRAC framework by ensuring that the adaptive law learns the modeling uncertainty by incorporating additional information [24, 12, 9]. It was shown in [8, 9] that for linearly parameterized uncertainties the requirement on persistency of excitation can be relaxed if online recorded data is used concurrently with instantaneous data for adaptation. In particular, for a linearly parameterized representations of the uncertainty, the following theorem can be proven [8, 9, 10]:

Theorem 1. *Consider the system given by (1), with the inverse law (2), and the reference model of (5). Assume that the uncertainty is linearly parameterizable using an appropriate set of bases over a compact domain D_x. For each online recorded data point i, let $\varepsilon_i(t) = W^T(t)\phi(x_i, \delta_i) - \hat{\Delta}(x_i, \delta_i)$, with $\hat{\Delta}(x_i, \delta_i) = \hat{\dot{x}}_i - \nu(x_i, \delta_i)$, where $\hat{\dot{x}}_i$ is the bounded estimate of $\dot{x}_i$, and consider the following update law*

$$\dot{W} = -\Gamma_W \phi(z) e^T PB - \frac{1}{p} \sum_{j=1}^{p} \Gamma_{W_b} \phi(x_i, \delta_i) \varepsilon_j^T, \tag{10}$$

where $\Gamma_{W_b} > 0$ is the learning rate for training on online recorded data. Let $Z = [\phi(z_1), ..., \phi(z_p)]$ and assume that $rank(Z) = m$. Furthermore, let B_α be the largest compact ball in D_x with radius α, and assume $x(0) \in B_\alpha$, define $\delta = \frac{2\|PB\|\bar{\eta}}{\lambda_{\min}(Q)} + \frac{p\bar{\eta}\sqrt{l}}{\lambda_{\min}(\Omega)}$ with $\Omega = ZZ^T$, and assume that D_x is sufficiently large such that $m_{rm} = \alpha - \delta$ is a positive scalar. If the states x_{rm} of the bounded input bounded output reference model of (5) remains bounded in the compact ball $B_m = \{x_{rm} : \|x_{rm}\| \leq m_{rm}\}$ for all $t \geq 0$ then the tracking error e and the weight error $\tilde{W} = W - W^*$ are mean-squared exponentially uniformly ultimately bounded. Furthermore, if the representation is exact over D_x, that is if $\sup_{z \in D_x} \|\bar{\eta}(z)\| = 0$, then the tracking error and weight error converge exponentially fast to a compact ball around the origin for arbitrary initial conditions, with the rate of convergence directly proportional to the minimum singular value of the history stack matrix Z.

Remark 1. The size of the compact ball around the origin where the weight and tracking error converge is dependent on the representation error $\bar{\eta}$ and the estimation error $\check{\varepsilon} = \max_i \|\dot{x}_i - \hat{\dot{x}}_i\|$. The former can be reduced by choosing appropriate number of RBFs across the operating domain, and the latter can be reduced by an appropriate implementation of a fixed point smoother. A fixed point smoother uses data before and after a data point is recorded to form very accurate estimates of $\hat{\dot{x}}_i$ using a forward-backward Kalman filter [18, 11]. Note that $\hat{\dot{x}}(t)$ is not needed at the *current* time instant t, which is a much more restrictive requirement. Therefore, an appropriate implementation of a fixed point smoother alleviates the time-delay often observed in estimating $\hat{\dot{x}}(t)$ with forward Kalman filter (or a low pass filter) only.

Remark 2. The history stack matrix $Z = [\phi(z_1), ..., \phi(z_p)]$ is not a buffer of the last p states. It can be updated online by including data points that are of significant interest over the course of operation. Theoretically, convergence is guaranteed as soon as the history stack becomes full ranked. New data points could replace existing data points once the history stack reaches a pre-determined size. It was shown in [10] that the rate of convergence of the tracking and weight error is directly proportional to the minimum singular value of Z. This provides a useful metric to determine which data points are most useful for improving convergence. Consequently, an algorithm for adding points that improve the minimum singular value of Z for the case of linearly parameterizable uncertainty was presented there.

Remark 3. The main limitation of the linearly parameterized RBF NN representation of the uncertainty is that the RBF centers need to be preallocated over an estimated compact domain of operation D_x. Therefore, if the system evolves outside of D_x all benefits of using adaptive control are lost. This can be addressed by evolving the RBF basis to reflect the current domain of operation. A reproducing kernel Hilbert space approach for accomplishing this was presented in [23]. However, when the basis is fixed, in order for the adaptive laws above to hold, the reference model and the exogenous reference commands should be constrained such that the desired trajectory does not leave the domain over which the neural network approximation is valid. Ensuring that the state remains within a given compact set implies an upper bound on the adaptation gain (see for example Remark 2 of Theorem-1 in [39]).

3 Feedback Linearization for MPC

The key enabling factor for the proposed switching CL-MPC architecture presented here is the guaranteed convergence property of CL-MRAC as established in Theorem 1. Once the approximation of the uncertainty is good enough the system shall change to the new MPC structure. Therefore a decision algorithm is implemented which tests for

$$\|x\| \neq 0 \;\text{ and }\; \|\hat{\dot{x}} - \nu - \nu_{ad}\| \leq \varepsilon_{\text{tol}}, \tag{11}$$

where $\varepsilon_{\text{tol}} \geq 0$ represents a tolerated approximation error. Note that due to Theorem 1 it can be shown that this guarantees an upper bound on $\tilde{W}(t_\sigma)$, where t_σ is a switching time. Note further that other automatic-switching algorithms, including those that approximate the switching surface probabilistically, are possible and expected to be investigated in our future work.

Once the weights converged to a neighborhood around their optimal values, as determined by the test in (11), the system switches to the model-based optimal controller. In this mode, the plant does not track a reference model but use the complete available control authority $\nu_{avail} \in \mathbb{R}^l$ after feedback linearization to track commands optimally. For the case that the system switched to the model-based optimal controller and with regard to equation (6) redefine the pseudo-control ν to be

$$\nu = \nu_{fb} + K_B \nu_{avail} - \nu_{ad}, \tag{12}$$

consisting of the linear feedback term $\nu_{fb} = K_M x$ with $K_M \in \mathbb{R}^{n\times n}$; a feedforward part $K_B \nu_{avail}$ with $K_B \in \mathbb{R}^{l\times l}$; and the adaptive part ν_{ad}. Let $B_m = BK_B$, then the feedback linearized system becomes

$$\dot{x}(t) = A_m x + B_m \nu_{avail} + B(\Delta - \nu_{ad}), \tag{13}$$

where the state space model (A_m, B_m) is in canonical form with the eigenvalues of A_m assigned by ν_{fb}. Choose the gains such that if $\Delta - \nu_{ad} = 0$, a unique solution to (13) exists and B_m satisfies assumption 1. Furthermore, the resulting matrices (A_m, B_m) need to be chosen such that a feasible optimal solution to the system (13) is known; one possibility is to choose (A_m, B_m) equal to the reference model, which was used during the exclusively adaptive case. The available control authority ν_{avail} is dynamically constrained by the physical maximum and minimum control allowed $\nu_{min/max}$, minus the adaptive part (ν_{ad}) of the pseudo control which is needed to cancel the uncertainty and the part (ν_{fb}) which is required in order for the feedback linearized system to recover the dynamics in 13. For each element of ν_{fb} we have

$$K_B^{-1}(\nu_{min} + \nu_{ad} - \nu_{fb}) \leq \nu_{avail} \leq K_B^{-1}(\nu_{max} + \nu_{ad} - \nu_{fb}). \tag{14}$$

Using equation (4), the last term in equation (13) is

$$\|\Delta(z) - \nu_{ad}(z)\| \leq \|\tilde{W}\| \|\phi(z)\| + \eta_{\text{sup}}. \tag{15}$$

Let $\beta(z) = \Delta(z) - \nu_{ad}$. The feedback linearized system can be written as

$$\dot{x}(t) = A_m x + B_m \nu_{avail} + B\beta(z) \tag{16}$$

Let t_σ be a time instant when the control architecture switches to using MPC. Due to Theorem 1 it follows that $\tilde{W}(t)$ approaches a neighborhood of zero exponentially fast, furthermore, since the algorithm switches to the optimal controller (MPC) only when $\|\dot{\hat{x}} - \nu + \nu_{ad}\| \leq \varepsilon_{\text{tol}}$ and $\|x\| \neq 0$, it follows that $\|\tilde{W}(t_\sigma)\|$ is small. Leveraging this fact, MPC design is performed on the ideal feedback linearized system with states $\bar{x}(t)$ given by

$$\dot{\bar{x}}(t) = A_m \bar{x}(t) + B_m \nu_{avail}(t), \tag{17}$$

assuming $\beta(z) = 0$.

3.1 Stability

Let $[t_\sigma, t_{\sigma+1}]$ be a finite interval where the algorithm has switched to using the optimal model based controller (e.g. MPC). It is clear that when the algorithm switches, although $\beta(z)$ is likely to be very small, it will probably not be zero. In this case, the question arises as to whether learning should continue or not. Since any possible initial transients have already passed, there seems no reason to continue to learn. In fact, such an approach can be thought equivalent to an assumption on allowable initial parameter error $\tilde{W}(0)$ for a non-switching based MPC [17, 1, 5]. One approach therefore, could be to continue to learn using a smaller learning-rate Γ and using estimates of model error only (not using also the tracking error as was the case in Theorem 1). The following lemma characterizes that in this case, a concurrent learning gradient descent law guarantees that the feedback linearization error $\beta(z)$ is exponentially bounded. To facilitate the analysis, it is assumed that a noise free estimate of $\dot{x}_i$ for all online recorded data points i is available. This assumption can be relaxed to yield mean squared exponential ultimate boundedness of $\tilde{W}$ instead of mean square exponential stability [27].

Lemma 1. *Consider the model error given by (3), $\varepsilon_i(t)$ as defined in (10) for the recorded data points, and the following gradient descent law*

$$\dot{W} = -\Gamma \sum \phi(x_i, \delta_i)\varepsilon_j^T). \tag{18}$$

Assume also that the history stack is full ranked, that is $rank(Z) = m$, then the parameter error is exponentially bounded as $\|\tilde{W}(t)\| \leq \exp^{-c_1 t} \|\tilde{W}(t_\sigma)\|$ for some $c_1 > 0$ dependent on Z and the parameter error $\tilde{W}(t_\sigma)$ at the instant the algorithm switches to model based optimal control. Furthermore,
$\beta(z(t)) \leq \exp^{-c_1 t} \|\tilde{W}(t_\sigma)\| \phi(z(t)) + \eta_{\sup}$ for all $t \in [t_\sigma, t_{i+1}]$.

Proof. Consider the quadratic function given by $V(\tilde{W}) = \frac{1}{2}\tilde{W}(t)^T \Gamma^{-1} \tilde{W}(t)$, and note that $V(0) = 0$ and $V(\tilde{W}) > 0 \ \forall \ \tilde{W} \neq 0$, hence $V(\tilde{W})$ is a Lyapunov function candidate. Since $V(\tilde{W})$ is quadratic, letting $\lambda_{\min}(.)$ and $\lambda_{\max}(.)$ denote the operators that return the minimum and maximum eigenvalue of a matrix, we have: $\lambda_{\min}(\Gamma^{-1})\|\tilde{W}\|^2 \leq V(\tilde{W}) \leq \lambda_{\max}(\Gamma^{-1})\|\tilde{W}\|^2$. Differentiating the Lyapunov candidate with respect to time along the trajectories of (18) we have

$$\dot{V}(\tilde{W}(t)) \leq -\tilde{W}(t)^T [\sum_{j=1}^{p} \phi(x_j)\phi^T(x_j)]\tilde{W}(t). \tag{19}$$

Let $\Omega = \sum_{j=1}^{p} \Phi(x_j)\Phi^T(x_j)$ and note that since $\phi(x(t))\phi^T(x(t)) \geq 0\, \forall \phi(x(t))$, $\lambda_{min}(\Omega) > 0$. Then it follows that

$$\dot{V}(\tilde{W}) \leq -\lambda_{\min}(\Omega)\|\tilde{W}\|^2 \leq -\frac{\lambda_{\min}(\Omega)}{\lambda_{\max}(\Gamma^{-1})}V(\tilde{W}). \tag{20}$$

Let $c_1 = \frac{\lambda_{\min}(\Omega)}{\lambda_{\max}(\Gamma^{-1})}$, then $\|\tilde{W}(t)\| \leq \exp^{-c_1 t}\|\tilde{W}(t_\sigma)\|$. It follows from the definition of $\beta(z)$ in (16) that for all $t \in [t_\sigma, t_{i+1}]$

$$\beta(z(t)) \leq \exp^{-c_1 t}\|\tilde{W}(t_\sigma)\|\phi(z(t)) + \eta_{\sup}. \tag{21}$$

□

The next theorem shows that $\tilde{x} = x - \bar{x}$, the difference between the ideal feedback linearized system (16) and the true feedback linearized system (17) is bounded. Therefore, applying the stabilizing feasible solution of the ideal system (17) to the true system (16) guarantees boundedness of the true system states.

Theorem 2. *Consider the true feedback linearized system in (16) and Lemma 1. Assume that a feasible optimal solution ν^*_{avail} exists for the ideal feedback linearized system of (17). Then, the states of the true system with the control ν^*_{avail} are uniformly ultimately bounded around those of the ideal system, and approach asymptotically a compact set that is a function of the representation error $\eta_{\sup}$ over every switching interval $[t_\sigma, t_{\sigma+1}]$ where MPC control is active.*

Proof. Note that

$$\dot{\tilde{x}} = A_m\tilde{x} + B\beta(z). \tag{22}$$

Let $V(\tilde{x}) = \frac{1}{2}\tilde{x}^T P_m\tilde{x}$, where P_m is the positive definite solution to $0 = Q_m + A_m^T P_m + P_m A_m$ for a positive definite Q_m, guaranteed to exist due to the feedback ν_{fb}, which is chosen such that A_m in (13) is Hurwitz. Hence $V(\tilde{W})$ is a radially unbounded quadratic Lyapunov function candidate with: $\lambda_{\min}(\Gamma^{-1})\|\tilde{W}\|^2 \leq V(\tilde{W}) \leq \lambda_{\max}(\Gamma^{-1})\|\tilde{W}\|^2$. It follows therefore that

$$\dot{V}(\tilde{x}) \leq -\tilde{x}^T Q_m\tilde{x} + \tilde{x}^T PB\beta(z). \tag{23}$$

Applying Lemma 1 we have

$$\dot{V}(\tilde{x}) \leq -\lambda_{\min}(Q_m)\|\tilde{x}\|^2 + \|\tilde{x}\|\|P_m B\|(\exp^{-c_1 t}\|\tilde{W}(t_\sigma)\|\phi(z(t))\| + \eta_{\sup}). \tag{24}$$

Let $c_2 = \|P_m B\|\|\tilde{W}(t_\sigma)\|$, and noting that the m basis functions are bounded by $\|\phi(.)\| \leq c_3$, we have

$$\dot{V}(\tilde{x}) \leq -\lambda_{\min}(Q_m)\|\tilde{x}\|^2 + \|\tilde{x}\|(mc_2c_3\exp^{-c_1 t} + \eta_{\sup}). \tag{25}$$

Therefore, outside of the compact set $\|\tilde{x}\| \geq \frac{mc_2c_3\exp^{-c_1t}+\eta_{\sup}}{\lambda_{\min}Q_m}$, $\dot{V}(\tilde{x}) \leq 0$. Therefore $\tilde{x}$ is uniformly ultimately bounded and approaches asymptotically the set $\|\tilde{x}\| \geq \frac{\eta_{\sup}}{\lambda_{\min}Q_m}$.

□

Corollary 1. *Assume that Theorem 2 holds and that an exact representation exists such that $\eta_{\sup} = 0$ in (4), then, the states of the true feedback linearized system asymptotically approach the states of the ideal feedback linearized system over every switching interval $[t_\sigma, t_{\sigma+1}]$ where MPC control is active.*

Proof. The proof follows by noting that (25) becomes

$$\dot{V}(\tilde{x}) \leq -\lambda_{\min}(Q_m)\|\tilde{x}\|^2 + \|\tilde{x}\|(mc_2c_3\exp^{-c_1t}), \quad (26)$$

hence, $V \to 0$ as $t \to \infty$.

□

4 Model Predictive Control

For the implementation of the MPC a discrete model of the feedback linearized system in equation (17) is formulated:

$$\bar{x}(k+1) = \bar{A}_m\bar{x}(k) + \bar{B}_m\nu_{avail}(k), \quad (27)$$

where $\bar{A}_m$ and $\bar{B}_m$ denote the discretized versions of the respective matrices in equation (17). Let $\Delta\nu_{avail}(k+1) = \nu_{avail}(k+1) - \nu_{avail}(k)$ be a future incremental control. The optimal control trajectory is captured by a sequence of incremental control signals:

$$\Delta U = \begin{bmatrix} \Delta\nu_{avail}(k) \\ \Delta\nu_{avail}(k+1) \\ \vdots \\ \Delta\nu_{avail}(k+N_c-1) \end{bmatrix}, \quad (28)$$

where N_c denotes the control horizon. Within the prediction horizon $N_p \geq N_c$ the MPC drives the state of the system $\bar{x}(k)$ onto the desired reference signal $r(k)$ by minimization of a quadratic cost function. Define the following matrices:

$$F = \begin{bmatrix} \bar{A}_m \\ \bar{A}_m^2 \\ \vdots \\ \bar{A}_m^{N_p} \end{bmatrix}, \Phi = \begin{bmatrix} \bar{B}_m & 0 & \dots\dots & 0 \\ \bar{A}_m\bar{B}_m & \bar{B}_m & 0\ \dots & 0 \\ \vdots & & & \\ \bar{A}_m^{N_p-1}\bar{B}_m & \bar{A}_m^{N_p-2}\bar{B}_m & \dots\dots & \bar{A}_m^{N_p-N_c}\bar{B}_m \end{bmatrix}, \quad (29)$$

where $F \in \mathbb{R}^{n\cdot N_p \times n}$ and $\Phi \in \mathbb{R}^{N_p\cdot n \times N_c\cdot n_s}$. Let $\Delta\bar{x}(k+1) = \bar{x}(k+1) - \bar{x}(k)$. Then the vector containing the predicted states $X \in \mathbb{R}^{n\cdot N_p}$ within the prediction horizon can be built by

$$X = F\Delta\bar{x}(k) + \Phi\Delta U. \tag{30}$$

In the MPC framework constraints can be formulated for the input and the states. The goal is to formulate the constraints dependent on the incremental control ΔU. For the control input we have

$$\nu_{avail,min} \leq M_1\nu(k-1) + M_2\Delta U \leq \nu_{avail,max}, \tag{31}$$

where $M_1 = \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix} \in \mathbb{R}^{Nc\cdot n_S \times n_S}$ and $M_2 = \begin{bmatrix} I\ 0 \ldots \ldots 0 \\ I\ I\ \ 0\ \ldots 0 \\ \vdots \\ I \quad \ldots \quad I \end{bmatrix} \in \mathbb{R}^{n_S\cdot N_c \times n_S\cdot N_c}$. The input constraints are placed by the vectors $\nu_{avail,min}$ and $\nu_{avail,max}$, each consisting of N_c elements of the minimum and maximum control input. Equation (31) can also be expressed in matrix form:

$$\begin{bmatrix} -M_2 \\ M_2 \end{bmatrix} \Delta U \leq \begin{bmatrix} -\nu_{avail,min} + M_1\nu_{avail}(k-1) \\ \nu_{avail,max} - M_1\nu_{avail}(k-1) \end{bmatrix}. \tag{32}$$

Similarly, constraints on the states can be placed by

$$X_{min} \leq F\Delta\bar{x}(k_i) + \Phi\Delta U \leq X_{max}. \tag{33}$$

Similar to equation (31), X_{min} and X_{min} are vectors containing the lower and upper constraints for the states. Written in matrix form we have

$$\begin{bmatrix} -\Phi \\ \Phi \end{bmatrix} \Delta U \leq \begin{bmatrix} -X_{min} + F\Delta\bar{x}(k) \\ X_{max} - F\Delta\bar{x}(k) \end{bmatrix}. \tag{34}$$

There exists a functional relationship between the predicted system state and the incremental control input $\Delta\nu_{avail}$. Using hard constraints on input and the states simultaneously can cause constraint conflicts. Introducing a slack variable $s > 0$ and relaxing the constraints $X_{min/max}$ solves this problem. Let $R_S \in \mathbb{R}^{n\cdot N_p}$ be a vector containing the reference command $r(k)$ with $R_s = [1,...,1]r(k)$ and define the following quadratic cost function, which reflects the control objective:

$$J = (R_s - X)^T\bar{Q}(R_s - X) + \Delta U^T\bar{R}\Delta U. \tag{35}$$

Here $\bar{R}$ and $\bar{Q}$ denote positive definite diagonal matrices. Inserting equation (30) into equation (35) the problem of model predictive control is finding the control sequence ΔU which minimizes the cost function

$$\begin{aligned} J = (R_s - F\bar{x}(k))^T\bar{Q}(R_s - F\bar{x}(k)) - 2\Delta U^T\Phi^T\bar{Q}(R_s - F\Delta\bar{x}(k)) \\ + \Delta U^T(\Phi^T\bar{Q}\Phi + \bar{R})\Delta U, \end{aligned} \tag{36}$$

subject to the inequality constraint

$$\begin{bmatrix} -M_2 \\ M_2 \\ -\Phi \\ \Phi \end{bmatrix} \Delta U \leq \begin{bmatrix} -\nu_{avail,min} + M_1 \nu_{avail}(k-1) \\ \nu_{avail,max} - M_1 \nu_{avail}(k-1) \\ -X_{min} + F\Delta\bar{x}(k) \\ X_{max} - F\Delta\bar{x}(k) \end{bmatrix}. \tag{37}$$

Note that since A_m is known a-priori, it may be possible to solve a significant portion of this problem off-line to create the optimal value-function which can be directly used on-line for an approximate optimal solution.

5 Trajectory Tracking in the Presence of Wing Rock Dynamics

Modern highly swept-back or delta wing fighter aircraft are susceptible to lightly damped oscillations in roll angle known as "Wing Rock". Wing rock often occurs at low speeds and at high angle of attack, conditions commonly encountered when the aircraft is landing (see [34] for a detailed discussion of the wing rock phenomena). Hence, precision control of the aircraft in the presence of wing rock dynamics is critical in order to ensure safe landing. In this section we use concurrent learning adaptive control and the proposed MPC framework to track a sequence of roll commands in the presence of wing rock dynamics. Let ϕ denote the roll attitude of an aircraft, p denote the roll rate and δ_a denote the aileron control input . Then a model for wing rock dynamics is ([26]):

$$\dot{\phi} = p \tag{38}$$

$$\dot{p} = \delta_a + \Delta(x), \tag{39}$$

where $\Delta(x) = W_0 + W_1\phi + W_2 p + W_3|\phi|p + W_4|p|p + W_5\phi^3$. The parameters for wing rock motion are partly adapted from [36, 38, 7]; they are $W_1 = 6.2314, W_2 = 2.1918, W_3 = -0.6245, W_4 = 0.0095, W_5 = 0.0214$. In addition to these parameters, a trim error is introduced by setting $W_0 = 0.8$. A simple inversion model has the form $\nu = \delta_a$. The linear part of the control law is given by $\nu_{pd} = -4\phi - 2p$ for the exclusive adaptive as well as the MPC part of the control framework of Figure 1. These values are chosen as they result in good baseline control performance without exciting high-gain oscillations. Furthermore, in the MPC part the feedforward gain is chosen to be $K_B = 4$. The reference model is chosen to be of second order with natural frequency of $2\,rad/sec$ and a damping ratio of 0.5. This choice results in reasonably smooth trajectories without large transients and without exceeding the constraints when baseline or CL-MRAC controllers are used. The learning rate is set to $\Gamma_W = 6$ for both the instantaneous update and the update based on stored data.

For the concurrent learning adaptive controller only points which increase the rank of the history stack are considered for storage. As long as the history stack does not contain at least as many linearly independent data points as the dimension of the regressor vector, a σ-modification term with gain $\kappa = 0.01$ is added to the update law. Once the history stack is full, an algorithm is employed which increases its minimum singular value ([10]).

The simulation runs for a total of $60\,s$ with a time step of $0.01\,s$. The reference signal $r(t)$ is comprised out of several step inputs. The first two steps start at $5\,s$ and $15\,s$, each having an amplitude of $30°$ and lasting $5\,s$. The next two step inputs occur after $25\,s$ and $35\,s$, each having an amplitude of $45°$ and also lasting $5\,s$. After $50\,s$ more aggressive commands shall be tracked. Therefore consecutive steps with an amplitude of $-45°$ or $45°$ are commanded, alternating every $3\,s$.

Figure 2 shows the performance of the proposed control architecture. During the first step the plant states still deviate from the reference model significantly. However, the tracking performance increases quickly, the plant tracks the reference model at the second step nearly perfectly. After about $30\,s$ the switching condition is met and the system automatically switches to the MPC part of the control framework. It can be observed that the performance increases drastically. This is attributed to the fact that the CL-MPC architecture leverages available control authority fully while simultaneously ensuring that the constraint on the roll rate is not violated.

Figure 3 shows the evolution of the adaptive weights. As soon as the history stack meets the rank condition after about $6\,s$ the weights start to converge to their optimal values, thus increasing the tracking performance significantly. At the switch to the MPC framework the weights have already nearly converged to their optimal weights. Still, the resulting parameter error is further reduced by the CL update law of Lemma 1 which learns only on stored data.

Figure 4 shows the trajectory of the system in the phase plane during the simulation. It can be seen that, once the MPC is switched on, the region the states reside in increases drastically. This is attributed to the fact that the full available control authority is used, thus increasing the roll rate in transient phases. In addition, despite the increased performance, the chosen state constraints on the roll rate are not violated.

Finally, Figure 5 shows the control input. As long as only the adaptive controller is used, the available control authority is not completely leveraged. Once the MPC is switched on, the complete available control authority is used. Additionally, the constraints placed on the input are not violated.

6 Conclusion

Initial transients often observed during online learning can result in undesirable performance of (receding horizon) online optimal control architectures such as Model Predictive Control. This could make it difficult to implement adaptive MPC on aerospace systems that have fast dynamics. We proposed an adaptive-optimal control architecture in which a concurrent learning adaptive controller is used to guarantee system stability while parameters are adaptively learned. The online-learned model is used to feedback linearize the system and transform its behavior to an ideal feedback-linearized system for which a feasible optimal MPC can be formu-

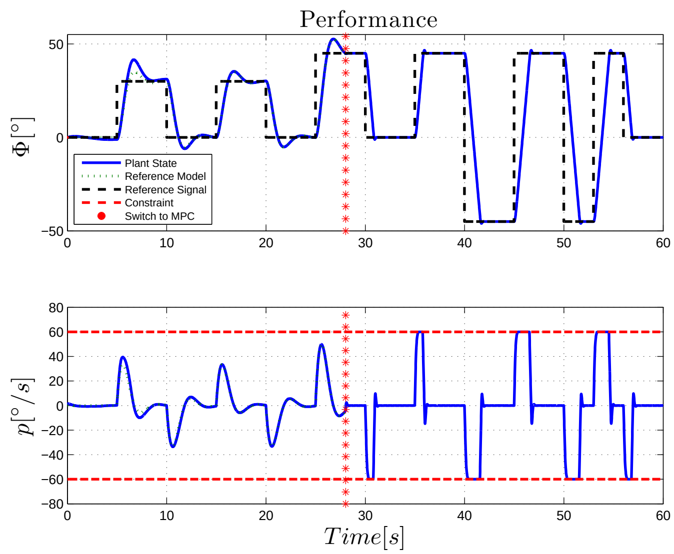

Fig. 2 Performance of the proposed control architecture. At the beginning of the simulation a distinct deviation tracking error is observed. Due to the concurrent learning adaptive controller the performance increases drastically over time. After the switch to the MPC framework, instead of tracking the suboptimal reference model, the plant tracks the command optimally. The constraints on the roll rate are not violated.

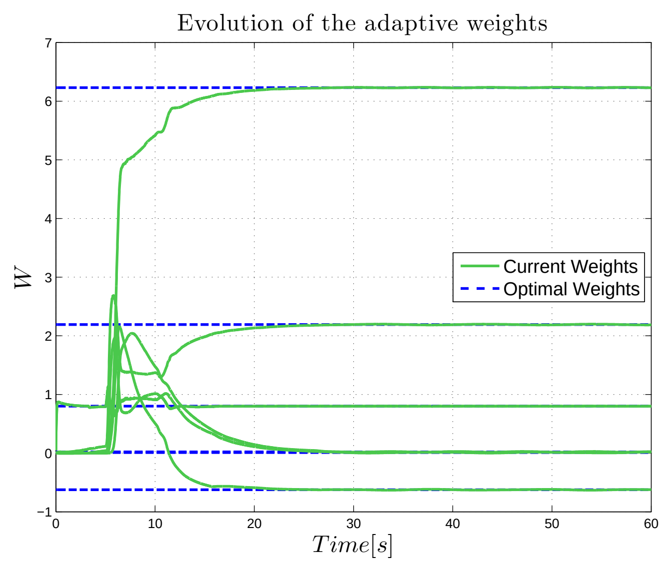

Fig. 3 Evolution of the adaptive weights. Once the history stack meets a condition of linearly independence on the stored data the weights start to converge to their optimal values. Even after the switch to the MPC framework learning based on stored data continues using the algorithm in Lemma 1 and the parameter error is further reduced.

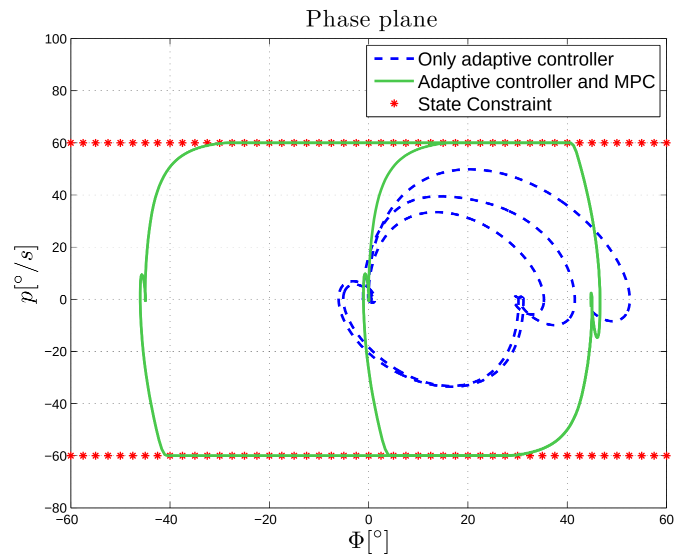

Fig. 4 Phase portrait of the system trajectory. Once the controller switches to the MPC framework the region in which the states evolve is drastically increased as the controller can execute optimal commands w.r.t. the constraints. In addition, MPC ensures that aggressive commands can be tracked without violating the constraints on the roll rate.

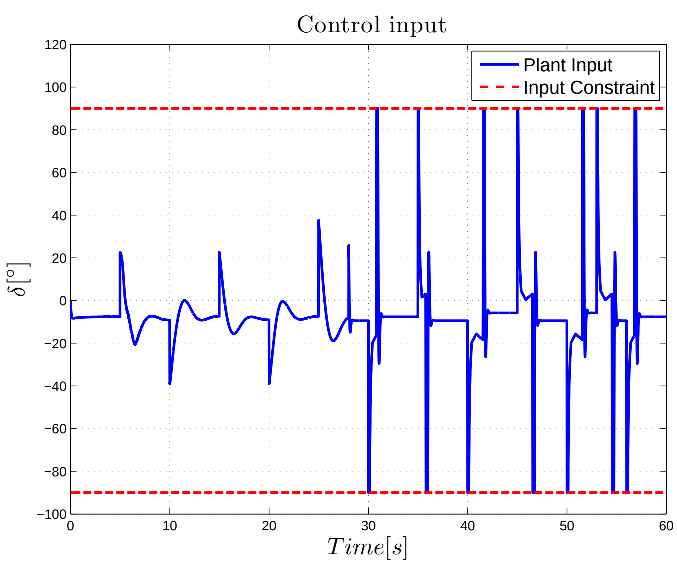

Fig. 5 Control input with constraints. In the beginning of the simulation the controller is concerned with letting the uncertain system behave like the reference model. For this, only a part of the available control authority is used as a conservative reference model is used to ensure constraints are not vioalated. Once the controller switches to the MPC framework the complete control authority is leveraged without violating input constraints.

lated. The MPC takes over after an online metric has gauged sufficient confidence in the learned parameters. It was shown that the states of the feedback linearized system stay exponentially mean square bounded around those of the ideal system, and sufficient conditions were provided to guarantee asymptotic convergence. Simulation results were presented on a wing-rock dynamics system with fast dynamics. These results establish the feasibility of the CL-MPC architecture. Furthermore, these results indicate that learning in adaptive controllers can be used to improve the performance of the system.

References

1. Veronica Adetola, Darryl DeHaan, and Martin Guay. Adaptive model predictive control for constrained nonlinear systems. *Systems and Control Letters*, 58(5):320 – 326, 2009.
2. Karl Johan Aström and Björn Wittenmark. *Adaptive Control*. Addison-Weseley, Readings, 2nd edition, 1995.
3. A. Aswani, Sastry S. S. Humberto, G., and C. Tomlin. Provably safe and robust learning-based model predictive control. arxiv arXiv:1107.2487v2, UC Berkley, 2011.
4. Alberto Bemporad and Manfred Morari. Robust model predictive control: A survey. In A. Garulli and A. Tesi, editors, *Robustness in identification and control*, volume 245 of *Lecture Notes in Control and Information Sciences*, pages 207–226. Springer Berlin / Heidelberg, 1999. 10.1007/BFb0109870.
5. P. Bouffard, A. Aswani, and C. Tomlin. Learning-based model predictive control on a quadrotor: Onboard implementation and experimental results. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 279 –284, may 2012.
6. Stephan Boyd and Sankar Sastry. Necessary and sufficient conditions for parameter convergence in adaptive control. *Automatica*, 22(6):629–639, 1986.
7. Anthony Calise and Tansel Yucelen. Enforcing a linear constraint in adaptive control: A kalman filter optimization approach. In *Proceedings of the AIAA Guidance Navigation and Control Conference, held at Chicago, IL*, 2009.
8. Girish Chowdhary. *Concurrent Learning for Convergence in Adaptive Control Without Persistency of Excitation*. PhD thesis, Georgia Institute of Technology, Atlanta, GA, 2010.
9. Girish Chowdhary and Eric N. Johnson. Concurrent learning for convergence in adaptive control without persistency of excitation. In *49th IEEE Conference on Decision and Control*, pages 3674–3679, 2010.
10. Girish Chowdhary and Eric N. Johnson. A singular value maximizing data recording algorithm for concurrent learning. In *American Control Conference*, San Francisco, CA, June 2011.
11. Girish Chowdhary and Eric N. Johnson. Theory and flight test validation of a concurrent learning adaptive controller. *Journal of Guidance Control and Dynamics*, 34(2):592–607, March 2011.
12. Manuel A. Duarte and Kumpati S. Narendra. Combined direct and indirect approach to adaptive control. *IEEE Transactions on Automatic Control*, 34(10):1071–1075, 1989.
13. W. Dunbar. Distributed receding horizon control of dynamically coupled nonlinear systems. *IEEE Transactions on Robotics*, 52:1249–1263, 2007.
14. W. B. Dunbar and R. Murray. Model predictive control of coordinated multi-vehicle formations. In *IEEE Conference on Decision and Control (CDC)*, pages 4631–4636, 2002.
15. W.B. Dunbar and R.M. Murray. Receding horizon control of multi-vehicle formation: A distributed implementation. In *IEEE Conference on Decision and Control (CDC)*, 2004.
16. W. Durham. Constrained control allocation. *AIAA Journal of Guidance, Control, and Dynamics*, 16:717–772, 1993.

17. Hiroaki Fukushima, Tae-Hyoung Kim, and Toshiharu Sugie. Adaptive model predictive control for a class of constrained linear systems based on the comparison model. *Automatica*, 43(2):301 – 308, 2007.
18. Arthur Gelb. *Applied Optimal Estimation*. MIT Press, Cambridge, 1974.
19. P. IOANNOU and J. SUN. Theory and design of robust direct and indirect adaptive-control schemes. *International Journal of Control*, 47(3):775–813, 1988.
20. E. C. Kerrigan. *Robust Constraint Satisfaction Invariant Sets and Predictive Control*. PhD thesis, University of Cambridge, Department of Engineering, November 2000.
21. Nakawan Kim. *Improved Methods in Neural Network Based Adaptive Output Feedback Control, with Applications to Flight Control*. PhD thesis, Georgia Institute of Technology, Atlanta Ga, 2003.
22. Y. H. Kim and F.L. Lewis. *High-Level Feedback Control with Neural Networks*, volume 21 of *Robotics and Intelligent Systems*. World Scientific, Singapore, 1998.
23. H. A. Kingravi, G. Chowdhary, P. A. Vela, and E. N. Johnson. Reproducing kernel hilbert space approach for the online update of radial bases in neuro-adaptive control. *Neural Networks and Learning Systems, IEEE Transactions on*, 23(7):1130 –1141, july 2012.
24. E. Lavretsky. Combined/composite model reference adaptive control. *Automatic Control, IEEE Transactions on*, 54(11):2692 –2697, nov. 2009.
25. D. Q. Mayne. *Nonlinear Model Predictive Control*, volume 26 of *Progress in Systems and Control Theory*, chapter Nonlinear Model Predictive Control: Challenges and Opportunities, pages 23–44. Birkhäuser, Basel-Boston-Berlin, 2000.
26. Morgan M. Monahemi and Miroslav Krstic. Control of wingrock motion using adaptive feedback linearization. *Journal of Guidance Control and Dynamics*, 19(4):905–912, August 1996.
27. M. Mühlegg, G. Chowdhary, and E. Johnson. Concurrent learning adaptive control of linear systems with noisy measurements. In *AIAA Guidance, Navigation, and Control Conference*, aug 2012.
28. Flavio Nardi. *Neural Network based Adaptive Algorithms for Nonlinear Control*. PhD thesis, Georgia Institute of Technology, School of Aerospace Engineering, Atlanta, GA 30332, nov 2000.
29. K.S. Narendra. Neural networks for control theory and practice. *Proceedings of the IEEE*, 84(10):1385 –1406, oct 1996.
30. G. D. Nicolao, L. Magi, and R. Scattolini. *Nonlinear Model Predictive Control*, volume 26 of *Progress in Systems and Control Theory*, chapter Stability and Robustness of Nonlinear Receding Horizon Control, pages 3–22. Birkhäuser, Basel-Boston-Berlin, 2000.
31. J.B. Rawlings. Tutorial overview of model predictive control. *Control Systems Magazine, IEEE*, 20(3):38 –52, June 2000.
32. A. Richards and J. P. How. Decentralized model predictive control of cooperating UAVs. In *IEEE Conference on Decision and Control (CDC)*, pages 4286–4291, Paradise Island, Bahamas, December 2004.
33. Arthur G. Richards. *Robust Constrained Model Predictive Control*. PhD thesis, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, Cambridge MA, February 2005.
34. Ahmed A. Saad. *Simulation and analysis of wing rock physics for a generic fighter model with three degrees of freedom*. PhD thesis, Air Force Institute of Technology, Air University, Wright-Patterson Air Force Base, Dayton, Ohio, 2000.
35. R.M. Sanner and J.-J.E. Slotine. Gaussian networks for direct adaptive control. *Neural Networks, IEEE Transactions on*, 3(6):837 –863, nov 1992.
36. S. N. Singh, W. Yim, and W. R. Wells. Direct adaptive control of wing rock motion of slender delta wings. *Journal of Guidance Control and Dynamics*, 18(1):25–30, Feb. 1995.
37. Gang Tao. *Adaptive Control Design and Analysis*. Wiley, New York, 2003.
38. Konstantine Volyanskyy and Anthony Calise. An error minimization method in adaptive control. In *Proceedings of AIAA Guidance Navigation and Control conference*, 2006.
39. Tansel Yucelen and Anthony Calise. Kalman filter modification in adaptive control. *Journal of Guidance, Control, and Dynamics*, 33(2):426–439, march-april 2010.

40. T. Zhang, S.S. Ge, and C.C. Hang. Direct adaptive control of non-affine nonlinear systems using multilayer neural networks. In *American Control Conference, 1998. Proceedings of the 1998*, volume 1, pages 515 –519 vol.1, jun 1998.