Least Squares Based Modification for Adaptive Control

Girish Chowdhary and Eric Johnson

***Abstract*— A least squares modification is presented to adaptive control problems where the uncertainty can be linearly parameterized. The modified weight training law uses an estimate of the ideal weights formed online by solving a least squares problem using recorded and current data concurrently. The modified adaptive law guarantees the exponential convergence of adaptive weights to their ideal values subject to a verifiable condition on linear independence of the recorded data. This condition is found to be less restrictive and easier to monitor than a condition on persistency of excitation of the reference signal.**

I. INTRODUCTION

In plants where the plant uncertainty can be linearly parameterized, the standard approach to the design of Model Reference Adaptive Controllers (MRAC) is to use a weight training law that attempts to estimate the ideal weights in order to cancel the uncertainty. If the weights do converge to their ideal values, the linear part of the MRAC tracking error dynamics dominate, greatly improving the performance and possibly allowing the use of linear metrics to characterize the adaptive controller. Most adaptive control methods are designed using a Lyapunov based approach and can be thought of as attempting to minimize a quadratic cost function by using a gradient descent type method. Gradient descent type methods however, are susceptible to local adaptation and weight drift. Furthermore, it can be shown that adaptive laws formulated using the gradient descent methodology are always at most rank 1 [5]. Boyd and Sastry have shown that in order to guarantee weight convergence for these adaptive laws the exogenous reference input must have as many spectral lines as the unknown parameters, a condition relating to Persistency of Excitation (PE) in the reference input [1]. The condition on PE states is required to guarantee parameter convergence for many classic and recent adaptive control laws as well (e.g. σ-mod [8], e-mod [11], Q-mod [16], and $L-1$ adaptive control [2]). However, it is hard to monitor whether a signal is PE, and using PE inputs only for weight convergence may waste control effort and cause undue stress.

Previously, we have suggested *Concurrent Learning* adaptive control, which uses past and current data concurrently for adaptation, for ensuring parameter convergence without requiring PE [5], [4]. In this paper we maintain the idea of using past and current data concurrently for adaptation, however, we use optimal least squares based approach rather than gradient descent for adaptation based on recorded data.

G. Chowdhary and Assoc. Prof. E. N. Johnson are with the Daniel Guggenheim school of aerospace engineering at the Georgia Institute of Technology, Atlanta GA, Girish.Chowdhary@gatech.edu, Eric.Johnson@ae.gatech.edu

Least squares, which offers the best linear fit for a set of data, has been widely studied for real time parameter estimation. The main contribution of this paper is the development of a modification term that brings the desirable properties of least squares to any MRAC gradient based adaptive law. The modified adaptive law ensures that the adaptive weights converge smoothly to an optimal unbiased estimate of the ideal weights. Furthermore, exponential tracking error and exponential weight convergence can be guaranteed if the recorded data meet a verifiable condition on linear independence, which is found to be less restrictive than PE. We use a simulation study of wing rock dynamics to demonstrate the effectiveness of the modified adaptive law.

II. MODEL REFERENCE ADAPTIVE CONTROL

This section discusses the formulation of model reference adaptive control (see e.g. [8], and [14]). Let $x(t) \in \Re^n$ be the known state vector, let $u \in \Re$ denote the control input, and consider the following system where the uncertainty can be linearly parameterized:

$$\dot{x} = Ax(t) + B(u(t) + \Delta(x(t))), \tag{1}$$

where $A \in \Re^{n \times n}$, $B \in \Re^n$, $B = [0, 0, ..., 1]^T$, and $\Delta(x) \in \Re(x)$ is a continuously differentiable function representing the scalar uncertainty. We assume that the system in 1 is controllable.

A reference model can be designed that characterizes the desired response of the system:

$$\dot{x}_{rm} = A_{rm} x_{rm}(t) + B_{rm} r(t), \tag{2}$$

where $A_{rm} \in \Re^{n \times n}$ is a Hurwitz matrix and $r(t)$ denotes a bounded reference signal. A tracking control law consisting of a linear feedback part $u_{pd} = K(x_{rm}(t) - x(t))$, a linear feedforward part $u_{crm} = K_r[x_{rm}^T, r(t)]^T$, and an adaptive part $u_{ad}(x)$ is proposed to have the following form

$$u = u_{crm} + u_{pd} - u_{ad}. \tag{3}$$

Define the tracking error e as $e(t) = x_{rm}(t) - x(t)$; with an appropriate choice of u_{crm} to satisfy the matching condition $Bu_{crm} = (A_{rm} - A)x_{rm} - B_{rm}r(t)$, the tracking error dynamics can be reduced to

$$\dot{e} = A_m e + B(u_{ad}(x) - \Delta(x)), \tag{4}$$

where the baseline full state feedback controller $u_{pd} = Kx$ is assumed to be designed such that $A_m = A - BK$ is a Hurwitz matrix. Hence for any positive definite matrix $Q \in \Re^{n \times n}$, a positive definite solution $P \in \Re^{n \times n}$ exists to the Lyapunov equation,

$$A_m^T P + P A_m + Q = 0. \tag{5}$$

The following two cases characterizing the uncertainty are considered:

Case 1: Structured Uncertainty: Consider the case where the form of the uncertainty is known, that is the mapping $\Phi(x)$ is known. The following assumption captures this case,

Assumption 1 The uncertainty $\Delta(x)$ can be linearly parameterized, that is, there exist a vector of constants $W = [w_1, w_2, ..., w_m]^T$ and a vector of continuously differentiable functions $\Phi(x) = [\phi_1(x), \phi_2(x), ..., \phi_m(x)]^T$ such that

$$\Delta(x) = W^T \Phi(x). \tag{6}$$

In this case letting $\hat{W}$ denote the estimate W the adaptive law can be written as

$$u_{ad}(x) = \hat{W}^T \Phi(x). \tag{7}$$

On examining equation 4 we see that if $\hat{W} = W$ then the tracking error dynamics are globally exponentially stable. Hence, we desire to design an adaptive law such that $\hat{W}(t) \to W$. For this case it is well known that the following adaptive law

$$\dot{\hat{W}} = -\Phi(x(t)) e^T P B \Gamma_W \tag{8}$$

where Γ_W is a positive definite matrix of appropriate dimensions results in $e(t) \to 0$ [14]. For this baseline adaptive law, it is also well known that a necessary and sufficient condition for guaranteeing $\hat{W}(t) \to W$ is that $\Phi(t)$ be persistently exciting [8],[14]. We will use Tao's definition of the excitation and the persistence of excitation of a bounded vector signal [14]:

Definition 1 A bounded vector signal $\Phi(t)$ is exciting over an interval $[t, t+T]$, $T > 0$ and $t \geq t_0$ if there exists $\gamma > 0$ such that

$$\int_t^{t+T} \Phi(\tau)\Phi^T(\tau) d\tau \geq \gamma I. \tag{9}$$

Definition 2 A bounded vector signal $\Phi(t)$ is persistently exciting if for all $t > t_0$ there exists $T > 0$ and $\gamma > 0$ such that

$$\int_t^{t+T} \Phi(\tau)\Phi^T(\tau) d\tau \geq \gamma I. \tag{10}$$

Hence, vector signals containing a step in every component are exciting, but not persistently exciting; whereas vector signals containing continuous periodic components are persistently exciting.

Case 2: Unstructured Uncertainty: In the more general case where it is only known that the uncertainty $\Delta(x)$ can be linearly parameterized over a compact domain, the adaptive part of the control law is represented using a Radial Basis Function (RBF) Neural Network (NN):

$$u_{ad}(x) = \hat{W}^T \sigma(x). \tag{11}$$

where $\hat{W} \in \Re^l$ and $\sigma = [1, \sigma_2(x), \sigma_3(x), ..., \sigma_l(x)]^T$ is a vector of known radial basis functions. For $i = 2, 3..., l$ let c_i denote the RBF centroid and μ_i denote the RBF width then for each i The radial basis functions are given as:

$$\sigma_i(x) = e^{-\|x - c_i\|^2/\mu_i} \tag{12}$$

Appealing to the universal approximation property of Radial Basis Function Neural Networks [12] we have that given a fixed number of radial basis functions l there exists ideal weights $W^* \in \Re^l$ and a real number $\tilde{\epsilon}$ such that

$$\Delta(x) = W^{*T} \sigma(x) + \tilde{\epsilon}, \tag{13}$$

where $\tilde{\epsilon}$ can be made arbitrarily small given sufficient number of radial basis functions.

III. Design of Adaptive Control Law Augmented with a Least Squares based Modification Term

In this section we propose a least squares based modification to the baseline adaptive law of equation 8 that guarantees exponential convergence of the parameter error vector $\tilde{W} = \hat{W} - W$ to zero without requiring persistency of excitation. In the following, we will consider the case of structured uncertainty, and point to ways in which the results can be generalized to the case of unstructured uncertainty.

A. *Online Least Squares Regression*

Let N denote the number of available state measurements at time t, and θ denote an estimate of the ideal weighs W. For a given data point $k \in 1, 2, ..., N$, the model error $\Delta(k)$ is

$$\Delta(k) = B^T[\dot{x}(k) - Ax(k) - Bu(k)]. \tag{14}$$

Since $A, B, x(k), u(k)$ are known, the problem of estimating system uncertainty can be reduced to that of estimation of $\dot{x}$ by using 14. When measurement for $\dot{x}$ is not available, for a given data point k, $\dot{x}_k$ can be estimated using a fixed point smoother [5].

Define the error $\epsilon(k) = \Delta(x(k)) - \Phi(x(k))^T\theta$, then the error for N discrete data points can be written in vector form as $\epsilon = [\epsilon(1), \epsilon(2), ..., \epsilon(N)]^T$. In order to arrive at the ideal estimate θ of the true weights W we must solve the following least squares problem:

$$\min_W \epsilon^T \epsilon. \tag{15}$$

Let $Y = [\Delta(1), \Delta(2), ..., \Delta(N)]^T$ and define the following matrix:

$$X = \begin{bmatrix} \phi_1(x(1)) & \phi_2(x(1)) & ... & \phi_m(x(1)) \\ \phi_1(x(2)) & \phi_2(x(2)) & ... & \phi_m(x(2)) \\ & & & \\ \phi_1(x(N)) & \phi_2(x(N)) & ... & \phi_m(x(N)) \end{bmatrix}. \tag{16}$$

A closed form solution to the least squares problem is given as [9]:

$$\theta = (X^T X)^{-1} X^T Y. \tag{17}$$

Equation 17 however, is not numerically efficient, and other methods such as the recursive frequency domain Fourier Transform Regression (FTR) have been developed [10]. The main benefits of the FTR are: 1. the matrix containing frequency domain information of recorded data has constant dimensions; 2. Unwanted frequencies can be implicitly filtered; 3. fixed point smoothing is not required for the estimation of the model error $\Delta(x)$. Let w denote the independent frequency variable, then the Fourier transform of an arbitrary signal $x(t)$ is given by

$$F[x(t)] = \tilde{x}(w) = \int_{-\infty}^{+\infty} x(t)e^{-jwt}dt. \tag{18}$$

Let N be the number of available measurements, and Δt denote the sampling interval, then the discrete Fourier Transform can be approximated as

$$X(w) = \sum_{k=0}^{N-1} x(k)e^{-jwk\Delta t}. \tag{19}$$

The Euler approximation for the Fourier transform in equation 18 is given by $\tilde{x}(w) = X(w)\Delta t$. This approximation is suitable if the sampling rate $1/\Delta t$ is much higher than any of the frequencies of interest w. This discrete version of the Fourier transform can be recursively propagated as follows

$$X_k(w) = X_{k-1}(w) + x(k)e^{-jwk\Delta t}. \tag{20}$$

This leads to a standard regression problem with complex data, where $\tilde{Y}(w)$ denotes the dependent variable, $\tilde{X}(w)$ denotes the independent variables, $\tilde{\epsilon}(w)$ denotes the regression error, and θ denotes the unknown weights:

$$\tilde{Y}(w) = \tilde{X}(w)\theta + \tilde{\epsilon}. \tag{21}$$

Given a measurement k and a given frequency range $\omega = 1..l$ the matrix of independent variables is given as

$$\tilde{X}(w) = \begin{bmatrix} \phi_1(x(1)) & \phi_2(x(1)) & ... & \phi_m(x(1)) \\ \phi_1(x(2)) & \phi_2(x(2)) & ... & \phi_m(x(2)) \\ \\ \phi_1(x(l)) & \phi_2(x(l)) & ... & \phi_m(x(l)) \end{bmatrix}. \tag{22}$$

The vector of dependent variables is given as $\tilde{Y}(w) = [\Delta(1), \Delta(2), ..., \Delta(l)]^T$. Using equation 14, and letting $x(w)$ and $u(w)$ denote the Fourier transform of the state and the input signals, the model error for a data point k in the frequency domain can be found as:

$$\Delta_k(w) = B^T[x_k(w)jw - Ax_k(w) - Bu_k(w)]. \tag{23}$$

The least squares estimate of the weight vector θ is then given by

$$\theta = [Re(\tilde{X}^*\tilde{X})]^{-1}Re(\tilde{X}^*\tilde{Y}), \tag{24}$$

where $*$ denotes the complex conjugate transpose. Forgetting factors can be used to discount older data while recursively computing the Fourier Transform [10].

B. Least Squares Based Modification Term

We now describe a method by which the least squares estimate of the ideal weights can be incorporated in the adaptive control law. Let $r^T = e^TPB$ where e, P, B are as in section II, Γ_W, Γ_θ are positive definite matrices denoting the learning rate, θ is the solution to the least squares problem (equation 17). The adaptive law for weight estimates $\hat{W}$ is chosen as

$$\dot{\hat{W}} = -(\Phi(x)r^T - \Gamma_\theta(\hat{W} - \theta))\Gamma_W. \tag{25}$$

In the above equation, the term $\Gamma_\theta(\hat{W} - \theta))$ denotes the least squares based modification to the adaptive law. For proceeding with stability analysis, consider the following condition on the recorded data.

Condition 1 Sufficient state measurements are recorded such that the matrix $\tilde{X}(w)$ of equation 22 has full column rank.

Recalling that the matrix $\tilde{X}(w)$ contains Fourier transform of the vector signal $\Phi(x(t))$ we note that condition 1 requires that the recorded data points be sufficiently different. The following theorem shows that condition 1 is sufficient to guarantee exponential stability when using the adaptive law of equation 25.

Theorem 1 Consider the control law of equation 3, the case of structured uncertainty, and the weight update law of 25, then all signals in the system 1 are bounded and $x(t)$ tracks x_{rm} exponentially. Furthermore, if condition 1 is satisfied, then $\hat{W}$ converges exponentially to the ideal weights W.

Proof: Let tr denote the trace operator, and consider the following Lyapunov candidate function

$$V(e, \tilde{W}) = \frac{1}{2}e^TPe + \frac{1}{2}tr(\tilde{W}^T{\Gamma_W}^{-1}\tilde{W}). \tag{26}$$

The time derivative of the Lyapunov candidate along the system trajectory 4 is

$$\begin{aligned}\dot{V}(e, \tilde{W}) = &-\frac{1}{2}e^TQe + r^T(\hat{W}^T\Phi(x) - W^T\Phi(x)) \\ &+ tr(\dot{\hat{W}}{\Gamma_W}^{-1}\tilde{W}^T).\end{aligned} \tag{27}$$

Let ϵ be such that $W = \theta + \epsilon$, adding and subtracting $(\hat{W}^T - \theta)^T\Gamma_\theta(\hat{W}^T - \theta)$ to equation 27 and using the definition of ϵ yields,

$$\begin{aligned}\dot{V}(e, \tilde{W}) = &-\frac{1}{2}e^TQe + r^T(\tilde{W}^T\Phi(x)) + tr(\dot{\hat{W}}{\Gamma_W}^{-1}\tilde{W}^T) \\ &+ \tilde{W}^T\Gamma_\theta(\hat{W} - \theta) - \tilde{W}^T\Gamma_\theta(\hat{W} - \theta).\end{aligned} \tag{28}$$

Rearranging yields:

$$\begin{aligned}\dot{V}(e, \tilde{W}) = &-\frac{1}{2}e^TQe \\ &+ tr((\dot{\hat{W}}{\Gamma_W}^{-1} + \Phi(x)r^T + \Gamma_\theta(\hat{W} - \theta))\tilde{W}^T) \\ &- \tilde{W}^T\Gamma_\theta(\hat{W} - \theta).\end{aligned} \tag{29}$$

Setting $tr((\dot{\hat{W}}{\Gamma_W}^{-1} + \Phi(x)r^T + \Gamma_\theta(\hat{W} - \theta))\tilde{W}^T) = 0$ yields the adaptive law. Consider the last term in 29,

$$\tilde{W}^T\Gamma_\theta(\hat{W}-\theta) = (\hat{W}-W)^T\Gamma_\theta(\hat{W}-\theta) = (\hat{W}-W)^T\Gamma_\theta(\hat{W}-W) + (\hat{W}-W)^T\Gamma_\theta\epsilon. \quad (30)$$

Using 24, the definition of ϵ, and *condition 3* yields:

$$\epsilon = W - [Re(\tilde{X}^*\tilde{X})]^{-1}Re(\tilde{X}^*\tilde{X})W = 0, \quad (31)$$

letting $\lambda_{min}(Q)$ and $\lambda_{min}(\Gamma_\theta)$ denote the minimum eigenvalues of Q and Γ_θ we have that equation 29 becomes:

$$\dot{V}(e,\tilde{W}) = -\frac{1}{2}\|e\|^2\lambda_{min}(Q) - \|\tilde{W}\|^2\lambda_{min}(\Gamma_\theta). \quad (32)$$

Hence $\dot{V}(e,\tilde{W}) < 0$ which shows that $e \to 0$ and $\hat{W} \to W$, since A_{rm} is Hurwitz, x_{rm} is bounded, therefore it follows that x is bounded. Since V is quadratic and radially unbounded, and since $\dot{V}(e,\tilde{W}) \leq \frac{max(\lambda_{min}(Q), 2\lambda_{min}(\Gamma_\theta))}{min(\lambda_{min}(P), \lambda_{min}(\Gamma_W^{-1}))}V(e,\tilde{W})$, the convergence is globally exponential (Thm. 3.2 in [6]). ■

Remark 1 The above proof guarantees exponential stability of the tracking error e and guarantees that $\hat{W}$ will approach the ideal weight W exponentially. This is subject to condition 1, which is a weaker and easier to monitor than the PE condition [4]. It is straight forward to see that if the signal is exciting over any finite time interval then data points can be recorded such that condition 1 is satisfied. Hence, the proof does not require persistency of excitation of $r(t)$.

Remark 2 The above proof can be easily extended to the case of unstructured uncertainty by using Radial Basis Function Neural Networks for approximating the uncertainty. For this case, since $Y = W^{*T}\sigma + \tilde{\epsilon}$ for some nonzero ϵ, the following adaptive law results in uniform ultimate boundedness of all states

$$\dot{\hat{W}} = -(\sigma(x)r^T - \Gamma_\theta(\hat{W}-\theta))\Gamma_W. \quad (33)$$

Furthermore, referring to equation 13 and noting that in this case $\epsilon = \tilde{\epsilon}$, it can be shown that the weights will approach a neighborhood of the best linear approximation of the uncertainty. Finally, for this case, the satisfaction of condition 1 is reduced to selecting distinct points for storage due to Micchelli's theorem [7].

Remark 3 Note that the term $\Gamma_\theta(\hat{W}-\theta)$ adds in as a modification term to the baseline adaptive. Since the above analysis is valid for any bounded initial condition and since the baseline adaptive law is known to be stabilizing, it is possible to set $\theta = 0$ until sufficient data is recorded to satisfy condition 1. This will result in a σ-modification like term until satisfaction of condition 1.

Remark 4 This proof can be modified to accommodate any least squares solution method, for example the standard least squares solution of equation 17 can be accommodated by replacing equation 31 with the following

$$\epsilon = W - (X^TX)^{-1}X^TXW = 0, \quad (34)$$

In this case, condition 1 requires that matrix X has full column rank.

Remark 5 The increased computational burden when using the adaptive law of equation 25 consists mainly of evaluating equation 24 to obtain θ. However, θ does not need to be updated as often as the controller itself. Recursive Least Square (RLS) method can also be used [3].

Remark 6 It is possible to imagine a switching approach in which the online estimate of the ideal weights θ is used in equation 7 by setting $\hat{W} = \theta$ when θ becomes available. However, this approaches looses the benefit of keeping the baseline adaptive law in the control loop, namely, the adaptive weights no longer take on values to minimize $V(t) = e^T(t)e(t)$.

Figure 1 shows the schematic of the presented adaptive control method with least squares modification.

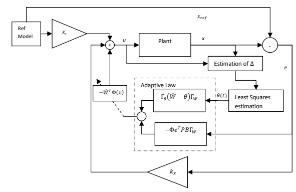

Fig. 1. Schematics of adaptive controller with least squares Modification

IV. SIMULATION RESULTS

In this section we evaluate the method of Theorem 1 in simulation study of regulation of of wing rock dynamics. Let ϕ denote the roll angle of an aircraft, p denote the roll rate, δ_a denote the aileron control input, then a model for wing rock dynamics is

$$\dot{\phi} = p \quad (35)$$

$$\dot{p} = \delta_a + \Delta(x). \quad (36)$$

where $\Delta(x) = 0.0 + 0.2314\phi + 0.6918p - 0.6245|\phi|p + 0.0095|p|p + 0.0214\phi^3$, the parameters of the wing rock model adapted from [13] and [15]. Initial conditions are chosen to be $\phi(0) = 1deg, p(0) = 1deg/s$. A second order reference model with damping ratio and natural frequency of 1 is chosen. We set $K = [-0.5, -0.3]$.

A. Case 1: Structured Uncertainty

Consider first the case where the structure of the uncertainty is known. Figure 2 shows the performance of the baseline adaptive control law of equation 8 without the least squares modification. For both the low gain ($\Gamma_W = 3$) and high gain ($\Gamma_W = 10$) case, it is seen that the performance of the controller is unsatisfactory. Figure 3 shows the phase portrait of the states when the adaptive law of Theorem 1 is used, and it is seen that the states follow a smooth trajectory to the origin. Furthermore, the performance for both the high gain and the low gain case is almost identical. Figure 4 shows the evolution of the adaptive control weights when only the baseline adaptive law of equation 8 is used. It is seen that the weights do not converge to the ideal values (W) and evolve in an oscillatory manner. In contrast, figure 5 shows the convergence of the weights when the adaptive law of Theorem 1 used. Figure 6 compares the reference model states with the plant states for the baseline adaptive law, while 7 compares the reference model and state output when the least squares modification based adaptive law is used. It is seen that the performance of the adaptive law with least squares modification is superior to the baseline adaptive law. Finally, figure 8 shows that the tracking error converges exponentially to the origin when least squares modification term is used. These results highlight the benefit of weight convergence.

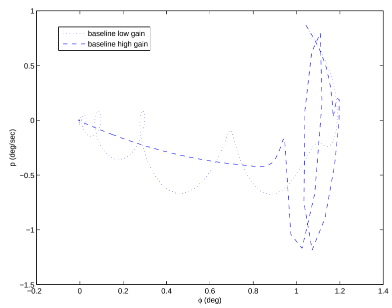

Fig. 2. Phase portrait of system states with only baseline adaptive control

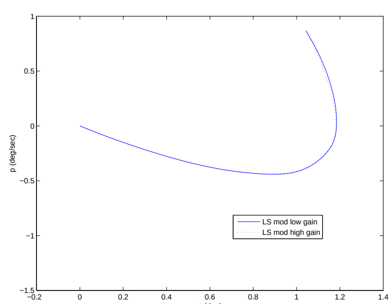

Fig. 3. Phase portrait of system states with least squares modification

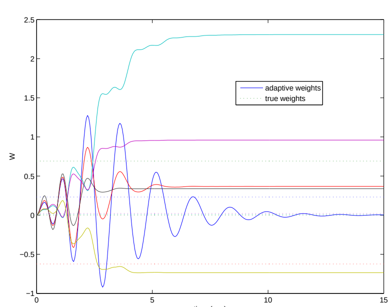

Fig. 4. Evolution of adaptive weights with only baseline adaptive control

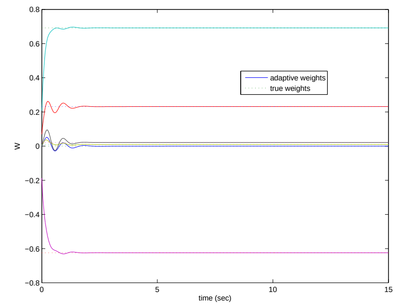

Fig. 5. Evolution of adaptive weights with least squares modification

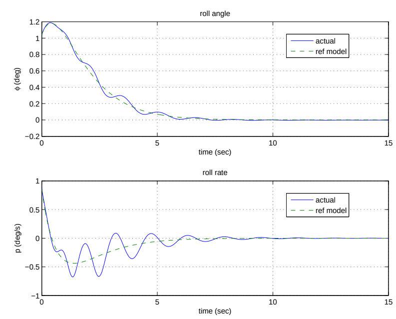

Fig. 6. Performance of adaptive controller with only baseline adaptive law

B. Case 2: Adaptation using RBF NN

For the results in this section we assume that it is only known that the structure of the uncertainty is unknown. Hence, RBF NN with 6 nodes and uniformly distributed centers over the expected range of the state-space are used to capture the model uncertainty. Figure 9 shows the trajectory of the system in the phase space when the baseline adaptive control law of equation 8 is used. The performance can be contrasted with smooth convergence to the origin seen in figure 10 when adaptive law with least squares modification is used. Since the ideal weights W^* in this case are not known, we evaluate the performance of the adaptive law by comparing the output of the RBF NN with the actual model uncertainty with weights *frozen* after the simulation run is over. Figure 11 shows the comparison. It is clearly seen that the NN weights obtained with the least squares modification

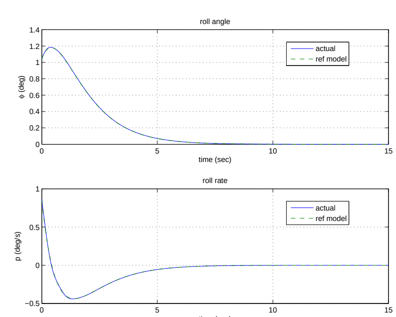

Fig. 7. Performance of adaptive controller with least squares modification

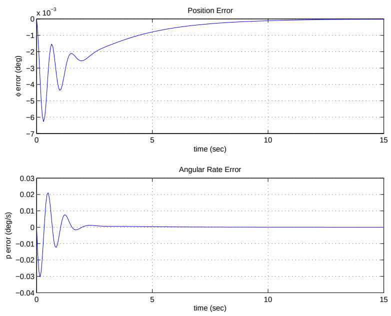

Fig. 8. Evolution of tracking error with least squares modification

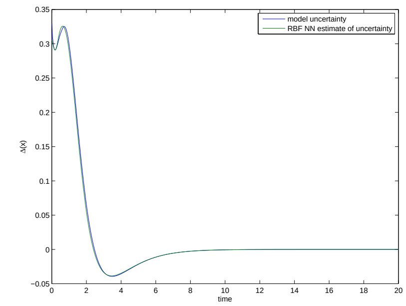

Fig. 11. RBF NN model uncertainty approximation with weights frozen post adaptation

based adaptive law are able to accurately capture the uncertainty, a clear indication that the weights are very close to their ideal values.

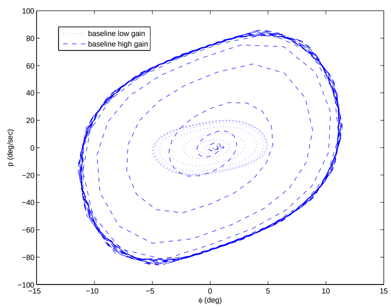

Fig. 9. Phase portrait of system states with only baseline adaptive control while using RBF NN

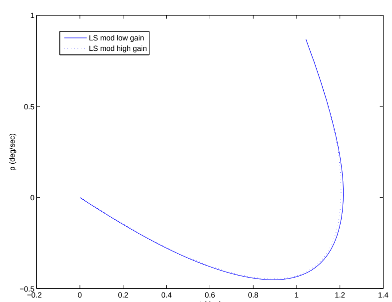

Fig. 10. Phase portrait of system states with least squares modification while using RBF NN

V. CONCLUSION

In this work we proposed a least squares based modification for adaptive control. This modification uses an least squares estimate of the ideal weights found by solving a linear least squares problem online by recording system states and estimating the model error directly in the adaptive law. The modified adaptive law uses both past and current data concurrently for adaptation and can guarantee exponential tracking error and parameter error convergence to zero if the recorded data meet a verifiable condition on linear independence.

VI. ACKNOWLEDGMENTS

This work was supported in part by NSF ECS-0238993 and NASA Cooperative Agreement NNX08AD06A.

REFERENCES

[1] Stephan Boyd and Sankar Sastry. Necessary and sufficient conditions for parameter convergence in adaptive control. *Automatica*, 22(6):629–639, 1986.

[2] Chengyu Cao and N. Hovakimyan. Design and analysis of a novel adaptive control architecture with guaranteed transient performance. *Automatic Control, IEEE Transactions on*, 53(2):586 –591, march 2008.

[3] Girish Chowdhary and Eric Johnson. Recursively updated least squares based modification term for adaptive control. In *American Control Conference 2010*, 2010.

[4] 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*, 2010.

[5] Girish V. Chowdhary and Eric N. Johnson. Theory and flight test validation of long term learning adaptive flight controller. In *Proceedings of the AIAA Guidance Navigation and Control Conference*, Honolulu, HI, 2008.

[6] Wassim M. Haddad and VijaySekhar Chellaboina. *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*. Princeton University Press, Princeton, 2008.

[7] Simon Haykin. *Neural Networks a Comprehensive Foundation*. Prentice Hall, USA, Upper Saddle River, 2 edition, 1998.

[8] Petros A. Ioannou and Jung Sun. *Robust Adaptive Control*. Prentice-Hall, Upper Saddle River, 1996.

[9] Ravindra V. Jategaonkar. *Flight Vehicle System Identification A Time Domain Approach*, volume 216 of *Progress in Astronautics and Aeronautics*. American Institute of Aeronautics and Astronautics, Reston, 2006.

[10] E. A. Morelli. Real time parameter estimation in the frequency domain. *Journal of Guidance Control and Dynamics*, 23(5):812–818, 2000.

[11] Kumpati S. Narendra and Anuradha M. Annaswamy. *Stable Adaptive Systems*. Prentice-Hall, Englewood Cliffs, 1989.

[12] J. Park and I.W. Sandberg. Universal approximation using radial-basis-function networks. *Neural Computatations*, 3:246–257, 1991.

[13] 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.

[14] Gang Tao. *Adaptive Control Design and Analysis*. Wiley, New York, 2003.

[15] Konstantine Volyanskyy and Anthony Calise. An error minimization method in adaptive control. In *Proceedings of AIAA Guidance Navigation and Control conference*, 2006.

[16] Kostyantyn Y. Volyanskyy, Wassim M. Haddad, and Anthony J. Calise. A new neuroadaptive control architecture for nonlinear uncertain dynamical systems: Beyond σ and e-modifications. *IEEE Transactions on Neural Networks*, 20(11):1707–1723, Nov 2009.