01 Jan 2007

Optimal Neuro-Controller Synthesis for Impulse-Driven System

Xiaohua Wang

S. N. Balakrishnan
Missouri University of Science and Technology, bala@mst.edu

Recommended Citation

X. Wang and S. N. Balakrishnan, "Optimal Neuro-Controller Synthesis for Impulse-Driven System," *Proceedings of the 2007 American Control Conference (2007, Marriott Marquis Hotel at Times Square New York City, NY)*, Institute of Electrical and Electronics Engineers (IEEE), Jan 2007.
The definitive version is available at https://doi.org/10.1109/ACC.2007.4282911

Proceedings of the 2007 American Control Conference
Marriott Marquis Hotel at Times Square
New York City, USA, July 11-13, 2007

Optimal Neuro-controller Synthesis for Impulse-Driven System

Xiaohua Wang and S. N. Balakrishnan, *Member, IEEE*

Abstract—**This paper presents a new controller design technique for systems driven with impulse inputs. Necessary conditions for optimal impulse control are derived. A neural network structure to solve the resulting equations is presented. The solution concepts are illustrated with a few example problems that exhibit increasing levels of difficulty. Two linear problems-one scalar and one vector-and a benchmark nonlinear problem-Van Der Pol oscillator-are used as case studies. Numerical results show the efficacy of the new solution process for impulse driven systems. Since the theoretical development and the design technique are free from restrictive assumptions, this technique is applicable to many problems in engineering and science.**

I. Introduction

Impulse control has been considered in many interesting applications, such as control of a periodically forced chaotic pendulum [1], biped robots [2], spaceship optimal fuelling [3], insect population [4], and financial systems [6-7].

There are three major reasons why impulse control is preferred over continuous control. One consideration is the nature of systems. In some cases, the plant is not amenable to using continuous control. For example, the interest rate changed by the federal bank which is used to control the money supply in the market cannot be changed continuously. In some other cases, use of impulse control is simply more efficient. For example, instantaneous changes to the density of the bactericide enables one to control the population of bacteria without enhancing their drug resistance, something which continuous control cannot avoid [5]. The other important consideration is that impulsive control could be the more practical and cheaper option compared with continuous control. For example, in spacecraft formation control problems, the optimal control forces are usually very small, about 250 μN in average, and are difficult to implement [8-9]. Yet it is important to use as little propulsion as possible because fuel weight takes away useful scientific or commercial payload. Thus development of optimal impulse schemes is extremely beneficial to many disciplines.

Origins of impulse control development can be traced to impulsive differential equation theory. Current research in impulse control focused on two aspects: the first issue is stability of impulse-driven systems and the second issue is the development of optimal controllers. Existence, uniqueness, and stability of the solutions to impulsive differential equations have been studied systematically in [10-11]. In impulse system literature, a common impulsive system model considered is

$$\begin{gathered}\dot{x}=f_c\left(t,x\right)+g_c\left(t,x,u_c\right)+g_d\left(t,x,u_d\right)\ \delta\left(t-\tau_i\right)\\ 0<\tau_1<\tau_2<\tau_3<\cdots<\tau_N<\infty\end{gathered} \tag{1}$$

where $x\left(t\right)\in D\subseteq\mathbb{R}^n$ are the system states, and D is an open set with $0\in D$. $u_c\subseteq\mathbb{R}^{m_c}$ and $u_d\subseteq\mathbb{R}^{m_d}$ are the continuous controls and the impulsive controls, respectively. $f_c:D\rightarrow\mathbb{R}^n$ is Lipchitz continuous. $g_d:D\rightarrow\mathbb{R}^{n\times m_d}$ and $g_c:D\rightarrow\mathbb{R}^{n\times m_c}$ are continuous differentialable functions. i indexes the moments when impulse is applied, $i=1\cdots N\in\mathbb{Z}^+$ and $0\le\tau_1<\tau_2<\tau_3<\cdots<\tau_N<\infty$ are instants where an impulse is applied. $\mathbb{R}$ represents the set of real number. $\mathbb{Z}^+$ represents the set of positive integer.

The above model in (1) can be written into the following form:

$$\begin{cases}\dot{x}=f_c\left(t,x\right)+g_c\left(t,x,u_c\right) & t\neq\tau_i\\ x^+=x^-+g_d\ \left(t,x,u_d\right) & t=\tau_i\end{cases} \tag{2}$$

where $i=1\cdots N\in\mathbb{Z}^+$. Superscript $+$ and $-$ denote the right limit and left limit with respect to the instants where an impulse is applied, respectively.

In order to consider optimality concepts with respect to an impulse-driven system, a performance index is given as

$$J=\min_{u\in U}\int_{t_0}^{t_f}L\left(x,u\right)dt \tag{3}$$

where $L\left(x,u\right)$ is a convex, non-decreasing, positive function, and U is the set of all the admissible controls.

Lyapunov theory is the most commonly used tool in determining system stability. Based on Lyapunov theory, reference [10] developed a comparison theorem. This theorem and its corollary are widely used in the stability characterization of impulsive systems. Reference [12] derived stability properties for a class of nonlinear impulsive systems based on the comparison theory of [10]. Reference [13] stabilized Chua's circuit and chaos in a phase-locked loop system using impulsive control. Reference [14] applied dissipativity theory to nonlinear dynamical systems with impulsive effects, and they generated extended Kalman-Yakubovich-Popov conditions for the study of impulse systems. Reference [15] considered necessary and sufficient conditions for controllability and observability of switched linear impulse systems.

Reference [16] derived the necessary conditions for both the fixed and variable time optimal impulse problems with different impulse model. Reference [16] solved the minimum

Manuscript received September 24, 2006. This work was supported in part by the NSF Grant ECS-0601706.

Xiaohua Wang is a PhD student with Department of Aerospace and Mechanical Engineering, University of Missouri, Rolla, MO 65401 USA. (Email: wxw98@umr.edu)

S. N. Balakrishnan is a professor with Department of Mechanical and Aerospace Engineering, University of Missouri, Rolla, MO 65401 USA. (Email: bala@umr.edu. Tel: 573-341-4675. Fax: 573-341-4607)

1-4244-0989-6/07/$25.00 ©2007 IEEE.

time control of a swing. Reference [17] considered the optimality of a nonlinear impulsive system in both finite horizon and infinite horizon cases. Two Hamilton functions were constructed to derive the optimality conditions. Optimal conditions for an autonomous linear impulsive system were also derived. Reference [18] provided a representation of generalized optimal solutions to nonlinear impulse problems in terms of differential equations with one measure. Their solution used a method called the method of discontinuous time change.

Though the literature on impulsive control is quite extensive, still there exists a need for the development of systematic impulsive control design methods. In critical application areas such as space and health, optimal impulse controller development will make the operations very cost-effective. This paper develops an optimal impulse-driven controller technique that satisfies those needs. Furthermore, the proposed neural network based technique is implementable.

The rest of the paper is organized as follows: Section II contains the derivation of necessary conditions for optimality and stability. Section III illustrates the special neural network scheme based on a structure called "single network adaptive critic(SNAC)". Section IV presents three illustrative problems and the simulation results. The case studies consist of results from a scalar linear problem, a vector linear problem, and a vector nonlinear problem-Van Der Pol oscillator. Section V provides the conclusions.

II. OPTIMAL IMPULSE CONTROL

A. Problem Formulation

In this paper, an autonomous fixed time impulse system is considered with the model given by

$$\dot{x} = f_c(x) + g_d(x)u_i\delta(t-\tau_i) \quad (4)$$

where $x \in \mathbb{R}^{n\times 1}$. $u_i \in \mathbb{R}^{m\times 1}$. $i = 1,2,3,\ldots,\mathbb{Z}^+$. $f_c(x) \in \mathbb{R}^{n\times 1}$ is Lipchitz continuous and $g_d(x)$ is continuous differentialable; δ is a dirichlet function; $0 \le \tau_i < \infty$ are known instants when an impulse is given, which are also mentioned as impulse instants in this paper.

The system in (4) can be written in the following form:

$$\begin{cases} \dot{x} = f_c(x) & t \ne \tau_i \\ x_i^+ = x_i^- + g_d(x_i^-)u_i & t = \tau_i \\ x_0 \equiv \text{known}, \; i = 1,2,3,\ldots,\mathbb{Z}^+ \end{cases} \quad (5)$$

where $i = 1,2,3,\ldots,\mathbb{Z}^+$. Superscript $+$ and $-$ denote the right limit and left limit with respect to the instant where an impulse is applied, respectively.

In this study, a fairly general cost function is considered for minimization as follows:

$$J = \Phi(x_f) + \sum_{i=1}^{k} L_d(u_i) + \sum_{i=1}^{k+1} \int_{t_{i-1}^+}^{t_i^-} L_c(x)dt \quad (6)$$

where $\Phi(x_f)$ is the constraint on the terminal states, $\sum_{i=1}^{k} L_d(u_i)$ is the penalty on control effort, and $\sum_{i=1}^{k+1} \int_{t_{i-1}^+}^{t_i^-} L_c(x_i)dt$ is the penalty on states. Note that $t_{k+1}^- = t_f$, where t_f is the final time.

B. Optimality Conditions

Theorem 1: Given the system dynamics as in (5) with known initial conditions, cost function as in (6), and assuming optimal control exists, by introducing the Hamiltonian [19]

$$H \triangleq L_c(x) + \lambda^T f_c(x) \quad (7)$$

necessary conditions for the optimality are presented in the following equations:

1) At $t = t_f$,

$$\lambda_f^T = \frac{\partial \Phi(x_f)}{\partial x_f} \quad (8)$$

2) For $t \in \left[t_i^+, t_{i+1}^-\right]$,

- State propagation equation is
$$\dot{x} = f_c(x) \quad (9)$$
- Costate propagation equation is
$$\frac{\partial H(x_i)}{\partial x_i} + \dot{\lambda}_i^T = 0 \quad (10)$$

3) Between pre impulse and post impulse, $t \in \left[t_i^-, t_i^+\right]$,

- State update equation is
$$x_i^+ = x_i^- + g_d(x_i^-)u_i \quad (11)$$
- Costate update equation is
$$\lambda^T(t_i^-) = \lambda^T(t_i^+)\left(I + \frac{\partial g_d(x_i(t_i^-))}{\partial x_i}u_i\right) \quad (12)$$
- Control equation is
$$\frac{\partial L_2(u_i)}{\partial u_i} + \lambda^T(t_i^+)g_d(x_i(t_i^-)) = 0 \quad (13)$$

Proof:

Details are not presented due to the page restriction.

Corollary 1: Let the linear system model, with the instants when an impulse is given as fixed, be described as

$$\begin{cases} \dot{x} = A_c x & t \ne \tau_i \\ x^+ = x^- + B_d u & t = \tau_i \end{cases} \quad (14)$$

where $A_c \in \mathbb{R}^{n\times n}$. $B_d \in \mathbb{R}^{n\times m}$. Let the quadratic performance index be given by

$$J = x_f^T S_f x_f + \sum_{i=1}^{k} u_i^T R u_i + \sum_{i=1}^{k+1} \int_{t_{i-1}^+}^{t_i^-} x^T Q x dt \quad (15)$$

where $S_f \in \mathbb{R}^{n\times n} \ge 0$, $R \in \mathbb{R}^{m\times m} > 0$, and $Q \in \mathbb{R}^{n\times n} \ge 0$. Furthermore, assume that the system is controllable and a semi positive definite matrix $P \in \mathbb{R}^{n\times n}$ exists such that

$$\dot{P} = -A_c^T P - PA_c - Q \qquad t \ne \tau_i \quad (16a)$$

$$P_i^- = P_i^+\left(I + B_d R^{-1} B_d^T P_i^+\right)^{-1} \qquad t = \tau_i \quad (16b)$$

$$P_f = S_f, \qquad t = t_f \quad (16c)$$

Then, an optimal impulsive feedback control is obtained as

$$u_i = -R^{-1} B_d{}' P_i^- x_i^- \quad (17)$$

Proof:
Details are not presented due to the page restriction.

Lemma 2 System (14) is impulsively controllable if and only if

$$\mathrm{Rank}\left[B_d, A_c B_d, A_c{}^2 B_d, \cdots, A_c{}^{n-1} B_d\right] = n \quad (18)$$

Proof: Proof is provided in [20] and is not presented here.

Theorem 2: Consider propagation equation for P as in (16a) -(16c), with assumptions

1) Final time t_f is large.
2) $Q = C^T C \geq 0$.
3) $P_f \geq 0$.
4) Interval between impulses is a constant, δt.
5) System (14) is impulsively controllable.
6) $\begin{bmatrix} C \\ CA_c \\ \vdots \\ CA_c{}^{n-1} \end{bmatrix}$ is full rank.

Then, the value of P_i^- reaches a unique positive definite steady state value.

$$\lim_{t \to \infty} P_i^- = P^- = const > 0 \quad (19)$$

where

$$P^- = \left(\left(e^{A_c^T \delta t} P^- e^{A_c \delta t} + \int_{\tau_{i-1}^+}^{\tau_i^-} e^{A_c^T \left(\tau_i^- - t\right)} Q e^{A_c \left(\tau_i^- - t\right)} dt \right)^{-1} + B_d R^{-1} B_d{}^T \right)^{-1} \quad (20)$$

Proof: Detailed proof is not given due to the page restriction.

Theorem 3: With the assumptions used in Theorem 2, the system in (23) is asymptotically stabilized by applying the following optimal control

$$u_i = -R^{-1} B_d{}' P^- x_i^- \quad (21)$$

where P^- is the steady state value of the pre impulse P_i^-.

Proof:
Proof of this theorem is carried out using Lyapunov's second method. Assume a Lyapunov function candidate as

$$V = \frac{1}{2} x^T P(t) x \quad (22)$$

where $P(t)$ satisfies (16a) – (16c). Note that $P > 0$. $P = 0$ is possible only when $t \to \infty$.

Now in the region between the impulses where $t \in \left[t_i^+, t_{i+1}^-\right]$,

$$\begin{aligned} \dot{V} &= \frac{1}{2} x^T \dot{P} x + x^T P \dot{x} \\ &= \frac{1}{2} x^T \left(-PA_c - A_c{}^T P - Q\right) x + x^T P A_c x \\ &= x^T \left(\frac{1}{2} PA_c - \frac{1}{2} A_c{}^T P - \frac{1}{2} Q\right) x \\ &= -\frac{1}{2} x^T Q x \leq 0 \end{aligned} \quad (23)$$

The relationship between the Lyapunov function values at the pre impulse(V_i^-) and post impulse(V_i^+) is given by

$$V_i^+ - V_i^- = \frac{1}{2}(x_i^+)^T P_i^+ x_i^+ - \frac{1}{2}(x_i^-)^T P_i^- x_i^- \quad (24)$$

To prove stability of the underlying system, it should be shown that the right hand side of(24) is less than zero. Note that the pre impulse and the post impulse states are related by

$$x_i^+ = x_i^- + B_d u_i \quad (25)$$

By replacing u_i from (21), this relationship becomes

$$\begin{aligned} x_i^+ &= x_i^- - B_d R^{-1} B_d{}^T P_i^- x_i^- \\ &= \left(I - B_d R^{-1} B_d{}^T P_i^-\right) x_i^- \end{aligned} \quad (26)$$

By replacing x_i^+ by x_i^- in (24) from (26), (24) becomes

$$\begin{aligned} & V_i^+ - V_i^- \\ &= \frac{1}{2}(x_i^-)^T \left(I - B_d R^{-1} B_d^T P_i^-\right)^T P_i^+ \left(I - B_d R^{-1} B_d^T P_i^-\right) x_i^- \\ &\quad - \frac{1}{2} x_i^{-T} P_i^- x_i^- \end{aligned} \quad (27)$$

In order to write P_i^+ as a function of P_i^- in (24), (16b) is manipulated to get

$$P_i^+ = P_i^- \left(I - B_d R^{-1} B_d^T P_i^-\right)^{-1} \quad (28)$$

The final step is to use (28) in (41) to get

$$\begin{aligned} & V_i^+ - V_i^- \\ &= \frac{1}{2} x_i^{-T} \left(I - B_d R^{-1} B_d^T P_i^-\right)^T P_i^- x_i^+ - \frac{1}{2} x_i^{-T} P_i^- x_i^- \quad (29) \\ &= -\frac{1}{2} x_i^{-T} B_d R^{-1} B_d^T P_i^- x_i^+ \leq 0 \end{aligned}$$

With the inequalities in (23) and (29), the Lyapunov function candidate has been shown to decrease with time, and consequently, the system is stable. However, since $x = 0$ is the only equilibrium point of the system, the system (14) is globally asymptotically stable with control given in (21).

III. Solution Technique: SNAC

This section introduces the single network adaptive critic(SNAC) technique which is used in this paper to solve the optimal impulse control problems. SNAC has been used in solving nonlinear control problems in [21-22]. This paper extends the SNAC scheme to impulse control problems.

A. Adaptive Critic Overview

The concept of adaptive critics is derived from the modeling of brain as a supervisor and an action structure [23] where the supervisor criticizes the action (controller) of the system to achieve a better overall goal. Several authors have used fixed structured multilayer-perceptron neural networks, [24-25], to solve nonlinear control problems arising in aerospace and power systems as well as other benchmark nonlinear control problems.

Novelty of this paper lies in using neural network structures, SNAC, to solve optimal impulse problems. Note that one can handle both the finite time and the infinite time problems using this structure. Only the infinite time case is presented in this paper.

B. Infinite Time Adaptive Critic Neural Network Scheme

For infinite time optimal control, the mapping between the states and the costates is not a function of time. Therefore, a single neural network can be used to capture the relation between the states and the costates and the costates are used to calculate the optimal control.

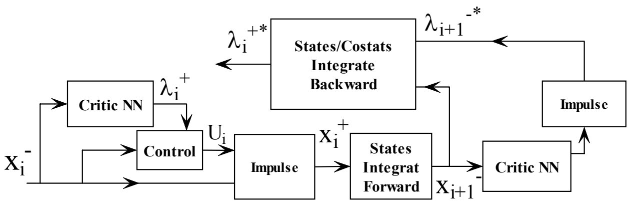


Figure 1. Architecture of Infinite Time SNAC

The idea of SNAC technique is to use the state and the costate propagation equations, the state and costate update equation, and the control expression in (8)-(12) to train a single neural network to obtain the optimal relation between the costates and the states. In this paper, the neural network is used to capture the function $\lambda^+\left(x_i^-\right)$ with x_i^- as inputs and λ_i^+ as outputs.

Fig. 1 gives the flowchart of optimal impulsive control synthesis using SNAC. The single neural network is called critic NN in the picture. X_i^- is a set of states chosen so that its values approximately span the domain of interest. The critic NN is initialized based on an initial stabilizing control design, with the function $\lambda_0\left(x_i^-\right)$. In the figure, "i" indexes the instants when an impulse is applied. "Impulse" in Fig. 1 represents the state or the costate update equations at the instants when an impulse is given.

The following are the steps used in the neural network training:

1)Input x_i^- to the critic neural network to obtain λ_i^+ as the output. With λ_i^+ and x_i^-, use (13) to calculate u_i.

2)Use the calculated u_i and x_i^- in the impulse state update equation (11) to get x_i^+. Use the state propagation equation (9) to get x_{i+1}^-. Input x_{i+1}^- to the critic neural network to get λ_{i+1}^+. Use λ_{i+1}^+ and x_{i+1}^- to get λ_{i+1}^{-*} through the impulse costate update equation(12). Use equations (9) and (10) to back propagate the states and costates and get the target λ_i^{+*}. Train SNAC with x_i^- as the input and λ_i^{+*} as the target output.

3)Stop training when the error between λ_i^{+*} and λ_i^+ is 'small enough' (within an error bound set by the control designer).

IV. Simulation Results

For concept illustration, a linear scalar system is considered first, followed by a linear multi-variable example. Finally, a benchmark nonlinear Van Der Pol oscillator problem is also solved using SNAC.

As discussed in Corollary 1, in linear system simulations, the states and costates can be retrieved backward from final time t_f, therefore, the optimal state and costate trajectories will be explicitly known. As long as t_f is large enough, the finite time solution should be very close to the infinite time optimal solution. This explicit solution is used to test the validity of the SNAC scheme in the linear case.

A. Scalar Case

Consider a system described by

$$\dot{x} = ax + bu\delta\left(t-\tau_i\right) \tag{30}$$

where $\tau = 1,2,3,4,...$ are the instants when an impulse is applied.

Consider a quadratic objective function of the form

$$J = \sum_{i=1}^{k+1} Ru_i^2 + \sum_{i=1}^{k} \int_{t_{i-1}^+}^{t_i^+} Qx^2 dt \tag{31}$$

where $R=1$, $Q=1$, and $k+1 \to \infty$.

Optimal conditions for the problem in (30) are as follows:

The state and costate propagation equations when $t \in \left[t_i^+, t_{i+1}^-\right]$ are given by

$$\begin{aligned} &\dot{x} = ax \\ &\dot{\lambda} = -qx - a\lambda \end{aligned}, \tag{32}$$

The state and costate update equations at the impulse instants are

$$\begin{aligned} &x^+ = x^- - \frac{b^2}{r}\lambda^+ \\ &\lambda^+ = \lambda^- \end{aligned} \tag{33}$$

The state and the costate in (32) and (33) evolve according to the following equation

$$\begin{bmatrix} x_{i+1}^- \\ \lambda_{i+1}^- \end{bmatrix} = \begin{bmatrix} e^{a\left(t-t_i\right)} & 0 \\ -\frac{q}{2a}\left(e^{a\left(t-t_i\right)} - e^{-a\left(t-t_0\right)}\right) & e^{-a\left(t-t_i\right)} \end{bmatrix} \begin{bmatrix} x_i^+ \\ \lambda_i^+ \end{bmatrix} \tag{34}$$

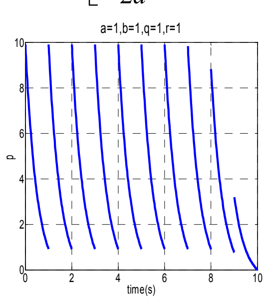


Figure 2. P History

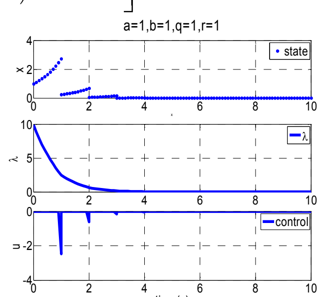


Figure 3. System Response History

Integrating $P_f = 0$ backward according to conditions from (16a)-(16c) from $t_f = 10$ seconds, which is large enough to approximate the infinite time problem, $P(t)$ is obtained. Fig. 2 shows the trajectory history of P when $a=b=q=r=1$. Fig. 3 presents the corresponding state, costate, and control histories. When the open loop system is stable $a=-1$, Figures 4 and 5 are the history of P and the system response, respectively.

In both stable $a=-1$ and unstable $a=1$ cases, the states go asymptotically to the origin with the optimal impulsive control. At the instants when an impulse is given, P reaches constant values, which in Fig. 2 and Fig. 4 are 0.908 and 0.322, respectively. It is easy to test that no matter which final value P_f is, as long as $P_f \geq 0$, P will converge to some constant value, as is shown in Theorem 2.

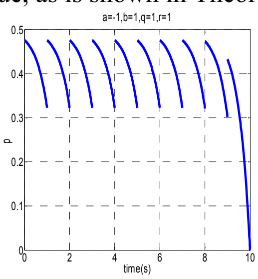


Figure 4. P History

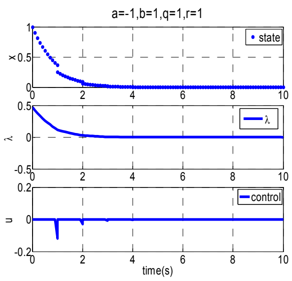
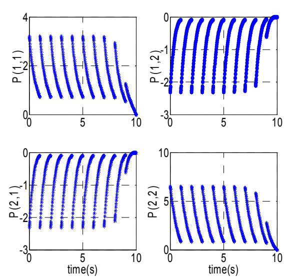


Figure 5. System Response History

To compare the NN performance, the SNAC is synthesized according to the algorithm in section III. Numerical results are presented for the system with parameters $a=b=q=r=1$. Fig. 6 shows the output of the training process starting from $\lambda_i^+ = P_{i\,0} x_i^-$, where $P_{i\,0} = .5$. It is easy to see from Fig. 6 that after three iterations, the slope of the line reaches 0.908, which is exactly the same value obtained through the explicit solution.



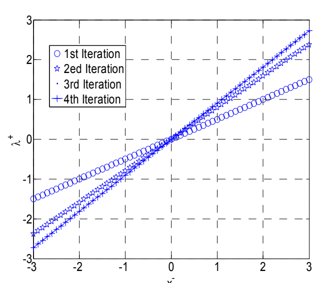


Figure 6. Neural Network Training Process

Figure 7. P vs. Time

By applying the trained neural network to calculate impulsive control, the system response and control are found to be the same as shown in Fig. 3, and therefore, these results are not presented here.

For this scalar problem, direct calculation of P^- using (20) is not difficult. This value is also found to be 0.908. Three methods, namely, a large but finite time approximation, SNAC, and closed form calculations generate the same result (value) for this problem, which validates the theorems presented. Also, it can be concluded that the SNAC technique generates optimal λ_i^+, which implies that SNAC generates optimal control.

B. Linear Case

Consider a simple 2-D system as shown in (35),

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} x + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} u\delta(t-\tau_i) \qquad (35)$$

Assuming fixed impulse instants at $\tau = 1,2,3,4,\dots$, the P matrix is calculated according to the conditions in (16a)-(16c). Assume quadratic objective function with the state weighting matrix Q and control weighting matrix R as identity matrices of dimension two. Fig. 7 shows the time history of the P matrix elements, $P(1,1)$, $P(1,2)$, $P(2,1)$, and $P(2,2)$. The simulation was carried for 10 seconds, which is considered large for this case.

From Fig. 7, it can be seen that with time, when an impulse is given, P reaches a steady state. Elements of the P matrix are given by

$$P = \begin{bmatrix} 0.715 & -0.0875 \\ -0.0875 & 0.841 \end{bmatrix} \qquad (36)$$

and satisfies $\lambda_i^+ = \lambda_i^- = \mathrm{P}\, x_i^-$.

In order to compare (36) to the SNAC solution, a neural network is trained using the SNAC scheme to output λ_i^+.

Using the least square approximation method to map the trained neural network, the gain matrix K satisfying the relation $\lambda^+ = Kx^-$ is calculated as

$$K = \begin{bmatrix} 0.715 & -0.0875 \\ -0.0875 & 0.841 \end{bmatrix} \qquad (37)$$

Note that the least square approximation gives exactly the same gain matrix as the one in (36). This result also validates the accuracy SNAC technique.

C. Van Der Pol Oscillator

Consider a Vander Pol oscillator system dynamics as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + \varepsilon\left(1-x_1^2\right)x_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \delta(t-\tau_i) \quad (38)$$

where ε is a constant parameter. In the simulation, $\varepsilon = 1$.

Assume a quadratic objective function with state weighting matrix Q and control weighting matrix R identity matrices of dimension 2.

Use the scheme shown in Fig. 1. Randomly choose states $\mathbf{x}_i^-$ from the range [-3,3]. Two neural networks are used to approximate the relation between $\mathbf{x}_i^-$ and $\lambda_1(x_1,x_2)_i^+$ and between $\mathbf{x}_i^-$ and $\lambda_2(x_1,x_2)_i^+$, respectively. Impulse instants are chosen as $\tau = 1,2,3,\dots,\infty$. These networks are initialized with stabilizing functions: $\lambda_1 = .9x_1$ and $\lambda_2 = .9x_2$.

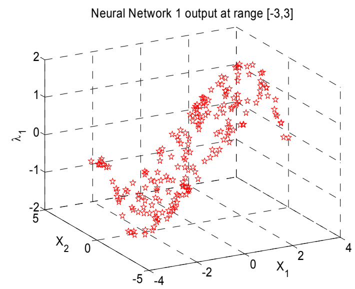


Figure 8. λ_1 vs. (x_1, x_2)

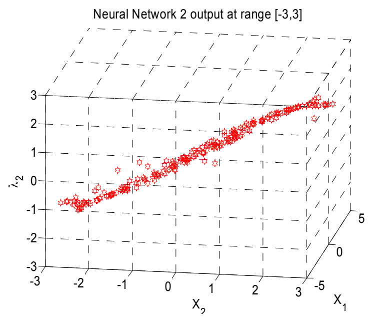


Figure 9. λ_2 vs. (x_1, x_2)

After training, outputs $\lambda_1(x_1, x_2)_i^+$ and $\lambda_2(x_1, x_2)_i^+$ of the trained networks are presented in Figs. 8 and 9, respectively. Note that as to be expected the relations between states and costates can be seen to be nonlinear. Closed loop system response using the trained neural networks is plotted in Fig. 10. From Fig. 10, it can be observed that the system is asymptotically stabilized. Fig. 11 is a comparison of the objective function values calculated with original control ($\lambda_1 = .9x_1$ and $\lambda_2 = .9x_2$) and with optimized control. It is easy to see that the objective function value with optimization is smaller than that without optimization.

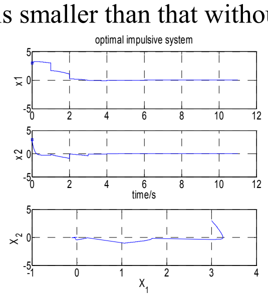


Figure 10. System Response

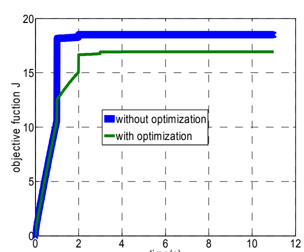


Figure 11. Objective Value

V. CONCLUSIONS

In this paper, necessary conditions for optimal impulse problems for systems with fixed impulse instants were derived. A single network adaptive critic (SNAC) method was developed to solve a fairly general class of nonlinear optimal impulsive control problems. Simulation results of a linear scalar impulse problem, a linear vector impulse problem, and a nonlinear Van Der Pol problem have shown the effectiveness of the SNAC scheme.

REFERENCES

[1] Guan, Z.R., Chen, G.R., Ueta, T. (2000), On impulsive control of a periodically forced chaotic pendulum system, IEEE Transactions on Automatic Control, 45:9, 1724 - 1727.

[2] Grizzle, J.W., Abba, G., and Plestan, F (2001), Asymptotically stable walking for biped robots: analysis via systems with impulse effects. IEEE Transaction on Automatic Control, 46:1, 51-64.

[3] Sun, J.T, Zhang, Y.P. (2004), Impulsive control of Lorenz systems. Fifth World Congress on Intelligent Control and Automation, 1:15-19, 71 – 73

[4] Gilbert, E., Harasty, G. (1971), A class of fixed-time fuel-optimal impulsive control problems and an efficient algorithm for their solution, IEEE Transactions on Automatic Control, 16:1, Feb, 1 - 11.

[5] Lu, Z.H. ; Chi, X.B. ; Chen, L.S, (2003), Impulsive control strategies in biological control of pesticide, Theoretical Population Biology. 64:1, 39-47.

[6] Sun, J.T, Qiao, F., Wu, Q.D. (2005), Impulsive control of a financial model, Physics Letters A, 335:4, 282-288.

[7] Case, J.H. (1979), Economics and the competitive process, New York, New York University Press.

[8] Vadali, S.S., Alfriend, K.T., Vadali, S.R. Sengupta, P. (2005), Formation establishment and reconfiguration using impulsive control. Journal of Guidance, Control, and Dynamics, 28:2, 262-268.

[9] Xin, M., Balakrishnan, S.N. and Pernicka, H. (2004), Deep-Space Spacecraft Formation Flying Using Theta-D Control, University of Missouri-Rolla, Rolla, MO, AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, Rhode Island, Aug. 16-19, 2004

[10] Lakshmikantham, V., Bainov, D.D, Simeonov, P.S. (1989), Theory of impulsive differential equations, World Scientific, Singapore; Teaneck, NJ.

[11] Samoilenko, A.M., and Perestyuk, N.A. (1995), Impulsive differential equations, (translated from the Russian by Yury Chapovsky), World Scientific, Singapore, River Edge, NJ.

[12] Yang. T (1999), Impulsive Control, IEEE Transactions on Automatic Control, 44:5, 1081-1083.

[13] Kozlov, A.K., Osipov, G.V., Shalfeev, V.D. (1997), Suppressing chaos in continuous systems by impulse control, Control of Oscillations and Chaos, Proceedings., 1st International Conference, 3, 578 - 581.

[14] Haddad, W.M., Chellaboina, V., and Kablar, N.A. (2001), Nonlinear Impulsive Dynamical Systems Stability and Dissipativity Part I, International Journal of Control, 74:17, 1631-1658.

[15] Xie, G.M., and Wang, L. (2004), Necessary and Sufficient Conditions for Controllability and Observability of Switched Impulsive Control Systems, IEEE Transactions on Automatic Control, 49:6, 960 - 966.

[16] Luo, J.C., Lee, E.B. (1998), Time-Optimal Control of the Swing using Impulse Control Actions, Proceedings of the American Control Conference, Philadelphia, Pennsylvania, 1, 200-204.

[17] Haddad, W.M.; Chellaboina, V.; Kablar, N.A.; Nonlinear impulsive dynamical systems. II. Feedback interconnections and optimality, Decision and Control, 1999. Proceedings of the 38th IEEE Conference on, 5, 5225 - 5234.

[18] Miller, B.M. (1996), The generalized solutions of nonlinear optimization problems with impulse control, SIAM Journal on Control and Optimization, 34:4, 1420-1440.

[19] Bryson, E., and Ho, Y.C (1975), Applied optimal control-optimization, Estimation and Control, Hemisphere Publishing Corp.

[20] Liu, X.Z. (1995), Impulsive control and optimization, Applied mathematics and computation, 73, 77-98.

[21] Padhi, R., Balakrishnan, S.N. (2006), Optimal management of beaver population using a reduced-order distributed parameter model and single network adaptive critics, IEEE Transactions on Control Systems Technology, 14:4, 628 - 640.

[22] Padhi, R., Balakrishnan, S.N. (2004), Development and analysis of a feedback treatment strategy for parturient paresis of cows, IEEE Transactions on Control Systems Technology, 12:1, 52 - 64.

[23] Santiago, R., and Werbos, P.J. (1994), New Progress towards Truly Brain-Like Control, Proceedings of WCNN'94, San Diego, CA, 27-33.

[24] Prokhorov, D., Wunsch, D. (1997), Adaptive Critic Designs, IEEE Transactions on Neural Networks, 8:5, 997 - 1007.

[25] Han, D.C., Balakrishnan, S.N. (2002), State-constrained agile missile control with adaptive-critic-based neural networks, IEEE Transactions on Control Systems Technology, 10:4, 481 - 489.