Data-Driven Coordinated Control of AVR and PSS in Power Systems: A Deep Reinforcement Learning Method

Oshnoei, Arman; Sadeghian , Omid ; Mohammadi-Ivatloo, Behnam ; Blaabjerg, Frede; Anvari-Moghaddam, Amjad

Published in:
2021 IEEE International Conference on Environment and Electrical Engineering

DOI (link to publication from Publisher):
10.1109/EEEIC/ICPSEurope51590.2021.9584640

Publication date:
2021

Document Version
Accepted author manuscript, peer reviewed version

Data-Driven Coordinated Control of AVR and PSS in Power Systems: A Deep Reinforcement Learning Method

Arman Oshnoei
Department of Energy
Aalborg University
Aalborg, Denmark
arman_oshnoei@yahoo.com

Omid Sadeghian
Faculty of Electrical and Computer Engineering
University of Tabriz
Tabriz, Iran
omidsadeghian@tabrizu.ac.ir

Behnam Mohammadi-Ivatloo
Faculty of Electrical and Computer Engineering
University of Tabriz
Tabriz, Iran
bmohammadi@tabrizu.ac.ir

Frede Blaabjerg
Department of Energy
Aalborg University
Aalborg, Denmark
fbl@energy.aau.dk

Amjad Anvari-Moghaddam
Department of Energy
Aalborg University
Aalborg, Denmark
aam@energy.aau.dk

***Abstract*— In this paper, a strategy based on deep reinforcement learning (DRL) as an intelligent coordinator for power system stabilizer (PSS) and automatic voltage regulator (AVR) in a two-are power grid is proposed. The proposed coordinator is developed to provide accurate online modification of the gains appearing in the structure of PSS and AVR which avoids unfavorable interactions between PSS and AVR under significant changes in the working point and thereby guaranteeing the stability of the power grid. A Markov decision manner is used to formulate the DRL problem and it is solved through a deep deterministic policy gradient approach with an actor-critic framework. Since the intelligent coordinator relies on the expert's science, some scaling coefficients are added to the coordinator body to achieve optimal performance. To confirm the effectiveness of the presented DRL approach, the design is conducted on Kundur's power grid. Simulations illustrate that the proposed DRL-based control can confirm the stability of the system and attain desired dynamic responses.**

Keywords—power system stability, deep reinforcement learning, coordinated control, interconnected power system

NOMENCLATURE

Power system terms

ω_i	Rotor speed.
ω_0	Synchronous speed.
$\Delta\omega_i$	Rotor speed change.
V_{Ti}	Terminal voltage change.
V_{refi}	Reference voltage of excitation system.
V_{ri}	Voltage measurement output.
V_{mi}	AVR's output.
V_{pi}	PSS's output.
V_{ni}	Stabilizing feedback loop.
E_{fqi}	Field voltage of excitation system.
$T_{ri}, T_{Ei}, T_{ni}, T_{ai}$	Time constants associated with excitation control.
K_{ai}	AVR's gain.
K_{fi}	Excitation' gain.
y_{1i}	Output of PSS's measurement block.
y_{2i}	Output of PSS's washout block.
y_{3i}	Output of PSS's lead-lag compensation block.
T_{ki}	PSS's time constants (k = 1, 2, 3, 4, 5, 6)
K_{pi}	PSS's gain.
i	Number of generators.

DRL coordinator terms

R_{AVR}	Scaling coefficient of AVR.
R_{PSS}	Scaling coefficient of PSS.
R_p	Positive reward.
R_n	Negative reward.
n	Number of iterations.
α	Learning rate.
γ	Discount rate.
r_c	Decay rate.
ϑ	Random noise.
a	Control action.
s	State.
t	Time step.

I. INTRODUCTION

Stability of a power grid is described as the capability of a grid to get back operating balance after being experienced abnormal conditions. The abnormal conditions may be considered as a large change in load amount and/or in output of renewable energy sources, the sudden outage of a generator or intense faults on the tie-lines. Thus, the power systems shall be designed in such a way that any cascading blackout caused by the occurrences can be avoided. This made a motivation for power engineers to evaluate the power system stability based on various feasible operating conditions. In this regard, transient and small-signal stabilities under fault situations should be dealt with carefully in order to retain the synchronism between generators [1]–[2]. Hence, a highly effective excitation system is needed to be designed to improve performance goals such as steady-state errors and small-signal and transient stabilities. The excitation system of synchronous generators is constituted of two main controllers: power system stabilizer (PSS) and automatic voltage regulator (AVR). The first one gives impetus voltage adjustment and improves the stability in the grid under sudden intense perturbations [3]. The latter heightens the stability of the grid after being experienced small perturbations. [4]. The revealed investigations in the area of the excitation model can be split into pair main classes. The first one relies on the combination of PSS and AVR using a two-stage design manner for the nominal condition. In this way, in order to reach the predetermined voltage adjustment, firstly the AVR takes action and the PSS is then devised to raise the stability of the small signal. However, simultaneous improvement of voltage and stability controls is a demanding duty as PSS and AVR use a single control signal of excitation system [5]. Thus, in the second class of researches, an integrated method has been developed to design PSS and AVR. This kind of investigations claims that as power grids frequently encounter variations in operating points, a coordinated PSS-AVR design could make the electricity grid robust versus intense perturbations.

A great deal of investigations has been done to enhance the grid' stability by the development of AVR and PSS controllers. Authors of [6] have used an adaptive control method to optimize the AVR performance in an interconnected power system. In [7], an optimized PID controller is utilized for AVR system to improve the dynamic response. In [8], the design procedure of multiple PSSs is examined to reduce the amplitude of the fluctuations and increase the stability. In [9], a consecutive conic programming method is proposed for the design of coordinated PSSs. These studies, however, concentrate on the optimal design of one of the controllers and do not use PSS and AVR simultaneously. In terms of the coordination among PSS and AVR, the authors of [10,11] have provided the coordination by getting a fixed parameter for PSS and AVR using analytical and robust techniques. Fixed parameters of the PSS and AVR may degrade the grid stability in faulty conditions. Artificial intelligence methods such as neural network, brain emotional learning, and fuzzy logic are recognized as potential options for coordination development among PSS and AVR in power grids [12-13]. The important feature of these techniques is the independent-model structures that allow techniques to control the power grid's uncertainty, intricacy, and nonlinearity. However, the aforesaid intelligent methods are generally only appropriate for a certain cycle period as suffering from the absence of the ability to learn online. By the fast growth in the field of machine learning, data-driven approaches based on reinforcement learning (RL), have received large attention and have become a strong mechanism in the development of intelligent networks [14]. The main concept of the RL is to acquire a policy along with the states and actions while obtaining maximum rewards through interacting with an agent with an environment. RL methods have attained remarkable success in intricate problems by combining them with a deep neural network, entitled deep RL (DRL). Deep Q-learning (DQL), as the most well-liked DRL method, is capable to give a fast forecast of the Q-values corresponding to each state/action couple, which considerably decreases the computational complexity in the conventional Q-learning [15]. Due to these advantages, the DQL has been used in various practical problems such as stochastic power grids [16], induction motor [17], and robotic [18]. However, this method employs discrete steps to make an estimate of the value function, which restricts its utilization for continuous space-based problems. To address this challenge, in the problems with the multi-dimensional state variables, a deep deterministic policy gradient (DDPG) algorithm can be used. Up to now, little research has been introduced using the DRL strategies to ensure stability and voltage control in power grids.

In this paper, an DRL-based intelligent structure DRL features is proposed to provide a coordination among AVR and PSS to guarantee the transient and dynamic stabilities of an interconnected power system. The DRL problem is solved by the DDPG algorithm using an actor-critic framework, which could update the parameters of PSS and AVR by providing online optimization in the face of severe disturbances. For implementation, each individual coordinator only requires local information associated with the synchronous generator including terminal voltage and rotor speed. The scaling coefficients in the intelligent coordinator are considered to attain an optimal performance. Some dynamic signals such as terminal voltages, rotor speed, rotor angle and acceleration of generators are illustrated to compare and approve the ability of the proposed intelligent coordinator. Simulations demonstrate that the DRL-based intelligent coordinator can get favorable dynamic results against large disturbances.

This paper is organized as follows. Section II illustrates the mathematical model for PSS and AVR. Section III explains the DRL coordinator and states the corresponding methodology. Section IV provides numerical simulations and discussion. Finally, concluding observations are provided in the last section.

II. MTHEMATICAL MODEL OF PSS AND AVR

Each synchronous generator is equipped with three controllers in power grids, including AVR, PSS, and governor-turbine. AVR offers adjustability for the generator voltage to hold it in a constant amount. Furthermore, AVR offers a steady performance of the grid if it experiences intense perturbations. In this study, the model of AVR with type DC4B is considered [19]. The dynamic equations associated with IEEE-DC4B excitation are expressed in (1-4).

$$\dot{V}_{ri}(t) = \frac{1}{T_{ri}}\big(V_{Ti}(t) - V_{ri}(t)\big) \tag{1}$$

$$\dot{V}_{mi}(t) = \frac{1}{T_{ai}}\Big((V_{refi}(t) + V_{Pi}(t) - V_{ri}(t) - V_{fi}(t))K_{ai} - V_{mi}(t)\Big) \tag{2}$$

$$\dot{V}_{ni}(t) = \frac{K_{fi}}{T_{Ei}T_{ni}}\Big(V_{mi}(t) - (K_{Ei} + S_{ei})E_{fqi}(t)\Big) - \frac{1}{T_{ni}}V_{ni}(t) \tag{3}$$

$$\dot{E}_{fqi}(t) = \frac{1}{T_{Ei}}\Big(V_{mi}(t) - (K_{Ei} + S_{ei})E_{fqi}(t)\Big) \tag{4}$$

where subscript *i* refers to the generator number.

A PSS operates to generate a suitable torque on the generator's rotor. PSS is responsible to compensate for the phase lag between the exciter input and electrical torque. The block diagram associated with the dynamic models of PSS and AVR can be found in [13]. The PSS-PSS1A model used in this work is described as [20]. Figure 1 depicts the diagram of a synchronous generator together with excitation control. As the figure illustrates, the PSS output is entered to adjust the voltage field.

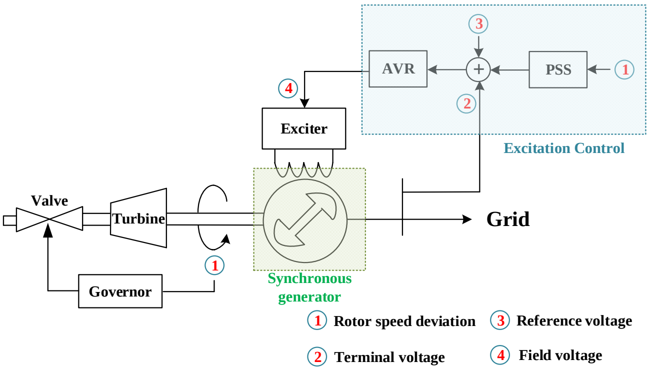


Fig. 1. Schematic veiw of a generator with excitation control.

$$\dot{y}_{1i}(t) = \frac{1}{T_{6pi}}\left(\omega_i(t) - \omega_0(t)\right) - \frac{1}{T_{6pi}} y_{1i}(t) \tag{5}$$

$$\dot{y}_{2i}(t) = \frac{K_{pi}}{T_{6pi}}\left(\omega_i(t) - \omega_0(t)\right) - \frac{K_{pi}}{T_{6pi}} y_{1i}(t) - \frac{1}{T_{5pi}} y_{2i}(t) \tag{6}$$

$$\dot{y}_{3i}(t) = \frac{K_{pi}T_{1pi}}{T_{2pi}T_{6pi}}\left(\omega_i(t) - \omega_0(t)\right) - \frac{K_{pi}T_{1pi}}{T_{2pi}T_{6pi}} y_{1i}(t) + \left(\frac{1}{T_{2pi}} - \frac{1}{T_{5pi}}\right) y_{2i}(t) - \frac{1}{T_{2pi}} y_{3i}(t) \tag{7}$$

$$\dot{V}_{pi}(t) = \frac{K_{pi}T_{1pi}T_{3pi}}{T_{2pi}T_{4pi}T_{6pi}}\left(\omega_i(t) - \omega_0(t)\right) - \frac{K_{pi}T_{1pi}T_{3pi}}{T_{2pi}T_{4pi}T_{6pi}} y_{1i}(t) + \frac{T_{3pi}}{T_{4pi}}\left(\frac{1}{T_{2pi}} - \frac{1}{T_{5pi}}\right) y_{2i}(t) + \left(\frac{1}{T_{4pi}} - \frac{T_{3pi}}{T_{2pi}T_{4pi}}\right) y_{3i}(t) - \frac{1}{T_{3pi}} V_{Pi}(t) \tag{8}$$

The generator is provided with a control command to convince the two conflicting control commands. In other words, the AVR and PSS raise the grid's stability including oscillation stability and terminal voltage control by a single control command. In general, AVR is provided with a high gain to give a quick reaction for raising stability indexes. In this case, the small-signal stability may be affected [21]. By contrast, although PSS increases the small-signal stability, the voltage regulation and oscillation stability improvement may be affected [22]. Thus, the AVR and PSS require coordinating their parameters for ensuring an appropriate performance in different operating points of the grid.

III. DRL BASED INTELLIGENT COORDINATOR

As mentioned in earlier section, developing a coordinator among AVR and PSS is required. For this end, in this paper, a DRL based intelligent control is developed as an intelligent coordinator among AVR and PSS to rehabilitate the deficiencies between them under substantial variations in the operating point of power systems. Figure 2 shows how the DRL coordinator acts to make coordination.

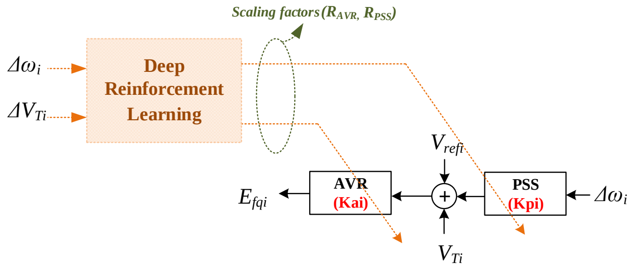


Fig. 2. Proposed coordinated control for AVR and PSS

As shown in Fig. 2 the coordinator is provided with two inputs including the terminal voltage (ΔV_{Ti}) and rotor speed ($\Delta\omega_i = \omega_i - \omega_0$) changes. The outputs are supplementary parameters to reform the constant parameters of the AVR and PSS against perturbations. Two scaling coefficients R_{AVR} and R_{PSS} are included as the outputs of the DRL coordinator. The scaling coefficients are obtained in a trial-and-error method to achieve an optimal system control. The coefficients are computed offline, hence the time and computational complexity are not of high significance.

The DRL problem can be expressed as a Markov decision procedure. The DRL agent is trained through interacting with the environment (i.e., power system) using rewards. The aim of an agent is to obtain efficient actions so that transient and small-signal stabilities can be ensured. At per time step t, according to the running state, the agent provides an action for the power system and takes a different state and reward. The agent retains repeating this process until it gets in a final state. The inputs (states) for the agent are as system data ($\Delta\omega_i$ and ΔV_{Ti}) which can be measured by phasor measurement units. In this paper, $\Delta\omega_i$ and ΔV_{Ti} are considered to train the DRL agent. The reward $R_{i.t}$ for each time step is calculated as

$$R_{i.t} = \begin{cases} postive\ reward(+R_{p.t}).\forall\ \Delta\omega_i.\Delta V_{Ti} \in stable\ area \\ negative\ reward(-R_{n.t}).\exists\ \Delta\omega_i.\Delta V_{Ti} \notin stable\ area \end{cases} \tag{9}$$

The final reward R_{Ti} of all iterations can be shown as

$$R_{Ti} = \sum_{t=1}^{n} R_{i.t}/n \tag{10}$$

In continue, the DDPG algorithm is employed to design the DRL agent. This algorithm comprises an actor network together with a critic network [23]. The critic network approximates action-value function $Q(s.a)$ using a Bellman equation, which is defined as follows:

$$Q_{t+1}^{(s.a)} = Q_t^{(s.a)} + \alpha[R_{i.t} + \gamma max Q_t^{(s'.a')} - Q_t^{(s.a)}] \tag{11}$$

A policy gradient theorem is used to update the actor-network. The gradient approximation for the coefficients of actor-network can be calculated as follows:

$$\nabla_{\theta^\mu} J \approx \frac{1}{N}\sum \nabla_a Q(s.a)|_{s=s_t.\ a=\mu(s_t)} \nabla_{\theta^\mu}\mu(s|\theta^\mu)|_{s=s_t} \tag{12}$$

where $\mu(s|\theta^\mu)$ represents a parameterized actor function; and θ^μ represents the policy coefficient. In DDPG, a deep neural network calculates directly the control action a so that a continuous state-action space is provided. Besides, to update the trained actor and critic networks slowly, a target network is created, which remarkably raises the learning stability [23]. Throughout the action exploration procedure, a random noise ϑ is added to develop the exploration policy μ' as

$$\mu'(s_t) = \mu(s_t|\theta_t^{\mu}) + \vartheta_t \tag{13}$$

where $\vartheta_{t+1} = \vartheta_t \times r_c$. In this paper, the actor-network is employed to assess the property of the actions. While the critic network is developed to generate the supplementary parameters for AVR and PSS to achieve minimum steady-state error for the power grid. The action network is provided with a vector state of the $\Delta\omega_i$ and ΔV_{ti} in time step t (i.e., $s_t = [\Delta\omega_{i.t}.\Delta V_{Ti.t}]$) as input, and gives a continuous action $\mu(s_t|\theta_t^{\mu})$ as output. The state s_t and action $\mu(s_t|\theta_t^{\mu})$ are then entered to the critic part and it produces a Q-value $Q(s_t.\mu(s_t|\theta_t^{\mu}))$ as output. Figure 3 shows the structure of the actor-critic network.

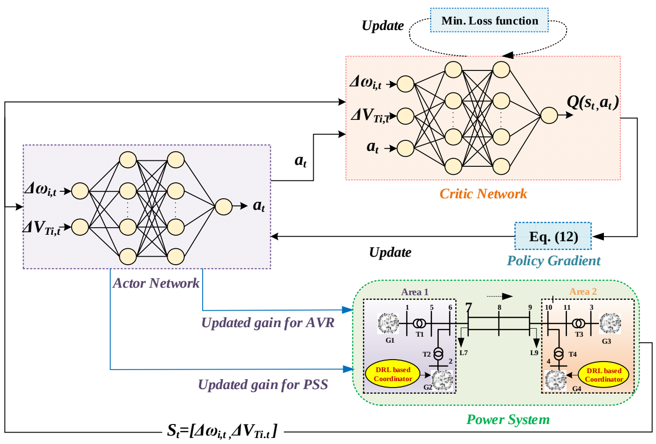


Fig. 3. The structure of the actor-critic network.

IV. Results and Discussion

The simulation analyses are accomplished on Kundur's power system. The system is divided into two control area, each of which includes two synchronous generators. The single-line view of the power grid is depicted in Fig. 3. The areas are connected to each other by two tie-lines. In normal performance, the power exchange between two areas is 413 MW. The load model in each area is a constant impedance model.

The system has a base frequency of 60 Hz. The values associated with base voltage and power for each generator are equal to 20 KV and 900 MVA, respectively. The AVR and PSS are installed on all the generators. The gains of AVR (K_a) and PSS (K_p) are 200 and 30, respectively. The detailed information of the system model can be found in [24]. It is assumed that generator 4 in the second area is provided with the DRL coordinator. It should be noted that the DRL coordinator can also be installed on other ones, which implies a multi-agent learning problem. However, to simplify the learning process, one agent is trained corresponding to generator 4. Table I summarizes the parameters of DRL agent.

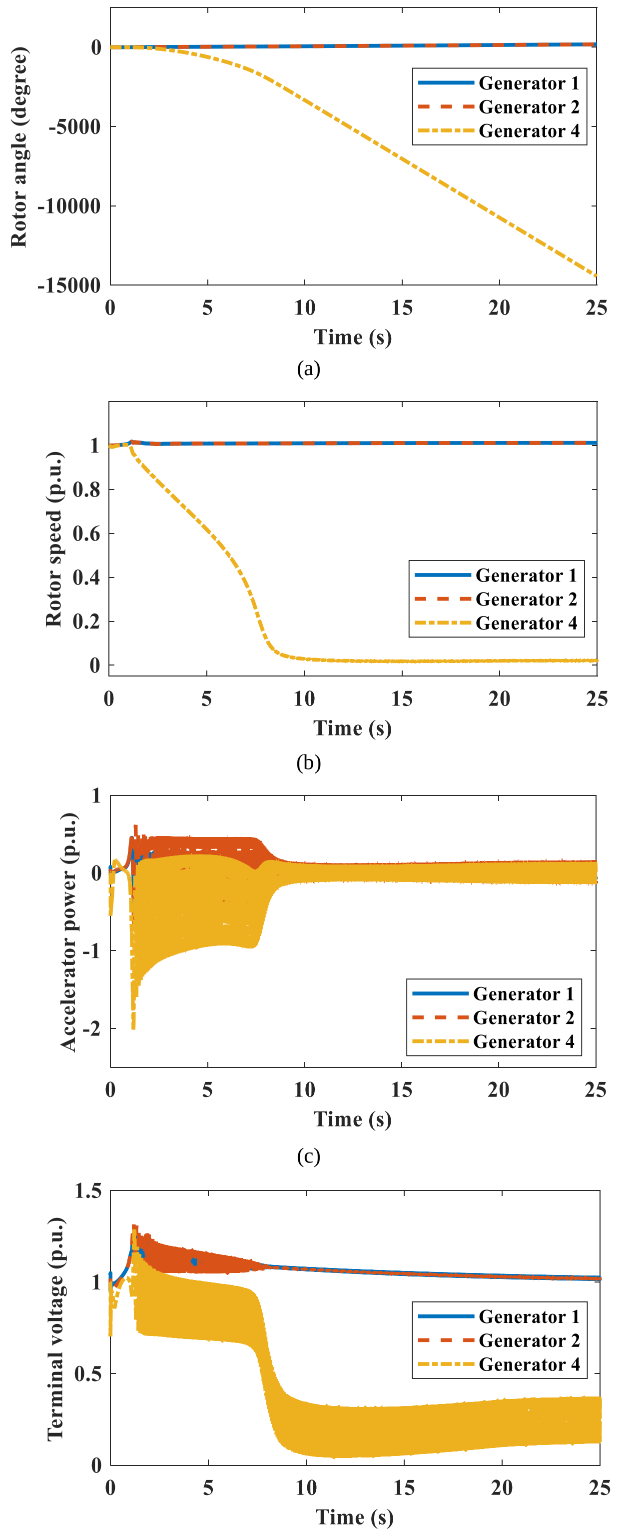


Fig. 4. Time-domain responses of rotor angle, rotor speed, accelerator power, and terminal voltage of the generators without DRL coordinator.

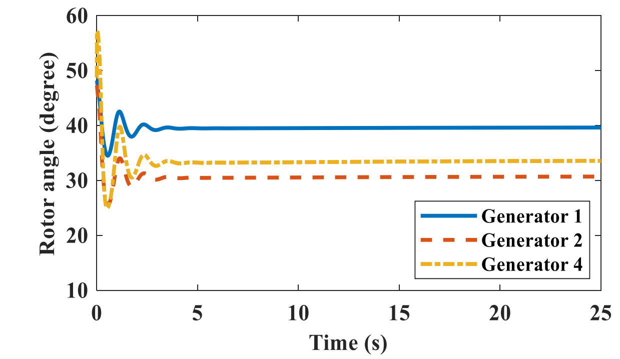
TABLE I. The parameters of DRL agent

Parameter	Value	Parameter	Value
n	5000	r_c	0.9995
α	0.001	γ	1

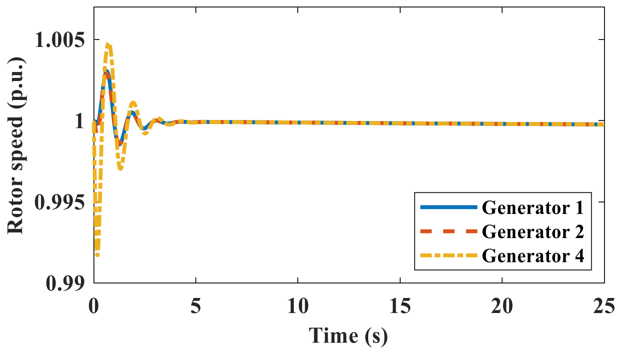
One fault scenario is investigated to prove the capability of the proposed DRL-based intelligent coordinator. In this scenario, generator 3 is disconnected which has the highest generation capacity in the system. The power grid operates in the steady condition before the fault. Fig. 4 illustrates the dynamic responses including rotor angle, rotor speed, accelerator power, and generators' voltage without DRL coordinator after the disconnection of generator 3. As the figure implies, the accelerator power fluctuates around zero so that a stable performance is not achieved. The rotor angle of generator 4 fluctuates with a big amplitude. Furthermore, the voltage of generator 4 is oscillating around zero. The rotor speed implies that area 2 is separated in the first seconds. That is, area 2 has been interrupted and the load amount of area 1 is only supplied by generators 1 and 2. Thus, the grid is unstable as it misses the total generation of area 2. Figure 5 shows the plot of the rotor angle, terminal voltage, rotor speed, and accelerator power with the DRL coordinator after the outage of generator 3. As seen, the operating conditions of the power grid are stable with the DRL coordinator. It implies that generator 4 continues its stable behavior in case of the fault and after. The accelerator power of steady-state becomes zero that represents a safe operation among the electrical and mechanical powers. The generators' rotor angles (Generators 1, 2, and 3) are parallel as a stable operation. In other words, the DRL intelligent coordinator has stabilized the power grid by updating the coefficients of PSS and AVR after the outage of the generator 3. It should be noted that as the design of the DRL coordinator or other intelligent methods relies on science knowledge about the controller and power grid, it might achieve better responses. For example, expanding the number of neurons or hidden layers in the structure of actor-critic network may yield improvement of the dynamic responses. However, the computational speed of the DRL coordinator may be affected which makes it inappropriate for dynamic decision making.

V. Conclusions

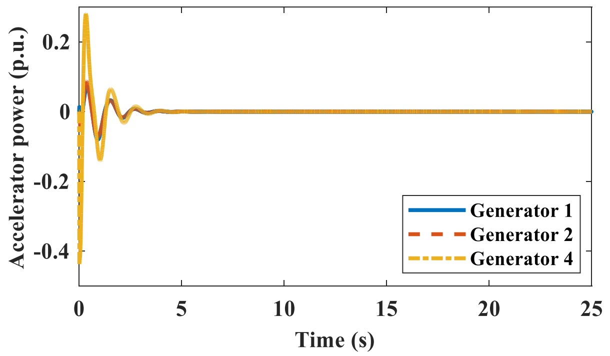
This paper proposed an intelligent method based on DRL for PSS and AVR in an interconnected power grid. The DRL coordinator was proposed to adjust the coefficients of PSS and AVR in an online manner. This coordinator was able to avoid unfavorable interactions among PSS and AVR under changes in the working point of the power grid. In the DRL problem, a DDPG algorithm based on an actor-critic framework was developed to produce control actions to guarantee secure and stable operation of the grid. Numerical simulations indicated that DRL intelligent coordinator could ensure the system stability under the fault condition. The stability studies can also be exercised by installing DRL intelligent coordinator on multiple generators (This implies a multi-agent learning problem) in a power grid, which can be considered as an extension of the studies of this paper.



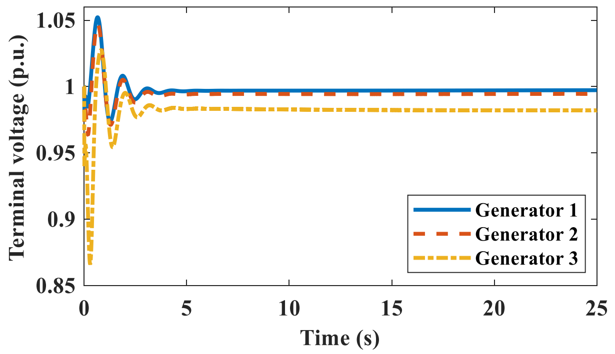
(a)



(b)



(c)



(d)

Fig. 5. Time-domain responses of rotor angle, rotor speed, accelerator power, and terminal voltage of the generators with DRL coordinator.

References

[1] Microgrids - Advances in Operation, Control, and Protection, A. Anvari-Moghaddam, H. Abdi, B. Mohammadi-Ivatloo, and N. Hatziargyriou (Eds.), Spring-er, 2021. ISBN: 978-3-030-59750-4, DOI: 10.1007/978-3-030-59750-4

[2] M. Mohiti, H. Monsef, A. Anvari-Moghaddam, H. Lesani, "Two-Stage Robust Optimization for Resilient Operation of Microgrids Considering Hierarchical Frequency Control Structure", IEEE Trans.

Industrial Electronics, vol. 67, no. 11, pp. 9439-9449, 2020. DOI: 10.1109/TIE.2019.2956417

[3] F.B. Carbajal, A.F. Contreras, I.L. Garcia, A.V. Gonzalez, J.C. Rosas-Caro, and V.M. Huerta, "Output feedback dynamic tracking excitation control of synchronous generators," *IET Gener. Transm. and Distrib.*, vol. 10, iss. 12, pp. 3041 – 3049, Sep. 2016.

[4] J. Ma, H. J. Wang, and K. L. Lo, "Clarification on power system stabiliser design," *IET Gener. Transm. and Distrib.*, vol. 7, iss. 8, pp. 973–981, Sep. 2013.

[5] R. Khezri, and H. Bevrani, "Voltage performance enhancement of DFIG-based wind farms integrated in large-scale power systems: coordinated AVR and PSS," *International Journal of Electrical Power and Energy System*, vol.73, pp. 400-410, Dec. 2015.

[6] Y. Batmani, and H. Golpıˆra, "Automatic voltage regulator design using a modified adaptive optimal approach," *Int. J. Electr. Power Energy Syst.*, vol. 104, pp. 349–357, Jan. 2019.

[7] M. Blondin, J. Sanchis, P. Sicard, J. Herrero, "New optimal controller tuning method for an AVR system using a simplified Ant Colony Optimization with a new constrained Nelder-Mead algorithm," *Appl. Soft Comput.*, vol. 62, pp. 216–229, Jan. 2018.

[8] G. Tu, Y. Li, J. Xiang, and J. Ma, "Distributed power system stabiliser for multimachine power systems," *IET Gener. Transm. and Distrib.*, vol. 13, iss. 5, pp. 603-612, Mar. 2019.

[9] R. A. Jabr, B. C. Pal, and N. Martins, "A sequential conic programming approach for the coordinated and robust design of power system stabilizers," IEEE Trans. Power Syst., vol. 25, no. 3, pp. 1627–1637, Aug. 2010.

[10] H. Bevrani, and T. Hiyama, "Power system dynamic stability and voltage regulation enhancement using an optimal gain vector," *Control Engineering Practice*, vol. 16, iss. 9, pp. 1109-1119, Sep. 2008.

[11] H. Golpira, H. Bevrani, and A. H. Naghshbandy, "An approach for coordinated automatic voltage regulator power system stabilizer design in large-scale interconnected power systems considering wind power penetration," *IET Gener. Transm. and Distrib.*, vol. 6, iss. 1, pp. 39 - 49, Jan. 2012.

[12] A. Oshnoei, M. Kheradmandi, and S. M. Muyeen, "Robust control scheme for distributed battery energy storage systems in load frequency control," IEEE Trans. Power Syst., vol 35, no. 6, pp. 4781-4791, 2020.

[13] R. Khezri, A. Oshnoei, A. M. Yazdani, and A. Mahmoudi, "intelligent coordinators for automatic voltage regulator and power system stabiliser in a multi-machine power system," *IET Gener. Transm. and Distrib.*, vol. 14, iss. 12, pp. 1751–8687, Dec. 2020.

[14] A. Anvari-Moghaddam, A. Rahimi-Kian, M.S. Mirian, J.M. Guerrero, "A Multi-Agent Based Energy Management Solution for Integrated Buildings and Microgrid System", Applied Energy, vol. 203, pp. 41-56, 2017. https://doi.org/10.1016/j.apenergy.2017.06.007

[15] V. Mnih et al., "Playing atari with deep reinforcement learning," 2013, arXiv:1312.5602.

[16] Z. Yan and Y. Xu, "Data-Driven load frequency control for stochastic power systems: A deep reinforcement learning method with continuous action search," IEEE Trans. Power Syst., vol. 34, no. 2, pp. 1653–1656, Mar. 2019.

[17] X. Qi, "Rotor resistance and excitation inductance estimation of an induction motor using deep-Q-learning algorithm," Eng. Appl. Artif. Intell., vol. 72, pp. 67–79, 2018.

[18] S. Phaniteja, P. Dewangan, P. Guhan, A. Sarkar, and K. M. Krishna, "A deep reinforcement learning approach for dynamically stable inverse kinematics of humanoid robots," in Proc. IEEE Int. Conf. Robot. Biomimetics, Macau, 2017, pp. 1818–1823.

[19] P. W. Sauer and M. Pai, "Power system dynamics and stability," Urbana, 1998.

[20] IEEE Recommended Practice for Excitation System Models for Power System Stability Studies, IEEE Standard 421.5-2005, Apr. 2006.

[21] G.J.W. Dudgeon, W.E. Leithead, A. Dysko, J. O'Reilly, and J.R. McDonald, "The effective role of AVR and PSS in power systems: Frequency response analysis," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1986–1994, Nov. 2007.

[22] A. Dysko, W.E. Leithead, and J. O'Reilly, "Enhanced power system stability by coordinated PSS design," *IEEE Trans. Power Syst.*, vol. 25, iss. 1, pp. 413-422, Feb. 2010.

[23] T. P. Lillicrap, J. J. Hunt, A. Pritzel et al., "Continuous control with deep reinforcement learning," arXiv preprint arXiv:1509.02971, 2015.

[24] P. Kundur, *Power System Stability and Control*. New York: McGraw Hill, 1994.