01 Jun 2008

Optimal Controller Synthesis of Variable-Time Impulsive Problems using Single Network Adaptive Critics

Xiaohua Wang

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

Optimal Controller Synthesis of Variable-Time Impulsive Problems Using Single Network Adaptive Critics

X. Wang and S.N. Balakrishnan

Abstract—**This paper presents a systematic approach to solve for the optimal control of a variable-time impulsive system. First, optimality condition for a variable-time impulsive system is derived using the calculus of variations method. Next, a single network adaptive critic technique is proposed to numerically solve for the optimal control and the detailed algorithm is presented. Finally, two examples-one linear and one nonlinear-are solved applying the conditions derived and the algorithm proposed. Numerical results demonstrate the power of the neural network based adaptive critic method in solving this class of problems.**

I. Introduction

Many dynamic processes are characterized by the fact that at certain moments of time they experience abrupt changes of the system states. These changes may seem instantaneous because the durations of these changes are negligible in comparison with that of the whole process. Therefore, it is natural to assume that these changes are in the form of impulses. Dynamic systems subjected to impulsive effects are defined as impulsive systems. In order to refer to the time instant, when an impulse happens, the term "impulse instant" is used hereafter in this paper.

There have been applications of impulsive control in chaotic systems [1-3], biped walking robots [4], optimal fixed time fueling process [5], biological control [6], financial and economics control [7-8], and satellite formation control [9-10]. There are many other examples in practice such as ultra high speed optical signals over communication networks, collision of particles, inventory control, government decisions, interest changes, and stock price changes etc.

Two types of impulse driven systems [11, 12] exist. One, which fixes the impulse instants, is called the fixed-time impulsive system. Optimal control of fixed-time impulsive systems has been studied by the authors in [13], where theorems are presented and linear/nonlinear examples are studied. The second type is the variable-time impulsive system. In this type of system, the impulse instants are not fixed but are functions of system states. This paper studies optimal control of a variable-time impulsive system, where an impulse instant is decided by a resetting condition.

Manuscript received December 20, 2007. This work was supported in part by the NSF Grant ECS-0601706.

Xiaohua Wang was a PhD student with Department of Aerospace and Mechanical Engineering, University of Missouri, Rolla, MO 65401 USA. Now she is with Emerson Electric Co (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).

Luo and Lee [14-15] derive the necessary conditions for a variable-time impulsive system using the calculus of variations method, but they do not consider the resetting conditions. Miller [16] uses a method called discontinuous time change to convert the system with impulses to an auxiliary system of conventional differential equations. Miller [17] also studies optimal impulsive control problem with a restricted number of impulses using the same method. Haddad et al. [18] present a hybrid version of Bellman's principle which provides necessary and sufficient optimality conditions for both fixed-time and variable-time impulsive systems. This method requires a Lyapunov function to serve as the optimal cost function at the same time, which would be difficult to find in practice.

Though the literature on impulsive control is quite extensive, still there exists a need for the development of a systematic design method. A major objective in this study is to develop a controller design algorithm for variable-time impulse problems to satisfy that need. Furthermore, the proposed neural network based technique does not require many assumptions and is implementable.

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

II. Optimality Condition For Impulsive Systems

A. Problem Formulation

In this paper, the following variable-time impulsive system is considered, with the system model given by

$$\begin{cases} \dot{x} = f_c(x) + g_c(x)\, u_c & G\left(x\left(\tau_i^-\right)\right) \neq 0 \\ x_i^+ = x_i^- + g_d\left(x_i^-\right) u_i & G\left(x\left(\tau_i^-\right)\right) = 0 \end{cases} \tag{1}$$

where $x(t) \in D \subseteq \mathbb{R}^n$ is the system state. D is an open set with $\mathbf{0} \in D$. $u_c \subseteq \mathbb{R}^{m_c}$ is the continuous control. $u_i \subseteq \mathbb{R}^{m_d}$ represents the impulsive control. $f_c : D \to \mathbb{R}^n$ is a Lipchitz continuous function. $g_d : D \to \mathbb{R}^{n \times m_d}$ is a continuous differentiable function. $g_c : D \to \mathbb{R}^{n \times m_c}$ is a continuous differentiable function. $\mathbb{R}$ represents the set of real numbers. n , m_c , and $m_d \in \mathbb{Z}^+$. $\mathbb{Z}^+$ represents the set of positive integers. $0 \le \tau_1 < \tau_2 < \tau_3 < \cdots < \tau_N < \infty$ are impulse instants. i

978-1-4244-1821-3/08/$25.00©2008 IEEE

indexes the impulse instant sequence. $i = 1 \cdots N \in \mathbb{Z}^+$. Superscript + and − denote the right and left limits with respect to the impulse instants. τ_i^- and τ_i^- are pre-impulse and post-impulse instants, respectively. $G: \mathbb{R}^n \to \mathbb{R}^p$ decides a state set S. $p \in \mathbb{Z}^+$. Initial states x_0 and initial time are assumed to be known.

As described in (1), system equation has two parts. When $G\left(x\left(\tau_i^-\right)\right) \neq 0$, system dynamics are continuous, characterized by ordinary differential equations. When $G\left(x\left(\tau_i^-\right)\right) = 0$, system dynamics are impulsive, with the behavior governed by an impulsive equation, where the states jump according to the impulsive control. $G\left(x\left(\tau_i^-\right)\right) = 0$ is referred to as a resetting condition hereafter.

In this study, a fairly general objective function in (2) is considered.

$$J = \Phi\left(x_f\right) + \sum_{i=1}^{k} L_d\left(u_i\right) + \sum_{i=1}^{k+1} \int_{\tau_{i-1}^+}^{\tau_i^-} L_c\left(x, u_c\right) dt \tag{2}$$

where $\Phi(x_f)$ is the constraint on the terminal states, $\sum_{i=1}^{k} L_d(u_i)$ is the penalty on the impulsive control, and $\sum_{i=1}^{k+1} \int_{\tau_{i-1}^+}^{\tau_i^-} L_c(x, u_c) dt$ is the penalty on the piecewise continuous states. Note that $\tau_{k+1}{}^- = t_f$, where t_f is the final time.

B. Optimality Conditions

Theorem 1: Given the system dynamics in (3) with a known initial time t_0, an initial state x_0, and the final time t_f,

$$\begin{cases} \dot{x} = f_c\left(x, u_c\right) & t \neq \tau_i \\ x_i^+ = x_i^- + g_d\left(x_i^-\right) u_i & t = \tau_i,\ i = 1, 2, 3, \ldots, \mathbb{Z}^+ \\ x_0 \equiv \text{known} & \end{cases} \tag{3}$$

A resetting condition is

$$G\left(x\left(\tau_i^-\right)\right) = 0 \tag{4}$$

and an objective function is as in (2), where the variable definitions are the same as before in (1). Assuming that optimal control exists, and introducing the Hamiltonian function [19],

$$H\left(x, u_c, \lambda\right) \triangleq L_c(x) + \lambda^T f_c(x, u_c), \quad \lambda \in \mathbb{R}^{n \times 1} \tag{5}$$

A continuous control u_c and an impulsive control series of u_i's that produce a stationary value of the objective function J are obtained by solving the following equations.

1) Between two consecutive impulses, $t \in \left[\tau_i^+, \tau_{i+1}{}^-\right]$,

- The state propagation equation is

$$\dot{x} = f_c\left(x, u_c\right) \tag{6}$$

- The costate propagation equation is

$$\frac{\partial H(x, u_c, \lambda)}{\partial x} + \dot{\lambda}^T = 0 \tag{7}$$

- The continuous control equation is

$$\frac{\partial H}{\partial u_c} = 0 \tag{8}$$

2) Between the pre-impulse and the post-impulse instants, $t \in \left[\tau_i^-, \tau_i^+\right]$,

- The state update equation is

$$x_i^+ = x_i^- + g_d\left(x_i^-\right) u_i \tag{9}$$

- The costate update equation is

$$\lambda^T\left(\tau_i^-\right) - \lambda^T\left(\tau_i^+\right)\left(I + \frac{\partial g_d(x_i(\tau_i^-))}{\partial x_i} u_i\right) - \mu_i^T \frac{\partial G}{\partial x_i^-} = 0 \tag{10}$$

- The resetting equation is

$$G\left(x\left(\tau_i^-\right)\right) = 0 \tag{11}$$

- The impulsive control equation is

$$\frac{\partial L_2(u_i)}{\partial u_i} + \lambda^T\left(\tau_i^+\right) g_d\left(x_i\left(\tau_i^-\right)\right) = 0 \tag{12}$$

- The jump equation is

$$H_i^+ = H_i^- \tag{13}$$

The boundary condition are split; i.e., some are given at $t = t_0$ and some are given at $t = t_f$:

$$x_0 \quad \text{given} \tag{14}$$

$$\lambda_f^T = \frac{\partial \Phi(x_f)}{\partial x_f} \tag{15}$$

Proof:

Using Hamiltonian function (5), the objective function (2) can be rewritten as

$$J = \Phi(x_f) + \sum_{i=1}^{k} L_d(u_i) + \sum_{i=1}^{k+1} \int_{\tau_{i-1}^+}^{\tau_i^-} \left(H - \lambda^T \dot{x}\right) dt \tag{16}$$

where $\tau_{k+1}{}^- = t_f$.

Suppose that u_i^* and u_c^* are the optimal control, continuous and impulsive, respectively, and x^* is the optimal state. Define $h \triangleq x_i^- + g_d\left(x_i^-\right) u_i$. Define J_0 as the optimal objective function. By introducing a set of Lagrange multipliers γ_i and μ, the optimal cost function can then be written as

$$\begin{aligned} J_0 = {} & \Phi(x_f) + \sum_{i=1}^{k}\left[L_d(u_i^*) + \mu G\left(x_i^{-*}\right)\right] \\ & + \sum_{i=1}^{k} \gamma_i\left[h(x_i^{-*}, u_i^*) - x_i^{+*}\right] + \sum_{i=1}^{k+1} \int_{\tau_{i-1}^+}^{\tau_i^-}\left[H(x^*, u_c^*) - \lambda^T \dot{x}^*\right] dt \end{aligned} \tag{17}$$

Now, perturb control by letting $u_i = u_i^* + \varepsilon n_i$ and $u_c = u_c^* + \varepsilon s$. Perturb impulsive instants by letting $\tau_i = \tau_i^* + \varepsilon \theta(t)$. Then the corresponding $x = x^* + \varepsilon \eta(t)$. The cost function after perturbation J_ε is

$$\begin{aligned} J_\varepsilon = {} & \Phi(x_f) + \sum_{i=1}^{k}\left[L_d(u_i) + \mu^T G\left(x_i^-\right)\right] \\ & + \sum_{i=1}^{k} \gamma_i\left[h(x_i^-, u_i) - \left(x_i^+\right)\right] + \sum_{i=1}^{k+1} \int_{\tau_{i-1}^+}^{\tau_i^-}\left[H(x, \lambda, u_c) - \lambda^T \dot{x}\right] dt \end{aligned} \tag{18}$$

The last term on the right hand side of (18) can be rewritten as

$$\sum_{i=1}^{k+1}\int_{\tau_{i-1}^{+\,*}+\varepsilon\theta_{i-1}}^{\tau_i^{-*}+\varepsilon\theta_i}\left[H(x^*+\varepsilon\eta,\lambda,u_c^*+\varepsilon s)-\lambda^T\left(\dot{x}^*+\varepsilon\dot{\eta}\right)\right]dt$$
$$=\sum_{i=1}^{k+1}\int_{\tau_{i-1}^{+\,*}+\varepsilon\theta_{i-1}}^{\tau_{i-1}^{+\,*}}+\int_{\tau_{i-1}^{+\,*}}^{\tau_i^{-*}}+\int_{\tau_i^{-*}}^{\tau_i^{-*}+\varepsilon\theta_i}\left[H(x^*+\varepsilon\eta,\lambda,u_c^*+\varepsilon s)-\lambda^T\left(\dot{x}^*+\varepsilon\dot{\eta}\right)\right]dt$$
$$=\sum_{i=1}^{k+1}\left(\int_{\tau_{i-1}^{+\,*}+\varepsilon\theta_{i-1}}^{\tau_{i-1}^{+\,*}}+\int_{\tau_i^{-*}}^{\tau_i^{-*}+\varepsilon\theta_i}L_c dt\right)+\int_{\tau_{i-1}^{+\,*}}^{\tau_i^{-*}}\left[H(x^*+\varepsilon\eta,\lambda,u_c^*+\varepsilon s)-\lambda^T\left(\dot{x}^*+\varepsilon\dot{\eta}\right)\right]dt$$
$$=\sum_{i=1}^{k+1}\left(\left(L_c\right)_i^-\varepsilon\theta_i-\left(L_c\right)_{i-1}^+\varepsilon\theta_{i-1}\right)+$$
$$\int_{\tau_{i-1}^{+\,*}}^{\tau_i^{-*}}\left[H(x^*+\varepsilon\eta,\lambda,u_c^*+\varepsilon s)-\lambda^T\left(\dot{x}^*+\varepsilon\dot{\eta}\right)\right]dt \tag{19}$$

Because the initial conditions are known, the state and time variations at the initial time are zeros. Rearrange the first term of the last line on the right hand side of (19),

$$\sum_{i=1}^{k+1}\left(\left(L_c\right)_i^-\varepsilon\theta_i-\left(L_c\right)_{i-1}^+\varepsilon\theta_{i-1}\right)=\sum_{i=1}^{k}\left(\left(L_c\right)_i^--\left(L_c\right)_i^+\right)\varepsilon\theta_i \tag{20}$$

Also, integrate the last term on the right-hand side of (20) by parts, equation (19) can be written as

$$\sum_{i=1}^{k+1}\int_{\tau_{i-1}^{+\,*}+\varepsilon\theta_{i-1}}^{\tau_i^{-*}+\varepsilon\theta_i}\left[H(x^*+\varepsilon\eta,\lambda,u_c^*+\varepsilon s)-\lambda^T\left(\dot{x}^*+\varepsilon\dot{\eta}\right)\right]dt$$
$$=\sum_{i=1}^{k+1}\int_{\tau_{i-1}^{+\,*}}^{\tau_i^{-*}}\left[\left(\left.\frac{\partial H}{\partial x}\right|_{x^*}+\dot{\lambda}^T\right)\varepsilon\eta+\left.\frac{\partial H}{\partial u_c}\right|_{u_c^*}\varepsilon s\right]dt \tag{21}$$
$$-\sum_{i=1}^{k+1}\varepsilon\lambda^T\eta\Big|_{\tau_{i-1}^{+\,*}}^{\tau_i^{-*}}+\sum_{i=1}^{k+1}\left(\left(L_c\right)_i^--\left(L_c\right)_i^+\right)\varepsilon\theta_i$$

Since $\delta J=\lim_{\varepsilon\to 0}\frac{J-J_0}{\varepsilon}$, the first order approximation of the perturbed cost J_ε can be written as

$$\delta J=\left.\frac{\partial\Phi(x_f)}{\partial x_f}\right|_{x_f^*}\eta_f-\left(\sum_{i=1}^{k}\left(\lambda_i^{+T}\eta_i^+-\lambda_i^{-T}\eta_i^-\right)-\lambda_f^{\ T}\eta_f\right)$$
$$+\sum_{i=1}^{k}\left[\frac{\partial L_d(u_i^*)}{\partial u_i^*}n_i+\mu^T\frac{\partial G\left(x_i^{-*}\right)}{\partial x_i^{-*}}\eta_i^-+\mu^T\frac{\partial G\left(x_i^{-*}\right)}{\partial x_i^{-*}}\dot{x}_i^{-*}\theta_i\right]$$
$$+\sum_{i=1}^{k}\gamma_i\left[\frac{\partial h}{\partial x_i^{-*}}\eta_i^-+\frac{\partial h}{\partial x_i^{-*}}\dot{x}_i^{-*}\theta_i+\frac{\partial h}{\partial u_i^*}n_i-\dot{x}_i^{+*}\theta_i-\eta_i^+\right]$$
$$+\sum_{i=1}^{k+1}\left(\left(L_c\right)_i^--\left(L_c\right)_i^+\right)\theta_i+\sum_{i=1}^{k+1}\int_{\tau_{i-1}^{+\,*}}^{\tau_i^{-*}}\left[\left(\left.\frac{\partial H}{\partial x}\right|_{x^*}+\dot{\lambda}^T\right)\eta+\left.\frac{\partial H}{\partial u_c}\right|_{u_c^*}s\right]dt \tag{22}$$

Rearranging terms, (22) can be written as follows:

$$\delta J=\left(\frac{\partial\Phi(x_f^{\ *})}{\partial x_f^{\ *}}-\lambda_f^T\right)\eta_f+\sum_{i=1}^{k}\left[\frac{\partial L_d(u_i^*)}{\partial u_i^*}+\gamma_i\frac{\partial h}{\partial u_i^*}\right]n_i$$
$$+\sum_{i=1}^{k}\left[\left(\mu^T\frac{\partial G\left(x_i^{-*}\right)}{\partial x_i^{-*}}+\gamma_i\frac{\partial h}{\partial x_i^{-*}}-\lambda_i^{-T}\right)\eta_i^-+\left(\lambda_i^{+T}-\gamma_i\right)\eta_i^+\right]$$
$$+\sum_{i=1}^{k+1}\int_{\tau_{i-1}^{+\,*}}^{\tau_i^{-*}}\left[\left(\frac{\partial H}{\partial x^*}+\dot{\lambda}^T\right)\eta+\frac{\partial H}{\partial u_c^{\ *}}s\right]dt$$
$$+\sum_{i=1}^{k}\left(\gamma_i\left(\frac{\partial h}{\partial x_i^-}\dot{x}_i^--\dot{x}_i^+\right)+\mu^T\frac{\partial G\left(x_i^-\right)}{\partial x_i^-}\dot{x}_i^-+\left(L_c\right)_i^--\left(L_c\right)_i^+\right)\theta_i \tag{23}$$

Since n_i, θ_i, s and η are independent, in order to eliminate them from influencing δJ, one can choose the multiplier $\lambda(t)$ such that the coefficients of n_i, θ_i, and η_i vanish. Consequently,

$$\gamma_i=\left(\lambda_i^+\right)^T \tag{24}$$

$$\dot{\lambda}^T=-\frac{\partial H}{\partial x^*} \tag{25}$$

$$\left(\lambda_i^-\right)^T=\mu\frac{\partial G\left(x_i^{-*}\right)}{\partial x_i^{-*}}+\left(\lambda_i^+\right)^T\frac{\partial h}{\partial x_i^{-*}} \tag{26}$$

$$\gamma_i\left(\frac{\partial h}{\partial x_i^{-*}}\dot{x}_i^{-*}-\dot{x}_i^{+*}\right)+\mu^T\frac{\partial G\left(x_i^{-*}\right)}{\partial x_i^{-*}}\dot{x}_i^{-*}+\left(L_c\right)_i^--\left(L_c\right)_i^+ \tag{27}$$

with the boundary condition

$$\lambda_f^T=\frac{\partial\Phi(x_f^{\ *})}{\partial x_f^{\ *}} \tag{28}$$

Now, (23) becomes

$$\delta J=\sum_{i=1}^{k}\left[\frac{\partial L_d(u_i^*)}{\partial u_i^*}+\gamma_i\frac{\partial h}{\partial u_i^*}\right]n_i+\sum_{i=1}^{k+1}\int_{\tau_{i-1}^{+\,*}}^{\tau_i^{-*}}\frac{\partial H}{\partial u_c^{\ *}}s\,dt \tag{29}$$

For δJ to be zero for any arbitrary n_i and s,

$$\frac{\partial L_2(u_i^*)}{\partial u_i^*}+\left(\lambda_i^+\right)^T g_d\left(x_i^{-*}\right)=0 \tag{30}$$

$$\frac{\partial H}{\partial u_c^{\ *}}=0 \tag{31}$$

Condition (27) can be rearranged and written as the jump condition,

$$H_i^+=H_i^- \tag{32}$$

All the equations in theorem 1 are validated. Equation (32) is referred to as the jump equation hereafter.

Remark 1: For $t\in\left[t_0,t_f\right]$, there are n unknown states $x\in\mathbb{R}^{n\times 1}$ and n unknown costates $\lambda\in\mathbb{R}^{n\times 1}$ decided by *2n* differential equations (6) and (7) for the continuous part, *2n* impulsive dynamics given by (9)-(10) at each impulse instant. There are *2n* boundary conditions given by (14)-(15) for the *2n* unknown states and costates. u_c and u_i are calculated by the optimality equations (8) and (12), respectively. At each impulse instant, an additional jump equation is given by (32) for deciding an optimal impulse instant τ_i together with the resetting equation (11). The number of unknowns is equal to the number of equations.

Remark 2: Since the boundary conditions are split, in order to solve for the optimal controls, the control design is faced with a two-point-boundary-value (TPBV) problem.

Remark 3: For an infinite horizon setting, the final constraint in the objective function (2) is removed and the boundary condition (15) is no longer needed. It is assumed that the optimal solution exists when $t_f \to \infty$.

III. Solution Technique: SNAC

This section introduces the single network adaptive critic (SNAC) technique to solve for the optimal control in the impulsive systems. SNAC has been used in solving nonlinear continuous control problems in [20-21]. This paper extends the SNAC scheme to impulsive control problems.

A. *Adaptive Critic Overview*

The concept of adaptive critic is derived from the modeling of the brain as a supervisor and an action structure [20] where the supervisor criticizes the action (controller) of the system to achieve a better overall goal. Several authors [21-23] have used multilayer-perceptron neural networks with a fixed structure to solve nonlinear control problems arising in aerospace and power systems, as well as other benchmark nonlinear control problems.

The novelty of this paper lies in using neural network structures, SNAC, to solve *optimal variable-time impulsive control* problems.

B. *Infinite Time Adaptive Critic Neural Network Scheme*

For infinite time optimal control, the mapping between the states and the costates does not depend explicitly on time. Therefore, a single neural network can be used to capture the relation between states and costates. In a control affine problem, the costates can be used to calculate the optimal control.

The idea of the SNAC technique is to use the state and the costate propagation equations, the state and costate update equations, the jump equation, and the control expression in Theorem 1 to train a *single* neural network to capture the optimal relation between costates and states at the impulse instant. In this paper, the neural network is trained to approximate the function $\lambda^+\left(x_i^-\right)$ with x_i^- as inputs and λ_i^+ as outputs. The solution presented here assumes that the number of resetting equation $p=1$, which means that the resetting condition is a scalar equation. In this way, the multiplier μ becomes a scalar and can be solved using the jump equation and the state/costate update equations.

Figure 1 gives the flowchart of the optimal impulsive control synthesis using "SNAC".

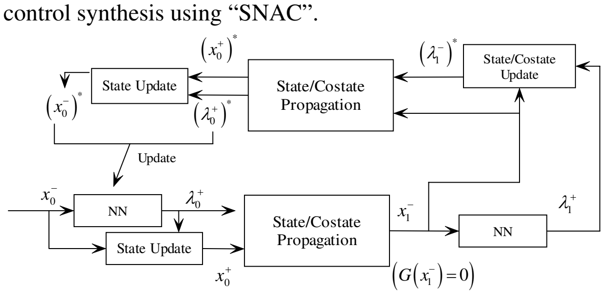

Fig.1 SNAC Scheme for an Impulsive Control Problem

Starting from the left bottom corner of the figure, x_0^- is a set of pre-impulse states satisfying the resetting condition, and it is chosen so that its values approximately span the domain of interest. The neural network is called "NN." λ_0^+ represents the post-impulse costates set, generated by the neural network NN. "State update" block represents the state update equation(9), where the impulsive control is calculated according to the optimal impulsive control equation (12). x_0^+ is the corresponding set of post-impulse states. "State/costate Propagation" block refers to the state and costate propagation equations (6) and (7) where the optimal continuous control is calculated according to the control equation(8). x_1^- is a set of pre-impulse states at the next impulse instant. "State/Costate Update" block refers to the set of equations at the impulsive instants, including (9)-(13). All the training is off-line. The following are the steps used for the neural network training:

1) Input x_0^- to the NN to obtain λ_0^+ as the output.

2) With λ_0^+ and x_0^-, use (12) to calculate u_i assuming λ_0^+ is optimal.

3) Use the calculated u_i and x_0^- in the impulse state update equation (9) to get x_0^+.

4) Integrate x_0^+ and λ_0^+ forward using the state and costate propagation equations (6) and (7) while the optimal continuous control u_c is calculated using (8). The propagation stops at the next impulse instant that satisfies the resetting equation (11). The value of the states at this instant are termed x_1^-.

5) After x_1^- is obtained, input x_1^- to the NN to ouput λ_1^+.

6) Use λ_1^+ and x_1^- to calculate $\left(\lambda_1^-\right)^*$ through the necessary condition set (9)-(13) at the impulse instant.

7) Use (6) and (7) to backpropagate x_1^- and $\left(\lambda_1^-\right)^*$ and get the target state $\left(x_0^+\right)^*$ and costate $\left(\lambda_0{}^+\right)^*$.

8) Use $\left(x_0^+\right)^*$ and $\left(\lambda_0{}^+\right)^*$ to calculate the target state $\left(x_0^-\right)^*$ through the state update equation (9) .

9) Input $\left(x_0^-\right)^*$ to the NN to output $\lambda_0{}^+\left(\left(x_0^-\right)^*\right)$

10) Train SNAC with $\left(x_0^-\right)^*$ as the input and $\left(\lambda_0{}^+\right)^*$ as the target output.

11) Stop training when the error between $\left(\lambda_0{}^+\right)^*$ and $\lambda_0{}^+\left(\left(x_0^-\right)^*\right)$ is 'small enough' (within an error bound set by the control designer).

12) Optimal impulsive control can be calculated in the closed loop form according to the optimal costate $\left(\lambda_0{}^+\left(x_0^-\right)\right)^*$.

IV. Simulation Results

For concept illustration, a linear vector system is considered first, followed by a nonlinear vector example with some similarity to the linear one.

A. Linear Vector Problem

The linear vector problem describes an oscillator with the following dynamics.

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_c + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_i \delta(t-\tau_i) \quad (33)$$

The impulse is activated when the following resetting condition (34) is satisfied.

$$x_1(\tau_i) = 0 \quad (34)$$

Choose the objective function J as

$$J = \sum_{i=1}^{k} \frac{1}{2}\left((u_d)_i\right)^2 + \sum_{i=1}^{k+1} \frac{1}{2} \int_{\tau_{i-1}^+}^{\tau_i^-} (x_1^2 + x_2^2 + u_c^2) dt \quad (35)$$

The state weighting matrix $Q = diag(1,1)$ and the control weighting matrix is a scalar, $R = 1$.

Assume $t_0 = 0$ and $x_0^+ = [0,3]^T$. This problem is considered as an infinite horizon problem by setting $k \to \infty$ and $t_f \to \infty$.

According to Theorem 1, system Hamiltonian function is

$$H = \frac{1}{2}(x_1^2 + x_2^2 + u_c^2) + \lambda_1 x_2 + \lambda_2(-x_1 + u_c) \quad (36)$$

Since $\frac{\partial H}{\partial u_c} = 0$, optimal continuous control $u_c = -\lambda_2$. The state and costate propagation equations are calculated according to (6)-(7). Substituting $u_c = -\lambda_2$, they become

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -x_1 - \lambda_2 \end{cases} \qquad \begin{cases} \dot{\lambda}_1 = -x_1 + \lambda_2 \\ \dot{\lambda}_2 = -x_2 - \lambda_1 \end{cases} \quad (37)$$

At the impulse instants, the necessary conditions are

$$\begin{cases} \lambda_1^+ = \lambda_1^- - \mu \\ \lambda_2^+ = \lambda_2^- \\ x_1^+ = x_1^- = 0 \\ x_2^+ = x_2^- - \lambda_2^+ \\ H_i^+ = H_i^- \end{cases} \quad (38)$$

From the jump equation (13), which implies the continuity of the Hamiltonian function at the impulse instants, considering $x_1^+ = x_1^- = 0$,

$$H^- = \frac{1}{2}(x_2^-)^2 + \lambda_1^- x_2^- - \frac{1}{2}(\lambda_2^-)^2 \quad (a)$$

$$H^+ = \frac{1}{2}(x_2^- - \lambda_2^+)^2 + \lambda_1^+(x_2^- - \lambda_2^+) - \frac{1}{2}(\lambda_2^+)^2 \quad (b) \quad (39)$$

$$H^+ = H^- \quad (c)$$

By substituting (39.b) and (39.a) into (39.c), λ_1^+ can be calculated as

$$\lambda_1^+ = \frac{H^- - \left(\frac{1}{2}(x_2^- - \lambda_2^-)^2 - \frac{1}{2}(\lambda_2^-)^2\right)}{(x_2^- - \lambda_2^-)} = \frac{2\lambda_1^- x_2^- + 2x_2^-\lambda_2^- - (\lambda_2^-)^2}{2(x_2^- - \lambda_2^-)} \quad (40)$$

By substituting (40) into (38), the relations between the post-impulse state/costate and pre-impulse state/costate are obtained as

$$\begin{cases} x_1^+ = x_1^- = 0 \\ x_2^+ = x_2^- - \lambda_2^+ \\ \lambda_1^+ = \dfrac{2\lambda_1^- x_2^- + 2x_2^-\lambda_2^- - (\lambda_2^-)^2}{2(x_2^- - \lambda_2^-)} \\ \lambda_2^+ = \lambda_2^- \end{cases} \quad (41)$$

Two 1-3-1 multilayer perceptron neural networks are chosen to approximate $\lambda_1^+(x_i^-)$ and $\lambda_2^+(x_i^-)$. Considering $x_1(\tau_i) = 0$ is a constant at the impulse instants, the relation between the costates and the states actually becomes $\lambda_1^+(x_2^-)$ and $\lambda_2^+(x_2^-)$ for this problem. The neural network is initially trained with a randomly chosen linear relation as $\lambda_1^+ = x_2^-$ and $\lambda_2^+ = x_2^-$.

Following the training process presented in the previous section, it is found that after three iterations, neural network output converges. The converged network input-output relation is plotted in Fig. 2. From Fig.2, one can observe that a linear relation exists.

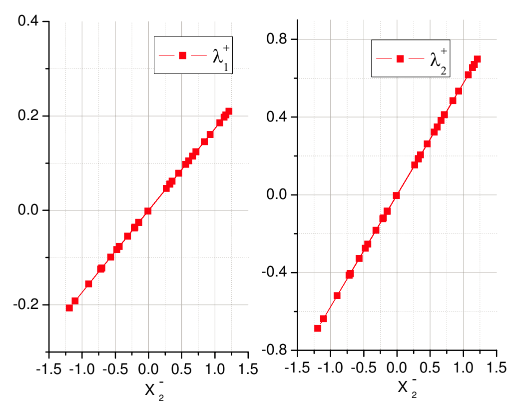

Fig.2. $\lambda_1^+(x_2^-)$ and $\lambda_2^+(x_2^-)$

Fig. 3 is the system response. Trajectories of x_1, x_2, λ_1, and λ_2 are presented for a test case with $x_0 = [0,3]^T$. The controls are calculated using the networks for impulsive control, $u_i = -\lambda_2^+$, and the relations $u_c = -\lambda_2$ for the continuous control. System states asymptotically go to zero in the picture. x_1 and λ_2 trajectories are continuous while x_2 and λ_1 trajectories have jumps at the impulse instants.

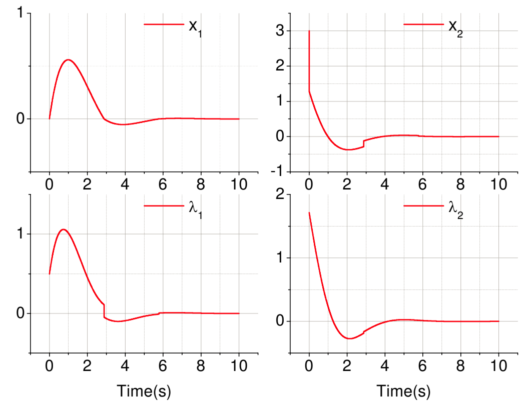

Fig.3. System Response with Controls Generated Using SNAC

Next application is a nonlinear variable-time impulsive problem.

B. Nonlinear Vector Problem

Consider a Van Der Pol oscillator system dynamics as follows:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}=\begin{bmatrix}x_2\\ -x_1+\left(1-x_1^2\right)x_2\end{bmatrix}+\begin{bmatrix}0\\ 1+x_1^2+x_2^2\end{bmatrix}u_c+\begin{bmatrix}0\\ 1\end{bmatrix}u_i\delta\left(t-\tau_i\right) \quad (42)$$

Assume a quadratic objective function with state weighting matrix $Q=I_{2\times2}$ and control weighting matrix $R=I_{1\times1}$.

The resetting condition is set as

$$x_1\left(\tau_i\right)=0 \quad (43)$$

Then the Hamiltonian function is

$$\begin{aligned}H=&\frac{1}{2}\left(x_1^2+x_2^2+u_c^2\right)+\lambda_1\left(x_2\right)\\&+\lambda_2\left(-x_1+\left(1-x_1^2\right)x_2+\left(1+x_1^2+x_2^2\right)u_c\right)\end{aligned} \quad (44)$$

The state and the costate propagation equations follow.

$$\begin{cases}\dot{x}_1=x_2\\ \dot{x}_2=-x_1+\left(1-x_1^2\right)x_2+\left(1+x_1^2+x_2^2\right)u_c\\ \dot{\lambda}_1=-\dfrac{\partial H}{\partial x_1}=-x_1+\lambda_2\left(-1-2x_1x_2+2x_1u_c\right)\\ \dot{\lambda}_2=-\dfrac{\partial H}{\partial x_2}=-x_2-\lambda_1+\lambda_2\left(\left(1-x_1^2\right)+2x_2u_c\right)\end{cases} \quad (45)$$

Optimal continuous control is calculated using

$$\frac{\partial H}{\partial u_c}=u_c+\left(1+x_1^2+x_2^2\right)\lambda_2=0 \quad (46)$$

which yields $u_c=-\left(1+x_1^2+x_2^2\right)\lambda_2$ (47)

By substituting (47) in (45), the state and the costate propagation equations become

$$\begin{cases}\dot{x}_1=x_2\\ \dot{x}_2=-x_1+\left(1-x_1^2\right)x_2-\left(1+x_1^2+x_2^2\right)^2\lambda_2\\ \dot{\lambda}_1=-x_1+\lambda_2+2\lambda_2x_1x_2+2\lambda_2^2x_1+2\lambda_2^2x_1^3+2\lambda_2^2x_1x_2^2\\ \dot{\lambda}_2=-x_2-\lambda_1-\lambda_2+\lambda_2x_1^2+2\lambda_2^2x_2+2\lambda_2^2x_2x_1^2+2\lambda_2^2x_2^3\end{cases} \quad (48)$$

From the necessary conditions at an impulse instant, the following equation result,

$$\begin{cases}x_1^+=x_1^-=0\\ x_2^+=x_2^-+u_d\\ \lambda_1=\lambda_1+\mu\\ \lambda_2=\lambda_2\\ u_d+\lambda_2=0\\ H_i^+=H_i^-\end{cases}\Rightarrow\begin{cases}x_1^+=x_1^-=0\\ x_2^+=x_2^--\lambda_2\\ \lambda_1=\lambda_1+\mu\\ \lambda_2=\lambda_2\\ H_i^+=H_i^-\end{cases} \quad (49)$$

Then the pre-impulse and post-impulse Hamiltonians become

$$H^-=\frac{1}{2}\left(x_2^-\right)^2-\frac{1}{2}\left(1+\left(x_2^-\right)^2\right)^2\left(\lambda_2^-\right)^2+x_2^-\left(\lambda_1^-+\lambda_2^-\right) \quad (50)$$

$$H^+=\frac{1}{2}\left(x_2^+\right)^2-\frac{1}{2}\left(1+\left(x_2^+\right)^2\right)^2\left(\lambda_2^+\right)^2+x_2^+\left(\lambda_1^++\lambda_2^+\right) \quad (51)$$

By equal H^+ to H^-,

$$\lambda_1^+=\frac{H^--\left[\frac{1}{2}\left(x_2^--\lambda_2^-\right)^2-\frac{1}{2}\left(1+\left(x_2^--\lambda_2^-\right)^2\right)^2\left(\lambda_2^-\right)^2-x_2^-\lambda_2^-\right]}{x_2^+} \quad (52)$$

Then the state and costate update equation (49) become

$$\begin{cases}x_1^+=x_1^-=0\\ x_2^+=x_2^--\lambda_2\\ \lambda_1^+=\dfrac{H^--\left[\frac{1}{2}\left(x_2^--\lambda_2^-\right)^2-\frac{1}{2}\left(1+\left(x_2^--\lambda_2^-\right)^2\right)^2\left(\lambda_2^-\right)^2-x_2^-\lambda_2^-\right]}{x_2^+}\\ \lambda_2^+=\lambda_2^-\end{cases} \quad (53)$$

Strong nonlinearity is observed in both the propagation equations in (48) and update equations in (53). Two 1-6-3-1 multilayer perceptron neural networks are used to capture the optimal relation $\lambda_1^+\left(x_i^-\right)$ and $\lambda_2^+\left(x_i^-\right)$.

The nonlinearity in the system brings numerical problems into the training process if the initial guess of the relation between the costates and the states is not proper. The states and costates may go to infinity in a finite time with an intial guess because of the nonlinearity.

In order to obtain a proper initial guess of $\lambda_1^+\left(x_i^-\right)$ and $\lambda_2^+\left(x_i^-\right)$, the infinite horizon relation optimal relation between the costate and state during the **continuous dynamic** part is used, which are denoted as $\lambda_1(x)$ and $\lambda_2(x)$ and used as the initial guess at post-impulse τ_i^+ instant. In this way, the initial guess of $\lambda_1^+\left(x^+\right)$ and $\lambda_2^+\left(x^+\right)$ is obtained. Then according to the state update equation, the initial guess of $\lambda_1^+\left(x^-\right)$ and $\lambda_2^+\left(x^-\right)$ can be calculated. The SNAC used to capture $\lambda_1(x)$ and $\lambda_2(x)$ will be referred to as the continuous SNAC hereafter. Details about the SNAC scheme for the infinite horizon costate and state relationship during the continuous dynamics part has been presented by Padhi [24].

After a proper guess of the initial guess of $\lambda_1^+(x_i^-)$ and $\lambda_2^+(x_i^-)$ is obtained, follow the training procedure presented in the previous section. Two 1-6-3-1 multilayer perceptron neural networks are chosen for the training. After 8 iterations, the training converges and the required error bound is achieved.

Fig.4 shows the difference between the initial relations and the trained relations of $\lambda_1^+(x^-)$ and $\lambda_2^+(x^-)$. The final trained relations $\lambda_1^+(x^-)$ and $\lambda_2^+(x^-)$ show strong nonlinearity in the pictures.

To compare impulsive SNAC with the infinite horizon SNAC relation of the continuous dynamics, the continuous SNAC relation $\lambda_1(x)$ and $\lambda_2(x)$ before and after training is depicted in Fig.5. From Fig.4 and 5, the relations captured by the impulsive SNAC and the continuous SNAC are different.

The reason that the relation captured by the impulsive SNAC and the continuous SNAC is different lies in that the relation captured by the impulsive SNAC is not an infinite horizon relation but a finite horizon relation. But in the optimal impulsive variable-time problem, a steady state relation $\lambda_1^+(x^-)$ and $\lambda_2^+(x^-)$ still exists because this finite horizon relation repeats between two impulses.

Fig.6 shows the system response using the costate and state relation captured by the neural network. It is seen from the picture that system states are asymptotically stable using the SNAC based control design. From the figure, x_1 and λ_2 trajectories are continuous while x_2 and λ_1 trajectories have jumps at the impulse instants. This phenomena is consistent with the state and costate update equations.

V. CONCLUSION

In this paper, the necessary conditions are derived for optimal control of the variable-time impulsive systems. A single neural network adaptive critic (SNAC) method is developed to numerically solve the linear/nonlinear optimal impulsive control problems. The simulation results of a linear variable-time impulsive problem and a nonlinear variable-time impulsive problem show the effectiveness of the SNAC scheme. A systematic scheme of the optimal control of variable-time impulsive systems has been developed.

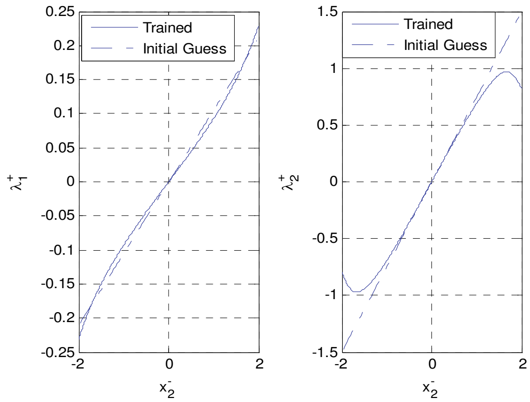

Fig.4. Impulsive SNAC Relation before Training and after Training

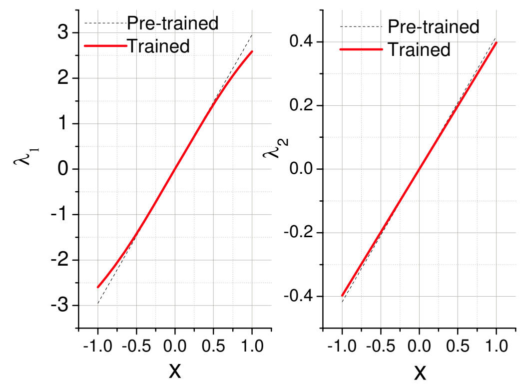

Fig.5. SNAC for the Continuous Dynamics before and after Training

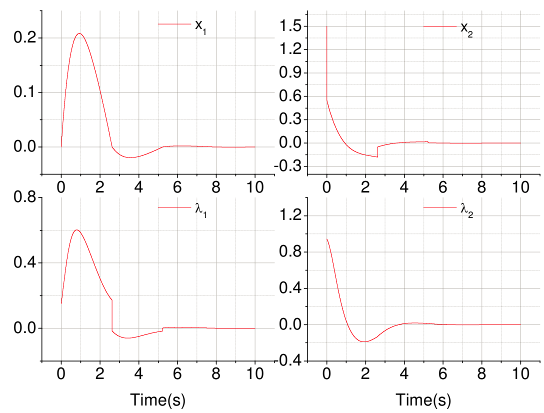

Fig.6. State and Costate Trajectory using SNAC

REFERENCES

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

[2] 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.

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

[4] 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.

[5] Gilbert, E. and 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):1-11.

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

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

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

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

[10] Xin. M., Balakrishnan, S. N. and Pernicka, H. (2004). Deep-Space Spacecraft Formation Flying Using Theta-D Control, AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, Rhode Island.

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

[12] Yang, T. (2001). Impulsive control theory, Berlin, New York , Springer Verlag.

[13] Wang, X. and Balakrishnan, S.N. (2007). Optimal Neurocontroller Synthesis for Impulse-Driven Systems, Proceedings of the 2007 American Control Conference, New York City, USA.

[14] Luo, J. and Lee, E. B. (1998). Time-Optimal Control of the Swing using Impulse Control Actions, Proceedings of the American Control Conference, Philadelphia, Page(s): 200-204.

[15] Luo, J. (2000). Optimal control for control actions with mixed impulse functions and bounded piecewise continuous functions, PhD dissertation, University of Minnesota.

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

[17] Miller, B. M. and Rubinovich, E. Y. (2003). Impulsive control in continuous and discrete-continuous systems, New York : Kluwer Academic/Plenum Publishers.

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

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

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

[21] Han, D. and 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.

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

[23] Padhi, R. and Balakrishnan, S.N. (2004). Optimal beaver population management using reduced order distributed parameter model and single network adaptive critics, American Control Conference, Proceedings of the 2004, 2(30):1598-1603.

[24] Padhi, R., Unnikrishnan, N., Wang, X., and Balakrishnan, S.N., A single network adaptive critic (SNAC) architecture for optimal control synthesis for a class of nonlinear systems (2006), Neural Networks, 19(10):1648-1660.