Single Network Adaptive Critic for Vibration Isolation Control ⋆

Jia Ma, Tao Yang, Zeng-Guang Hou, and Min Tan

Key Laboratory of Complex System and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing, 100080 China, (Tel: 86-10-82614502; e-mail: jia.ma@ia.ac.cn).

Abstract: Vibration isolation control is the critical issue to guarantee the performance of various vibration-sensitive instruments and sensors in practical engineering systems. In this paper, single network adaptive critic (SNAC) based controllers are developed for vibration isolation applications. The SNAC approach differs from the typical action-critic dual network structure in adaptive critic designs (ACDs) by eliminating the action network, which leads to substantial computational savings. Two training methods, i.e., the off-line and online methods are proposed to adapt the SNAC controllers respectively. In contrast with the existing off-line SNAC training method, the off-line method proposed in this paper adopts the least mean square (LMS) training algorithm with variable learning rate to make the training procedure converge faster. Furthermore, for real-time control purpose, the online learning method is presented for tuning the weights of the critic networks along the real-time state trajectories of the isolation system. Additionally, the "shadow critic" training strategy used in the online method further improves the convergence rate. Simulation results have shown that the developed SNAC controllers using the different training methods can converge to the continuous-time optimal control solution at satisfactory speed. Moreover, the designed SNAC controllers alleviate vibration disturbance more effectively and have better control performance in comparison with the passive isolator.

1. INTRODUCTION

Mechanical vibration exists in all locations where equipments work, which may cause many problems. In high-precision manufacturing industry such as the integrated circuit (IC) industry (Chen et al. [2005]), in order to meet the stringent accuracy and performance requirements, its production machinery and measuring devices must have high sensitivity. However, they are very sensitive to vibration noise simultaneously. Thus a small amount of undesirable vibration may cause poor accuracy and greatly shorten the lifetime of the machines. And in vehicle engineering (Al-Holou et al. [2002]), vibration isolation has always been a focus of research for years. Much work has been undertaken to improve the riding comfort and reduce the jerk effect on the body and components of the car. In addition, vibration isolation problems in space applications have attracted increasing interests in recent years. The major adverse effect of vibration in space is degraded performance of various sensitive instruments. For example, for a satellite camera used to image objects on the ground, a small vibration on the spacecraft may result in significant image degradation (Anderson et al. [2000]). Hence, attention in this paper is paid to vibration isolation control, which aims to mitigate the effect of vibration on sensitive components.

Many vibration isolation methods have been discussed in the literature (Chen et al. [2005], Spanos et al. [1995], Miller et al. [1995]). Among them, there are two most effective approaches for attenuating undesirable vibration; that is

- Isolation of the disturbance source, called "noisy side" isolation
- Isolation of the sensitive payload, called "quiet side" isolation

"Quiet side" isolation has the advantage of protecting the critical component from all vibration sources. So this isolation method is adopted in this paper. Specifically, an isolator is equipped between the sensitive payload and sources of vibration to reduce the transmission of external disturbances. Additionally, "quiet side" vibration isolation system can be generally divided into two types: passive and active ones. A passive system requires no continuous input of power. But due to the constraints of physics and mechanical components, it has limited performance, especially at lower frequencies . Thus in order to achieve the desired performance of vibration control, an active vibration system is considered in this work although external power has to be provided.

The active vibration control problem discussed in this paper is formulated as an optimal control problem. Dynamic programming is a very useful tool in solving the latter problem. However, direct solution of the Hamilton-Jacobi-Bellman (HJB) equation, which is the centerpiece of the dynamic programming algorithm to obtain the optimal cost function, is computationally intense due to the "curse of dimensionality". To overcome the numerical

⋆ This work was supported in part by the National Natural Science Foundation of China (Grants 60775043 and 60635010).

complexity, approximate dynamic programming (ADP) was proposed (Barto et al. [1983], Werbos [1977]), in which neurocomputing is used to approximate dynamic programming solution. In (Werbos [1992]), Werbos has suggested a family of adaptive critic designs (ACDs) to implement the idea of ADP, which generally includes: heuristic dynamic programming (HDP), dual heuristic programming (DHP), and global dual heuristic programming (GDHP). A typical structure of ACDs consists of two neural networks, namely a critic neural network and an action neural network, both of which are used to approximate the HJB equation. By iteratively training the two networks until they are consistent to each other, the optimal or suboptimal control law for the system can be obtained. So far, ACDs methods have been successfully used to synthesis neurocontrollers for optimal control problems such as aircraft autolanding (Gaurav and Balakrishnan [1997]), inverted pendulum (Si and Wang [2001], Liu et al. [2001]).

This paper attempts to handle the vibration isolation problem using the single network adaptive critic (SNAC) approach. This approach, firstly proposed by Padhi in 2004 (Padhi et al. [2004]), is a significant improvement to the DHP architecture. In conventional DHP approach, the critic network estimates the derivatives of the cost-to-go function (the function J of Bellman's equation in DP) with respect to the plant states and the action network is responsible to generate control signals. In contrast to DHP, the SNAC approach adopts only the critic network instead of the action-critic dual network. Thus the iterative training loops between the action and critic networks are no longer required, which leads to significant computational savings. In this paper, some improvements have been made in the off-line training method which is used in (Padhi et al. [2004]). Specifically, the least mean square (LMS) algorithm with variable learning rate is introduced to accelerate the training process. Based on the modified off-line training method, the online training method is proposed to adapt the SNAC controller, which makes the SNAC approach possible to deal with dynamic systems with uncertainties on-line. In this method, the "shadow critic" training technique is adopted to make the training procedure converge faster.

The rest of this paper is organized as follows. Section 2 describes an active-passive vibration isolator, and presents its mathematical model for optimal control purpose. Section 3 presents a SNAC based control approach. And two methods of training SNAC controllers, i.e., off-line training and online training methods are discussed in detail. Section 4 describes some simulation experiments which are carried out to evaluate the performance of the designed vibration controllers. Conclusions are summarized in Section 5.

2. SYSTEM DESCRIPTION AND THE MATHEMATICAL MODEL

2.1 System Description

For suppression of vibration originating on the base, the "series" active-passive architecture is designed as shown in Fig. 1. In this diagram, an active stage is composed of a stiff spring K_s, a damper C_s, and a soft actuator u. In order to actively minimize vibration transmission, a voice coil motor is adopted as the soft actuator in this design, which generates a controlling force based on the motion measurements of the payload mass M_s and the intermediate mass M_u. Under the active stage, a spring K_t is used as a passive isolator. Thus, the active portion of the system reduces vibration transmission at low frequencies and the passive portion attenuates high frequency inputs.

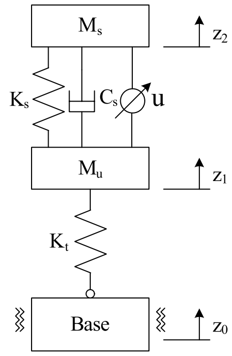

Fig. 1. An active vibration isolation system.

2.2 Mathematical Model

By employing the Newton's second law of motion, the governing equations can be deducted as follows:

$$\begin{aligned} M_u\ddot{z}_1 &= C_s(\dot{z}_2-\dot{z}_1)-K_t(z_1-z_0)+K_s(z_2-z_1)-u \\ M_s\ddot{z}_2 &= -C_s(\dot{z}_2-\dot{z}_1)-K_s(z_2-z_1)+u \end{aligned} \quad (1)$$

where z_1, $\dot{z}_1$ and $\ddot{z}_1$ respectively denote the displacement, velocity and acceleration of the intermediate mass M_u, z_2, $\dot{z}_2$ and $\ddot{z}_2$ respectively represent the movement parameters of the payload mass M_s, the disturbance input z_0 is the displacement of the base excited by external vibration sources, u is the control force generated by the voice coil motor, K_t and K_s are the spring coefficients, C_s is the damper coefficient.

By letting the state vector $X(t) \triangleq [z_1(t), \dot{z}_1(t), z_2(t), \dot{z}_2(t)]^T$, (1) can be written as

$$\dot{X}(t) = AX(t) + Bu(t) + Ez_0(t) \quad (2)$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{K_s+K_t}{M_u} & -\frac{C_s}{M_u} & \frac{K_s}{M_u} & \frac{C_s}{M_u} \\ 0 & 0 & 0 & 1 \\ \frac{K_s}{M_s} & \frac{C_s}{M_s} & -\frac{K_s}{M_s} & -\frac{C_s}{M_s} \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & -\frac{1}{M_u} & 0 & \frac{1}{M_s} \end{bmatrix}^T,$$

$$E = \begin{bmatrix} 0 & \frac{K_t}{M_u} & 0 & 0 \end{bmatrix}^T.$$

Now the optimal control problem can be formulated to drive $X(t) \to 0$ with minimization of the following cost-to-go function J:

$$J = \frac{1}{2}\int_0^\infty (X^T Q_w X + R_w u^2)dt \quad (3)$$

where $Q_w > 0$ and $R_w > 0$ are weighting matrices for state and control respectively.

3. SNAC CONTROLLER DESIGNS

Although the SNAC approach eliminates the action network, it retains the main features of the ACDs and still relies on the principle of ADP. According to the idea of ADP, the costate recursion is derived in the following (Padhi et al. [2004]), which is the base of training the critic network in SNAC.

In discrete-time formulation, the system can be described by

$$X(k+1) = f_x(X(k), u(k)). \tag{4}$$

The utility function is defined as

$$U(k) = f_U(X(k), u(k)). \tag{5}$$

Since the vibration control problem is an infinite horizon problem, the cost-to-go function $J(k)$ from time step k is given by

$$J(k) = \sum_{n=0}^{\infty} \gamma^n U(k+n) \tag{6}$$

where γ is a discount factor ($0 < \gamma < 1$). In this paper, γ is assumed to be 0.95.

Then $J(k)$ can be split into

$$J(k) = U(k) + \gamma J(k+1). \tag{7}$$

The 4×1 costate vector at time step k is defined as

$$\lambda(k) = \frac{\partial J(k)}{\partial U(k)}. \tag{8}$$

Thus,

$$\begin{aligned} \frac{\partial J(k)}{\partial u(k)} &= \frac{\partial U(k)}{\partial u(k)} + \gamma \frac{\partial J(k+1)}{\partial u(k)} \\ &= \frac{\partial U(k)}{\partial u(k)} + \gamma \left[\frac{\partial X^T(k+1)}{\partial u(k)} \right] \left[\frac{\partial J(k+1)}{\partial x(k+1)} \right] \\ &= \frac{\partial U(k)}{\partial u(k)} + \gamma \left[\frac{\partial X^T(k+1)}{\partial u(k)} \right] \lambda(k+1). \end{aligned} \tag{9}$$

To minimize the cost-to-go function, the necessary condition for optimality as follows must be satisfied

$$\frac{\partial J(k)}{\partial u(k)} = 0. \tag{10}$$

Then the optimal control equation can be written as

$$\frac{\partial U(k)}{\partial u(k)} + \gamma \left[\frac{\partial X^T(k+1)}{\partial u(k)} \right] \lambda(k+1) = 0. \tag{11}$$

The costate equation is derived as follows:

$$\begin{aligned} \lambda(k) = \frac{\partial J(k)}{\partial X(k)} &= \frac{\partial U(k)}{\partial X(k)} + \gamma \frac{\partial J(k+1)}{\partial X(k)} \\ &= \frac{\partial U(k)}{\partial X(k)} + \frac{\partial u(k)}{\partial X(k)} \frac{\partial U(k)}{\partial u(k)} \\ &+ \gamma \left[\frac{\partial X^T(k+1)}{\partial X(k)} + \frac{\partial u(k)}{\partial X(k)} \frac{\partial X^T(k+1)}{\partial u(k)} \right] \frac{\partial J(k+1)}{\partial X(k+1)} \\ &= \frac{\partial U(k)}{\partial X(k)} + \gamma \frac{\partial X^T(k+1)}{\partial X(k)} \lambda(k+1) \\ &+ \frac{\partial u(k)}{\partial X(k)} \left[\frac{\partial U(k)}{\partial u(k)} + \gamma \frac{\partial X^T(k+1)}{\partial u(k)} \lambda(k+1) \right]. \end{aligned} \tag{12}$$

Hence, by using the optimal control equation (11), the costate equation (12) on the optimal path can be simplified to

$$\lambda(k) = \frac{\partial U(k)}{\partial X(k)} + \gamma \frac{\partial X^T(k+1)}{\partial X(k)} \lambda(k+1). \tag{13}$$

3.1 Off-line Neural Network Training

The critic neural network used in SNAC aims to capture the functional relationship between $X(k)$ and $\lambda(k+1)$, whereas in ACDs the critic network capture the relationship between $X(k)$ and $\lambda(k)$. Hence, the SNAC approach is applicable to problems where the optimal control equation (11) is explicitly solvable for the control variable $u(k)$ in terms of the state vector $X(k)$ and costate vector $\lambda(k+1)$, while eliminating the action network.

To train the critic network, whose output is $\lambda_a(k+1)$, a variable called $\lambda_t(k+1)$ has to be calculated as the desired output. In (Padhi et al. [2004]), $\lambda_t(k+1)$ was calculated from the costate equation (13) by using $X(k+1)$ and $\lambda(k+2)$. Then the Levenberg-Marquardt backpropagation scheme was used to train the critic network. The learning rate of the critic network cannot be adjusted directly using this technique, whereas this training parameter is very important to the learning loop dynamics. Specifically, the learning rate may determine whether or not the process will converge and how fast the convergence rate is if it converge. Hence, in order to schedule the learning rate value for faster convergence, the LMS algorithm (Prokhorov and Wunsch [1997]) is applied to update the weights of the critic network in this paper.

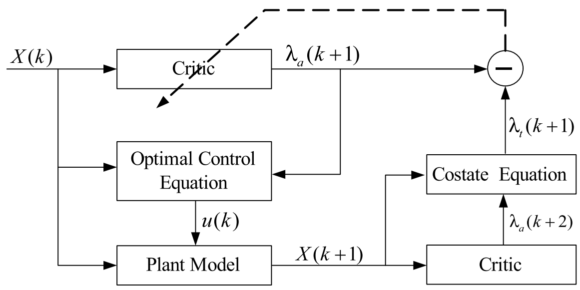

Fig. 2. SNAC off-line training scheme.

The off-line training process is illustrated in Fig. 2. Details are given in the following.

The error measure over time is defined as

$$E_c = \frac{1}{2} \sum_k e_c^T(k) e_c(k) \tag{14}$$

where

$$e_c(k) = \lambda_a(k) - \frac{\partial U(k)}{\partial X(k)} - \gamma \frac{\partial J(k+1)}{\partial X(k)} = \lambda_a(k) - \lambda_t(k), \tag{15}$$

$\lambda_a(k)$ is the actual output of the critic network, $\lambda_t(k)$ is the desired output of the critic network, which can be calculated by (13).

The necessary condition for (14) to be minimized is given as follows:

$$\frac{1}{2} \frac{\partial}{\partial W_c} \left(e_c^T(k) e_c(k) \right) = e_c(k) \frac{\partial e_c(k)}{\partial W_c} = e_c(k) \frac{\partial \lambda_a(k)}{\partial W_c} = 0 \tag{16}$$

where W_c is the weights of the critic network.

According to the LMS algorithm, the update rule for the critic network's weights can be written as

$$\Delta W_c = -\alpha \left(\lambda_a(k) - \lambda_t(k)\right) \frac{\partial \lambda_a(k)}{\partial W_c} \tag{17}$$

where α is the learning rate of the critic network ($0 < \alpha < 1$).

During the training process, the learning rate starts from a larger value and then anneals downwards as learning proceeds.

Note that state generation is an important part of training procedures for SNAC. In order to satisfy persistent excitation, a sufficient varied set of training samples should be repeated often enough. For this purpose, define $S = \{X : X \in \text{Domain of interest}\}$ on which the critic network has to be trained. This set should cover a large number of points of the state space where the admissible trajectories lie.

The training procedure continues until convergence conditions for the critic network are met. Firstly, check the convergence on the training set. Let the outputs from the critic network be $\lambda_a(k)$ and the calculated outputs obtained from the costate equation (13) be $\lambda_t(k)$. A tolerance value *tolc* is used to test the convergence of the critic network. The relative error $e_{rc}(k)$ is defined as

$$e_{rc}(k) = \| \lambda_t(k) - \lambda_a(k) \|_2 \,/\, \| \lambda_t(k) \|_2 \,.$$

When $\| e_{rc}(k) \|_\infty < tolc$, the convergence condition for the current training episode is met.

After successful training the critic network on the current training set, a set of new states are randomly selected as the test set from S. Their corresponding costate values $\lambda_a(k)$ and $\lambda_t(k)$ are used to obtain $e_{rc}(k)$. If $\| e_{rc}(k) \|_\infty < tolc$, the training procedure stops.

The total off-line training procedure of the SNAC network can be summarized as follows.

(1) Initialize k=0.
(2) Randomly select N states as a training set S_t from the set S.
(3) For each element $X(k)$ of S_t, do the following.
    (a) Apply $X(k)$ to critic network to obtain $\lambda_a(k+1)$.
    (b) Calculate $u(k)$ by the optimal control equation (11) using $X(k)$ and $\lambda_a(k+1)$.
    (c) Apply $X(k)$ and $u(k)$ to the plant to obtain $X(k+1)$.
    (d) Apply $X(k+1)$ to the critic network to obtain $\lambda_a(k+2)$.
    (e) Calculate $\lambda_t(k+1)$ by the costate equation (13) using $X(k+1)$ and $\lambda_a(k+2)$.
(4) Update the weights of the critic network for all $X(k) \in S_t$ in this episode by adopting the LMS algorithm.
(5) Check the convergence of the critic network. If the convergence conditions are satisfied, terminate the training procedure. Otherwise, repeat Steps 2-4.

3.2 Online Neural Network Training

An online training strategy which stems from the *strategy 3a* in (Lendaris and Paintz [1997]) is adopted in this paper. Details are given in the following.

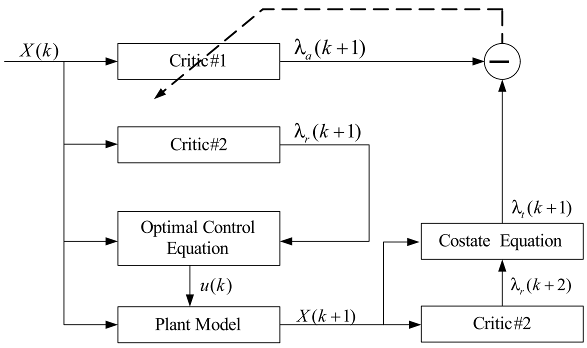

Fig. 3. SNAC online training scheme.

A copy of the critic network called "shadow critic" is introduced in this training strategy. The original network is called CriticNN#1 and the copy is called CrticNN#2. Let the outputs from the CriticNN#1 be $\lambda_a(k)$. Let the outputs from the CriticNN#2 be $\lambda_r(k)$. In each episode, the CriticNN#2's parameters are hold constant and the CriticNN#2 is used to calculate the target costate states $\lambda_t(k)$. Then the target values obtained are applied for training the CriticNN#1 on-line. Only at the end of the episode, the CriticNN#2 is made identical to the adapted CriticNN#1 and the convergence condition is checked. This strategy can provide the target values generated from a stable platform during each training episode while adapting the CriticNN#1. In this way, the overall speed of convergence can be improved. The online training process is illustrated in Fig. 3.

Since the training procedure is performed on-line, the convergence condition is checked on the states along the trajectory; that is, the states generated in the next episode (assuming each episode continues N time steps) are used to check the convergence condition.

Assume the current time step is k. The relative error $e_{rc}(k)$ is defined as

$$e_{rc}(k+n) = \| \lambda_t(k+n) - \lambda_r(k+n) \|_2 \,/\, \| \lambda_t(k+n) \|_2$$

where $\lambda_t(k+n)$ can be calculated by the costate equation (11) using $X(k+n)$, $\lambda_r(k+n+1)$ and $n = 1, 2, \cdots N$.

$$e_{rc} \triangleq \{e_{rc}(k+n), n = 1, 2, \cdots, N\}.$$

The tolerance value *tolc* is also used as a convergence criterion for the critic network in the online training procedure. When $\| e_{rc} \|_\infty < tolc$, the critic network is considered convergent.

The total online training procedure of the SNAC network can be summarized as follows.

(1) Initialize $k = 0$ and $X(0)$.
(2) For each episode i ($i = 1, 2, \cdots$), where an episode contains N time step, do the following.
    (a) Apply $X(k)$ to the CriticNN#1 to obtain $\lambda_a(k+1)$.
    (b) Apply $X(k)$ to the CriticNN#2 to obtain $\lambda_r(k+1)$.

(c) Calculate $u(k)$ by the optimal control equation (11) using $X(k)$ and $\lambda_r(k+1)$.
(d) Apply $X(k)$ and $u(k)$ to the plant to obtain $X(k+1)$.
(e) Apply $X(k+1)$ to the CriticNN#2 to obtain $\lambda_r(k+2)$.
(f) Calculate $\lambda_t(k+1)$ for the CriticNN#1 by the costate equation (13) using $X(k+1)$ and $\lambda_r(k+2)$.
(g) Update the weights of the CriticNN#1 by adopting the LMS algorithm.
(h) If time step $k < i \times N$, increase k and go to a).
(i) Set CriticNN#2=CriticNN#1.
(j) Check the convergence of the critic network. If the convergence condition is satisfied, terminate the training procedure. Otherwise, go to 2).

4. COMPUTATIONAL EXPERIMENT

4.1 Discrete model of the vibration isolation system

Since the SNAC approach described in this paper is based on a discrete system plant, the continuous model of the vibration isolation system and the cost-to-go function J given in Section 2 should be discretized. Their discrete forms can be written as follows:

$$X(k+1) = X(k) + \Delta t\,[AX(k) + Bu(k) + Ez_0(k)] \quad (18)$$

where A, B, E have been defined in (2),

$$J(k) = \sum_{n=0}^{\infty} \frac{1}{2}\gamma^n \left(X^T(k+n)Q_w X(k+n) + R_w u(k+n)^2\right)\Delta t. \quad (19)$$

Then the utility function can be obtained as

$$U(k) = \frac{1}{2}\left(X^T(k)Q_w X(k) + R_w u(k)^2\right)\Delta t. \quad (20)$$

4.2 Simulation Parameters

The specifications of the vibration isolation system are given below:

$$M_s = 320\ \text{kg}, \quad M_u = 49\ \text{kg}, \quad C_s = 1286\ \text{N}\cdot\text{s/m},$$
$$K_t = 190000\ \text{N/m}, \quad K_s = 15300\ \text{N/m}.$$

The following design parameters were chosen:

$$Q_w = diag[1143, 10, 1200000, 100000], \quad R_w = 10^{-5},$$
$$tolc = 10^{-3}, \quad \text{the sampling time} \quad \Delta t = 10^{-3}s.$$

In synthesis of the SNAC vibration controller, the off-line training and online training methods were adopted respectively. In the off-line training procedure, one neural network which had a 4-6-4 structure was used as the critic network. In this network, linear functions for both the input and hidden layers and the output layer served as the activation functions. The learning rate started from 0.9 and then decreased as learning proceeded until it reached 0.005 and stayed at 0.005 thereafter during the current training episode. The domain of states $S = \{X: \parallel X \parallel_\infty \leq 1\}$ was used, from which 500 points were randomly selected for training the critic network in each episode.

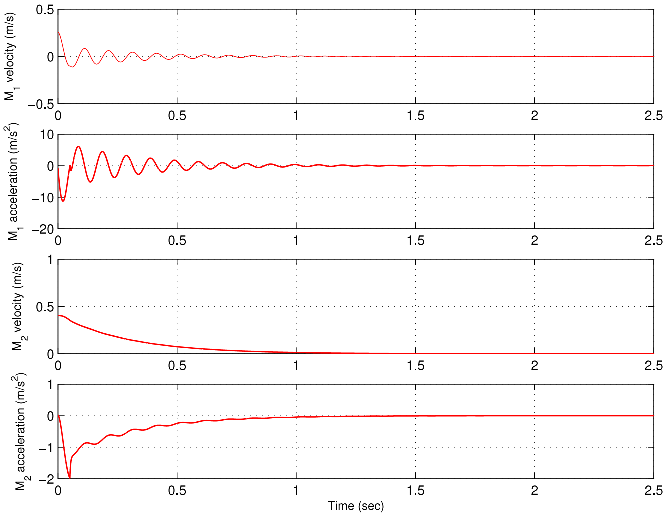

Fig. 4. State trajectories of the active vibration isolation system during the online training procedure

In the online training procedure, the structure of the critic networks was same with the one adopted in off-line training. One set of initial conditions used was $X(0) = [0, 0, 0, 0]^T$.

Since the weights of the critic networks are tuned on-line and the control signals are obtained based on the output values of the critic networks, randomly initialization of the critic networks may lead the system to be unstable. Thus it is necessary to adopt a stabilizing controller which generates the on-line training set for the critic networks in the first training episode. In this paper, a state feedback controller was used to guarantee the system to be stable at the beginning. The state feedback gain matrix $K = [1, 1, 1, 1]$. Then the controller $u = -KX$ was used in the first training episode. In this paper, each episode continued 500 time steps.

Fig. 4 shows the trajectories of the states during the online training procedure.

4.3 Simulation Results

In this paper, the acceleration of the payload were selected to assess the vibration isolation performance of the SNAC controllers while the disturbance signal z_0 was exerted to the system.

In the first numerical experiment, for generating a periodic disturbance, the displacement of the base z_0 was set as

$$z_0(t) = 0.1\sin(2\pi f t)$$

where f was the disturbance frequency and the sampling time was also 10^{-3}s.

At disturbance frequencies from 1 Hz to 25 Hz, the simulation frequency responses of the passive system and the active system are shown in Fig. 5. On the whole, the SNAC controllers alleviate the disturbance more effectively than the passive isolator. And at the aspect of the magnitude responses, they approximate the optimal controller designed based on the continuous model using the linear quadratic regulator (LQR) method.

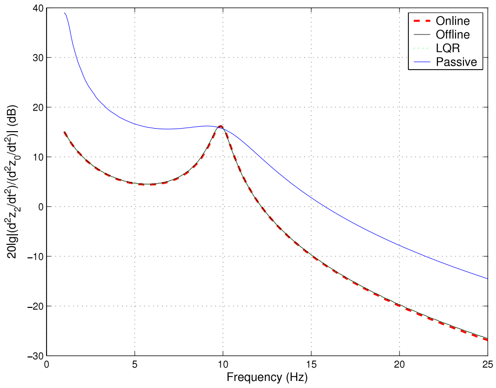

Fig. 5. Frequency response diagram of the vibration isolator, between payload acceleration $\ddot{z}_2$ and base acceleration $\ddot{z}_0$.

Another simulation was carried out at the following bump-like disturbance set $z_0(t)$

$$z_0(t) = \begin{cases} a(1-\cos(8\pi t))/2, & \text{if } 0.50 \leq t \leq 0.75 \text{ and } 1.25 \leq t \leq 1.50 \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

where a=0.025m was the bump amplitude.

Fig. 6 illustrates the dynamic performance of the SNAC controllers, which have similar dynamics with the LQR controller. For this bump-like disturbance, the SNAC controllers have apparently better robustness and have better control characteristics such as lower overshoot, shorter settling time than the passive isolator.

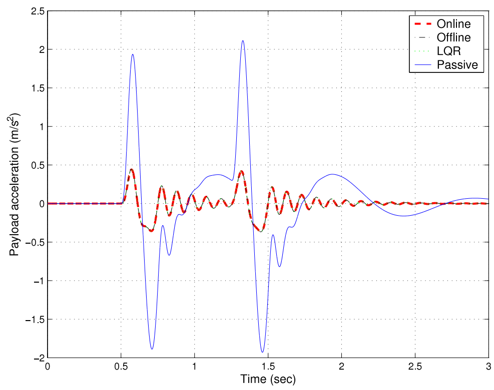

Fig. 6. Acceleration of mass M_s with a bump-like disturbance input.

5. CONCLUSION

In this paper, the SNAC approach was applied to the vibration isolation problem. The two training methods, i.e. off-line training and online training, were presented to adapt the SNAC optimal controllers respectively. Simulation results have shown that the designed SNAC controllers approximate the optimal controller designed in continuous-time domain using the LQR method. Hence, the effectiveness of the training methods presented in this paper can be verified. In addition, in comparison with the passive isolation system, the designed SNAC controllers can alleviate vibration disturbances more effectively and have better control performance. In conclusion, the SNAC approach can be considered effective in handling the vibration isolation problem under the assumption that the system model is known. More work will be pursued to design more robust and adaptive SNAC controller for the practical vibration isolation system with uncertainties by adopting the on-line training method discussed in this paper.

REFERENCES

N. Al-Holou, T. Lahdhiri, D.S. Joo, J. Weaver, and F. Al-Abbas. Sliding mode neural network inference fuzzy logic control for active suspension systems. *IEEE Transactions on Fuzzy Systems*, 10(2):234–246, 2002.

E.H. Anderson, J.P. Fumo, R.S. Erwin, C.S.A.E. Inc, and M.V. CA. Satellite ultraquiet isolation technology experiment (SUITE). In *Proceedings of the IEEE Aerospace Conference*, volume 4, pages 299–313, 2000.

A.G. Barto, R.S. Sutton, and C.W. Anderson. Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics*, 13(5):835–846, 1983.

Yi-De Chen, Chyun-Chau Fuh, and Pi-Cheng Tung. Application of voice coil motors in active dynamic vibration absorbers. *IEEE Transactions on Magnetics*, 41(3): 1149–1154, 2005.

S. Gaurav and S.N. Balakrishnan. Adaptive critic based neurocontroller for autolanding of aircrafts. In *Proceedings of the American Control Conference*, volume 2, pages 1081–1085, 1997.

G.G. Lendaris and C. Paintz. Training strategies for critic and action neural networks in dualheuristic programming method. In *Proceedings of International Conference on Neural Networks*, volume 2, pages 712–717, 1997.

Derong Liu, Xiaoxu Xiong, and Yi Zhang. Action-dependent adaptive critic designs. In *Proceedings of International Joint Conference on Neural Networks*, volume 2, pages 990–995, 2001.

L.R. Miller, M. Ahmadian, C.M. Nobles, and D.A. Swanson. Modelling and performance of an experimental active vibration isolator. *Journal of vibration and acoustics*, 117(3):272–278, 1995.

R. Padhi, N. Unnikrishnan, and S.N. Balakrishnan. Optimal Control Synthesis of a Class of Nonlinear Systems Using Single Network Adaptive Critics. In *Proceedings of the American Control Conference*, volume 2, pages 1592–1597, 2004.

D.V. Prokhorov and D. Wunsch. Adaptive critic designs. *IEEE Transactions on Neural Networks*, 8(5):997–1007, 1997.

J. Si and Yu-Tsung Wang. Online learning control by association and reinforcement. *IEEE Transactions on Neural Networks*, 12(2):264–276, 2001.

J. Spanos, Z. Rahman, and G. Blackwood. A soft 6-axis active vibration isolator. In *Proceedings of the American Control Conference*, volume 1, pages 412–416, 1995.

P.J. Werbos. Advanced forecasting methods for global crisis warning and models of intelligence. *General Systems Yearbook*, 22:25–38, 1977.

P.J. Werbos. Approximate dynamic programming for real-time control and neural modeling. *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*, 15:493–525, 1992.