Combining a Stability and a Performance-Oriented Control in Power Systems

Mevludin Glavic, *Member, IEEE*, Damien Ernst, *Member, IEEE*, and Louis Wehenkel, *Member, IEEE*

***Abstract*—This paper suggests that the appropriate combination of a stability-oriented and a performance-oriented control technique is a promising way to implement advanced control schemes in power systems. The particular approach considered combines control Lyapunov functions (CLF) and reinforcement learning. The capabilities of the resulting controller are illustrated on a control problem involving a thyristor-controlled series capacitor (TCSC) device for damping oscillations in a four-machine power system.**

***Index Terms*—Control Lyapunov functions, power system stability control, reinforcement learning.**

I. Introduction

The concept of control Lyapunov functions (CLF) provides a powerful tool for studying stabilization problems [1]. Most nonlinear control design methods based on CLF provide strong guarantees of stability but do not directly address important issues of control performances.

Reinforcement Learning (RL) emerges as an attractive learning paradigm that offers a panel of methods that allow controllers to learn a goal-oriented control law from interactions with a system or its simulation model [2], [3] but do not provide stability guarantees when employed online.

In principle, a feedback control system should try to optimize some mix of stability and performance [4], [5], and this paper suggests a combination of RL and CLF as a way to implement an advanced control scheme for the system stability control. Results of controlling a thyristor-controlled series capacitor (TCSC) to damp power system oscillations are presented.

II. RL and CLF: A Short Description

RL is a general algorithmic approach to solve stochastic optimal control problems by trial and error. Dynamic programming is the underlying formal framework [2]. An RL agent (controller) interacts with its environment (system) as follows [2], [3]: At each discrete time step, the agent receives (through measurements) s_t, which is a representation of the system state, and the agent selects an action u_t from the set of actions available in the state. As a consequence of taking the action, the agent receives (through measurements) a numerical reward $r_t = r(s_t, u_t)$ that the agent tries to maximize (minimize) over time. In the usual setting, the discounted sum of rewards $\sum_{t=0}^{\infty} \gamma^t r(s_t, u_t)$ is to be maximized. A model-based RL method, known as prioritized sweeping [2], is used here.

In the concept of CLF, for a nonlinear system $\dot{x} = f(x, u)$, one selects a Lyapunov function candidate $V(x)$ and finds a stabilizing feedback control law $u(x)$ that renders $\dot{V}(x)$ negative definite. In short, the CLF approach, assuming a $V(x)$ has been found, allows the search for stabilizing inputs. The results from [6] on application of a CLF concept to control a TCSC device are largely followed.

Manuscript received July 2, 2004. Paper no. PESL-00143-2003.

The authors are with the Electrical Engineering and Computer Science Department, University of Liège, 4000 Liège, Belgium (e-mail: glavic@montefiore.ulg.ac.be; ernst@montefiore.ulg.ac.be; lwh@montefiore.ulg.ac.be).

Digital Object Identifier 10.1109/TPWRS.2004.841146

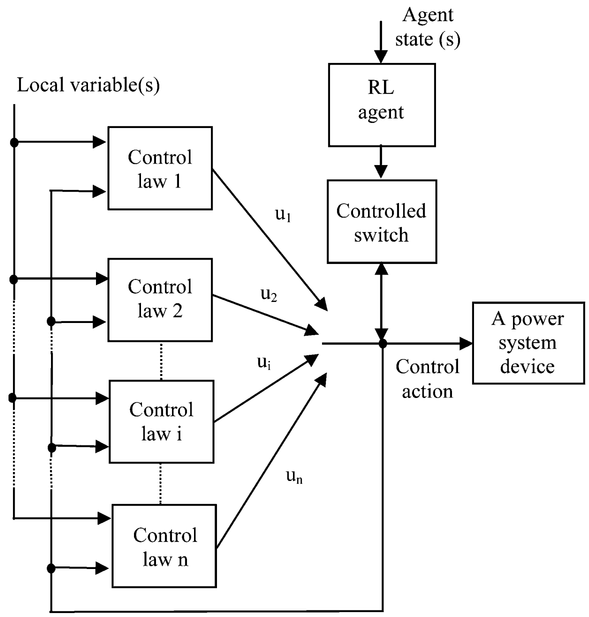

Fig. 1. Conceptual diagram of the proposed control.

III. Proposed Control Scheme

The learning by trial and error is not justified when one intends to apply it online and when the primary concern is the system stability. In this case, the "exploration-stability" tradeoff is to be resolved. On the other hand, in the CLF concept, assuming a Lyapunov function candidate has been found, there are many control laws that render its derivative negative definite, and the problem is to choose the most appropriate one. This paper introduces a control scheme that tries to make use of the advantages of the two control techniques and, at the same time, to overcome the mentioned difficulties in applying them alone (assure safe learning in online mode by limiting available control actions to be stabilizing ones and, at the same time, to improve over stabilizing controls by the minimizing prespecified cost).

The idea is to employ a RL method (in online mode) in order to compute an approximation of the optimal sequence comprising basic control laws derived from the concept of CLF (for stability guarantees of the sequence of stabilizing control laws, see [7]), as illustrated in Fig. 1. For the particular case considered, each individual control law is derived as a stabilizing continuous control law that renders a common (global) Lyapunov function candidate decreasing. The aim of having the active control signal as an input to each of the control laws is to assure a "jumpless" switch from one control law to another. Furthermore, the control scheme as a whole is assumed to rely on strictly locally measurable information. Although only the combination of the RL and CLF concept is considered in this short note, the proposed control scheme is not limited to these control techniques. Any control with stability guarantees can be combined with the RL and any heuristic search technique can be combined with the CLF control in a similar way.

IV. Test Results

To illustrate the capabilities of the proposed control scheme to control a TCSC aimed to improve damping of power system oscillations, a four-machine power system model, as described in [8], is used. For the simulation purposes, four linear, continuous with "hard" limits, basic control laws [6]

$$u = X_{\mathrm{ref}} = F(V_{lm}\dot{V}_{lm}) = k \cdot (V_{lm}\dot{V}_{lm}) \tag{1}$$

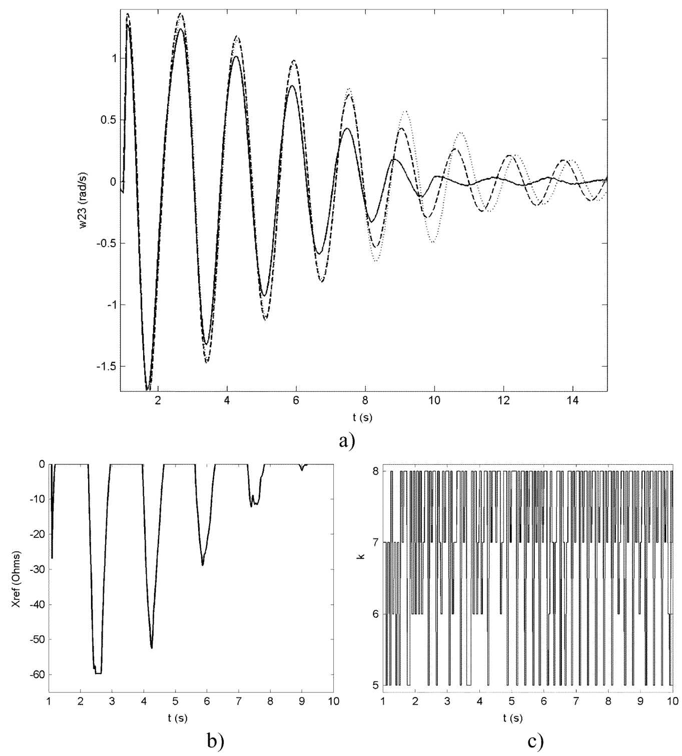

Fig. 2. Controlled system response. (a) $\boldsymbol{k} = \mathbf{5.0}$ (dotted), $\boldsymbol{k} = \mathbf{8.0}$ (dashed), after 300 learning scenarios (solid). (b) Control efforts. (c) Sequence of $\boldsymbol{k}$.

are used. V_{lm} is the locally inferred magnitude of the voltage drop across the line l-m, where the TCSC is installed [6] and $X_{\mathrm{ref}}([-61.57, 0]\ \Omega)$ is the FACTS reactance reference. $\dot{V}_{lm}$ can be inferred based upon tracking of V_{lm}. The values of the parameter considered are $k = (5.0, 6.0, 7.0, 8.0)$ (as in [6]). In this setting, the sequence of parameter k is to be learned.

The active power flow through the line l-m is chosen as the measurement. A pseudo state (due to partial observability), used inside the RL algorithm, is defined from the history of the measurements and actions taken

$$s_t = (P_{et}, P_{et-1}, P_{et-2}, k_{t-1}, k_{t-2}). \tag{2}$$

The aim of the control is to maximize damping of the electrical power oscillations in the line. In this respect, the reward at time t is defined by

$$r_t = \left\{ -\frac{|P_{et} - \bar{P}_e|}{10^6} - c \cdot u_t \right. \tag{3}$$

where $\bar{P}_e$ represents an online estimate of the steady-state value of the electric power transmitted through the line

$$\bar{P}_e = \frac{1}{1200} \sum_{i=0}^{1199} P_{et+1-i} \tag{4}$$

which is a moving average over the last $1200 * 50\ \mathrm{ms} = 60$ s, and parameter c penalizes the value of the control.

The system becomes unstable through growing interarea oscillations of about 0.72 Hz when subjected to the 100-ms three-phase self-cleared short circuit on one of the two lines. The learning period is partitioned into different scenarios. Each scenario starts with the power system being at rest and is such that at 1 s, the short circuit occurs. The simulation then proceeds in the post-fault configuration until t is greater than 60 s. The electrical power transmitted in the line is discretized in 100 values within interval $[-250, 250]$ MW, and parameters are set to $\gamma = 0.98$ and $c = 0.5$. The controlled system responses when subjected to the fault with the TCSC controlled by CLF-based control for $k = 5.0$ and $k = 8.0$ are given in Fig. 2(a) (in terms of a relative generator speed). In the same figure, the response of the system being controlled by the proposed control scheme after 300 learning scenarios is presented. Note that the proposed control scheme improves the oscillations damping in initial as well as subsequent swings, and the settling time is quite smaller than in the case of standard CLF-based controls. The "jumpless" nonexcessive control law is achieved [see Fig. 2(b)]. The sequence of the values of k is illustrated in Fig. 2(c).

V. Conclusion

The initial results of our research efforts in combining stability-oriented and performance-oriented control techniques to solve power system control problems are presented in this paper. It is demonstrated by the combination of the RL algorithm and the concept of the CLF applied to the problem of power system oscillations damping.

References

[1] P. Kokotovic and M. Arcak, "Constructive control: A historical perspective," *Automatica*, vol. 37, pp. 637–662, 2001.
[2] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.
[3] D. Ernst, M. Glavic, and L. Wehenkel, "Power system stability control: Reinforcement learning framework," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 427–436, Feb. 2004.
[4] P. J. Werbos. (1998) Stable Adaptive Control Using New Critic Designs. [Online]. Available: http://arxiv.org/html/adap-org/9 810 001
[5] G. K. Venayagamoorthy, R. G. Harley, and D. C. Wunsch, "Implementation of adaptive critic-based neurocontrollers for turbogenerators in a multimachine power system," *IEEE Trans. Neural Netw.*, vol. 14, no. 5, pp. 1047–1064, Sep. 2003.
[6] M. Ghandhari, G. Andersson, and I. A. Hiskens, "Control Lyapunov functions for controllable series devices," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 689–694, Nov. 2001.
[7] T. J. Perkins and A. G. Barto, "Lyapunov design for safe reinforcement learning," *Mach. Learning Res.*, vol. 3, pp. 803–832, 2002.
[8] P. Kundur, *Power System Stability and Control*. New York: McGraw-Hill, 1994.