*Proceedings of the 7th Annual GRASP Symposium, Wichita State University, 2011*

# Single Network Adaptive Critic Aided Nonlinear Dynamic Inversion

Geethalakshmi S. Lakshmikanth
Faculty: John M. Watkins* and James E. Steck**

** Department of Electrical Engineering and Computer Science, ** Department of Aerospace Engineering*

**Abstract.** A Single Network Adaptive Critic (SNAC) [1] architecture has been shown to have much faster convergence than Adaptive Critic (AC) and offers an optimal control design method for state regulation. Dynamic Inversion (DI) [2], on the other hand, is a very popular nonlinear control design method for both output regulation and command tracking that offers a closed form expression for control but lacks the interpretation of optimal control design. Here, we propose a systematic technique where a pre-synthesized SNAC network aids the DI control so that the resulting control can be used for both near optimal regulation and command following. The potential of this new methodology is demonstrated by considering a benchmark nonlinear system.

## 1. Introduction

Designing a controller for nonlinear systems is most commonly a challenging task. When the need to optimize a cost function is also necessary, control design becomes very complicated. Approximate Dynamic Programming (ADP) approach offers a systematic method of optimal control design, using a dual neural network architecture called the Adaptive Critic (AC). Even though advantages of SNAC have been harnessed very well for regulatory (states of the system are driven to zero) applications, there has been no evidence of using the SNAC architecture for command following (states are required to track/follow a command signal or target signal) applications. A hybrid technique is proposed, which maintains the advantages of SNAC and NDI making SNAC extendable to command following applications.

## 2. Experiment, Results, Discussion, and Significance

The idea and its formalization are explained by considering a simple scalar nonlinear system, described by the following state equation:

$$\dot{x} = x - x^3 + u \tag{1}$$

where $x$ and $u$ are the state and control variables respectively (both being scalars). With the goal to drive the states to zero, a cost function is defined as $J = \frac{1}{2}\int_0^\infty \left(Q_c x^2 + R_c u^2\right)dt$. Following the procedure for optimal control design, we obtain the optimal control equation:

$$u^* = -R_c^{-1}\lambda \tag{2}$$

and the costate equation: $\dot{\lambda} = -\left[Q_c x + \lambda\left(1 - 3x^2\right)\right]$. SNAC network consisting of a single critic network, outputs $\lambda$ with $x$ as the input. It is synthesized offline according to the procedure outlined in [1]. We now evaluate the nonlinear DI control and the output is defined by $y = x$. From the system dynamics in (1), we see that the system has a relative degree of 1. Following the first order error dynamics, we arrive at the nonlinear DI [2] control given by:

$$u = \dot{y}^* - k_{p1}\left(y - y^*\right) - x + x^3 \tag{3}$$

where $y^*$ is the target or command signal. Pseudo control is defined: $u_s = -k_{p1}x$ and the optimal pseudo control, $u_s^* = x - x^3 + u^*$ where, $u^*$ is defined in (2). The goal is to evaluate a closed form expression for $k_{p1}$ in (3) using (2) that will make the NDI response guided by SNAC, a near-optimal one. We now define a time varying cost function,

$$J_{DI} = \frac{1}{2}\begin{bmatrix} \left(u_s - u_s^*\right)^T r_1 \left(u_s - u_s^*\right) \\ +r_2\left(k_{p1} - \bar{k}_{p1}\right)^2 \end{bmatrix} \tag{4}$$

where $r_1$ and $r_2$ are scalar positive weights and $\bar{k}_{p1}$ is the value of $k_{p1}$ at the beginning of every cycle. The constrained optimization problem $\min\limits_{k_{p1}} J_{DI}$ is solved to arrive at the new value of $k_{p1}$. It is very important to keep the weight $k_{p1}$ positive for stability of the system. The value of weight obtained is used in the appropriate nonlinear DI equation for control, depending on whether output regulation or command following is desired. Fig. 1 shows that the gains of NDI calculated above, steer the NDI response very close to that of SNAC as a regulator and with a unit step input, it generates a near optimal response.

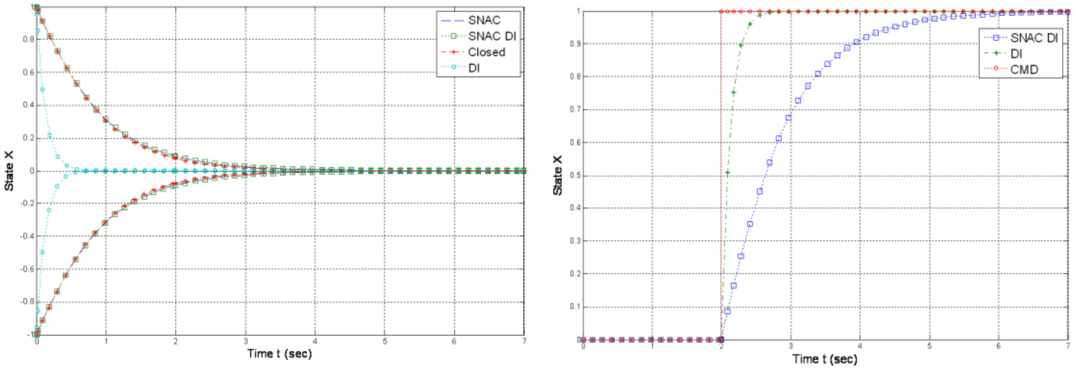

Fig. 1: Comparison of responses using SNAC aided NDI control and NDI control as a regulator and command follower.

## 3. Conclusions and Future Work

A very novel technique of near-optimal control design has been proposed utilizing the advantages of both SNAC and NDI for a class of nonlinear systems. The work presented will be extended to a more complex nonlinear system, such as, 3 Degree of Freedom (DOF) airplane (Longitudinal plane) to control forward velocity and pitch angle.

## 4. Acknowledgements

This material is based upon work supported by NASA under award number NNXO9AP20A. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NASA.

**References**

[1] Padhi, R.," Optimal Control of Distributed Parameter Systems Using Adaptive Critic Neural Networks," Ph.D. Dissertation, University of Missouri Rolla, 2001.
[2] Fisher, J.R.," Aircraft Control using Nonlinear Dynamic Inversion in Conjunction with Adaptive Robust Control," Master's Thesis, Texas A&M University, 2004.