Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Supplementary Damping Controller Design using Direct Heuristic Dynamic Programming in Complex Power Systems

Chao Lu*, Jennie Si**

**Department of Electrical Engineering, Tsinghua University, Beijing, 100084 China (e-mail: luchao@tsinghua.edu.cn)*
*** Department of Electrical Engineering, Arizona State University, Tempe, AZ 85048 USA (e-mail: si@asu.edu)}*

Abstract: In modern, large scale interconnected power grids, low-frequency oscillation is a key roadblock to improved power transmission capacity. Supplementary generator control, flexible AC transmission system (FACTS), and high voltage direct currents (HVDC) are engineered devices designed to damp such low frequency swings. In this paper a neural network-based approximate dynamic programming method, namely direct heuristic dynamic programming (direct HDP), is applied to power system stability enhancement. Direct HDP is a learning and approximation based approach to addressing nonlinear system control problems under uncertainty, and it is also a model-free design strategy. The action and critic networks of the direct HDP are implemented using multi-layer perceptrons; learning is carried out based on the interactions between the controller and the power system. For this design approach, real time system responses are provided through wide-area measurement system (WAMS). The controller learning objective is formulated as a reward function that reflects global characteristics of the power system under low frequency oscillation, as well as tight coupling effects among system components. Direct HDP control design is illustrated by case studies, which are also used to demonstrate the learning control performance. The proposed direct HDP learning control is also developed as a new solution to a large scale system coordination problem by using the China Southern Power Grid as a major test bed.

1. INTRODUCTION

Power system is a complex network composed of diverse components. The most important task of power system control is to maintain stable operation of the system. Exciter and governor control of generators are traditional methods. High voltage direct current (HVDC) and flexible AC transmission system (FACTS) equipment are modern control devices that are frequently adopted in recent years. However, system nonlinearity and uncertainty, plus coordination of multiple controllers make up the three major challenges for a controller design in a complex system of immense scale.

Many nonlinear phenomena are apparent in a power system including dead zone and control limits, to name a few. In addition, some special characteristics of nonlinear systems, bifurcation and chaos for example, have also been observed in power systems (Cutsem & Vournas, 1998). Even more challenging is the fact that many power system nonlinearities are difficult to model mathematically.

For a nonlinear control system design, two approaches are usually adopted: linearization of a system model and energy function based design. The former includes linearization around the operating point (Yu, Vongsuriya, et al, 1970) and exact feedback linearization (EFL) (Lu, Sun et al, 2001), then linear design principles can be applied, however, EFL depends greatly on the accuracy of the system model. In the second approach, energy storage functions or Hamiltonian functions are constructed based on the dissipation or Lyapunov stability theory to obtain the nonlinear control law (Webster, 1999), but the energy function cannot be easily constructed such that only simplified power system elements can be taken into consideration.

As regional grids continue getting merged into larger networks, the challenges arising from system uncertainty have become even greater. The models and parameters deviate from reality due to simplifications in modelling and their variations with time (Kosterev, Taylor, et al, 1999). The power system operating conditions vary constantly, but the controller design is only based on typical operating conditions. Robust control can guarantee the stability of linear time-invariant (LTI) system (Ohtsuka, Taniguchi, et al, 1992) if the uncertainties can be modelled correctly. Adaptive control is another approach to address the uncertainty problem; it can also be applied to power system stability control (Pierre, 1987).

Most classic control designs are based on mathematical models, but the limitations of these models seriously affect the practical control performance. Several approximate dynamic programming (ADP) methods (Si, Barto, et al, 2004) are proposed based on real system responses, which reflect important system dynamics, and using these methods, the constraints of nonlinearities and uncertainties can be

The research was supported by National Natural Science Foundation of China under project numbers 50595413 and Tsinghua Basic Research Foundation. The second author's research was supported by NSF under grants ECS-0002098.

potentially handled through learning. The approximation is used to solve the curse of dimensionality of dynamic programming. It is implemented using the estimation of cost function instead of computing the exact performances of all combinations of states and controls. The controller calculations through interactions with system also reduce the complexities introduced by the system model. Generator control using dual heuristic programming (DHP) was achieved via computer simulations and physical experiments (Venayagamoorthy, Harley, et al, 2003). In this approach, a system model implemented using a neural network was pretrained to predict system responses in the next time step, and the training quality of the model have great impact on the performance of the controller. In this paper, we consider a model-independent ADP approach (Si & Wang, 2001), which can be viewed as a model-free version of the action-dependent heuristic dynamic programming in adaptive critic designs (ACD); this will be referred to as direct HDP in the followings. In (Enns & Si, 2003), direct HDP was employed to control the flight of an Apache helicopter—a complex, continuous state/control, MIMO nonlinear system with uncertainty. The learning performance of the direct HDP application in multi-machine power systems is studied in (Lu, Si, et al, 2004). These results suggest the potential of adaptive critic designs for scalable complex system control applications.

In this paper, the direct HDP method is employed to solve the low-frequency oscillation problem in multi-machine power system. The swing period is usually greater than 1 second; therefore, the controller has sufficient time to adapt.

2. FRAMEWORK OF DIRECT HDP CONTROL

The basic control framework of ADP is shown in Fig.1, where u is the control signal, and X is the state vector. The effect of u is evaluated by using a cost function and is used to update the control policy. The system/envirionment is usually represented by a set of differential equations. Under direct HDP, a time-domain computer simulation, or the real dynamic system itself, with realistic nonlinearities and uncertainties are used in learning control. Controller parameter updates are based on real-time system responses. The controller performance is evaluated by the cost function, which can be used to reflect the global system dynamics and coordination among different controllers.

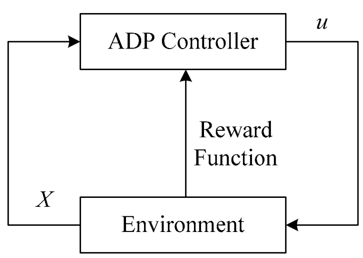

Fig. 1. Schematic diagram of ADP control

Direct HDP control comprises two main parts: an action network and a critic network. The former produces control signals according to a learning policy represented by the approximating network, while the latter approximates the cost/reward function J by minimizing the Bellman error. These two parts are usually implemented by neural networks because of their universal approximation capability and the associated simple back-propagation learning algorithm.

In power systems, real-time system dynamics fed back to the direct HDP controller are provided by a wide-area measurement system (WAMS) as illustrated in Fig. 2, which is based on synchronized phasor measurement techniques and modern digital communication networks. Through this system, some remote key system variables can be gathered and used as feedbacks to improve system performance (Chaudhuri, Majumder, et al, 2004).

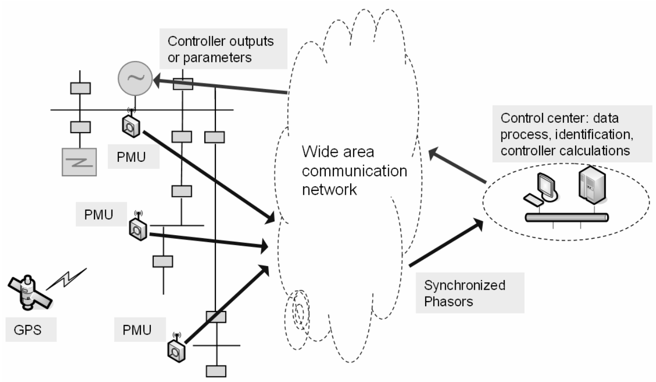

Fig. 2. Framework of online coordinated control based on the direct HDP method and WAMS

The signals that reflect system dynamics can be collected by phasor measurement units (PMU), tagged with an exact time stamp obtained from global positioning systems (GPS), and then transmitted to the WAMS center through a wide area communication network. After data pre-processing, the information can be used as inputs to online direct HDP design programs. During each time step, new control signals or parameters for different controllers are produced through the interaction between direct HDP and the environment (power system). Since cost functions are used in controller design, the resulting controllers are coordinated.

Fig.3 shows the schematic representation of direct HDP control (Si & Wang, 2001). The reinforcement signal *r(t)* is obtained from the external environment.

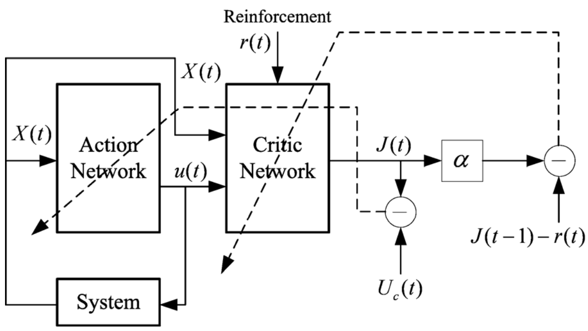

Fig. 3. Schematic diagram for implementation of direct HDP. The solid lines represent signal flows, and the dashed lines represent the paths for parameter tuning

During online learning, the controller is "naive" when it starts to control, that is, both the action and critic neural networks are randomly initialized for their weights. Once a system state is observed, an action will be subsequently produced based on the parameters in the action network. A "better" control output under the specific system state is rendered as a result of optimizing the principle of optimality. In the action network, this set of system operations is reinforced and the control value is adjusted by tuning the weights.

The output of the critic network (the J function) approximates the discounted total reward-to-go. Specifically, it approximates *R(t)* given by

$$R(t) = \sum_{k=1}^{\infty} \alpha^{k-1} r(t+k) \tag{1}$$

where *R(t)* is the future accumulative reward-to-go value at time t and α is a discount factor for the infinite-horizon problem ($0 < \alpha < 1$).

The critic network is trained to approximate the "value function" *J(t)* by minimizing the objective function, which represents the balance of the principle of optimality, as follows

$$E_c(t) = \{\alpha J(t) - [J(t-1) - r(t)]\}^2 / 2 \tag{2}$$

The principle behind adapting the action network is to backpropagate the error between the desired ultimate objective, denoted by U_c, and the cost function *R(t)*.

The weight updated in the action network adjusts its weights to minimize the following objective function

$$E_a(t) = [J(t) - U_c(t)]^2 / 2 \tag{3}$$

3. LEARNING AND ADAPTATION ABILITY OF DIRECT HDP CONTROLLER

3.1 A Two-area System and SVC Supplementary Controllers

The learning ability of the direct HDP controller is demonstrated in a 2-area system. Fig. 4 is the system diagram (Kundur, 1994) including four generators located in two areas, among which are two parallel tie lines from bus 7 to bus 9.

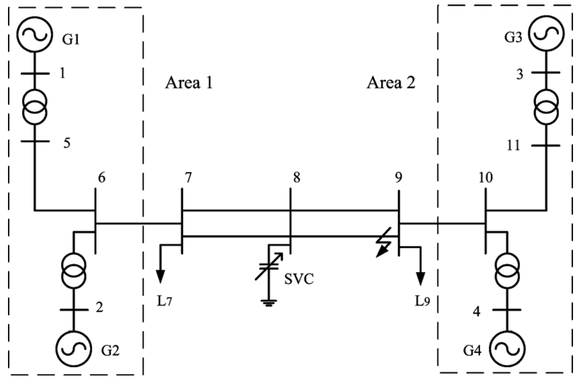

Fig. 4. Single line diagram of the two-area test system

An SVC is placed in the middle of the tie lines to support voltage and suppress oscillations. Traditional proportional-integral (PI) method is used for regulation, but this method by itself cannot guarantee the stability after severe faults. Therefore a supplementary controller is needed. In (Kundur, 1994), the design is based on the conventional pole-placement method (C1), which is shown in Fig. 5.

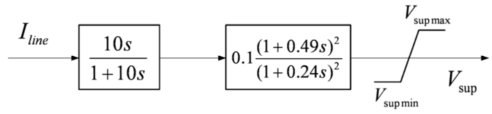

Fig. 5. Block diagram of the supplementary controller C1

In Fig. 4, V_{supmax}=－V_{supmin}=0.1, and V_{sup} represents the output. The input signal I_{line} is the magnitude of the current in the transmission line between buses 9 and 10, and it is chosen as input based on the system's observability.

In order to compare control performance of the direct HDP controller and that of C1, the same input signal I_{line} is used. The phase adjustment of the supplementary control input signal is critical with regard to the oscillation damping problem, and it is implemented through the two phase-shift blocks in C1. However, if only one variable is inputted, the ordinary MLP neural network is not capable of changing the phase to what we want. An additional input—the differential of I_{line}—is necessary, and an approximate differentiator as that given in (4) is employed to reduce noise amplification.

$$\frac{1}{(\tau_2 - \tau_1)} * \left(\frac{1}{\tau_1 s + 1} - \frac{1}{\tau_2 s + 1}\right) \qquad 0 < \tau_1 < \tau_2 \tag{4}$$

The structure of the supplementary damping controller using the direct HDP method is shown in Fig.6.

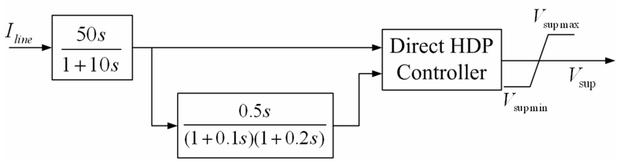

Fig. 6. Block diagram of supplementary control using the direct HDP method

3.2 Direct HDP Algorithm Implementation

In the direct HDP approach, the reward function is the only evaluation function on the task known by the controller, in this case, prior knowledge about the power system is applied explicitly to the reward function. For the 2-area system, the reinforcement signal *r(t)* is given as follows:

$$r(t) = -(b_1 \Delta\omega_{\text{inter-area}}^2 + b_2 \Delta\omega_{local1}^2 + b_3 \Delta\omega_{local2}^2) \tag{5}$$

$\Delta\omega_{inter-area}$, $\Delta\omega_{local1}$ and $\Delta\omega_{local2}$ are rotor speed deviations corresponding to different oscillation modes. By adjusting the weights b_i (i=1, 2, 3), the most possible destabilizing oscillation mode is chosen to be suppressed first. During normal operations, all generators are synchronized and the three delta omega variables are zero. Once the system undergoes a disturbance, *r(t)* may be viewed as an index of the kinetic energy of the entire system oscillation, which reflects how far the current system is away from the equilibrium point. In this paper, SVC is used to damp the inter-area power swing, as such, b_1=0.8, and b_2=b_3=0.1.

In the direct HDP controller shown in Fig. 3, there is one hidden layer in the action or critic neural network. The number of neurons in the input layer varies according to the input signals. There are 6 neurons in the hidden layer and 1 neuron in the output layer. The controller can be easily expanded to multi-input multi-output cases, only the neuron numbers in the input and output layers need to be altered. The update algorithm uses gradient descent, and learning of the neural networks is carried out using a Task-Run-Trial setup (Lu, Si, et al, 2007).

The system stability is a prerequisite when used in all simulation runs because instability results in system states with abnormal values and the neural network learning in this situation is hazardous. The instability criterion is placed on the generator angle difference, which is impossible to exceed 180 degrees in a real power system.

3.3 Simulation Results

Two supplementary damping control loops are considered in the study: (1) C1－PSS-type controller proposed by Kundur; (2) C2－direct HDP control with the tie line current I_{line} as the input.

The basic scheme used in the simulation is the same as in (Kundur, 1994), i.e., the same load flow calculation results and the same disturbance in which a three-phase short circuit occurs near bus 9 and is cleared by tripping the line between buses 8 and 9 after 74ms. This case is also used to train the direct HDP controller from randomly initialized action and critic weights. Fig. 7 is an example of two trials in the learning process of the C2 controller. Without optimizing the direct HDP algorithm, the calculation time of one trial is about 5 seconds, most of which are consumed by the power system simulation.

In trial 1, since initialization was random, controller parameters during the first several seconds were not in accordance with the desired actions. The system lost stability after 4 seconds, and the simulation was aborted because the failure criterion was reached. No proper mapping relationship between the inputs and the outputs can be observed. But this trial, characterized by a failure, provided the direct HDP controller with much useful information about what state-action pairs may have brought instability and thus should be avoided. In trial 2, a refined and accurate control strategy is clearly observed as shown in Fig. 7(e). After serious faults, the output of the controller switches between the lower and upper bounds in order to use the full control power, much like a bang-bang controller. When oscillation reduces, so does the control magnitude. Continuous training based on trial 2 can increase the speed of controller response further. In Fig. 8, the performance of learned direct HDP controller is compared with the traditional one.

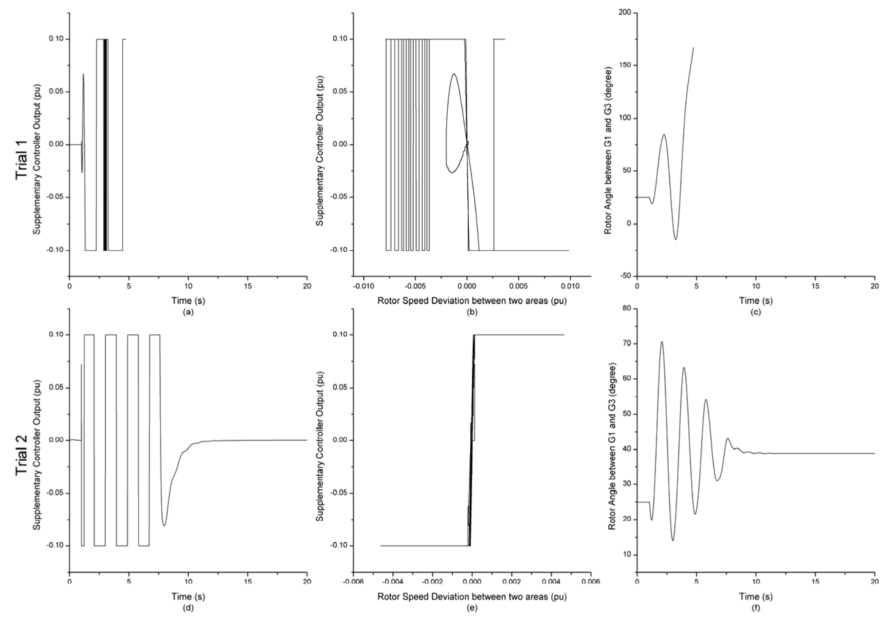

Fig. 7. Learning process of the direct HDP controller including two trials. (a), (d) output of the controller; (b), (e) mapping relationships between inputs and outputs; (c), (f) rotor angle between generators G1 and G3.

To evaluate controller performance under different conditions, the following case was simulated. The loads at bus 9 are reduced from 1769 MW to 1569 MW and correspondingly an active power of 200 MW injected into the network from generator 2 is reduced, and the line between buses 7 and 8 is out of service and no line tripping occurs after the instantaneous short circuit fault.

In this case, the system model, the load level, and the system structure have all been changed, and the controller designed on the basis of the basic case cannot guarantee system stability. Fig. 9 shows the learning process of the C2 controller to maintain system stability under new conditions.

In the beginning, the equivalent gains of the two current signals are positive, but after learning is completed, the gains are negative (Fig. 9 (a)). However, these are the correct controller parameters under new conditions. The adapted direct HDP controller damps the oscillations well (Fig. 9 (b)) compared to the unstable swings with fixed C1 controller.

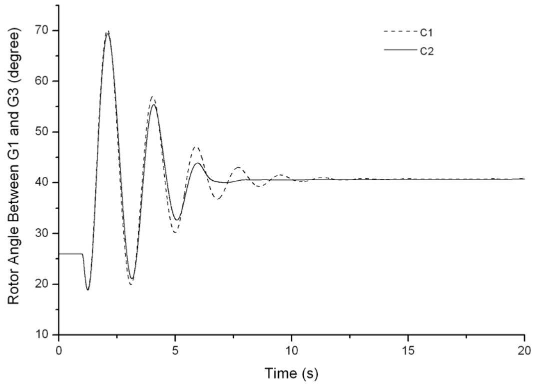

Fig. 8. Rotor angle between G1 and G3 after disturbances.

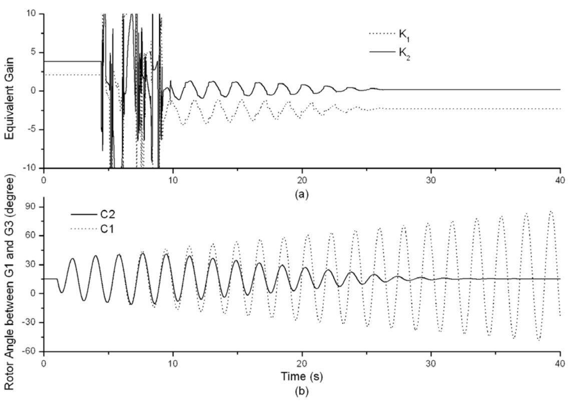

Fig. 9. Learning process when the system configurations change significantly.

4. DIRECT HDP COORDINATED CONTROL IN A COMPLEX SYSTEM

4.1 China Southern Power Grid (CSG)

CSG is an AC/DC hybrid power network, and mainly composed of four provincial grids. The distance of transmission from west to east is over 1000km. High capacity power transmission takes place through 5 AC lines and 2 DC links (Guiguang and Tianguang) in parallel in 2006. The grid structure is shown in Fig. 10. Upon completion of connecting regional grids, low frequency oscillation has become a prominent problem. Power modulation control of HVDC is an attractive approach to enhance the stability. The structure of modulation controller is as that shown in Fig. 5, and the input signals are chosen as the active power on the parallel AC lines (Tianshengqiao-Yulin, Nanning-Pingguo) based on the analyses of observability and dynamic relative gain array.

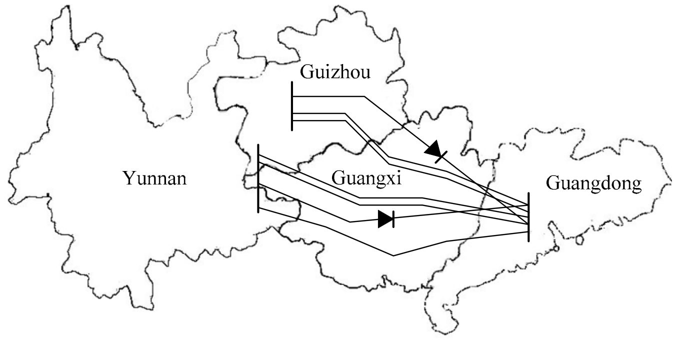

Fig. 10. Simplified Diagram of South China Grid structure

4.2 Independent Design of HVDC Modulation Controllers

Since the terminals of two DC links are not far from each other in terms of electricity distance, their controllers must be coordinated. In section III, the learning ability of direct HDP controller is validated, and in this section, its coordinating design considering practical engineering constraints will be demonstrated.

If only one DC power modulation controller is used, the parameters can be tuned to achieve desired control performance, however, if these independently designed controllers work together, the interaction between any two DC links deteriorates system damping capability, as shown in Fig. 11. In the multi-infeed dc system, the supplementary dc control must be designed on a coordinated basis to avoid unexpected excitation of new poorly damped mode (Pilotto, Szechtman, et al, 1995).

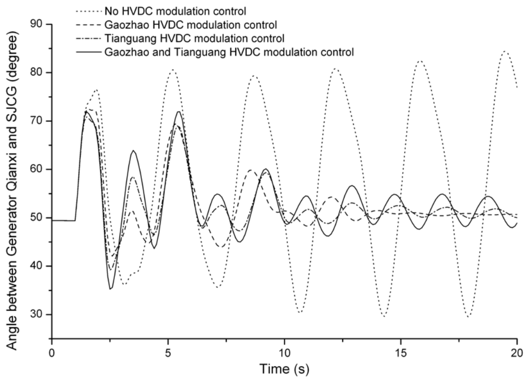

Fig. 11. Rotor angle between Generator Qianxi in Guizhou and SJCG in Guangdong.

4.3 Online Coordinated Control using Direct HDP

For direct HDP controller, the control law update is instructed by the reward function. Its definition is the weighted sum of squares of relative rotor speed differences among three approximate inertia centers: Guangdong, Guizhou and Yunan. This function reflects stability of the entire system, if only one oscillation mode is suppressed, the reward is not minimal, and the controller parameters will be adjusted continuously.

The basic algorithm of HVDC direct HDP supplementary damping controllers is not changed, however, the structure is improved for efficiency enhancement (Lu, Si, et al, 2006). Starting from the independently designed Guiguang and Tianguang DC power modulation controller parameters, the new direct HDP controller can obtain a set of coordinated values. Fig. 12 illustrates the learning process, where the variations are within acceptable bounds for real devices.

After coordination, the oscillation with several frequencies is damped well, as shown in Fig. 13.

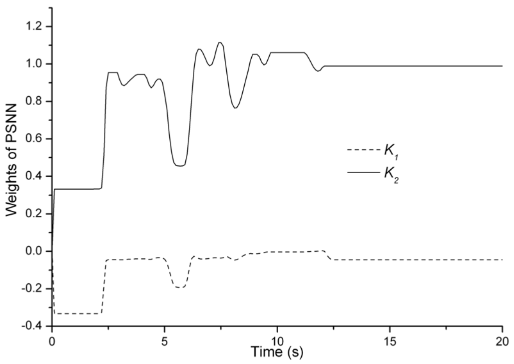

Fig. 12. Learning process of Guiguang modulation controller

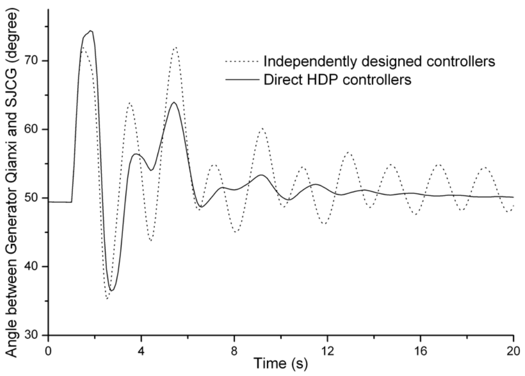

Fig. 13. Rotor angle between Generator Qianxi in Guizhou and SJCG in Guangdong.

5. CONCLUSIONS

The three major challenges arising from system nonlinearities, uncertainties, and lacking of coordinated design make the stability control of a large-scale power system difficult. In this paper, the direct HDP method is employed to damp low-frequency oscillations. Distributed controllers are chose to use one common cost function so that the design and adaptation can be easily coordinated. The performance of the direct HDP controller is validated in two systems. The first case is the four-machine two-area system, where the training that begins from random initialization and the adaptation under various operating conditions demonstrate good learning abilities of the direct HDP controller. Further, the coordinated design approach has been validated through two DC power modulation controllers in CSG, and the advanced performances compared with the traditional control approach are presented.

REFERENCES

Chaudhuri, B., R. Majumder, B. C. Pal. (2004). Wide-Area Measurement-Based Stabilizing Control of Power System Considering Signal Transmission Delay, *IEEE Trans. Power Systems*, **19(4)**, 1971-1979.

Cutsem, T.V., and C. Vournas. (1998). *Voltage stability of electric power systems*. Boston: Kluwer Academic Publishers.

Enns, R., J. Si. (2003). Helicopter Trimming and Tracking Control Using Direct Neural Dynamic Programming, *IEEE Trans. Neural Networks*, **14(4)**, 929-939.

Kosterev, D.N., C. W. Taylor, W. A. Mittelstadt. (1999). Model Validation for the August 10, 1996 WSCC system outage, *IEEE Trans. Power Systems*, **14(3)**, 967-979.

Kundur, P. (1994). *Power System Stability and Control*, New York: McGraw-Hill.

Lu, C., J. Si, X. Xie, et al. (2004). SVC supplementary damping control using direct neural dynamic programming. In *Proceedings of the 2004 IEEE International Symposium on Intelligent Control*, 270-274.

Lu, C., J. Si, X. Wu, et al. (2006). Approximate Dynamic Programming Coordinated Control in Multi-infeed HVDC Power System. In *Proceedings of 2006 Power Systems Conference and Exposition*, 2131-2135.

Lu, C., J. Si, X. Xie, et al. (2007). Convergence of Direct Heuristic Dynamic Programming in Power System Stability Control. In *Proceedings of 20th International Joint Conference of Neural Networks*.

Lu, Q., Y. Sun, S. Mei, (2001). Nonlinear *Control Systems and Power System Dynamics*, Boston: Kluwer Academic Publishers.

Ohtsuka, K., T. Taniguchi, T. Sato, et al. (1992). An H-infinity optimal theory-based control system. *IEEE Transactions on Energy Conversion*, **7(1)**, 108-115.

Pierre, D.A. (1987). A perspective on adaptive control of power systems. *IEEE Transacions on Power Systems*. **2(2)**, 387-396.

Pilotto, L.A.S., M. Szechtman, A. Wey, et al. (1995). Synchronizing and damping torque modulation controllers for multi-infeed HVDC systems, *IEEE Transactions on Power Delivery*, **10(3)**, 1505-1513

Si, J., A. Barto, W. Powell, D. Wunsch. (2004). *Learning and Approximate Dynamic Programming: Scaling up to the Real World*, New York: John Wiley & Sons.

Si, J., Y. T. Wang. (2001). Online Learning Control by Association and Reinforcement, *IEEE Trans. Neural Networks*, **12(2)**, 264-276.

Venayagamoorthy, G.K., R. G. Harley, D. C. Wunsch. (2003). Dual Heuristic Programming Excitation Neurocontrol for Generators in a Multimachine Power System, *IEEE Trans. Industry Applications*, **39(2)**, 382-294.

Webster, J.G. (1999). *Encyclopedia of electrical and electronics engineering*. New York: John Wiley.

Yu, Y.N., K. Vongsuriya, L. N. Wedman. (1970). Application of an optimal control theory to a power system. *IEEE Transactions on Power Apparatus and Systems*, **89(1)**, 52-62.