TRANSIENT STABILITY ENHANCEMENT OF MODERN POWER GRID USING PREDICTIVE WIDE-AREA MONITORING AND CONTROL

by

Reza Yousefian

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Electrical Engineering

Charlotte

2016

Approved by:

_____
Dr. Sukumar Kamalasadan

_____
Dr. Yogendra Kakad

_____
Dr. Valentina Cecchi

_____
Dr. Brett Q. Tempest

ABSTRACT

REZA YOUSEFIAN. Transient stability enhancement of modern power grid using predictive wide-area monitoring and control. (Under the direction of DR. SUKUMAR KAMALASADAN)

This dissertation presents a real-time Wide-Area Control (WAC) designed based on artificial intelligence for large scale modern power systems transient stability enhancement. The WAC using the measurements available from Phasor Measurement Units (PMUs) at generator buses, monitors the global oscillations in the system and optimally augments the local excitation system of the synchronous generators. The complexity of the power system stability problem along with uncertainties and nonlinearities makes the conventional modeling non-practical or inaccurate. In this work Reinforcement Learning (RL) algorithm on the benchmark of Neural Networks (NNs) is used to map the nonlinearities of the system in real-time. This method different from both the centralized and the decentralized control schemes, employs a number of semi-autonomous agents to collaborate with each other to perform optimal control theory well-suited for WAC applications. Also, to handle the delays in Wide-Area Monitoring (WAM) and adapt the RL toward the robust control design, Temporal Difference (TD) is proposed as a solver for RL problem or optimal cost function. However, the main drawback of such WAC design is that it is challenging to determine if an offline trained network is valid to assess the stability of the power system once the system is evolved to a different operating state or network topology. In order to address the generality issue of NNs, a value priority scheme is proposed in this work to design a hybrid linear and nonlinear controllers. The algorithm so-called supervised RL is based on mixture of experts, where it is initialized by linear controller and as the performance and identification of the RL controller improves in real-time switches to the other controller. This work also focuses on transient stability and develops Lyapunov energy functions for synchronous generators to monitor

the stability stress of the system. Using such energies as a cost function guarantees the convergence toward optimal post-fault solutions. This energy energy functions are developed on inter-area oscillations of the system identified online with Prony analysis. Finally, this work investigates the impacts of renewable energy resources, in specific Doubly Fed Induction Generator (DFIG)-based wind turbines, on power system transient stability and control. As the penetration of such resources is increased in transmission power system, neglecting the impacts of them will make the WAC design non-realistic. An energy function is proposed for DFIGs based on their dynamic performance in transient disturbances. Further, this energy is augmented to synchronous generators' energy as a global cost function, which is minimized by the WAC signals. We discuss the relative advantages and bottlenecks of each architecture and methodology using dynamic simulations of several test systems including a 2-area 8 bus system, IEEE 39 bus system, and IEEE 68 bus system in EMTP and real-time simulators. Being nonlinear-based, fast, accurate, and non-model based design, the proposed WAC system shows better transient and damping response when compared to conventional control schemes and local PSSs.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Sukumar Kamalasadan who gave me this great opportunity to be a part of his research group. His energy, motivation, support, and advice were a valuable assist not only to develop this work, but also to grow my personality during these five years. I would also like to thank Dr. Valentina Cecchi, Dr. Yogendra Kakad, and Dr. Brett Q. Tempest for being part of my graduate committee and for letting me learn from their experience in the field.

I am thankful for all the amazing educators I have worked alongside. I am also thankful to the family of EPIC for providing the friendly atmosphere for us to grow in and learn. I would like to appreciate all my lab mates for their collaboration, help, guidance, hint, and feedback: Rojan Bhattarai, Amirreza Sahami, SeyyedMahdi Moghadasi, Rabindra Maharjan, Saeed Mohajeryammi, Iman Naziri Moghadam, Masoud Davoudi, and many others.

I am enormously thankful for my wonderful family: my parents who always trusted me, and been a constant source of comfort and support during an education, which almost took a lifetime; my siblings for all their support and encouragement; and to my wonderful wife, Yasaman Kamyab Hessary who was always there for me, motivated me to be ambitious, hard worker, and grateful for the things I have in my life.

Finally, I thank the National Science Foundation, EPIC, ECE department, and graduate school for providing me the financial assistance to accomplish this project.

DEDICATION

To my beloved family

My parents: Mehdi Yousefian, Masoumeh Norouzian

My siblings: Amin, Faezeh

And my beloved wife: Yasaman Kamyab Hessary

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: OVERVIEW

1.1 Introduction

In general, ensuring secure and stable operation of power systems exposed to different disturbances is among the most formidable challenges that power engineers face today. Modern power system is considered to be more complex and nonlinear than before due to significant integration of distributed energy resources, severe transmission congestion and growth of energy markets deregulation. Such power systems typically exhibit multiple electro-mechanical oscillation modes. These oscillations can be classified as local or related to different areas. Local oscillations are mainly due to single generator swinging against the rest of the network, whereas area oscillations are associated with groups of generators oscillating relative to each other. With growing demand that includes nonlinear loads and increasing penetration of distributed energy resources power transfer capability of the electric network is constantly decreasing. Modern grid management also leads to closer operation of network to its limits resulting in stressed network with smaller stability margins. During such conditions, loosely damped inter-area modes of oscillations is enough to make the system unstable.

Local oscillations are typically damped by generation excitation control using power system stabilizers (PSSs). However, these local damping controllers were not originally designed to address the modern power system's uncertainties and unpredictability introduced by integration of power systems [2]. These shortages of existing power system controllers, together with desire to smarten the power grid, makes it necessary to design and develop a Wide Area Control (WAC) in the power grid. The WAC coordinates the actions of a number of distributed agents using supervisory control

data acquisition, Phasor Measurement Unit (PMU) or other sources of Wide Area Monitoring (WAM) [3].

Power systems in the presence of drastic system changes can encounter transient instability due to large excursions of generator rotor angles and losing synchronism of generators. Conventionally, algebraic and differential equations are used to describe the behavior of power systems dynamics and to create mathematical models to represent these process. In general, a model is purposeful simplification of a system for solving a particular problem. However, the complexity of the power system stability problem along with uncertainties and nonlinearities makes the modeling non-practical or inaccurate. In this type of problems the progress of the power system transient and oscillations are monitored, and prediction of stability of swings is done to classify the fault in the form of Transient Stability Assessment (TSA). However, the WAM temporal information can further be used in WAC designs to perform real-time transient stability enhancement, which can improve the power transfer capability of a transmission system and prevent the system from generation or load disconnection, or catastrophic failure following a sequence of disturbances in the system.

With respect to control designs in wide-area power system, multiple linear model-based adaptive and hierarchical control schemes for damping post-disturbance oscillations have been reported [4–6]. Global PSS, Adaptive control, Linear Matrix Inequality (LMI), and H∞-based damping controls have been effectively used. All these classical designs require a linear model of the system, which is not obtainable with an acceptable degree of accuracy in practice. These control systems can perform well enough during smaller disturbances or small signal stability analysis of the grid. An adaptive controller which covers wide area system and is capable of learning and tracking the extreme nonlinear and unpredictable dynamic behavior of the system and can optimally coordinates actions of generating units as agents is a most ideal type of control and decision making in the modern power grid.

Machine learning-based or Artificial Intelligence (AI)-based control architecture utilized under the umbrella term measurement-based techniques has been long proposed to overcome some of the aforementioned issues. Most notable of them is the method of Neural Networks (NNs), which is based on mimicking the intelligence by which the human brain represents information.

It is known that, learning controllers such as Neural Networks (NNs) can improve tracking performance through repeated trials of interaction with environment and overcome the nonlinearity problem of the system [7]. Reinforcement Learning (RL), considered as a distinct method in this area, is often applied to sequential dynamics problems and optimization of action in the course of the process [8] with approximated model of the system. This method uses NNs to develop optimal controllers from measurements of available system inputs and outputs and has shown more reliability and stability than classical control methods [9–17]. NN-based approaches, being strongest benchmark for RL tasks [16], has utilized features of back propagation to model and solve the cost function of the system [17]. Further, NN structure can be investigated to gain more efficiency in terms of modeling the nonlinearity of the system dynamics and uncertainty regarding the wide area monitoring and communications. This work mainly focuses on the optimality and generality concerns of NNs in the WAC of power system. Such a design can be implemented in real-time environment, which along with transient energy functions, can monitor the stability level of the system and optimally solve for control actions.

1.2 Contribution

This thesis shows that WAM measurements around a power system can track bulk stresses after disturbances and NN-based WAC in control center can mitigate or damp the corresponding transient oscillations. One important advantage of this approach is that the controller can mitigate global transient oscillations. That is, our formulation in terms of NN-based regression allows the WAC to model and control nonlinear

dynamics of the power system. Furthermore, through adaptation of RL the control scheme is employed as a multi-agent modeling, which ensures optimal control solution in real-time.

There are mainly two concerns when employing such a RL-based control technique as a WAC system, in general, with AI techniques. First, the design should reach to optimal solution in time; i.e. as the control actions in transient domain are needed to be applied fast (usually less than 1 sec) the NN-based controller should provide timely convergence of the weights in new operating regions. This problem becomes more complex in WAC domain as they are highly dependent on the performance of the communication infrastructure, without which the functionality of them will not be achieved. Knowing that time delay affects wide-area power system stabilizer design, the consequence of delayed input data and output signals in WAC systems should be considered and modeled. Also, the performance requirements posed on by time availability for decision making and control actuation needs to be elicited. For first swing transient stability, control actions must be taken prior to the peak of the forward inter area angle swing (around 1s) [18, 19]. This thesis shows that, RL algorithms are well suited to tackle these uncertainties in the system. Temporal Difference (TD) method is proposed to solve the RL problem, widening the window of prediction and accounting for communication network delays as in Model Predictive Control (MPC).

Being a learning-based technique, RL algorithm is not guaranteed to lead to global optimal solutions until the learning is complete. Also, heuristic nature of these types of controllers still causes stability issues thus cannot be fully trusted in real-life implementation even after being tested using real-time simulations [10]. Due to these reasons, till now, conventional linear local controllers like PSS or adaptive control has been used as a solution to damp oscillations that occur during transient instability of power system, even though being designed to address small signal stability. The need

to develop a hybrid global controller taking advantage of both linear and intelligent designs that can be performed during linearized and nonlinear system transitions, and yet provide stable control actions under unforeseen disturbances is thus very clear.

Generality is another challenge when using such AI-based designs. Several works have tackled this issue by designing a hybrid controller as linear and nonlinear based controllers for power system stabilization. We propose a technique based on value prioritization of controllers in the form of mixture of experts or supervised RL method. This technique guarantees the control system convergence to optimal solution, eliminating the necessity of offline training. It also can mitigate the problem of over-fitting in NNs when the amount of training data is limited. The main purpose of this design is using the known linear model to mainly control the system, and exploit the NN designs to be activated in the case of nonlineairty and uncertainties. Additionally, WAC may show destabilizing interactions with local controllers which makes the system unstable during certain contingency situations due to difference in policy or objective. However, we propose a unified objective function as a performance and identification criteria for controllers to indicate level of activation in the system. In such design, parametric and functional adaptation has been performed, and a hybrid controller that can be integrated with conventional PSS or WAC is designed and illustrated.

Also, in contrast to other research that considers linear objective or cost function, our method, getting the benefits of nonlinear approximation, estimates the actual nonlinear energy function in real-time and optimally minimizes such a global function. Usually, energy functions are used to provide stability index in transient domain; i.e. used for stability classification or prediction. However, by employing our method, such energy function could be further incorporated to real-time WAC design to damp the oscillations. Such a method has several advantages as it is not dependent on post-fault operating points; does not rely on parameters of the system as it can be impractical with conventional techniques if the system topology changes after a

fault; and it is predictive. The method is mainly based on energy increments or energy function damping. To monitor such a function online there is need for online eigenvalue analysis, which is performed by Prony analysis in this work.

Another consideration in this thesis is the transient dynamics of Doubly Fed Induction Generator (DFIG)-based wind generators. In general, renewable energy resources are capable of maintaining their local stability while system is exposed to transient inter-area oscillations, if these oscillations are not severe enough to activate their protection devices. Further, this work proposes exploitation of DFIG in the form of potential, reactive, and inertial energies to consume/support some energy from/to the system. The system dynamics is represented in coupled oscillatory platform with synchronous generators as coherent group of generators containing potential, kinetic, and field energies. Through adaptation of the proposed multi-agent WAC, wind farms power references along with synchronous generators excitation voltage references are controlled to optimize the overall energy exchange in the system.

Finally, we summarize the contributions in terms of new formulation, analysis, testing and practical application as follows:

- In terms of formulation, this thesis expands the existing RL problem to wider window of optimization. Further, it introduces new formulation to find out the better actions with respect to performance and identification of hybrid controllers. Finally, a novel energy function is developed for renewable energy resources, which along with energy function of synchronous generators, is minimized forward-in-time.

- In term of analysis, this thesis establishes useful measures for rotor-angle stability, in general dynamic performance of the WAC, including damping performance, energy functions, critical-clearing time, overshoots, and so on. These analysis is performed in several case studies such as with/without communication delays, data drop out, tie-line short circuit, cascading failures, sudden load

change, trained/partially-trained NNs.

- In term of testing, this thesis uses several power system case studies including 2-area 5-machine system, IEEE 39-bus 10-machine system, and IEEE 68-bus 16-machine system which is representative of interconnected New-England New-York power grid. Furthermore, the later is modified with DFIG-based wind farms to study the transient behaviour of such units. These dynamic models are developed in detail in several software tools including PSCAD, RSCAD, Power System Toolbox (PST), PSS/E, Matlab, and so on.

- In term of practical application, this thesis proposes a novel real-time benchmark for interconnection of the WAC with power system model. The WAC has been implemented in TI Micro-controller and is connected to Real-Time Digital Simulator (RTDS) through hardware in the loop to study the practical application challenges.

1.3 Dissertation Organization

WAC designed on NNs are a way to mitigate and damp inter-area oscillations by means of WAM measurements. We are interested in transient stability dynamics of power system including synchronous generators and renewable energy resources to define the optimization cost function. Further, this function is minimized in real-time addressing the robustness, optimality, nonlinearity, and generality issues of the design. The control actions are augmeneted to the local controllers of these generating units providing extra damping of the energy functions and leading to convergence to the global post-fault equilibrium points.

Chapter 2 provides a brief reviews of previous work for rotor angle stability, WAC designs, transient stability assessment techniques, and explores the possible directions that can be further investigated.

Chapter 3 develops the WAC design on real-time benchmark addressing the opti-

mality and robustness RL algorithm.

Chapter 4 proceeds the design to a hybrid controller in order to gain generality for NNs. The design is applied and tested as a hybrid local and global design and a hybrid linear and nonlinear design.

Chapter 5 focuses on transient stability enhancement problem. The carefully developed energy function of synchronous generators is linked to the RL problem. This chapter studies the power system rotor angle stability from eigenvalue, coherency, and transient stability perspectives.

Chapter 6 investigates and develops energy functions for renewable energy resources, in specific DFIGs, in transient disturbances. This energy is augmented to the global cost function to study the impacts of these units in transient stability enhancement.

Chapter 7 discusses possible future work, summarizes, and concludes the thesis.

CHAPTER 2: REVIEW OF LITERATURE

2.1 Rotor-Angle Stability

Initially, we elaborate the power system stability problem as the ability of power system to regain the operating equilibrium states after being subjected to a physical disturbance, with most system variables bounded so that practically the entire system remains intact [1].

In the context of rotor angle stability, the dynamics of each synchronous generator bus can be can be characterized by the complex terminal voltage $V_t\angle\delta$, where δ is the rotor angle with respect to synchronously rotating reference frame. The rotor speed is given by $\omega = \dot{\delta}$. Disturbances on power system components, e.g. power lines causes the system to move away from the pre-fault equilibrium point and experience a transition toward the post-fault dynamics. Rotor angle stability as one of the main classifications of power system stability, refers to the ability of synchronous machines to remain in synchronism after being subjected to the fault. Instability occurs in the form of increasing angular swings of some generators leading to their loss of synchronism with other generators. This behavior could be assessed by using a simplified 2nd order synchronous generator model as [1],

$$\Delta\dot{\delta} = \Delta\omega \tag{2.1}$$

$$M\Delta\dot{\omega} = P_m - P_e - D\Delta\dot{\delta} \tag{2.2}$$

where, $\Delta\delta = \delta - \delta^*$ is the rotor angle deviation, $\Delta\omega = \omega - \omega^*$ the speed deviation, the symbol "$*$" denotes the post-fault operating point, M the inertia constant of the synchronous generator, D is the damping coefficient, P_m the mechanical power, and

Table 2.1: Rotor angle stability [1]

Characteristic	Small-signal stability	Transient stability
Disturbance	Small	Severe
Causes	Varying load	Short Circuit on Transmission line
Linearity	Linear	Non-Linear
Dependence	Initial States	Initial States, Disturbance severity
Stability form	Lack of damping (Periodic)	Lack of synchronization (Aperiodic)
Nature	Inter-Area, Local	Inter-Area, Local
Time frame	10-20 sec	3-5 sec

P_e the electrical power as,

$$P_{ej} = \sum\nolimits_{k \in N_j} B_{jk} V_j V_k \sin(\delta_{jk}) \tag{2.3}$$

where, N_j is the set of neighboring buses of the j^{th} bus. This power represents the power flow through transmission lines into the power network, which follows highly nonlinear behavior.

In general, swing equation (2.2) can be resolved into two components: synchronizing power component in phase with $\Delta\delta$ and damping power component in phase with $\Delta\omega$. The stability depends on the existence of both components of torque for each of the synchronous machines. Rotor angle stability can be characterized mainly into two categories: small signal stability and transient stability. The consideration is based on the size of disturbance, time span and involved devices as presented in Table 2.1.

2.1.1 Rotor Angle Stability Assessment

Rotor angle stability assessment can be categorized into Small Signal Stability Assessment (SSAT) and Transient Stability Assessment (TSA). In general, small signal stability as shown in Table 2.1 is concerned with the ability of the power system to maintain its synchronism under small disturbances, which allows the linearization of system equations for purposes of analysis. The stability of the system is assessed by the characteristics of the eigenvalues of system matrix. This SSAT is usually

performed with the purpose of improvement in damping performance of the system through employment of frequency-based continuous controllers such as Power System Stabilizers (PSSs).

Transient instability occurs when the power system is subjected to a severe disturbance. The resulting system response involves large excursions of generator rotor angles and is influenced by the nonlinear power-angle relationship. This type of instability is usually due to insufficient synchronizing torque, manifesting as first swing instability [1]. TSA is performed offline at pre-fault stage or online during the fault. The literature in this category is respectively rich, with various methods utilized, [20–30]. TSA mainly focuses on the critical clearing time of the faults in power system. A comprehensive simulation of faults provides useful information regarding the vulnerable points of the system and makes sure of safe stability margin. Besides, these evaluations can provide helpful database for real-time analysis to perform preventive or emergency control actions. In this type of problems the progress of the power system transient and oscillations are monitored, and prediction of stability of swings is done to classify the fault. Various conventional approaches for TSA have been proposed in the power systems literature which is listed briefly in Table. 2.2. More information regarding the TSA, control techniques and methods is provided in Appendix A.

2.2 Power System Control Hierarchy

Rotor angle stability problems may be either local or global in nature. Local oscillations are usually associated with rotor angle oscillations of a single generator against the rest of the power system; while, inter-area oscillations involve a group of generators in one area swinging against another group [31]. The characteristics of local and inter-area modes depends on the strength of the transmission system, generator excitation, control, and plant output.

Table 2.2: Transient Stability Assessment (TSA) techniques

Technique	Characteristic
Numerical	• Weak performance in real-time implementation, as they require accurate information of the power network topology [29]. • It requires the post-fault system simulation to conclude the stability status [20].
Direct	• Based on direct calculation of the Transient Energy Function (TEF) [20]. • The numerical integration is required only on the fault-on state trajectory. • Analytical TEF with detailed device models cannot be derived for multi-machine power system. Hence, such problem formulation may lead to excessive simplifications.
Hybrid	• These approaches are mainly based on equivalent modeling, which can be integrated much faster than real-time [30]. • Their algorithms is complex.
Intelligent	• These approaches are fast and simple. • Excessive offline training.

2.2.1 Local Controllers

Local control is the first control level in the control hierarchy and has the fastest response [32]. This control responds to local system dynamics and ensures that the variables track their reference values. Synchronous generators supply most of the electrical energy in the power system, and are mainly responsible for maintaining the stability of the electrical network. Therefore, effective control of these devices is very critical [18]. Local continuous feedback controls are depicted in Fig. 2.1.

2.2.2 Secondary and Tertiary Controllers

In order to maintain the power quality and stability of the generating units for longer term variables, the next level controller is deployed to determine the set points for the primary control. The secondary control works as a centralized Automatic Generation Controller (AGC) and compensates the steady-state errors of the voltage and

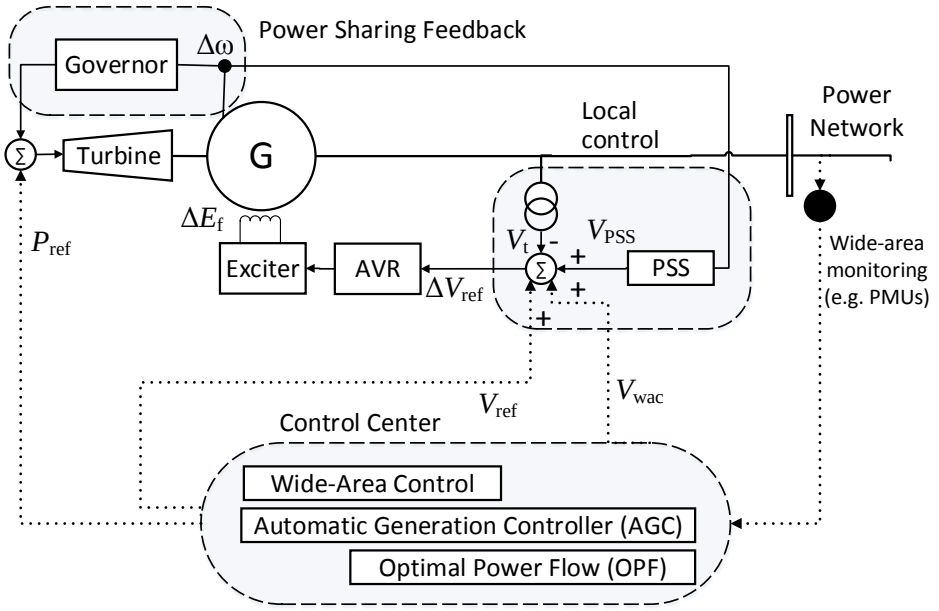


Figure 2.1: Schematic of synchronous generator local, secondary, tertiary, and wide-are control

frequency. This controller makes use of communications and Wide-Area Monitoring (WAM) systems to coordinate the action of all the generation units within a given area. The time response of this control level is in the range of minutes, thus having a slow dynamic if compared with the local control. Finally, the tertiary control level could be utilized for optimizing the operation of the system [32].

2.2.3 Wide-Area Controllers

Generally, Wide-Area Control (WAC) coordinates the actions of a number of distributed agents using Supervisory Control and Data Acquisition (SCADA), Phasor Measurement Unit (PMU), or other sources of WAMs [3,18,19,33,34]. Several control architecture is designed ([35–38]) for such applications in the form of hierarchical designs, distributed control methods, central designs, or multi-agent-based techniques. Overall such designs have better controlability on inter-area modes of oscillations due to better observability than local controllers. Most of them use frequency domain methods and root-locus criteria with signals including changes in tie line flows, inter-area angles and/or machine speeds. For instance, in [4] a global PSS has been

presented with a supplementary input from PMUs, geographically spread over coherent areas of power system. This design is built on top of the existing local controller, resulting in a hierarchical control architecture with significant advantages in terms of reliability and operational flexibility. Several other conventional techniques has also been used to be applied for this application, such as robust techniques [39] and optimal control methods [31].

In addition, it is known that transient instability may happen in further swings due to lack of damping torques [1]. The continuous control actions can also be used as the input to the excitation system, to make the problem dynamic mitigating and damping control. This type of control schemes are designed not only to provide a stable final state but also minimize state excursions along the trajectory and increase the power system stability margins [40]. Due to nature of this problem, nonlinear designs should be applied for this application [41]. In [41] a new structures for stability enhancing excitation controllers is designed using a nonlinear multi-machine system model and Lyapunov's direct method. The controller is design to ensure the negativity of the derivative of Lyapunov function defined on third-order model of synchronous generator.

2.3 Machine Learning Algorithms

In general, the term "learning" means adjusting the parameters or in the case of NNs, weights, to reduce the error between the target outputs and the actual outputs. In this thesis, two main methods of SL and RL are investigated for their application in power system control. Initially, a brief overview of these methods is presented:

2.3.1 Supervised Learning (SL) Algorithm

SL methods are referred to problems involving static I/O mappings and minimization of a vector error signal, with no explicit dependence on how training examples are gathered. In this category full knowledge of the problem context is available.

The system learns to perform its designated task with assistance of a *teacher*. Availability of data pairs as an input and desired outputs helps the system to update its parameters in order to minimize the error. Generally, SL can be categorized as:

- Regression problem
- Classification Problem

These two types are discussed later for their application in power system stability.

2.3.1.1 Neural Network (NN) Classes

There are mainly two classes of NNs with several types each investigated in power system control scheme:

- Static Neural Network
  - Multi-Layer Perceptron Neural Network (MLPNN)
  - Radial Basis Functions Neural Network (RBFNN)
  - Functional link (FNL)
- Dynamic Neural Network
  - Recurrent Neural Network (RNN) (Fig. 2.2)
  - Simultaneous Recurrent Neural Network (SRN)
  - Time Delay (TDL) Neural Network (Fig. 2.3)

Static NNs are characterized by node equations that are memoryless; However, Dynamic NNs can be described by differential equations. As a static network, it is shown that the RBF is superior in classification and pattern recognition problems, while the MLP is more efficient in function approximation [42].

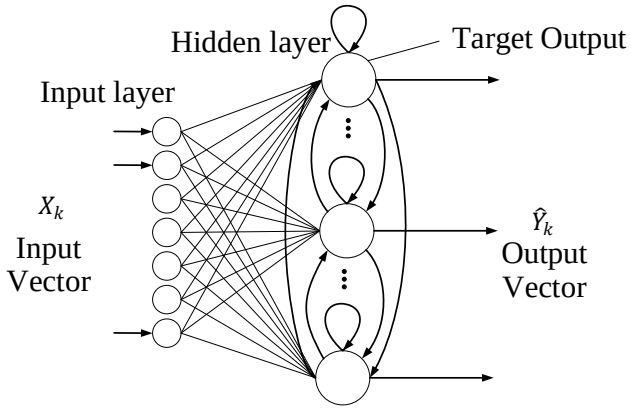


Figure 2.2: Structure of Recurrent Neural Network

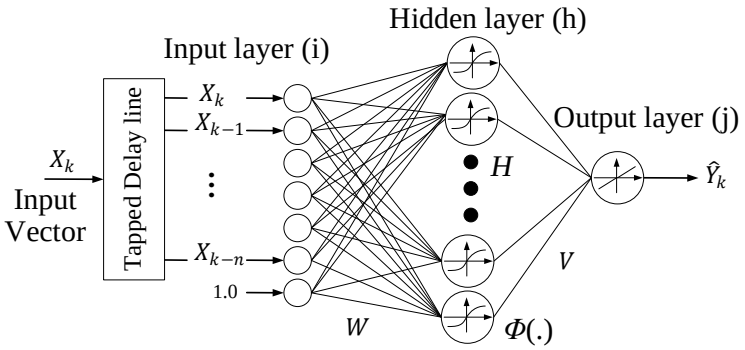


Figure 2.3: Structure of Time Delay Neural Network

In general, MLP is the simplest type of SL consisting of Feed Forward Neural Networks (FFNNs) constructed with three main layers of input, hidden, and output layers, each consisting of input, feature, and decision units in the form of *perceptron* architecture [43]. The FFNN is based on feed forward activation, in which units of each layer pass on their activation to next layer, until the output layer where the actual response to the input layer is generated. Then, these outputs are compared with the desired responses in the form of training patterns. If we keep the first-layer weights of FFNN fixed and with only the second-layer weights tuned, the NN has only one layer of tunable weights. Then, the output values are computed by inner

product between the weight vector W and a state-dependent feature vector $\Phi(.)$ as,

$$\hat{Y} = \hat{W}^T \Phi(X) + \epsilon \tag{2.4}$$

where, $\Phi(.) \in \Re^h$ is the corresponding nonlinear mapping function of the states, $\hat{W} \in \Re^{h,j}$ is the parameter vector of approximated weights of the FFNN, with $h, j \in N$ are dimensional space of the feature vector representing the hidden and output layers of the NN, and ϵ is the NN functional approximation error.

2.3.1.2 Training

The SL is based on the back-propagation of the error through the NN. In order to start the training process, all the weights are initialized with small values. The output values is computed based on the inputs in training set and the weights. Further, the error is calculated in the form of norm 2 of output error

$$e(t) = \hat{Y}(t) - Y^*(t) \tag{2.5}$$

Then, the error at the output layer is used to compute for the error at the hidden layer. Error back-propagation learning uses the gradient decent algorithm since the activation functions used are differential. Gradient descent means that the parameter vector is updated along the negative gradient direction of the mean square error. As a result, change of each NN weights can be derived from the deviation of NN's output to its optimal value e by means of gradient descent via back-propagation through the NN model. This training could be in the form of

- Online learning:

$$\Delta W(t) = \alpha e(t) \Phi(X(t)) \tag{2.6}$$

- Batch learning:

$$\Delta W = \alpha \sum_{t} e(t)\Phi(X(t)) \tag{2.7}$$

where, α is small step size learning parameter. In the first one, the training data becomes available in a sequential order and is used for training at each time t. On the other hand, in batch learning the training is performed once the the entire training data set is available.

2.3.2 Reinforcement Learning (RL) 0Algorithm

RL is often applied to problems involving sequential dynamics and optimization of a scalar performance objective, with online exploration of the effects of actions [7,8,44,45]. The key feature of RL is that training information from the environment is used as an evaluative signal. This method has taken the attention in power system control studies since it can achieve optimal solution requiring no prior knowledge about the system and it can adapt itself to fit the changing environment [44]. In spite of SL, there is no desired output available in this category. However, interacting with environment can provide evaluative feedback, which can further be used to update the learning system toward improving its quality of performance. In other words, action system maps from states to actions that optimizes some performance criterion. The goal in RL is to find a single input value that maximizes the total amount of rewards over the sequence of decision [45]. Taking actions in RL, the agent has to balance two conflicting objectives, exploitation and exploration.

2.3.2.1 Adaptive Critic Designs (ACDs)

ACDs are common approach to handle RL, which are capable of optimization over time and under conditions of noise and uncertainty. Since actions should be taken at each time step and their effect is not known until the end of the sequence, it is not possible to design an optimal controller using the traditional SL. Several research

works have been done in this area, proposing different types of critics [7]. In essence, the adaptive critic method determines optimal control policy for a system by successively adapting two NNs, called critic NN and action NN which learn respectively the desired cost function and desired control value based on the cost function. These two NNs approximate the Hamilton Jacobi Bellman (HJB) equation associated with optimal control theory [7]. The cost-to-go function is given as follows:

$$J(t) = \sum_{k=0}^{\infty} \gamma^k U(t+k) \tag{2.8}$$

where, U is the utility function used for reward or punishment in terms of RL concept or incremental cost function. This function can be represented as,

$$U(t) = -\Delta x(t)^T Q \Delta x(t) - u(t)^T R u(t) \tag{2.9}$$

where, x is the states of the system, u is the control action, the weighting matrix $Q > 0 \in \Re^{m \times m}$ is required to be positive-definite, and $\Delta x = x - x^*$, x^* being the operating points derived from the reference model, and γ is the discount factor needed to maintain the solution as a finite horizon problem with a limit on the upper bound of the solution. By selecting an appropriate value of $\gamma \in (0, 1]$, we can weight the future values of the utility function and affect the convergence process [46].

2.3.2.2 Solutions to the Problem

Dynamic Programming (DP) has gained much attention from many researchers in order to obtain approximate solutions of the RL problem and the HJB equation [47–49]. Various techniques have been presented in this area which use the (2.8) or derivatives of that as optimization goal to be solved. An alternative way of distinguishing ACD methods is to consider the role of system models in the training loops of each method [7, 48]. In general, ACD methods are categorized based on the critic

training methods and effect of system model in the training process. Various versions are proposed which are Heuristic Dynamic Programming (HDP), Dual Heuristic Programming (DHP), Action Dependent HDP and DHP (ADHDP, ADDHP).

All these provided structures can realize the same function that is to obtain the optimal control policy, while the computation precision and running time are different from each other. The model based methods have been shown to be much more efficient for training neuro-controllers and to produce superior designs to non-model based methods. Generally speaking, the computation burden of HDP is low but the computation precision is also low. On the other hand, DHP and ADDHP have an important advantage over the simple ACDs since their critic networks build a representation for derivatives of through by being explicitly trained on them and area of model-based control we usually have a sufficiently accurate and well-defined model network. Ref. [7] provides a full description and analysis on these models.

2.3.2.3 Training

Considering HDP as a main method of ACD technique, critic network represents the cost function by being trained explicitly and directly based on the states of the system. Three NNs are implemented in this method: identifier NN, critic NN, and action NN. General scheme of the ACD controller is provided in Fig. 2.4. In particular, the training process of the critic NN is based on DP, which estimates J by updating its policy with respect to error, e_C, with elements of the rewards obtained from the environment, $U(t)$, the cost functions at current time step, $J(t)$, and future time step, $J(t+1)$. This can be written as,

$$e_C(t) = \hat{W}_C^T \Phi_C(x(t)) - \gamma \hat{W}_C^T \Phi_I(x(t+1)) - U(t) \tag{2.10}$$

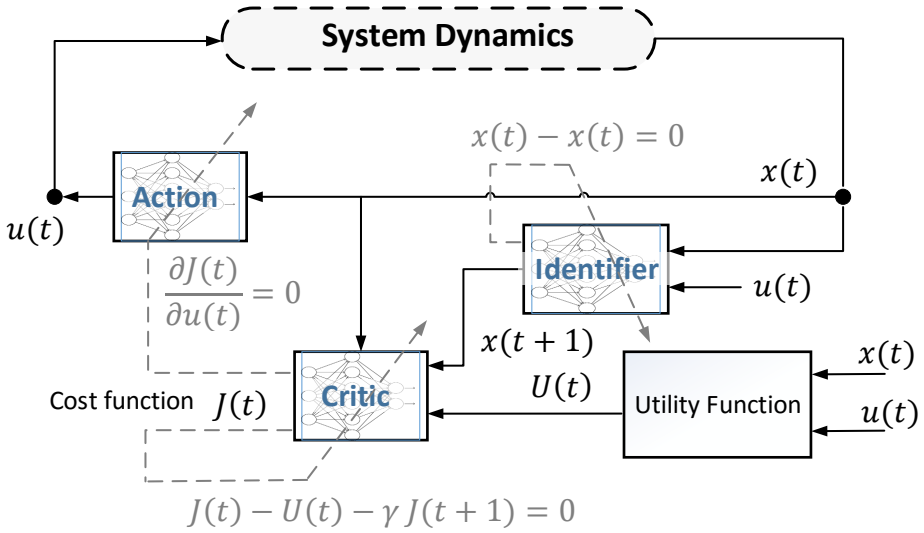


Figure 2.4: Structure of the Adaptive Critic Design

where, critic NN future outputs is based on predicted states derived from identifier NN. Training of the identifier NN is derived as,

$$e_I(t) = x(t) - \hat{W}_I^T \Phi_I(x(t)) \tag{2.11}$$

and, the action training is based on minimizing the derivative of the cost function to the chosen action. The purpose is to have the action error asymptotically goes to zero in an iterative process. This can be derived as,

$$e_A(t) = \frac{\partial U(t)}{\partial u(t)} + \gamma \frac{\partial J(t+1)}{\partial x(t+1)} . \frac{\partial x(t+1)}{\partial u(t)} \tag{2.12}$$

It should be noted that back-propagation is one of the main computational algorithms required to effectively train NNs in this scheme. Essentially, this algorithm uses the chain rule for calculating derivatives within the elements of the NN. It allows the error existing in the NN output introduced above to be used to correctly adjust the weights of the NNs. As the learning procedure progresses with respect to iterations, it should be able to achieve better representation of the model or policy that

is being implemented.

Remark: Optimal convergence of HDP or Approximate Dynamic Programming (ADP) in the case of general nonlinear systems has been discussed and proved in [49]. Additionally, It has been shown [50] that RL could be used as an optimal controller guaranteeing global optimal conditions for a non-convex functions.

2.4 AI-based Controller

Intelligent controllers are capable of learning and modifying their behavior while interacting with the system in form of *regression*. Here, We will focus specifically upon NNs as a type of AI-based controller that is capable of learning and controlling. With sufficient neurons and training process, NN is able to learn and represent any function [51]. This type of nonlinear control problem has been shown in literature that has considerable potential in power system stability study.

2.4.1 Application of Supervised Learning (SL)

The main application of SL-based NN control designs is the modeling of nonlineaities inherited in power system network equations. As a SL algorithm, the target or desired value for a controller is known at the time. This desired control action can be derived in the context of model reference adaptive control [52]. In the work of [52], two schemes of nonlinear control system have been presented:

- Direct methods
- Indirect methods

In direct method, the weights of NN controller is *directly* adjusted by the error of plant output and the reference. In the later, the controller training is performed based on the estimated parameters of the plant. Similar method has been employed in [53] to design an adaptive NN-based PSS. The input to the controller is the rotor speed of the generator and the output is the damping control fed to the excitation system. Controller is trained based on the NN identifier output error from desired speed, which

is constant. In [54] the idea has been applied to one-step ahead predicted output error. The same training is then performed by the chain rule through identification and control NNs. It should be noted that several NN models have been proposed in this area such as Generalized Neuron (GN) [54] or RNN [55] to address the generality, robustness, and efficiency requirements of such designs.

Another sets of designs in this area is devoted to employing the NNs to tune the conventional synchronous generators, such as in [56] for PSS. In [56] a RBFNN is used to generate the desired PSS parameters in real-time based on input vector of generator real power, reactive power output and terminal voltage. This type of adaptive design is trained offline for excessive number of operating points.

2.4.2 Application of Reinforcement Learning (RL)

The key advantage of RL-based controllers over SL methods is that the new knowledge in the online process can improve the training for further events or recursively at each iteration [45]. This feature manifests itself in the stability problem as a damping control schemes in addition to the case of cascading failures. Additionally, they are well-suited to perform optimal control algorithms, especially when multiple agents are involved. This technique has been center of attention for the wide range of power system applications in recent years. However, the research in this area is still weak due to reliability and practicality issues. RL-based designs such as ACDs have been shown to be more robust for wider operating regions and contingencies in comparison to the classical methods. In general, their main characteristics are:

- Requiring partial model of the system.
- I/O measurements of the system being sufficient for designing the controller.
- Wide operating regions and disturbances without prior knowledge.
- Depending on disturbance measurements that are not readily available.
- System measurements can be used but noisy measurements requires an extra

attention.

- The offline phase requires higher computational efforts.

The application of RL has been mainly investigated in wide-area control domain when coordination of multiple units is needed. This is presented in next section.

2.5 AI-based Wide-Area Control (WAC)

The only feasible way to implement the WAC has been to monitor and communicate states and control signals between each local substation and the control system through Wide-Area Monitoring (WAM) system [33]. Integration of monitoring and controlling systems in the power grid electricity infrastructure, has shown promising possibilities for more advanced stability control schemes based on timely detection of disturbances as they propagate through the network. With these improvements, new ideas on power system stability control specially in large scale oscillations has emerged [57]. In recent years, monitoring requirements, itself, has gained considerable attention by researchers [58, 59]. In these works, the requirements and challenges of WAC designed on WAM systems and the challenges in implementation is discussed.

Overall, WAC coordinates the actions of several units in power system including synchronous generators, FACTS devices, or renewable resources to gain global stability. It is very difficult for an analytical control techniques to perform a continuous supervisory level control of the system. This is partly due to

- Classical control schemes depend on a mathematical model of the plant, and this model is often based on linearization at a specific operating condition.
- Non-accurate component model parameters such as lines [60] or exciters [61] may be prone to variation due to the fault and have various impacts on the response of the system.
- As the control scheme moves from a local controller to a wide-area scheme, the complexity of the control system is increased. This makes such designs more

impractical.

- WAC performance is highly dependent on the quality of data provided by WAM systems. Wide-area designs are adversely affected by uncertain communication channels, packets drop, time delays and lack of global power system models and their time constants. Most of the works proposed in this area ([10,12,62]) treat the unknown parameters as constants.

In general, for classical methods stability and generality is not guaranteed. On the other hand, AI-based techniques can handle such problems, which will be discussed next.

2.5.1 Application of Reinforcement Learning

RL has been widely proposed in literature for Wide-area power system stabilization. For instance, [63] has applied ACD for a UPFC that provides auxiliary signals to the real and reactive power references of a UPFC series inverter in order to achieve enhanced damping of system oscillations. Further, this approach is used in [11] as a computational tool as an optimal damping controller for a gate-controlled series capacitor. The controller, using WAMs, for a gate-controlled series capacitor is used to provide damping of system modes. The design yields a fixed weight nonlinear controller, which is easier to implement in practical systems in comparison to conventional controllers. Further, [14] has focused on the design of the controller based on development time and hardware requirements for real-time implementation. A wide-area nonlinear damping controller is designed in this work using an existing Static Var Compensator (SVC). Particle swarm optimization is applied to tune the parameters of the SVC external damping controller but based on some linearized mathematical models of power systems. In [15] Q-Learning based real-time decentralized control scheme is proposed based on WAMs for excitation control of generators. Overall in these works, the transient stability of the power system is enhanced by mitigating

angle instability, meanwhile damping of power system oscillations is improved.

2.5.2 Wide-Area Identification

Application of NNs in nonlinear system identification, in particular in WAM-based designs is reviewed here. The efficient and accurate training of NNs to approximate functions has been an open topic for many years. It depends on several parameters such as choosing appropriate model, learning rate, noise in the data, size of the database, training algorithm, minimization algorithm, and so on. For instance in [10] an annealing learning rate scheme is used for NNs in order to ensure that NNs adapt themselves to the plant dynamics quickly and converge to new operating points.

As mentioned in section 2.3.1, the main characteristic of MLP is that there are no connections between the neurons on the same layer [64]. Additionally, in such a network every output error has a direct impact on all the weights of the input weight matrix, i.e. the outputs of MLP interact with one another and the error in each one affects the others. Hence, a control network designed using NNs creates unwanted interactions between the controllers, whereas, the idea behind the WAC is to augment each local controller based on the effect of only that controller on the global cost function or objective function [10]. Ref. [10] has tackled this issue by using a FNL NN.

The error back-propagation algorithm can be utilized to solve RL problems. Usually, FFNN is used that can only gain a static mapping of the I/O space. In order to be able to model dynamic systems, development of a NN that is able to store internal states is required. These types of NNs are called RNN under the umbrella term Dynamic NN, and their main characteristic is the internal feedback or time-delayed connections. Although RNN is biologically more realistic than FFNN, it is more difficult to train them due to the problems of exploding or vanishing gradients. The difference between a conventional RNN and SRN is that the feedback in RNNs are time-delayed, whereas, in SRNs they are instantaneous [65]. The SRN uses much

higher sampling rates to emulate instantaneous feedback. In order to implement a SRN in real-time, the simultaneous recurrences have to be carried on several times within one time step of the measurements. The SRN-based WAC system have been implemented with a new training algorithm and two step training approach in [65]. Further, [66] has expanded the design to a novel four dimensional scalable multi-rate cellular NN architecture as WAMs. RNN is used as computational engine for each cell as they have dynamic memory. By using information from PMUs that are optimally located in a power system, each layer predicts a state variable for one or more time steps.

2.5.3 Scalable Designs

One of the main challenges of WAC designs, in specific AI-based systems, is the problem of scalability and dimensionality. Depending on the number of WAM measurements and the signals to be analyzed, the volume of data for a typical system could be enormous. Various techniques has been proposed in power system research groups to tackle this problem based on clustering and model reduction techniques to overcome the problem of excessive data analysis for model reduction. Using Feature selection or unsupervised learning techniques signals can be grouped according to their resemblance to each other reducing the computational cost. Due to the nature of rotor angle stability, most of the WACs proposed are designed based on center of inertia (COI), center of angle (COA), or center of speed (COSP) [67].

The common coherency method based on pre-specified number of areas is also used for the purpose real-time ACD based WACs. For instance, in [13], a new concept called a "virtual generator" is introduced which is simplified representations of groups of coherent synchronous generators in a power system. It allows WACs to exploit the realization that a group of coherent synchronous generators in a power system can be controlled as a single generating unit for achieving wide-area damping control objectives. This implementation is made possible by the availability of WAMs

from PMUs. Also, in [68] we have used the COI-based signals to monitor and control the pre-defined areas of the system, which limits the observability to inter-area oscillations.

2.5.4 WAM Constraint Consideration

WAC systems are highly dependent on the performance of the communication infrastructure, without which the functionality of them will not be achieved. Knowing that time delay affects wide-area power system stabilizer design, the consequence of delayed input data and output signals in WAC systems should be considered and modeled. Also, the performance requirements posed on by time availability for decision making and control actuation needs to be elicited. The characteristics of time delays could be constant, bounded, or even random, depending on the network protocols adopted, distance, and the chosen hardware and could be in the range of 7ms to 1s. PMU data delay in WAM systems and their nature has been analyzed in several research works; see [18, 19, 69, 70] and references therein.

RL algorithms are well suited, as mentioned before, to tackle these uncertainties in the system. Ref. [10] has designed an ACD-based controller to improve the damping of the rotor speed deviations of the synchronous machines by providing auxiliary reference signals for the AVR of the generators as well as the line voltage controller of the STATCOM. RBFNN-based identifier is presented in this work to predict the states in real-time in presence of transport lags associated with the present communication technology for WAM. The results provided indicates that the proposed WAC improves the damping of the rotor speed deviations of the generators during large scale disturbances. Ref. [12] take advantage of ACDs in including the communication delays in implementation of real-time WAC design with a single SRN. The NN serves a dual purpose of continuous identification of the power system dynamics and generation of appropriate damping control signals. Through such design damping of several modes of oscillations in power system is provided.

2.5.5 Real-time Implementation

Real-time implementation aims at bringing these AI-based designs one step closer to practical applications. This technique has been successfully developed in real-time benchmark in several works [10, 12, 71]. For instance, in [71] the RL-based neuro-controllers for turbo generators in a multi-machine power system has been implemented on the Innovative Integration M67 card consisting of the TMS320C6701 processor. The results showed robustness in presence of system operation changes. Overall, these works have encountered some challenges including data communication quality, WAC calculation speed, and so on.

2.5.6 Transient Stability Enhancement Controller

As discussed in previous sections TSA is mainly used for corrective or emergency control actions. It is shown in literature that, the WAM temporal information can further be used in WAC designs to perform real-time transient stability enhancement, which can improve the power transfer capability of a transmission system and prevent the system from generation or load disconnection, or catastrophic failure following a sequence of disturbances in the system. Article [72] have used the RL for preventing cascading failure (CF) and blackout in smart grids by acting on the output power of the generators in real-time. This article makes use of the state-action policy update feature of RL algorithm, as it can learn from interactions with the system.

2.6 AI-based Techniques in Modern Power System

With the increased penetration of renewables in transmission system, the effective inertia of the system will be reduced and system rotor-angle stability following large disturbances could significantly be affected [73]. Several works have been done to address the impacts of renewable resources on power system stability [73–79]. Generally, it is believed that the renewable *type* do not significantly affect the power system oscillations. Rather, the *penetration level* will have a damping effect due to reduction

in the size of synchronous generators that engage in power system oscillations [77].

The majority of these resources are interfaced to the grid using Voltage-Sourced Converter (VSC) units. Voltage control of a VSC in the dq-framework can be achieved in a nested loop based on an inner current control loop and an outer voltage control loop. The controller of the inner loop regulates the converter current, and controller of the outer loop regulates the output voltage [80–82]. Furthermore, Maximum Power Point Tracking (MPPT) is usually applied to the generation control to extract the maximum allowable power from the wind turbines or PV arrays [83], along with power sharing controllers.

The type of power control employed for the renewable generation directly affects the rotor angle and speed of synchronous generators [80]. For instance, [84] shows that when active power flows change, the way that the wind turbine provides reactive power support to the system is critical in maintaining rotor angle stability of conventional units in the system and minimizing the deviation of field voltage. Same applies to the active power control, since the oscillations are produced by active power differences between generation and consumption. Therefore, the implementation of appropriate control strategies in renewable sources, particularly the terminal voltage control, can lessen the power requirements of conventional synchronous units and help to mitigate large rotor angle swings. There are several techniques applied for this purpose such as optimal control methods, robust methods, Energy function methods, and so on. Additionally, AI-based designs have been investigated for the renewable-integrated power system rotor angle stability control.

2.6.1 Supervised Learning (SL) Control

In the area renewable control, AIs such as NN and FL methods have been also successfully applied in different applications [85–90]. Generally, in the area of renewable energy resources, NNs are mainly used as the prediction tool for generation forecasting along with load prediction in microgrid application [85]. A comprehensive

review on the application of NN in renewable energy systems can be found in [86]. As a controller, in [87], a neuro-fuzzy gain tuner is proposed to control a laboratory DFIG. The input for each neuro-fuzzy system is the error value of generator speed, active or reactive power. In [88], a method to design an adaptive fuzzy system for for maximum energy extraction from variable speed wind turbines is proposed and tested. The proposed control techniques have low memory occupancy and high learning capability, having advantage over classical control methods; thus, could be well implemented on a micro-controller. In [89], NN has been employed as rotor and grid side convertor controllers gaining better dynamic characteristics in comparison to conventional PID controllers. In the [90] application of NNs to control a grid-connected rectifier/inverter is investigated. A NN-based control strategy is presented and tested in this work in a more practical nested-loop control condition. The NN implements a DP algorithm and is trained by using back-propagation through time. Overall, NNs has been shown in these works that have better performance in comparison to conventional techniques regarding system's dynamic responses.

2.6.2 Reinforcement Learning (RL) Control

RL has also been subject of renewable control in recent years. The application of NNs as an intelligent control algorithm has been shown in [91] in microgrid with multiple renewable resources. In general RL can provide predictive, optimal, adaptive control designs for renewable-integrated power systems. A DHP-based control system in a system wide adaptive predictive WAC scheme is used to ensure the dynamic performance and voltage dynamics of the micro grid as the system operation conditions change. Ref. [80] has proposed controller based on ADP techniques on the bench mark of NNs to approximate the optimal control policy according to the interaction between the controller and the power plant. The method is developed for the DFIG-based wind farm to improve the system transient stability under fault conditions and has shown effective results. Furthermore, in [92], a RBFNN is designed

for WAM that identifies the I/O dynamics of the nonlinear power system with power system stabilizers, a large wind farm, and multiple flexible ac transmission system (FACTS) devices. the proposed WAC design has shown better performance during transient events, but without considering either multiple wind farms or communication time-delay compensation.

CHAPTER 3: OPTIMAL AND ROBUST WIDE-AREA CONTROL DESIGN

3.1 Introduction

Wide-Area Control (WAC) systems, as mentioned in previous chapter, are highly dependent on the performance of the communication infrastructure, without which the functionality of them will not be achieved. Knowing that time delay affects wide-area power system stabilizer design, the consequence of delayed input data and output signals in WAC systems should be considered and modeled. Also, the performance requirements posed on by time availability for decision making and control actuation needs to be elicited. PMU data delay in WAM systems and their nature has been analyzed in several research works; see [18, 19, 69, 70] and references therein. The aim of this chapter is to design a WAC that is robust to such delays and data quality, predictive, and optimal as it involves several agents.

In general, Reinforcement Learning (RL) problems are well suited to tackle aforesaid issues. This chapter focuses on the solving method for such problem in order to achieve optimal and robust performance of WAC system. Unlike the works done by other researchers in power system for RL-based WAC designs [5, 10, 12, 17, 71, 93–95] which focuses on Dynamic Programming (DP) to solve the sequential optimization problem, this chapter uses average Temporal Difference (TD(λ)) method [96–100] to achieve the cost function forecasting and learning requirements of WAC design. TD methods are considered as a bridge between DP and Monte Carlo (MC) based designs. By adjusting coefficients of future returns, λ, TD(λ), constitutes this bridge by controlling the bias/variance trade-off, and significantly speeding up the learning. It is understood that, larger values for the TD parameter suffers larger variance in the updates, but also enjoys lower bias [101]. This technique, shown to be implementable

online in an inexpensive manner [102], makes the approximation of the cost function naturally converge faster to optimal value in forward view. On the other hand, Eligibility Trace (ET) method, as in backward view, uses a memory parameter to mark states that are eligible for learning, reduces the process load. Thus together with ET and TD, the learning speed, priority weights for states, and convergence of learning can be optimized.

In this chapter, we present a unified optimization method, where the cost function is approximated by the online TD algorithm. A new online method is proposed in Adaptive Critic Design (ACD) context to achieve the same updating algorithm as of offline TD(λ) in order to do the estimation during the episode. Chapter focuses on ACD controllers with new critic scheme based on moving target learning approach. The theoretical and application framework of this design for WAC applications in power grid is the main contribution of this chapter. A new implementation method is discussed based on the proposed architecture and the framework is evaluated using an experimental test bed with real-time digital simulator (RTDS).

The chapter is organized as follows. Section 3.2 provides a brief mathematical preliminaries and problem formulation and section 3.3 discuss the proposed methodology. In section 3.4, a WAC architecture is illustrated and Section 3.5 discusses the implementation method on 2-area system. Section 3.6 presents the real-time implementation and case studies and section 3.7 concludes the chapter.

3.2 Mathematical Preliminaries and Problem Formulation

Consider a general class of nonlinear system at time t of the following form:

$$\dot{x}(t) = af(x, W, t) + bg(x, W, t)u(t) + d(t) \tag{3.1}$$

$$u(t) = h(x(t)) \tag{3.2}$$

where, $x \in \Re^n$ is the state vector, $f(.) \in \Re^n$, $g(.) \in \Re^{n \times m}$, $h(.) \in \Re^m$ are unknown nonlinear functions, W is the unknown parametric change, $u \in \Re^m$ is the control input, a and b are parametric coefficients that are not changing, and d is unknown bounded disturbance ($\|d\| \leq d_m$). The objective is to track the system output to a reference command.

$$\dot{x}^*(t) = a^* x^*(t) + b^* r(t) \tag{3.3}$$

where, r is the input command and a^* and b^* are reference coefficient matrices. The task of the controller is to track a desired state vector as specified by the reference input with the state vector remaining bounded.

Based on RL approach, it is desired to find the control action which minimizes the cost-to-go function given as

$$J(x(t)) = \sum_{k=0}^{\infty} \gamma^k U(x(t+k), u(t+k)) \tag{3.4}$$

where, $\gamma \in (0, 1]$ is the discount factor, and,

$$U(x(t), u(t)) = \Delta x(t)^T Q \Delta x(t) + u(t)^T R u(t) \tag{3.5}$$

where $\Delta x(t) = x(t) - x^*(t)$.

3.2.1 Solutions to the Reinforcement Learning Problem

In RL-based control, the objective is to develop an optimal control action u^* with convergence to the optimal policy or cost function, J^*. In general, TD estimates J^* by updating its policy with respect to estimation error designing e in the direction of

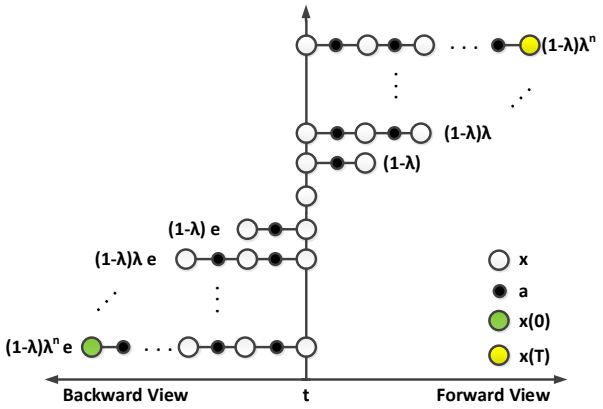


Figure 3.1: Return weights in TD(λ)

the n-step return of x, denoted as $R^{(n)}$ as,

$$\Delta J(x(t)) = \alpha e(t) \tag{3.6}$$

$$e(t) = R^{(n)}(x(t)) - J(x(t)) \tag{3.7}$$

$$R^{(n)}(x(t)) = \sum_{k=0}^{n-1} \gamma^k U(x(t+k), u(t+k)) + \gamma^n \hat{J}(x(t+n)) \tag{3.8}$$

where, α is a constant step-size parameter. In particular, DP is equivalent to 1-step TD, considering just one time step ahead prediction, and n-step TD ($n = T$ as the terminal stage) is the same as MC using the latest possible prediction as the target which allows less dependency on system model. Fig. 3.1 shows the parameter λ characterizes how fast the exponential weighting in each time steps TD error falls off in backward and forward views.

$$\Delta J(x(t)) = e(t) \tag{3.9}$$

$$e(t) = U(x(t), u(t)) - \gamma J(x(t+1)) \tag{3.10}$$

3.2.1.1 Forward View

Forward view of TD(λ) could be used for providing backups toward any average of n-step returns.

$$R^f = (1-\lambda)\sum_{n=1}^{T-n-1}\lambda^{n-1}R^{(n)} + \lambda^{T-n-1}R^{(T-t)} \tag{3.11}$$

By rearranging the above equation, change of policy in TD(λ), denoted as, ΔJ^f is estimated by TD error, $e(k)$, in forward view as

$$\Delta J^f(x(t)) \approx \alpha\sum_{k=t}^{T-1}(\lambda\gamma)^{k-t}e(k) \tag{3.12}$$

$$e(k) = U(x(t+k), u(t+k)) + \gamma J(x(t+1+k)) - J(x(t+k)) \tag{3.13}$$

3.2.1.2 Backward View

In the backward view, the ET for state $x(t)$ is denoted as $et(x(t))$, and is defined as,

$$et(x(t)) = \gamma\lambda et(x(t-1)) + 1. \tag{3.14}$$

leading to,

$$\Delta J^b(x(t)) = \alpha et(x(t))e(t) \tag{3.15}$$

At each step, the ET for states that are not visited are decayed by $\lambda\gamma$. The traces are said to indicate the degree by which each state is eligible for undergoing learning. The reinforcing events are the moment-by-moment one-step TD errors. Full description of forward and backward views of TD(λ) can be found in [8].

3.3 Proposed Moving Reference TD(λ)

Let us split the cost function derivative as,

$$\Delta J(x(t)) = \Delta J^{fb}(x(t)) + \Delta J^{i}(x(t)) \tag{3.16}$$

where, the first element is the update due to forward and backward views, and the second one is with respect to action and its relation to the estimated cost function toward optimal value. For on-line implementation, we hypothesize that the forward view can be related to the backward view if a moving reference of optimization horizon is considered. With this property, the reference for policy update can be set at any time step before the current time and can be added to the forward view from the current time. This is similar to a Receding Horizon Control (RHC) problem, but with a moving window that can be enabled forward and backward in time. With this in mind from (3.11), a truncated R^f at time k, denoted as R^f_k can be defined as

$$R^f_k(x(t)) = (1-\lambda)\sum_{n=1}^{k-n-1} \lambda^{n-1} R^{(n)}_{t+n-1}(x(t)) + \lambda^{T-n-1} R^{(k-t)}_{k-1}(x(t)) \tag{3.17}$$

where, k is the time step that TD(λ) is truncated and $R^{(n)}_k(x(t))$ is defined as the n-step return of U and J corresponding to the episode k as

$$R^{(n)}_k(x(t)) = \sum_{j=0}^{k-1} \gamma^j U(x(t+j), u(t+j)) + \gamma^k J_k(x(t+k)) \tag{3.18}$$

where, J_k is the estimated cost function at step k, set as terminal state of the optimization window. The policy used in the n-step returns is now a specific time index. By setting k, the current time step, the truncated λ-returns from all previous time steps are updated, such that they are now truncated at the current time step. Thus, reference of the policy can be adjusted at any time step before the current time

step, making the learning from the backward view more practical. Rearranging the equations above, leads to change of policy, as

$$e(k) = U(x(t+k), u(t+k)) \quad + \quad \gamma J_k(x(t+1+k)) - J_{k-1}(x(t+k)) \tag{3.19}$$

$$\Delta J^f(x(t)) \quad \approx \quad \alpha \sum_{k=t}^{T-1} (\lambda\gamma)^{k-t} e(k) \tag{3.20}$$

Using (3.14) to refer to et, we can get the error of estimation over time for backward view as follows:

$$\Delta J^b(x(t)) = \alpha e(t) \sum_{k=0}^{t} (\lambda\gamma)^{t-k} I_{x(k)} \tag{3.21}$$

where, $I_{x(t)}$ is an identity indicator function, equal to 1 if $x = x(t)$ and equal to 0 otherwise. The combination of forward and backward views at the time t leads to:

$$\begin{aligned} \Delta J^{fb}(x(t)) &= \alpha \sum_{k=t}^{T-1} (\lambda\gamma)^{k-t} e(k) \sum_{k'=0}^{t} (\lambda\gamma)^{k'-t} I_{x(k')} \\ &= \alpha \sum_{k=0}^{t} \sum_{k'=t}^{T-1} (\lambda\gamma)^{k'-k} e(k) I_{x(k')} \end{aligned} \tag{3.22}$$

After reaching to the terminal state, the next update sequence is started as iteration for action calculation. The iteration scheme is in a form of incremental optimization that requires iterating between a sequence of action policies u^i that optimizes the J^i. Algorithm is initialized with action policy, u^0, derived from previous time step. This can be represented as

$$\Delta J^i(x(t)) = J^{i+1}(x(t)) - J^i(x(t)) \tag{3.23}$$

$$J^{i+1}(x(t)) = U(x(t), u^i(t)) + J^i(x(t+1)) \tag{3.24}$$

$$u^{i+1}(t) = \arg min_u (J^{i+1}(x(t), u)) \tag{3.25}$$

Note that i is the value iteration index, whereas k is the time index.

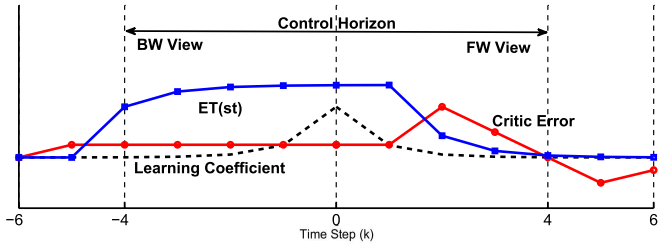


Figure 3.2: Elements of forward and backward views in proposed architecture

The architecture can be compared with RHC, in the sense that it performs the optimization of the control sequence over a finite time horizon length instead of infinite time span [103]. Considering forward and backward view a window around the current time, t_0, can be developed with the proposed method. Fig. 3.2 shows the key elements of the proposed control scheme in a moving reference framework with predicted errors, ET, include the impact of this approach on learning algorithm. Viewing T as a control window, and current time step, t at the center of this window, $(\lambda\gamma)^k$ would be coefficients of the error at each time step. Upon analyzing the figures, it can be noted that the proposed method can provide weighted values for the value function based on the moving reference. In other words, time steps closer to current time would have more influence on learning than farther time values. Also, as the time horizon can be changed along with the time step, the method has more bandwidth to move from temporal to spatial complexity architectural patterns.

3.4 Proposed TD-based WAC Design

In this section, NN design for RL problem is illustrated, proposing the backward and forward views in TD on a finite horizon around the current time step. The architecture discussed in this section is evolved from the ACD approach. The fundamental equations and algorithm in this architecture is developed based on the moving target TD(λ) methodology in previous section towards a WAC design for power grid with better applicability. Three NNs have been implemented in the WAC: Wide-Area NN Identifier (WANNID), Critic NN, and Action NN. Each network training and control

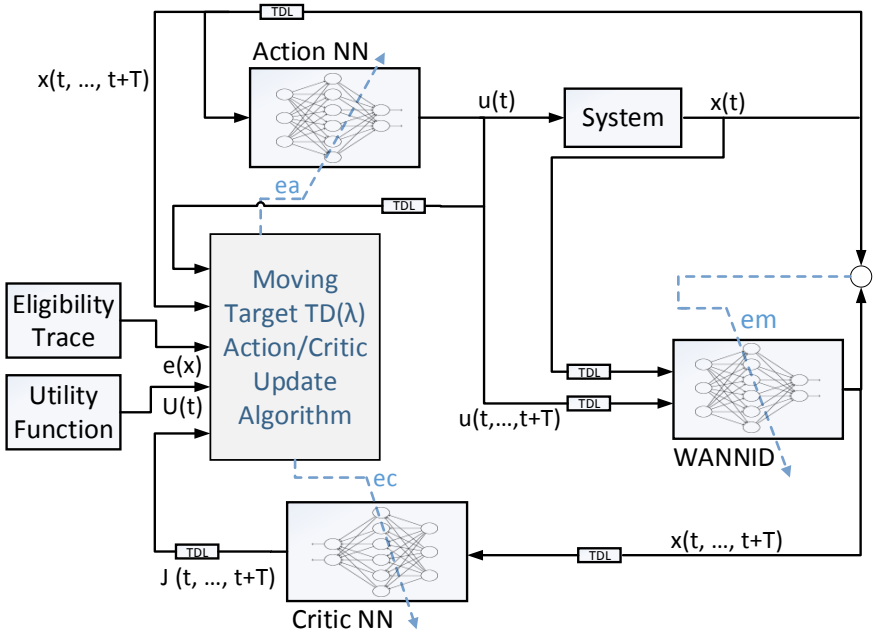


Figure 3.3: Proposed WAC architecture based on moving reference TD(λ)

functions are discussed next. Then the training algorithm, and their relationship them as depicted in Fig. 3.3 will be discussed in detail.

The purpose of the WANNID is to model the equivalent power grid of the area of interest with multiple generators and estimate the desired states through WAM. WANNID is used to predict the future behavior of the system which is then used to control the system through cost function approximation. To start, WANNID is trained off-line with a random set of inputs. Further, it is trained on-line tracking the actual output of the system, and updating the weights based on identification error, as

$$e_m(t) = x(t) - f_M(x(t-1), u(t-1)) \tag{3.26}$$

The Critic NN learns the objective function that is used by Action NN to optimize the overall goal. The goal is expressed as a form of Lyapunov stability function for the system, in this case, the power system. The inputs to the Critic NN are action

and states, and the output is the cost function. The Critic learns this approximation by minimizing the error expressed by the summation of

$$||E|| = \sum\nolimits_t e(t)^T e(t) \tag{3.27}$$

over time period t, where the error value of e at each time period using the proposed TD(λ) based on (3.22) is given by,

$$e(t) = \alpha \Delta J^{fb}(x(t)) \tag{3.28}$$

As a result, change of Critic weights, $\Delta W_C(t)$, at time t can be achieved from the change of policy rate by means of gradient descent via back-propagation through the NN. This,

$$\Delta W_C(t) = e(t) \frac{\partial J^{fb}(x(t))}{\partial W_C(x(t))} \tag{3.29}$$

where, $W_C(t)$ is the weight vector of the Critic NN, which estimates the cost function at time t. Further, considering FFNN structure,

$$\Delta W_C(t) = e(t) \hat{\Phi}(x(t)). \tag{3.30}$$

From (3.22) it can be seen that, we are capturing the changes in the estimated states and back up state transitions, with respect to state returns, priorities of the state, and state predictions. The weight update feature is proportional to a decaying ET, where the backward update for ET can be defined as,

$$et(x(t)) = \lambda \gamma et(t-1) + \Phi_C(x(t)) \tag{3.31}$$

Applying (3.22) in (3.30), the weight vector of the current time step $W_{C|k}(t)$ is de-

termined by sequentially performing TD backups and can be incrementally defined as,

$$\begin{aligned} W_{C|k+1}(t) &= W_{C|k}(t) \\ &+ \alpha et(x(t))\{U(x(t+k),u(t+k)) + \gamma[W_{C|k}(t)]^T\Phi_C(x(t+1)) \\ &- [W_{C|k-1}(t)]^T\Phi_C(x(t))\}\Phi_C(x(t)) \end{aligned} \tag{3.32}$$

where, k is to indicate the time step when the true on-line TD(λ)-returns are truncated. This represents the moving reference TD(λ) backward view of the truncated λ-return algorithm.

The goal of the Action training is to minimize the sum of the derivative of the costs derived from the Critic NN, with respect to the Action network. This can be represented as:

$$\frac{\partial J(t)}{\partial u(t)} = 0 \tag{3.33}$$

Hence, error is defined as,

$$e_A(t) = (1-\lambda)\sum_{n=0}^{N}(\gamma\lambda)^k\left[\gamma\frac{\partial J(x(t+k+1))}{\partial u(t)} + \frac{\partial U(x(t+k),u(t+k))}{\partial u(t)}\right] \tag{3.34}$$

$$e_A(t) = (1-\lambda)\sum_{n=1}^{N}\left[\gamma^k\lambda^{k-1}(\frac{\partial J(t+k)}{\partial u(t)} + \frac{\partial U(t+n)}{\partial u(t)}) + \frac{\partial U(t)}{\partial u(t)}\right] \tag{3.35}$$

where,

$$\frac{\partial U(x(t+1),u(t+1))}{\partial u(t)} = 2x(t+1)\frac{\partial x(t+1)}{\partial u(t)} \tag{3.36}$$

Applying the chain rule for derivatives results in

$$\frac{\partial J(x(t+k))}{\partial u(t)} = \prod_{k=0}^{N}(\frac{\partial J(t+k)}{\partial x(t+k)} + \sum_{m=1}^{M}\frac{\partial J(x(t+k))}{\partial u_m(t+k)}\frac{\partial u_m(t+k)}{\partial u(t+k)})(\frac{\partial x(t+1)}{\partial u(t)}) \tag{3.37}$$

where, m is the number of the outputs of the Action NN, and N is the window length of the forward view, and, $\partial x(t+1)/\partial u(t)$ and $\partial J(x(t))/\partial x(t)$ are respectively, derived through back-propagation of WANNID and Critic NN. The update to the weights of the Action NN, applying least mean square minimization, may be written as:

$$W_A^{i+1}(t) - W_A^i(t) = -\alpha(\frac{\partial u(t)}{\partial W_A^i(t)})e_A(t) \tag{3.38}$$

where, α is a small positive learning rate, and W_A^i is the Action weight at iteration i in action-cost function sequence leading to new cost function.

In order to perform the TD-based ACD Critic and Action training, an iterative procedure is used (Table 3.1), where the Critic weights at each iteration, $W_{C|k}$, are adjusted while holding the Action's weights, W_A, fixed, and vice-versa. It should be noted that, conventional TD(λ) is an update target that is based on the full experience of sequence, while, the proposed TD(λ) algorithm is an inexpensive incremental method, and doesn't require storage of all observed states and rewards based on ET. The entire process can be defined as a simultaneous optimization problem. Here, an iterative procedure is used to get the optimization of the Critic and Action to converge.

3.5 Implementation as a WAC

In this section implementation of the proposed WAC is assessed on a two area test system shown in Fig. 3.4. As a WAC, the proposed architecture is tested to reduce local and inter area oscillations in the power system generators.

3.5.1 Two Area Test System

A 5-machine 8-bus power system with dynamic equivalent frequency dependent generators is modeled using electro-magnetic transient simulation (EMTP) software PSCAD. This network consists of five generating units such that three generators, G_1, G_2, and G_4 are of larger size than G_3 and G_5. All generators are equipped with

Table 3.1: Neural Network training procedure based on moving reference TD(λ)

Step	Action	Comment
1	$i = 0$, J_0, u_0, $et(0) = 0$	Initialization
2	$k = 0$	Initialization of time horizon
3	$U(k) = U(x(t+k), u(t+k))$	Utility function at time $t+k$ (3.5)
4	$u^i(k) = f_A(x(t+k), W^i_{A\|k})$	Action output at time $t+k$
5	$J^i(k) = f_C(x(t+k), W^i_{C\|k-1})$	Critic output at time $t+k$
6	$x(t+k+1) = f_M(x(t+k), u^i(t+k), W_M)$	Identifier output at time $t+k$
7	$J^i(k+1) = f_C(x(t+k), W^i_{C\|k})$	Critic output at time $t+k$
8	$e(k)$	Critic errors by (3.19)
9	$et(x(t+k))$	ET update by (3.14)
10	$\Phi(x(t+k))$	State function of FFNN
11	$W^i_{C\|k}$	Update Critic Weights by (3.32)
12	$\partial J^i(k)/\partial u^i(k)$	Equation (3.37)
13	$\partial U(k)/\partial u^i(k)$	Equation (3.36)
14	e_A	Action errors by (3.35)
15	$W^i_{A\|k}$	Update Action Weights
16	$k = k+1$	Repeat until $k = T$
17	$i = i+1$	Repeat the loop from step 2

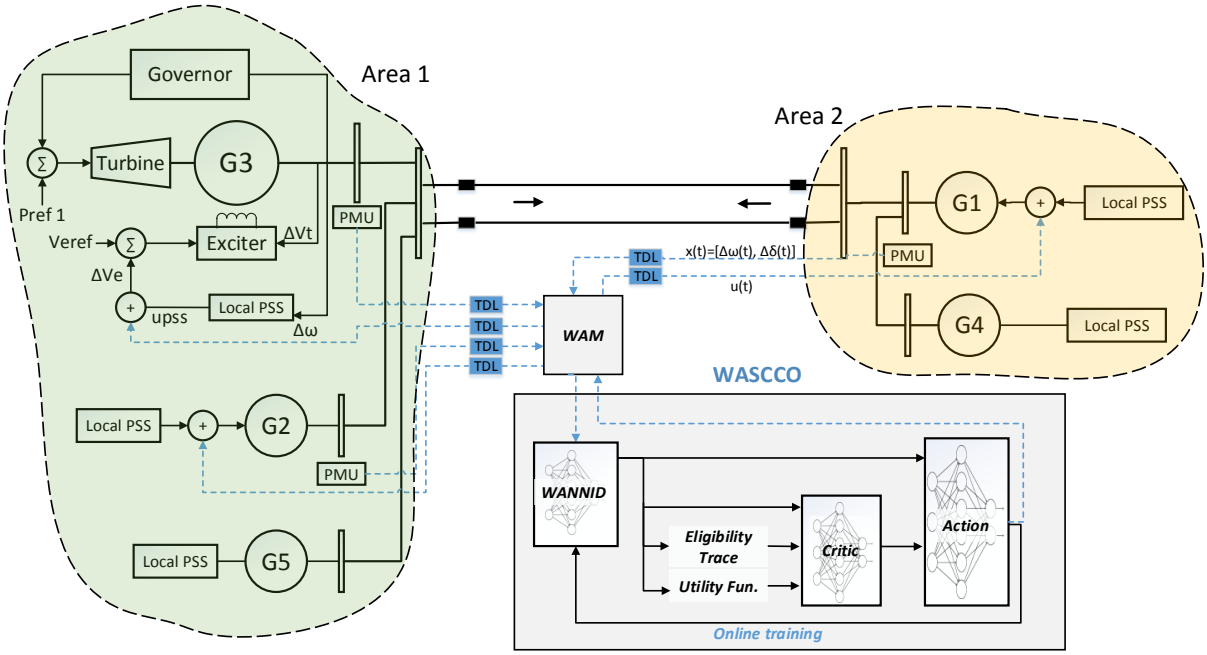


Figure 3.4: Diagram of two area power system with TD(λ)-based WAC

Table 3.2: 2-Area system PSS parameters

Generator	K_s	T_s	T_1	T_2	T_3	T_4
G_1, G_2	1	3	0.3	0.1	0.3	0.1
G_3	0.07	3	0.07	0.03	0.07	0.03

governors, exciters, AVRs, and conventional PSS. Parameters of all generators, governors, exciters, AVRs, transmission lines, loads, and operating conditions are given in [9]. An important consideration in this test is that, three of the five generators, G_2, G_3, and G_5 are considered to form one area, while generators G_1 and G_4 form the second area. Under normal conditions, each area serves its own load and is almost fully loaded with a small load flowing over the tie-lines. The consideration and procedure used for the PSS are similar to those found in [9] and the parameters are given in Table 3.2 with the transfer function of,

$$u_{pss} = K_s \frac{sT_s}{1+sT_s} \frac{1+sT_1}{1+sT_2} \frac{1+sT_3}{1+sT_4} \Delta\omega(s) \tag{3.39}$$

The generator, itself, is modeled as a fifth-order dynamic model. Details of the generator model is presented in [104].

The main focus of this architecture is to show the ability to damp inter-area mode oscillations for WAC system in presence of delays and nonlinearities of power system. This power system is built on the infrastructure established on the WAM system with consideration of time delay. All PMU equipped generators are monitored with rotor angle (δ) and the rotor speed (ω) measured and transmitted to the WAC center via the PMU data collection system, Phasor Data Concentrator (PDC). The horizon of backward and forward views of the truncated λ-return algorithm is set based on the delays in the WAM system. Considering receiving measurement signals being delayed for t_d, and sending action signals through the WAM system encountering time delay t_d as well, one can set the horizon of the optimization as [-t_d; t_d]. Signals are selected to monitor the the rotor angle stability. The utility function for quantifying the

Table 3.3: Configuration of Neural Networks for 2-Area TD(λ)-based WAC

NN	Inputs	Delays	Hidden	Outputs	Input Signals	Output Signals
WANNID	9	3	25	6	w,δ,u	w,δ
Action	6	3	15	3	w,δ	u
Critic	6	3	20	6	w,δ	J

stability level of the power system is proposed as summation of utility function for each monitored generator as,

$$U(t) = \sum_{i=1}^{3} U_i(t) \tag{3.40}$$

where,

$$U_i(t) = -0.8(\Delta\delta_i(t))^2 - 0.4(\Delta\omega_i(t))^2 \tag{3.41}$$

where, i is the generator index.

3.5.2 WAC Training Methodology and Assessment

Fig. 3.4 shows the block diagram for the proposed controller architecture. Training method is as follows.

3.5.2.1 Offline Training

Offline training is performed to set the initial weights for online implementation. For this a batch learning structure with random inputs and related outputs of the power system model is captured. The training of Action is initialized with the target of local PSS. Next, Critic NN is updated based on the proposed method, followed by Action update, alternatively. Once the Critic NN's and Action NN's weights have converged, the Action NN is connected to the generator's exciter to replace the PSS. Table 3.3 provides the NNs parameters used which are identified in a heuristic manner.

3.5.2.2 Online Training

The process of online training starts by monitoring the states to incrementally train the WANNID followed by Critic and Action update iterations (Fig. 3.4). For all the simulations, a real-time experimental test bed is used, which will be presented later. Parameters of the moving reference TD update method such as et, λ, and γ are set, initially to positive constants. Then, future states and actions are calculated by the trained WANNID and Action NN, for the use in the forward view of the Critic. Subsequently, these variables, alongside with calculated utility functions and ET as the backward view are fed to the Critic training process. Future cost functions are calculated and Action training is performed to optimize the cost function over the moving horizon. Process terminates when the Action and Critic errors are within acceptable range.

In order to assess the performance of the proposed method, first a mechanical torque change has been implemented on G_3 at $t = 6s$ in PSCAD causing the states to oscillate as shown in Fig. 3.5-a. As mentioned in section IV, Critic training in backward view has been implemented using ET to eliminate the redundant computation. At each time step, the ET for all states are decayed by $\lambda\gamma$, and the ET for the active states are incremented by $\Phi(x(t))$ (Fig. 3.5-b). In this simulation, λ is set to 0.7, and γ is 0.4. As it is shown, when there is a local fault in the area 1, the states related to this area are activated. On the other hand, forward view in Critic training has predictive feature. Considering the time step of 2.5 ms for the controller process, and total transmission delays of 0.02 s, the horizon of 4 time steps ahead and before has been used to develop proper time horizon. In Fig. 3.6, returns for current time step and 3 next steps has been shown. As it can be seen, return for further steps can predict the cost function, J^*, better than one step prediction. In order to improve the accuracy of n-step prediction, TD(λ) for predicted and previous states are considered. Critic and Action's update errors of the conventional Critic of WAC, TD(0), and the proposed

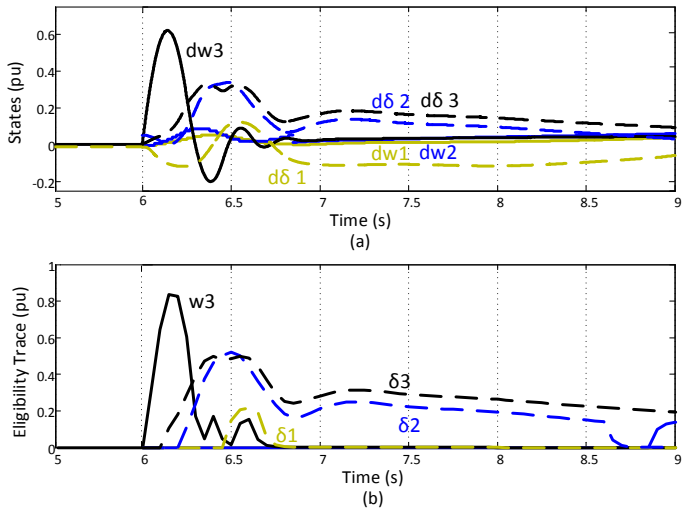


Figure 3.5: Mechanical torque change of G_3 at t=6 s causing (a) oscillation of the states (b) eligibility trace of the states

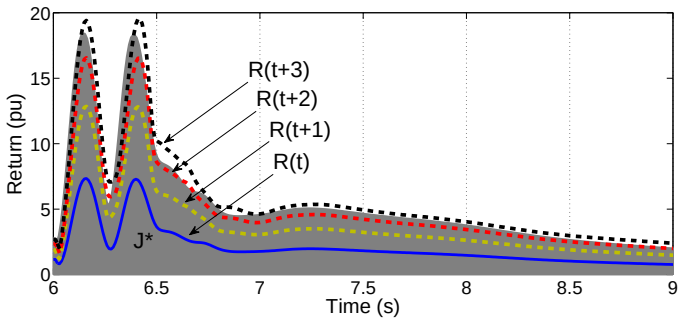


Figure 3.6: Forward view of returns and the optimal cost function

one for a horizon window from optimal values are presented in Fig. 3.7. J^* has been derived offline for 1000 epochs for this case study, and Critic and Action NN optimal weights, W^*, has been calculated based on this value. Fig. 3.8 depicts the deviation of the weights in conventional and proposed method from the optimal value, W^*. As it can be seen, the proposed method showed less error in weight changes.

3.6 Real-Time Implementation

In order to study the response of the system during inter-area mode of oscillations, we next adapted the controller in real-time to the two power system models. In previous section, simulations were performed for a control architecture in a MATLAB coding environment and power system modeling in PSCAD (Path4 in Fig. 3.9). To

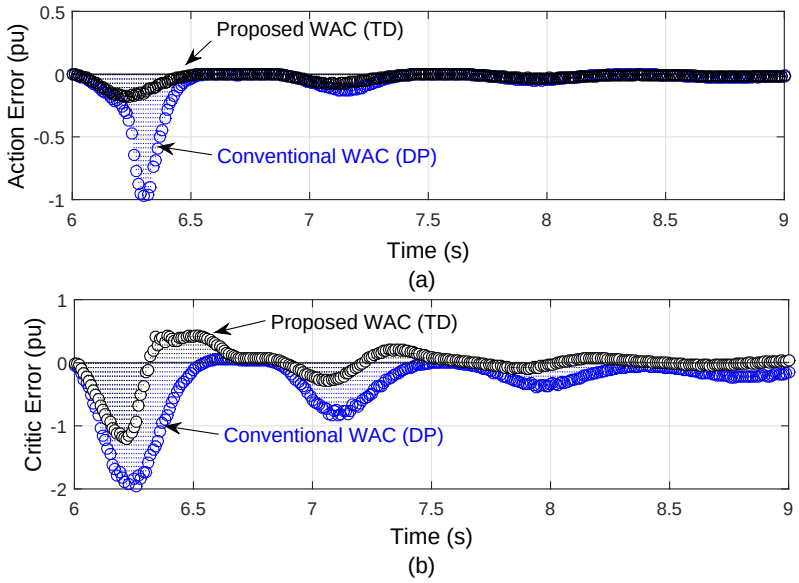


Figure 3.7: Update error (a) action NN (b) critic NN

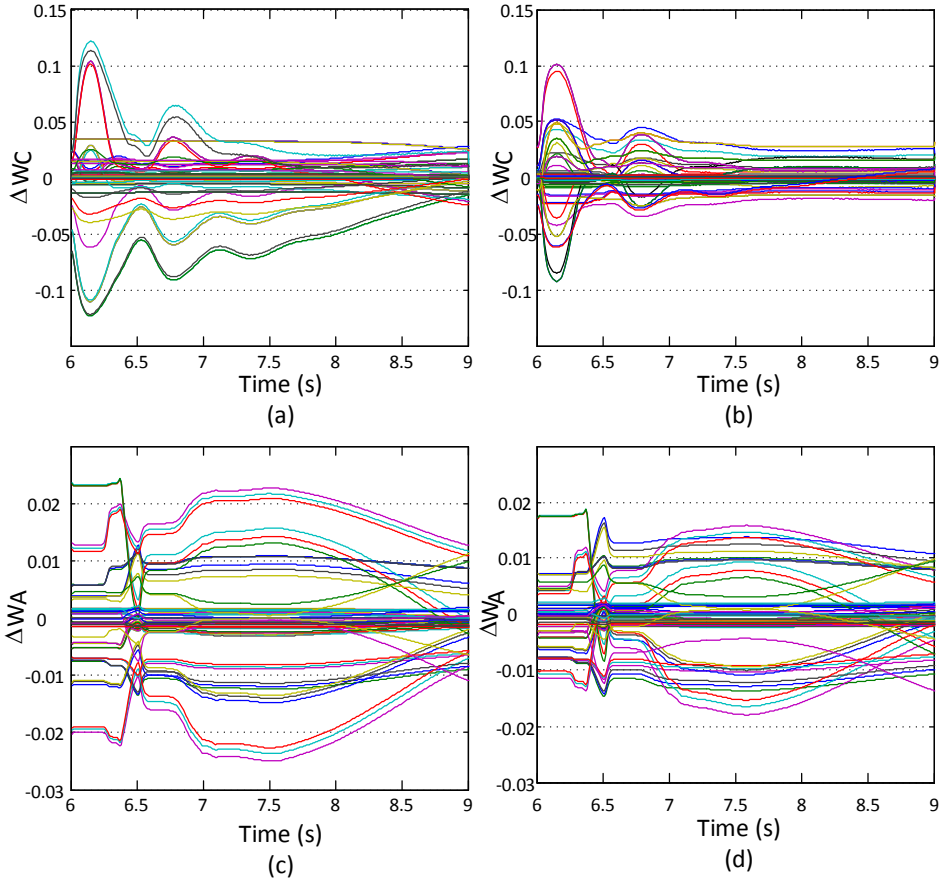


Figure 3.8: $\Delta W = W - W^*$ (a) conventional critic NN (b) proposed critic NN (c) conventional action NN (d) proposed action NN

perform simulations in real-time benchmark, power system modeling has been done in RSCAD (Path3). The modeled power systems is then connected with exciter acting as a 'nominal' controller. The input of the exciter is then connected to the control architecture implemented in Texas Instrument (TI) Controller board, Piccolo C28335. For this, MATLAB codes was first converted to SIMULINK (Path1) and then to C language by means of Code Composer Studio software (Path2) which is then deployed to TI controller board (Path3).

In the modern power grid, WASCCO can be located at the energy management center, and the communication to local controllers can be developed using secured protocol. As in Fig. 3.10, the signals from generators are communicated to a TI board running the WASCCO algorithm via LAN in the analog format (Path8). This has been accomplished using 12 channel Gigabit-Transceiver Analogue Output (GTAO) block provided in RSCAD. The control signals after the computation are adopted using TI in the format of discrete and binary signals that are sent back to the power system model represented in RSCAD (Path7). Then, Gigabit-Transceiver Digital Input (GTDI) block, digital input processor, provided in RSCAD is set to get the binary control signals for the generators these are then converted to real values. The WANNID, Action and Critic training has been initially performed off-line in PSCAD and then applied online in the real-time simulator. The signals implemented to augment the PSS in each of the three generators are then analyzed for damping the inter-area mode oscillations. The test bed consists of a real-time digital simulator (RTDS) controlled by a controller board in which control algorithms resides. The power system model is running in real-time at 50 micro seconds time stamp. The generator data is sampled and sent to the controller during this sample time and the controller performs the action every 20 micro seconds.

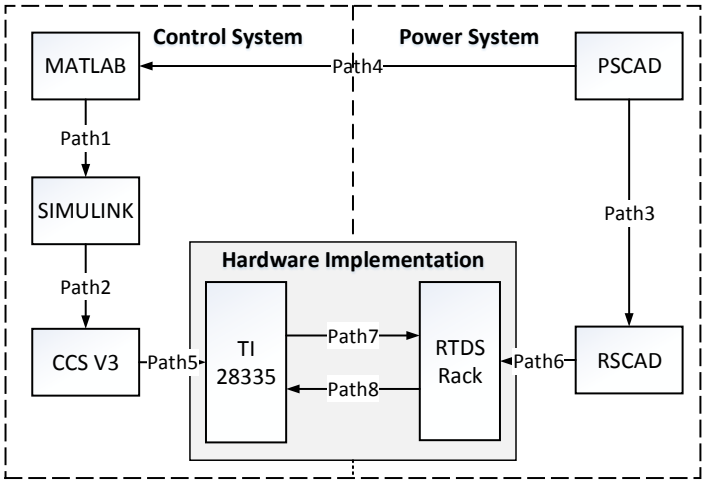


Figure 3.9: Real-time implementation flowchart

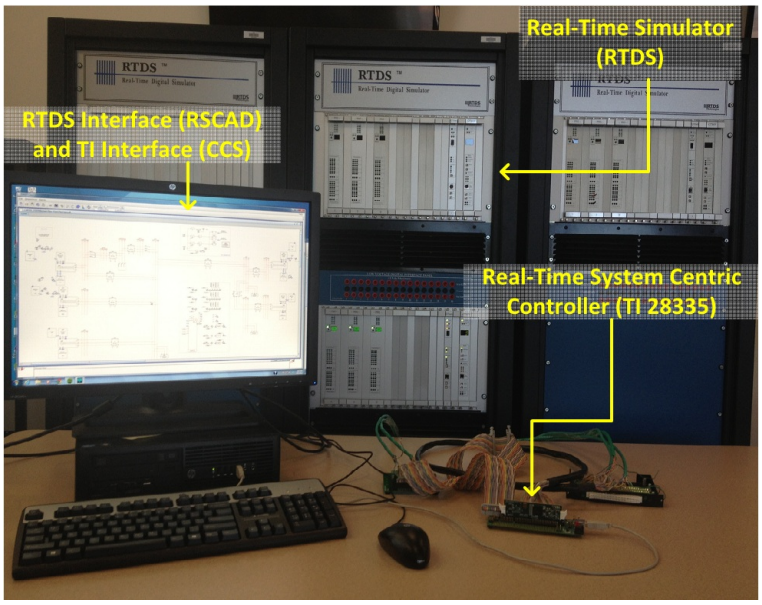
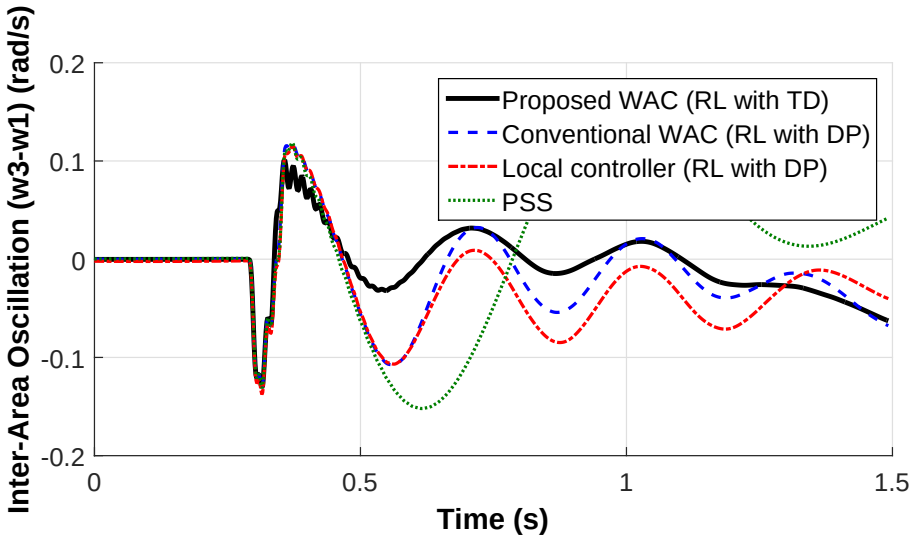


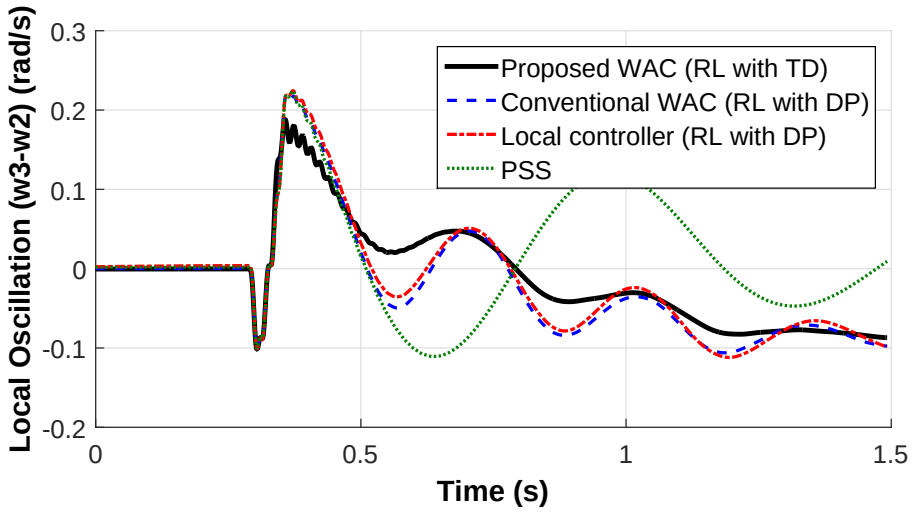
Figure 3.10: Real-time implementation benchmark

3.6.1 Real-Time Test Results on a Two Area Power System

Using the described real-time platform, the proposed WAC performance in two non-linear time domain case studies is simulated. In these simulations, four controller algorithms including local PSS, local RL-based controller, RL-based WAC and the proposed WAC have been implemented and tested. Local and inter-area oscillations have been monitored to analyze the performance of the controllers.



(a) Case A.1: comparison on speed deviations between G_1 and G_3 in the case of short circuit



(b) Case A.1: comparison on speed deviations between G_2 and G_3 in the case of short circuit

Figure 3.11: Case A. two area power system response comparison of the proposed TD-based WAC to local and conventional controllers

3.6.1.1 Case A.1: Short Circuit

In this case study, a 100ms three-phase short circuit at one of the transmission lines is simulated. In order to access the performance of the controllers, the line is disconnected by means of breakers on the sides of line. This case study effects all oscillatory modes of the system and changes the topology of that as well. Inter-area and local oscillations in this case are presented in Fig.3.11a and 3.11b. As it can be

seen, a notable damping improvement is gained, when the proposed control algorithm has been used. In addition, overshoots and undershoots of oscillation are reduced as well.

3.6.1.2 Case A.2: Power Transfer Capability

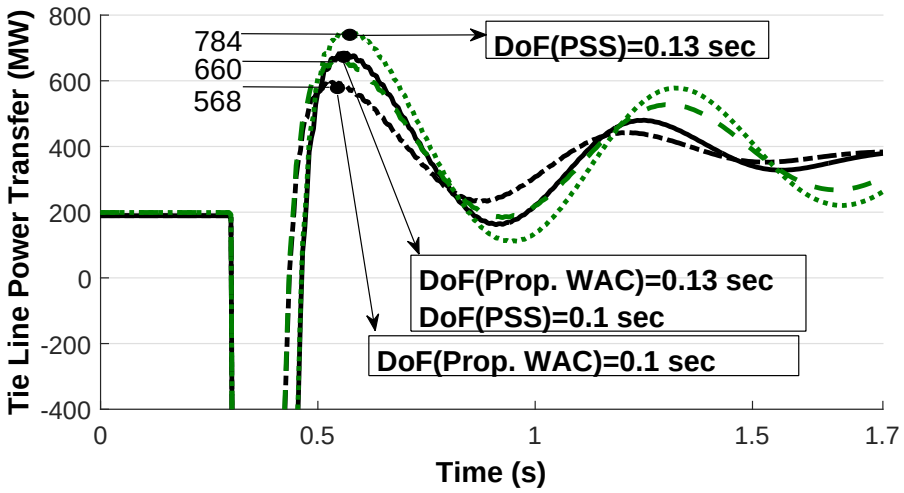
In further studies, power transfer capability has been the subject of the analysis. When there is a fault on one of the lines, the parallel line would encounter a power flow oscillation. As depicted in Fig. 3.12a, by utilizing this controller, power transfer margin can be increased by 124 MW. In addition, duration of fault (DoF) is increased for the proposed WAC to gain the same power transfer margin as PSS, which indicates increasing critical fault clearing time by applying the proposed controller.

3.6.1.3 Case A.3: Sudden Torque Change

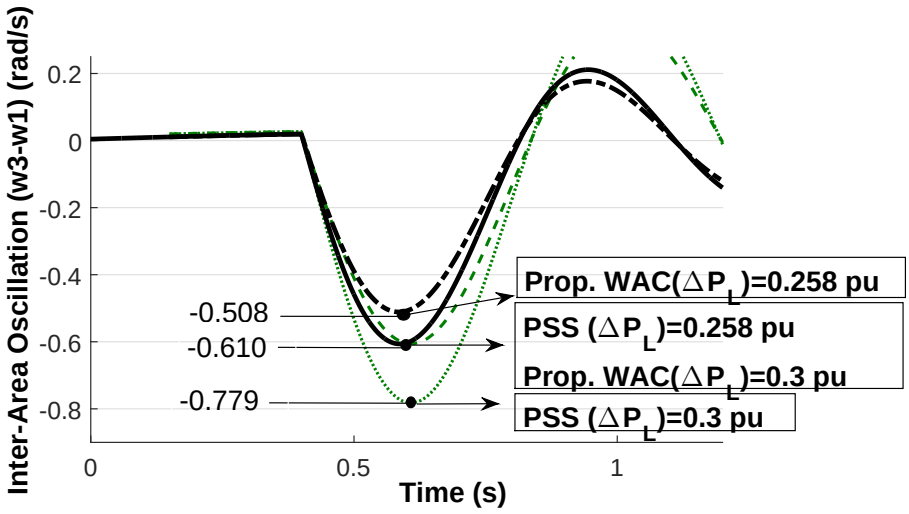
In this case, frequency regulation capability in presence of sudden torque change has been studied. A 0.258 pu torque step decrease, ΔP_L, has been generated the terminal of G_3 in the system with conventional PSS. This action causes a mismatch between mechanical and electrical powers of generator which leads to frequency increment. The inter area oscillations would have same undershoot as the system with proposed controller with 0.3 pu torque change. In addition, with the same amount of ΔP_L for PSS the inter area oscillation undershoot is increased by 38% (Fig. 3.12b).

3.6.2 Real-Time test Results on a 39 bus Power System

In order to investigate the capability of the proposed control architecture in damping the inter area oscillations, the 10-machine 39-bus New England test system was implemented in real-time test bed (Fig. 3.13). The importance is to illustrate that the approach can be implemented on a large scale system, since the dynamic equivalent areas are considered for control. The network reduction and dynamic equivalence model development can be achieved by means of model order reductions or similar methods in WAMs [105]. In addition, one of the main focuses of this method is to



(a) Case A.2: power transfer capability as a function of duration of fault (DoF)



(b) Case A.3: inter area oscillation as a function of load change on G_3

Figure 3.12: Case A. two area power system response comparison of the proposed TD-based WAC to local PSS

reduce the training load by means of ET. The states that are not activated are considered as redundant measurements and would not participate in the training, which results in less training data processing for NNs. Based on the dynamic data of the generators and the parameters of the transmission systems [106], the generators can be grouped into a desired number of groups. In this simulation, the test system has been grouped as 3 areas as depicted in Fig. 3.13 based on slow coherency grouping

method [107]. It is assumed that each bus or substation has a PMU sensor that transmits signal to the local Phasor Data Concentrators (PDC).

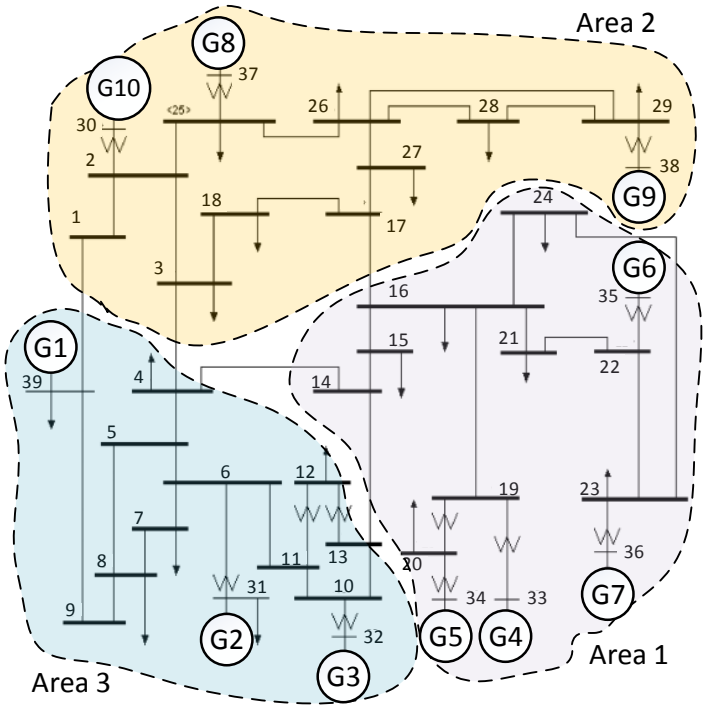


Figure 3.13: New England 39-bus test system.

A utility function composition approach is adopted that helps speeding up the training process of the NNs by commonly using center of angles and center of speed. For each area, an average angle and speed is found from measurements through the PMUs in that area as,

$$\delta_j^{COI} = \frac{\sum_i \delta_i H_i}{\sum_i H_i}, \quad \omega_j^{COI} = \frac{\sum_i \omega_i H_i}{\sum_i H_i} \tag{3.42}$$

where, j is representative of the areas and i denotes the generator in that area. The WANNID, action and critic NNs are trained on the aggregated area signals, and control signals are calculated according to the energy of each area. The training is

based on utility function of each area defined as,

$$U_j^{COI}(t) = -0.8(\Delta\delta_j^{COI}(t))^2 - 0.4(\Delta\omega_j^{COI}(t))^2. \tag{3.43}$$

Three transient case studies has been performed to evaluate the effectiveness of the proposed damping controller in comparison to local PSS and conventional WAC.

3.6.2.1 Case B.1: Short Circuit Test

100 ms self-clearing 3-In this case, phase fault in the middle of the tie line between buses 16 and 17 connecting areas 1 and 2 at $t = 1s$ has been performed.

3.6.2.2 Case B.2: Sudden Load Change Test

In this case, Load at bus 20 has been decreased to 50% of the base value 628 MW at $t = 1s$. This case evaluates the validity of the proposed controller, even though the fault is inside an area.

To evaluate the validity of this grouping, actual utility function trajectory of each generator is presented in Fig. 3.14, which confirms the similarity of attributes of the generators in each group for these case studies. The simulation results in Fig. 3.15 demonstrates the performance of the proposed WAC during inter-area oscillation between areas 1 and 2. Speed deviation for the area 1 is depicted as well, which shows the performance of the WAC in comparison to the local controller, in this case the PSS. As it can be seen, the system is capable of reaching to a new stable operating point quickly and with less oscillations and higher damping.

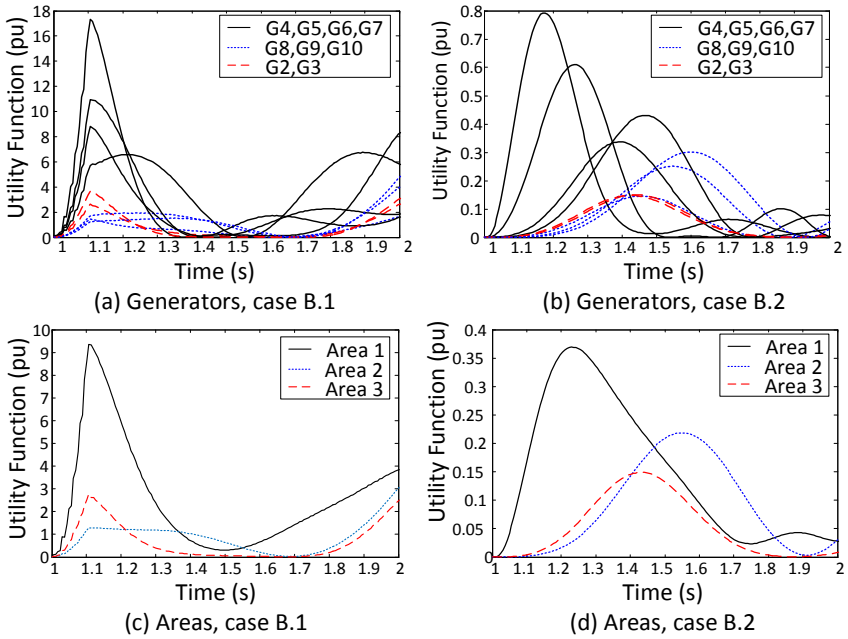
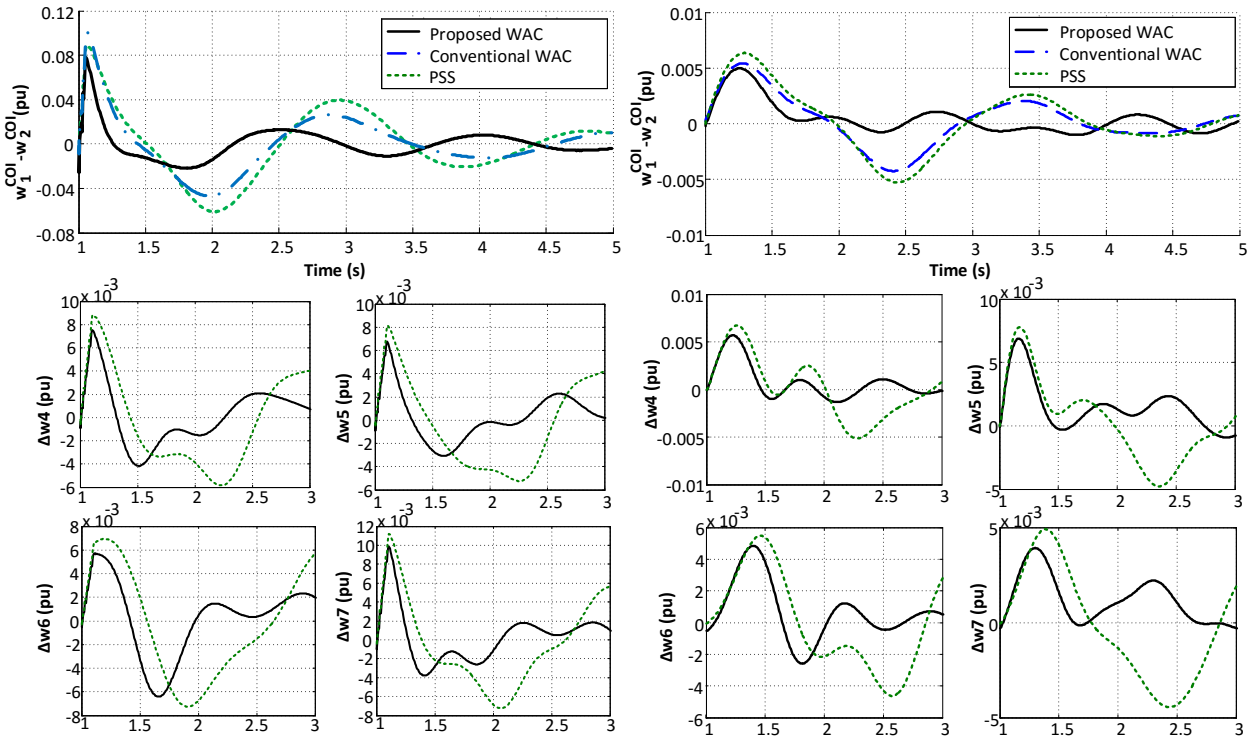


Figure 3.14: Comparison on utility function of the in IEEE 39-bus system generators (U_i) and areas (U_j^{COI}) for the case studies



(a) Comparison on speed response in case B.1. (b) Comparison on speed response in case B.2.

Figure 3.15: Case B. IEEE 39-bus power system inter-area and local speed response comparisons between the proposed TD(λ)-based controller and local and conventional controllers (a) Case B.1 (b) Case B.2

Table 3.4: Case B. IEEE 39-bus system transient Energy Index Comparisons

	Case B.1			Case B.2		
Energy (pu)	Prop. WAC	Conv. WAC	PSS	Prop. WAC	Conv. WAC	PSS
Area 1-2	0.16	0.5	1	0.34	0.66	1
G4	0.48	0.66	1	0.50	0.75	1
G5	0.46	0.68	1	0.50	0.70	1
G6	0.58	0.87	1	0.25	0.625	1
G7	0.44	0.60	1	0.28	0.71	1

For the transient tests carried out, the transient energies of individual generators of area 1 (G_4, G_5, G_6, and G_7) and inter-area energy between area 1 and area 3 within the first 4 s after the disturbances are calculated by accumulation of the utility function in the course of time. This value as a energy or cost function is the identification of performance of the controllers. Table 3.4 provides this value for the controllers employed in the case studies with PSS value as the base. the results shows that in the presence of proposed controller, transient energy is reduced for inter area and local oscillations indicating better damping when compared to the conventional controllers.

3.6.2.3 Case B.3: Inter Area Bulk Power Transfer

In this case study, the proposed architecture is applied to real-time power sharing scenario between the areas. In normal conditions, there is a power exchange of 180 MW between area 1 and 2, 380 MW between area 1 and 3, and 160 MW between area 2 and 3. The WAC monitoring the inter area speed is able to balance the power transfer to mitigate any frequency deviation between the areas. This could be highlighted in the case of disturbance in the system effecting the power exchange between the areas. Let us assume a renewable source is connected to the area power grid. Assuming that bus 36 is connected to 560 MW renewable source with the load of 460 MW, a three phase fault has been applied to tie line connecting area 1 to 2 causing oscillations on the other tie lines as well (Fig. 3.16). As it can be seen, the proposed controller has lead to lower overshoots and less oscillations improving the power transfer between

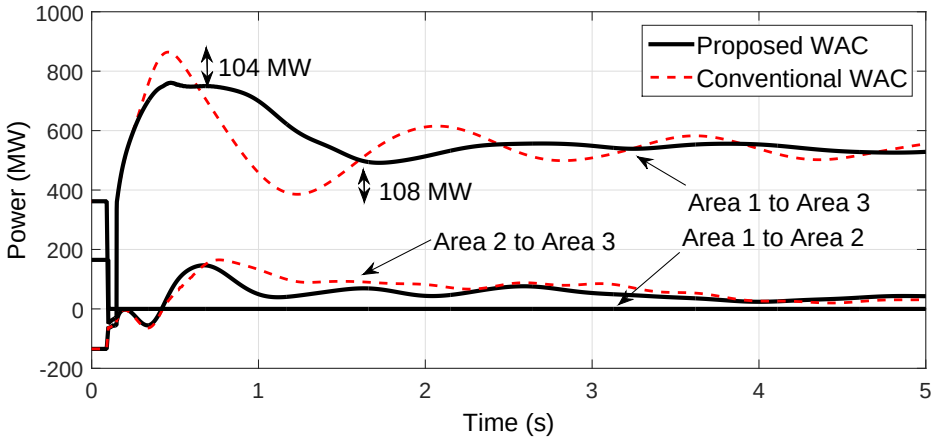


Figure 3.16: Case B.3: IEEE 39-bus system inter area disturbance and power transfer oscillations

the areas during dynamic changes. This difference (104 MW) could be critical, since this may cause a the renewable source mismatch between the generation and load leading to switch it to islanding mode. This incident can produce a cascading effect mismatch which leads to further decrease in generator power eventually leading to load-shedding.

3.7 Conclusion

In this chapter, an intelligent controller for real-time power system stabilization and transient voltage control is being proposed and implemented. The proposed architecture develops a hybrid reinforcement and temporal difference learning framework which allows the system to learn from interactions and predict future states. The developed algorithm is evaluated using real-time digital simulation on a benchmark two area power system model and tested on a real feeder using hardware. The main advantage of this design is its ability to learn from the past using eligibility traces and predict the optimal trajectory through temporal difference learning. The control algorithm is evolved from ACD performed on-line at a finite horizon through backward and forward view. Results shows that the proposed method can be implemented in real life and have better capability to damp oscillations than conventional schemes.

CHAPTER 4: SUPERVISED REINFORCEMENT LEARNING

4.1 Introduction

As discussed in previous chapters, local oscillations are typically damped by generation excitation control using power system stabilizers (PSS) [1]. However, these local damping controllers cannot damp and stabilize inter-area oscillation modes unless tuned for that purpose [108]. Inappropriate offline tuning of local controllers can also adversely affect the system operation. For instance, such tuning may lead to high gain feedback for some network but this may destabilize some other networks [109]. Wide Area Control (WAC) is often proposed to mitigate such oscillations.

There are mainly two major issues that affects the performance of a WAC:

- First, such WAC shows destabilizing interactions with local controllers, which makes the system unstable during certain contingency situations.
- Second, WAC actions are adversely affected by uncertainties, e.g. communication channels, lost packets, time delays, and lack of accurate and detailed global power system models and their time constants.

Ideally, a WAC design should mitigate the aforementioned issues and as such should augment the local controllers, thus contributing towards inter-area damping.

Classical WAC models mainly based on frequency domain methods usually treat the unknown parameters as constants [12,62]. In general, for such methods stability is not fully guaranteed. The second set of designs are measurement based. Mostly these designs uses machine learning techniques that can be trained using WAM. In previous chapters we discussed about the RL-based WAC design, which had the capability of dealing with uncertainties. RL, different from both the centralized and

the decentralized control schemes, employs a number of semi-autonomous agents to collaborate with each other to perform optimal control theory. Although, RLs look promising in theory, host of problems make them difficult to implement in real life. These problems roots in heuristic nature of such controllers, which causes reliability issues and cannot be fully trusted to implement. These designs are thus not guaranteed to possess global optimal properties.

In general, Neural Network (NN) as a distinct learning-based function approximator has been effectively implemented as a power system intelligent controller in several works, and their ability to adapt during nonlinear transient conditions have been discussed [10, 66, 68, 71, 110]. These architectures use NNs in the form of supervised learning as an intelligent PSS for damping generator oscillations. However, majority of these works have used the intelligent controller by itself. For such designs perfect training of NNs are required for overall stability which may not be achievable all the time.

One solution will be to develop a hybrid architecture that uses a local controller along with a learning-based architecture as a global controller. Several works have focused on designing a hybrid linear and nonlinear based control architecture [110–113] and robust control techniques [111], for power system stabilization. Combination of a linear and NN-based nonlinear adaptive controller through switching law is studied in [113]. In [110], we have used linear adaptive and a NN-based controller as local controllers with explicit neuro-identifier to augment the performance of the adaptive controller. In this work a new architecture that prioritize the local and global controller and at the same time ensures robustness, stability, and optimality of the closed loop system is proposed.

In this chapter, the controller scheme of 2-level combination of local and global parts is proposed initially as: (a) The supervised local feedback, (b) The RL-based global wide-area signal. The global control feedback is designed to enhance the wide-

area behavior of the system that suppress inter-area oscillations based on Lyapunov energy function.

The main advantages of the proposed architecture are:

- A value priority structure that makes sure that the closest global optimal solution is achieved for a given time horizon.
- The structure ensures quality of performance, stability and identification using an energy function that is stable.
- The design monitors the system with WAM constraints such as transmission delay and data dropout and makes sure that the overall system stability is maintained.
- The method adapts based on system changes and ensures stability when the NN controller fails to perform.

Next, the value priority scheme is applied to two global controller: (a) Adaptive control, and (b) learning-based control. Conventional adaptive control [6] can provide damping of inter-area oscillation modes. However, it is a well known fact that the response of a dynamic nonlinear system cannot be tracked using a linear model-based controller. Such a design is capable of *adapting* to the normal varying operating points only when the variations can be represented as a linear parametric set [6] as these controllers are not designed for nonlinear uncertainties especially in the WAC designs.

Another method proposed in this chapter is the augmentation of global linear controller as a supervisor with global nonlinear WAC as a RL. The approach inspired by [44, 114] is based on the combination of experts to enhance optimality of the control and generality of NN learning. Our choice of performance criterion is on the basis of direct Lyapunov stability analysis that encourages minimization of the composite

identification and closed-loop errors. The uniqueness of the proposed method is its softmax value priority network that makes sure that the closest stable optimal solution at a given time horizon is achieved. The linear MRAC works well especially when the control objective is to make the output follow a particular path, and the NN controller is employed in tracking the unknown nonlinearity and uncertainity of the system. The learning-based controller seamlessly evolves based on system changes from a simple supervised structure to a full RL-based controller. Moreover, additional source of exploration is provided that eliminates the necessity of offline training for RL. The proposed method allows the system to dynamically shift between linear and intelligent controllers and thus can be effectively utilized on any practical set up. The global controller can also interact between two or more local controllers thus operating as a wide-area control system.

The chapter is organized as follows. Section 4.2 provides a brief overview of mathematical preliminaries and theoretical design of local and global control. In section 4.3, implementation aspects of the design is covered. Test results on a 8-bus and 68-bus 16-machine power systems are presented in Section 4.4. Furthermore, sections 4.5 to 4.9 provide the hybrid design for adaptive and learning based control designs, and finally section 4.10 concludes the paper.

4.2 Hybrid Global and Local Control Structure

In this section, a methodology to construct a stabilizing trajectory, $k(t)$, as a value priority on control output is proposed and is presented in Fig. 4.1. With k as a value indicator, a hybrid control action can be defined as,

$$u(t) = (1 - k(t))u^l(t) + k(t)u^g(t) \tag{4.1}$$

where $k \in [0, 1]$ is an interpolation parameter that determines the level of control or autonomy on the part of each controller. u^l is the local controller that estimates the

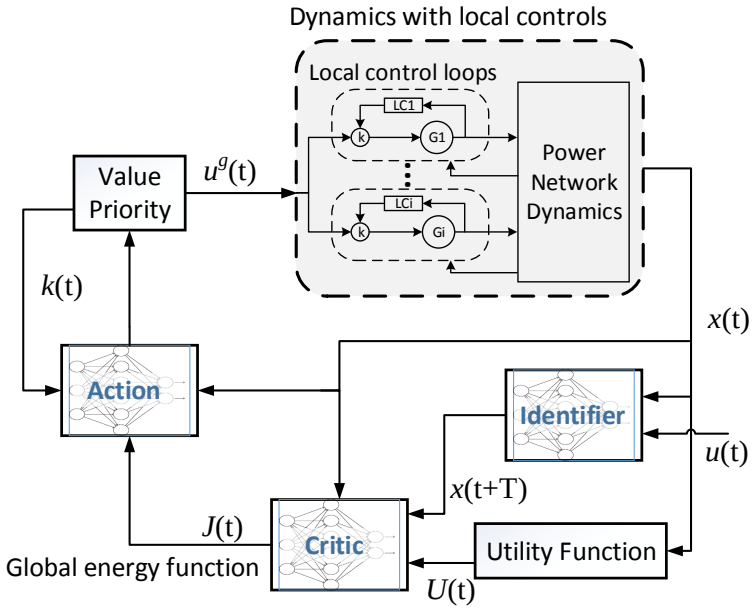


Figure 4.1: Structure of the wide-area global control center

action based on linear parametric design, which in this section is the conventional PSS

$$u_i^l(t) = K_{si}\frac{sT_{si}}{1+sT_{si}}\frac{1+sT_{1i}}{1+sT_{2i}}\frac{1+sT_{3i}}{1+sT_{4i}}\frac{1+sT_{5i}}{1+sT_{6i}}\dot{\delta}_i(t) \tag{4.2}$$

where, $u_i^l(t)$ is the local control of generator i at time t. The area j control signal u_j^g is the global action of RL controller based on nonlinear adaption of NN as,

$$u_j^g(t) = W_A(t)^T\Phi_A([\dot{\delta}_1(t), ..., \dot{\delta}_{N_j}(t)]) + \epsilon_A \tag{4.3}$$

where, $\dot{\delta}$ is the rotor speed and N_j is total number of areas.

4.2.1 Proposed Value Priority Function

To qualify and distinguish the controllers, first a value priority function $\dot{J}_{vp}$ and a system space in the form of Ω-regions for the controllers are defined. Let $\Omega^g \cup \Omega^l$ covers the state space related to $\dot{J}_{vp}^g$ and $\dot{J}_{vp}^l$ with $\Omega^g \cap \Omega^l \neq 0$. The problem entails searching for a Lyapunov-like function for controllers whose associated Ω-regions covers the state

space. Additionally, the problem formulation requires that whenever there is a system trajectory movement towards Ω^g, then $\dot{J}^g_{vp} < \dot{J}^l_{vp}$, meaning that the RL has learned the optimal solution. The following hybrid shifting rule is initialized at t_0 by $k = 0$ as a representation of local control activating the system so that $x_0 \in \Omega^l$; Then,

$$\dot{k}(t) = k(t) + \beta(\dot{J}^l_{vp}(t) - \dot{J}^g_{vp}(t)) \tag{4.4}$$

The RL starting with the nonlinear region eventually advances to cover the linear part as well, as long as $\dot{J}^g_{vp} < \dot{J}^l_{vp}$, which reduces the supervised region.

A value priority function $\dot{J}_{vp}$ to prioritize local and global controllers is proposed here, which is derived from the Lyapunov energy function of the system, $\dot{J}$. The proposed function is,

$$\dot{J}_{vp} = \dot{J}_{id} + \dot{J}_p \tag{4.5}$$

where, $\dot{J}_{id}$ is representative of the identification of the global controller and $\dot{J}_p$ is performance criteria for both controllers.

4.2.1.1 Identification of the global controller

If the identifier predicts the output with a smaller error, this index decreases, and hence the RL-based global controller may be chosen to control the system. The significance of identifier is that it plays an important role in predicting the future energy function and estimating the optimization policy. For this purpose, ϵ_{id} is defined as aggregation of e_{id} in backward window with decay factor of ζ_{id} as,

$$\epsilon_{id}(t) = \int_0^t exp(-\zeta_{id}(t-\tau))\|e_{id}(\tau)\|d\tau \tag{4.6}$$

$$\dot{J}_{id}(t) = \frac{1}{2}\epsilon^T_{id}(t)Q\epsilon_{id}(t) \tag{4.7}$$

4.2.1.2 Performance of the controllers

The performance value priority function of the controllers is defined based on the action error in a backward view as,

$$\epsilon_p(t) = \int_0^t exp(-\zeta_p(t-\tau))\|u(\tau) - u^*(\tau)\|d\tau \tag{4.8}$$

$$\dot{J}_p(t) = \frac{1}{2}\epsilon_p^T(t)R\epsilon_p(t) \tag{4.9}$$

In order to derive this element, considering the current time as the terminal state, the performance of each controller with respect to the previous time steps can be found. Calculations are done backward for u^* and J^*, and forward for x^* for $t = [t_0 - kT, t_0]$ with t_0 being the current time such that

$$u^*(t) = \min_u(J(t)) \tag{4.10}$$

$$J^*(t) = U(x(t), u^*(t)) + J(x(t) + \Delta x(x(t), u^*(t))) \tag{4.11}$$

$$x^*(t+T) = x^*(t) + \Delta x(x^*(t), u^*(t)) \tag{4.12}$$

$\Delta u(t)$ is then calculated as,

$$\begin{aligned} \Delta u(t) &= \frac{\partial J(t)}{\partial u(t)} = \sum\nolimits_{j=i}^{k} \gamma^{j-i}\frac{\partial U_j}{\partial u(t)} \\ &\approx R.u(t) + \gamma\frac{\partial x(t)}{\partial u(t-T)}Q\Delta x(t-T) \end{aligned} \tag{4.13}$$

4.2.2 Value Priority Implementation Algorithm

Algorithm 1 presents the the value priority implementation algorithm and its link to local and global controllers. The value priority computes the action, u, as a weighted sum of the actions received by both component policies, u^g and u^l, based on (4.1). In this method, the value priority parameter, k, not only provides u, but also plays

an important role to modify the RL's policy, J^g. The RL policy is given by (2.10), uses an update parametric law of (2.6) derived for critic NN weights, W_C. After each state transition, the parameters of RL-based global controller is updated according to the update rule derived below. This feature leads the RL to merge faster towards supervised local controllers, during the period when the RL is still not fully capable of performing optimal control action based on the value priority criteria. This is written as,

$$\dot{W}_C = \alpha[ke_C + (1-k)(J^l - J^g)]\nabla_{W_C} J^g \tag{4.14}$$

$$\dot{W}_A = \alpha[ke_A + (1-k)(u^l - u^g)]\nabla_{W_A} u^g \tag{4.15}$$

where J^g and J^l are critic NN output with control inputs of u^g and u^l, respectively. Upon converging to a solution for RL, global controller performance is first compared to the local controller with respect to its action chosen, $u^g(t)$ and $u^l(t)$, during previous time steps. Furthermore, an identification index is calculated with respect to Lyapunov function and input-output stability of the system. The parameters of value priority function, Q and R, are set offline based on linear characteristic equation of the system and the reference model. More accurate calculation of these coefficients, based online monitoring is possible, though this would affect only the final numerical results not the methodology presented. The shift rate between local and global controller is taken as $\beta = 0.01$.

4.2.3 Stability of the Proposed Method

Lemma: Let J_n^g and u_n^g be the RL sequence at iteration n that performs weights updates per (4.14) and (4.15) as

$$J_{n+1}^g(t) = k(t)[U(t) + {}^3J_n^g(t+T)] + (1-k(t))[J^l(t)] \tag{4.16}$$

$$u_{n+1}^g(t) = k(t)[\arg\min J_n^g(t)] + (1-k(t))[u^l(t)] \tag{4.17}$$

Algorithm 1. Value priority

1) Start of the value priority, t_0.
2) Initialization: $k(t_0) = 0, \dot{J}^g_{vp}(t_0) = 0, \dot{J}^l_{vp}(t_0) = 0$.
3) Local controller: $u^l(t), J^l(t)$.
4) RL with value priority:
   a) Critic:$\dot{W}_C = \alpha[ke_C + (1-k)(J^l - J^g)]\nabla_{W_C} J^g$.
   b) Action:$\dot{W}_A = \alpha[ke_A + (1-k)(u^l - u^g)]\nabla_{W_A} u^g$.
   c) if $|\dot{W}_C|\&|\dot{W}_A| < \epsilon$ proceed to 5.
   d) if $|\dot{W}_C|\&|\dot{W}_A| > \epsilon$: Update W_A and W_C and proceed to 4.a.
5) At global controller:
   a) Calculate $\dot{J}^g_{vp}(t)\&\dot{J}^l_{vp}(t)$, Eq. (4.5-4.16).
   b) Update $k(t) = func(k(t-1), \dot{J}^g_{vp}(t), \dot{J}^l_{vp}(t))$ Eq. (4.4).
   c) $u(t) = (1-k(t))u^l(t) + k(t)u^g(t)$.
6) next time step, $t = t + T$.

Assuming,

$$J^g_0 = 0 \leq J^l \leq J^* \leq J^{max} \tag{4.18}$$

as $n \to \infty$, $J^g_n \to J^*$ and $u^g_n \to u^*$. Hence, the sequence of J^g_n converges to the solution of the HJB.

Proof: Knowing that critic training is based on the combination of errors of self exploration and supervised action as defined in (4.14), this feature may imply the possibility of contradiction between the policies, while the boundedness is yet proven. Here we prove that, the update method will always converge to the optimal value. Based on [49] it is proven that RL policy J^g_n stand alone is a non-decreasing sequence for HDP. This means $J^g_{n+1} \geq J^g_n$. Knowing that, the boundary conditions of k, $k = 0$ and $k = 1$ are stable, since they have one target to update, the focus is the stability of the controller when $0 < k < 1$. We investigate three cases:

1. If $J^g_{n-1} \leq J^l$, then, based on assumption (4.16) the target update is in the same direction as the incremental update. This makes the output to converge to optimal value faster, and k remains zero.

2. If $J^g_{n-1} > J^l$ and $|e_C| \leq |(J^l - J^g_n)$, then, k is increased, and since $J^g_{n+1} \geq J^g_n$ is

non-decreasing, then for the following time sequence $J_n^g > J^l$ holds.

3. If $J_{n-1}^g > J^l$ and $|e_C| > |(J^l - J_i^g)$, then the RL policy will get closer to the optimal value after few iterations of delay till $k = 1$ and then latches to the RL controller.

The same method implies action convergence since they are directly connected to each other.

Theorem 1: Let the optimal control u^* be bounded by $|u^*| < u^{max}$ on a compact set Ω, where $u^{max} > 0$. Then, for power system (2.1)-(2.2) with PSS and HDP as the controllers with adaptation law of (2.10) for NNs, and value priority trajectory of (4.4) and value priority function defined in (4.5), all the signals in the closed-loop switching system described above are bounded. Therefore, as $t \to \infty$, $\Delta u(t) \to 0$ and $\Delta x(t) \to 0$.

Proof: Considering the state space of the error vector as,

$$\dot{E} = Ae + B\tilde{u}(x) \tag{4.19}$$

where, $e = x - x^*$ is the tracking error, E is the error vector, $\Delta u(x) = u(x) - u^*(x)$, and based on (4.3) $\Delta u(x) = \tilde{W}_A \Phi_A(x) + \epsilon_A$ with $\Delta W_A = W_A - W_A^*$. Lyapunov stability function can be defined as,

$$V = \frac{1}{2} e^T P e + \frac{k\gamma}{2} \Delta W_A^T \Delta W_A \tag{4.20}$$

where,

$$AP + PA = -Q, \quad P b_c e^T = P_1 e^T \tag{4.21}$$

P is a symmetric matrix, and $P_1 > 0$ is the solution to the Lyapunov equation for some $Q1 > 0$, a positive definite matrix. Let P_1 be the last column of P and derivative

of Lyaunov function is gained,

$$\dot{V} = -e^T Qe - e^T P_1 b(k\epsilon_A + (1-k))\tilde{u}^l) + k\gamma \Delta W_A^T \Delta \dot{W}_A \tag{4.22}$$

Now, let $e_A = f(u(t), x(t+T), J(t+T))$ based on (2.12), then $\Delta W_A = f(\epsilon_A, \epsilon_I, \epsilon_C, \epsilon_{id} Q \epsilon_{id}, \epsilon_p R \epsilon_p)$. The first three elements are the action, identifier, and critic approximation errors, which using approximation theory can be reduced by changing the number of neurons in the network hidden layer. Moreover, error of $J(t+T)$ is due to the identifier and action's error in previous time steps, which is projected in the value priority function. And finally, based on the previous Lemma,

$$\Delta W_{An}^T \Delta \dot{W}_{An} = (W_{An} - W_A^*)^T (W_{An+1} - W_{An}) < 0 \tag{4.23}$$

Hence, it can be assumed that $\Delta W_{An}^T \Delta \dot{W}_{An} = -|\epsilon^g(x)|$, where $\epsilon^g = \max\{\epsilon_A, \epsilon_I, \epsilon_C\}$. If the desired states are bounded and in a compact set, ϵ^g is finite for a given set of neurons, and, boundedness of the local controller is given by $\epsilon^l(x)$, then, $\dot{V}$ can be demonstrated to be negative such that,

$$\begin{aligned} \dot{V} \leq & -|e|\lambda_{min} Q|e| - |e|Pb\lambda_{max} \max\{\epsilon_A(x), \epsilon^l(x)\} - |\epsilon^g(x)| \\ \leq & -|e|\lambda_{min} Q|e| - |e|Pb\lambda_{max} \max\{\epsilon_A(x), \epsilon^l(x)\} \end{aligned} \tag{4.24}$$

and the upper bound will be,

$$|e| \leq \frac{\lambda_{max} Pb}{\lambda_{min} Q} \max\{\epsilon_A(x), \epsilon_l(x)\} \tag{4.25}$$

This completes the proof.

4.3 Implementation Aspects of the Proposed WAC

The proposed 2-level value prioritized control structure is a hybrid combination of local/decentralized controllers which are conventional PSS and the wide-area central controller. As a measurement-based control design, performance of the proposed WAC is highly dependent on the WAM infrastructure. In this section we investigate implementation aspects and challenges of such design considering currently used centralized communication architecture, the supervisory control and data acquisition (SCADA) system [10], PMUs and other support systems.

4.3.1 WAM Infrastructure

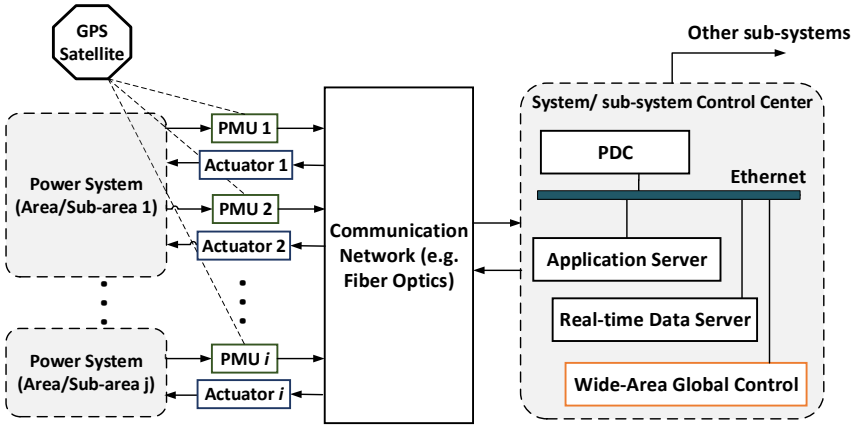
In the modern power grid infrastructure, WAM system consists of three main elements: PMU, Phasor Data Concentrator (PDC), and communication network (see Fig. 4.2a).

4.3.1.1 PMUs

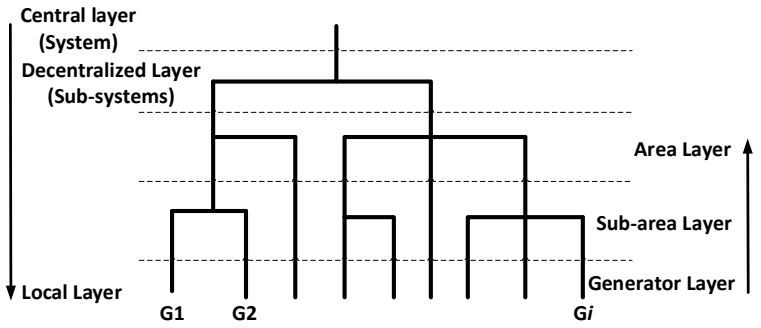
We assume that each generator located at different areas of power system has a PMU that transmits remote signals to the PDC located at the control center over a communication channel. Generally, such units can transmit time-tagged phasor measurements to WAC at the speed of 100 frames/s; therefore, for the proposed control architecture time step is considered as T =10 ms.

4.3.1.2 Communication Network

The global control can only be calculated after the PMU signals are transferred to the control center. This may incur certain time-delay and potential missing data information and may degrade the performance of the proposed control scheme jeopardizing even the closed loop system stability. Allowable time for control actions for a transient stability problem considering inter-area oscillation frequency prior to the first swing are usually considered to be less than 1 s. Depending on the distance and the network media, a time delay of today's wide-area communication networks (prop-



(a) Wide-area monitoring scheme.



(b) Coherency based generator grouping and system layering

Figure 4.2: WAC design considerations

agation delay) can be in the range of 6 ms to 1 s. For instance, fiber optic latency for the BPA system is considered to be less than 26ms [18]. Also, Synchronous Optical Network (SONET) architecture deployed by many utilities is capable of delivering messages from one area to multiple nodes on the system in approximately 6 ms [57]. It is worth noting that, by using dedicated channels the communication delay could be minimized.

4.3.1.3 Control Center

At the control center, PDC is responsible for collecting the PMU measurements and sorting them into a single time-synchronized data-set [115]. Subsequently, PDC sends the data in the form of packets to the WAC using local Ethernet. After processing

Algorithm 2. Implementation Design

1) Each PMU transmits the signals to the control center via communication network (Fig. 4.2)
2) At control center: Communication
   a) Calculate transmission delay (t_d) using PDC.
   b) Compute transmission delay threshold value (t_d^{max}).
   c) If $t_d < t_d^{max}$: proceed with the existing communication.
   d) If $t_d > t_d^{max}$: divide system into several sub-systems based on coherency grouping (Fig. 4.2b).
   e) Implement distributed architecture.
   f) Locate PDC at each sub-system and communicate with other PDCs or a Super PDC.
3) At control center: Control
   a) Calculate computational delay in WAC (t_c).
   b) Set computational delay threshold value (t_c^{max}) as packet delay (10 ms).
   c) If $t_c < t_c^{max}$: proceed with the existing WAC.
   d) If $t_c > t_c^{max}$: use aggregate signals for proposed controller based on coherency grouping (Fig. 4.2b, Eq. 4.26).

and performing control action, the WAC outputs in the form of auxiliary signals to the voltage references of generators is then sent to the substations. This may also include delay due to computational time [18].

4.3.2 Implementation Considerations

Considering this WAM infrastructure, there are mainly two constraints in implementing the proposed WAC system. First, the controller should be robust enough to accommodate the communication delays and losses and at the same time should be stable. Second, the system stability should be evaluated for scalable models and the models should be validated for controller implementation. Each of these challenges and proposed solution methods are discussed next.

4.3.2.1 Robustness

In the implementation design stage, transmission delays are taken into account to find the practical requirements of WAM infrastructure. As presented in Algorithm 2, at first the design is developed considering a centrally controlling WAC. Then the

delay in the present communication network is considered. To confirm that the delays are within the limit, threshold is found empirically based on inter-area energy deviation. Further, the system is divided into separate sub-systems with a potential communication media between them if the delays are found above the threshold level. This results in a distributed platform. Considering the total number of PMUs as $i = m \times n$, m signals are sent to the sub-system PDC throughout this platform, n being the number of sub-systems. Then, the sub-system PDC is linked to communicate with other PDCs or a super PDC. Overall, in this approach, based on the length of communication channel the time delays would be the same, if the central controller is located between the sub-systems. However, the number of data points will be decreased to m for each sub-system, and n for the overall system assuming all of them are connected to a super PDC. Thus, this method makes sure that the controller is robust and at the same time implementable in large system keeping the number of signal and delays within the threshold. Recently several works have focused on such a distributed design [19]. There are issues in implementing such a communication architecture such as designing multiple PDC's and communication channels. These issues are out of the scope of this thesis and is left for future investigations.

Given such a WAM infrastructure, WAC design should be robust for online implementation. For this, in this chapter, delay-free channels are assumed in WAC design stage, and constant unknown transmission delays are taken into account in the simulations to demonstrate the robustness of the design. Ref. [10] has shown the applicability of the NNs for compensating known static delays. However, we adopt the value priority mechanism of the two-level control strategy to guarantee system nominal operation in the case of communication failures. If data packet is lost, missing sensor sends zero values to identifier NN leading to an increase in e_{id}. In addition, this packet will be held until the next packet arrives. In the case of delay compensation, this block estimates states for $2t_d$ to calculate the e_{id}, accounting for both

sending and receiving delays. The goal of identifier NN is to predict the states at each time step $t + (t_d + 1.T)$ with given data at time step $t - t_d$. The difference of time steps between input and output of identifier NN is $2t_d + 1.T$, which implies that more the delay, more the block waits for calculation of the error, yielding higher values. This process thus proves the validity of the proposed model and robustness of the control architecture considering communication delays.

4.3.2.2 Scalability

Scalability could be viewed from two stand-points. First, a WAM with scalable communication infrastructure needs to be established that can accommodate the signal transactions and control especially with appropriate speed, bandwidth and required accuracy. Through the current state-of-the-art centralized communications 16 PMUs are able to stream data to a super PDC handling about 6250 data points per second [19]. Beyond which distributed approach as discussed in the above subsection could be utilized as a future solution. Second main challenge is the curse of dimensionality in RL and in general with any NN as the high number of state variables increases the learning and approximation time needed for optimal control policies. This issue also imposes some computational delay on WAC. However, in our approach, due to the supervision of local controllers, the heuristic nature of the RL is not of a concern. The reason is as follows. The architecture does not fully depend on the model except for the initial conditions or not-trained condition of the learning controller. However, one of the challenges in hybrid control design is compromises and trade-offs between conflicting objectives. Considering this, the proposed value priority scheme is designed to provide global performance, identification, and stability objective which are in the same direction of minimization of overall Lyapunov function. The stability is achieved if the learning weights error of NNs is minimized over specified time period, which is projected as a value priority metric.

Moreover, the nature of the WAC and coherent areas grouping can reduce the size

of data set. In this chapter, conventional slow coherency technique is used which is based on eigenvalue analysis and further transformation of the original state variables into a set of area and local variables [107]. Hence, the commonly used center of area speed can be found from WAMs sensors to construct the utility function as,

$$\dot{\delta}_j^{COI} = \sum_i \dot{\delta}_i M_i / \sum_i M_i \tag{4.26}$$

where, j is representative of the areas and i is the generator index in that area. The Wide-Area NN Identifier, action and critic NNs are trained using the aggregated area signals, and control signals are calculated according to energy transfer in each area. The training is based on utility function of all the areas defined as,

$$U(t) = -\sum_{l=0}^{3}\sum_{j=1}^{nj} Q_j(\Delta\dot{\delta}_j^{COI}(t-l))^2 + R_j(u_j(t-l))^2 \tag{4.27}$$

where, l is the time delay index, j area index, nj the number of monitored coherent areas or sub-areas.

Remark: This aggregation method addresses the modeling dimensionality issue and reduces the computational time of WAC to packet delay. It should be noted that, this approach also allows seamsless integration of the proposed WAC in a distributed mode as presented in Algorithm 2 and Fig. 4.2b; i.e. by cutting off the hierarchical clustering dendrogram in any level based on the infrastructure of the system, and providing the necessary channels between the sub-system.

4.4 Simulation Results on Test Power Systems

In order to analyze the evaluate the performance of the proposed 2-level controller, we will consider two multi-machine power systems: an 8-bus 5-generator power system (Fig. 4.3) and 68-bus 16-generator of New England/New York power system (Fig. 4.4). For analysis, dynamic equivalent frequency dependent models of the generators is

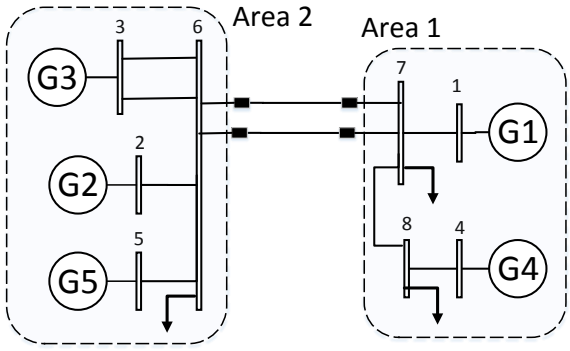


Figure 4.3: Diagram of 8-bus power system

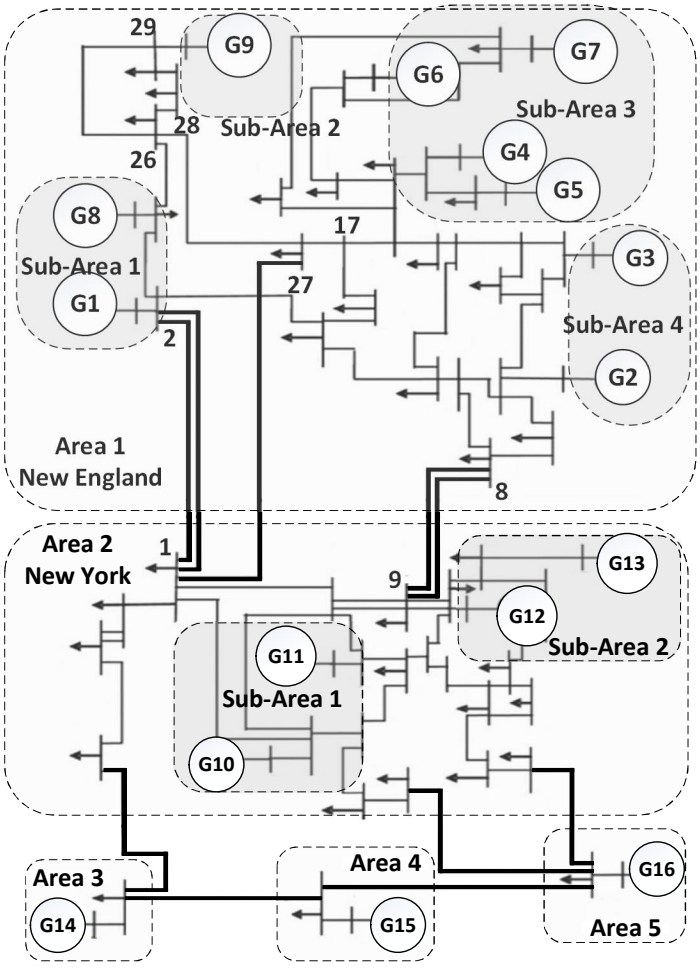


Figure 4.4: Single line diagram of the 68-bus 16-machine New England/New York power system with areas and sub-areas.

Table 4.1: NN parameters for the test power systems

NN	8-bus system Nodes			68-bus system Nodes			Input	Output
	Input	Hidden	Output	Input	Hidden	Output		
Identifier	13	25	2	47	80	9	$\dot{\delta}$,u	$\dot{\delta}$
Action	7	15	2	28	50	9	$\dot{\delta}$	u^g
Critic	7	15	2	28	50	9	$\dot{\delta}$	J

developed using electromagnetic transient simulation (EMTP) software PSCAD^{TM} and transient stability program PST^{TM}. Four control schemes, the proposed value prioritized global controller, RL-based controller acting alone and conventional local and global controllers are deployed on the test power systems in order to evaluate the efficiency of the proposed method. The parameters of the local controllers and for the two systems are provided in previous chapters. Parameters of the global controller is presented in Table 4.1.

4.4.1 Case A.1. 8-bus system: Identification Component

The purpose of this study is to show that, proposed value priority will merge to global stabilizer as soon as it has identified the system and optimal transient energy policy effectively. A 0.2 p.u. torque step decrease has been generated in G3 at $t = 4s$ and returned to initial point after 4s. At the beginning of first period, $t = 4s$, RL with partially random weights (50%) is utilized, and at the beginning of second period, $t = 8s$, the trained one is used (Fig. 4.5a). It can be seen from the Fig. 4.5a that there is significant weight change at 4s when random weights are used, and small change at 8s when the weights are trained. It should be noted that each time step consists of maximum 500 epochs of supervised RL training, and Wide-Area Identifier NN (WANNID) is able to reach to close to local optimal solution (in terms of weight) for this specific case study.

The inter-area speed response in Fig. 4.5d shows that in comparison to local controller, oscillation is less damped in the first period using the global controller with partially-trained WANNID weights, and opposite in the second period with the

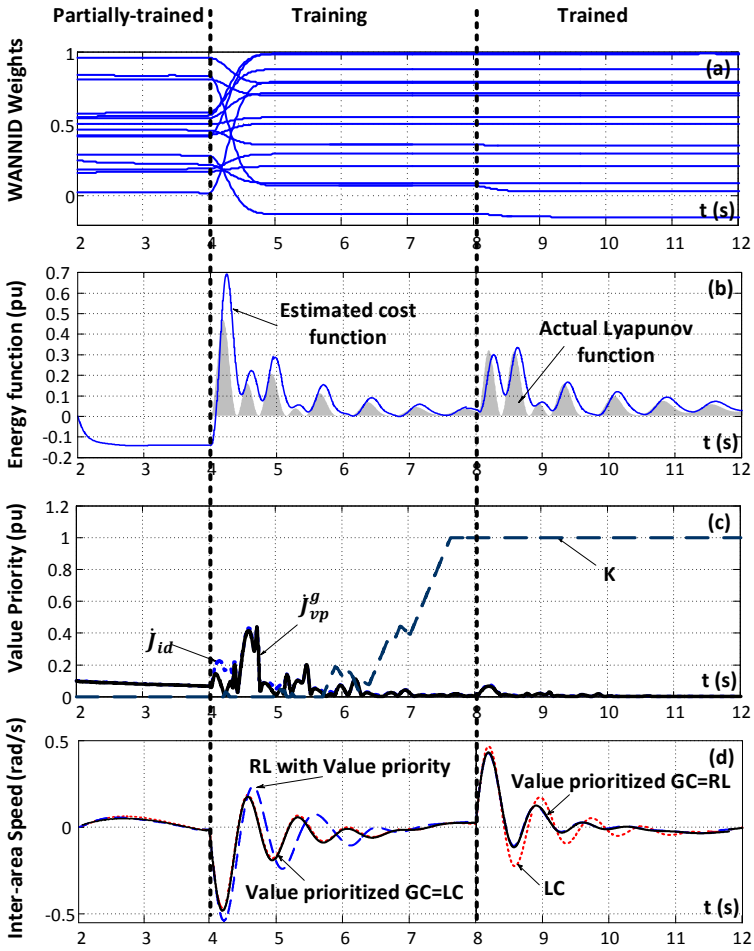


Figure 4.5: Case A.1. results. (GC= Global controller, LC=Local controller)

trained WANNID weights. As discussed before, improvement in the state dynamic prediction that is gained by WANNID, leads to better approximating the cost function by means of Critic NN. The value priority design, captures the error in the form of $\dot{J}_{id}$ and $\dot{J}^{g}_{vp}$, and computes the value priority parameter k. It can be concluded from the figure that, when the identification error is reduced, value priority switches to the global controller. In this case the proposed value priority design has been validated that can make sure that the closest global optimal solution is achieved.

4.4.2 Case A.2. 8-bus system: Performance Component

In this case study, a 100 ms three-phase short circuit in the middle of one of the tie-lines between area 1 and 2 is simulated. The line is disconnected by means of breakers

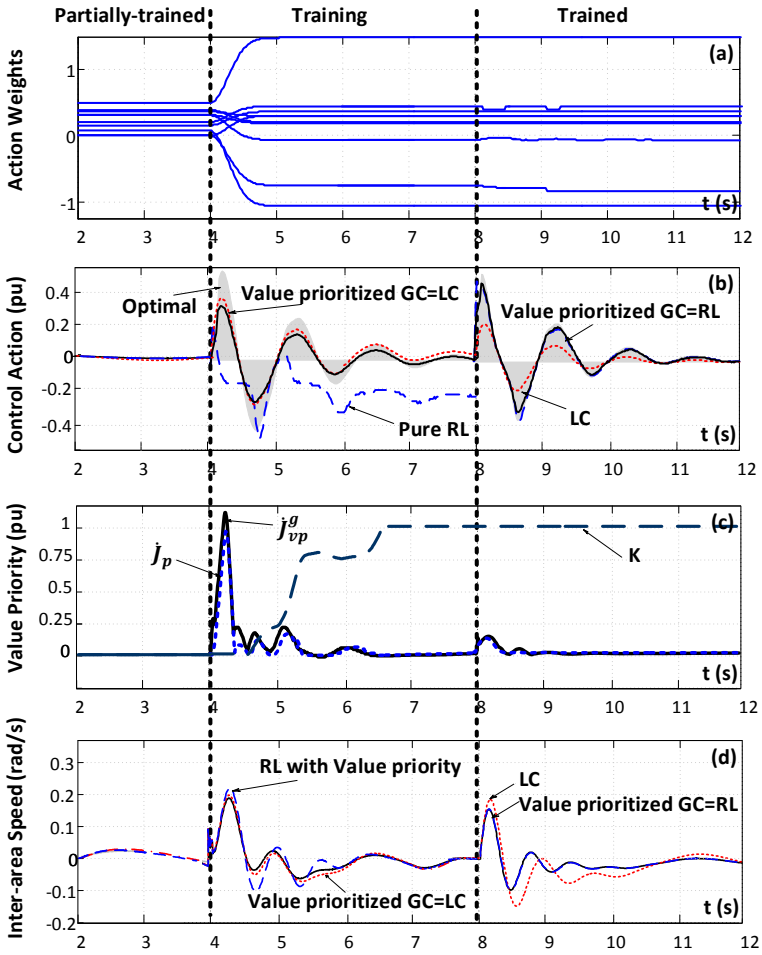


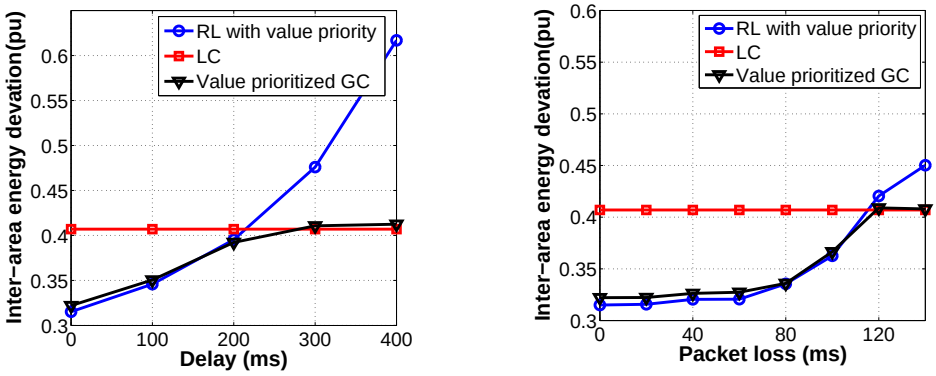
Figure 4.6: Case A.2. results. (GC=Global controller, LC=Local controller)

located on either ends and then reclosed after 4s. In this case, Action NN weights are initialized as partially random at $t = 4s$ and further trained online, making two periods of partially-trained and trained Action (Fig. 4.6). Like the previous test case, the trained Action NN has led to better performance in comparison to local PSS with respect to inter-area oscillation when the trained weights are deployed. It can be noted from the figure that the global controller's action is closer to the optimal value in second period, though, RL alone was unstable in the first period due to partially-trained weights. As it can be seen, the proposed architecture performs as one of the controllers operating alone at each time period with a transition value of k. At 4s the value priority sets the action target as the local controller, and further, in the course of time and training, when $\dot{J}_p$ and $\dot{J}^g_{vp}$ is reduced, value priority made a gradual shift

from full supervision to full autonomy of the global controller. This case validates the effectiveness of the value priority scheme, as it not only maintains the stability of the system, but also improves the learning capability of the RL.

4.4.3 Case B.1. 68-bus system: Data transmission delay and loss

In this case study, a self-clearing 50 ms three-phase short circuit at the middle of the tie-line connecting buses 26 and 29 between sub-area 1 (G9) and sub-area 2 (G1 and G8) in area 1 is simulated. This event activates oscillatory mode between these two regions, causing the energy of these sub-areas to change. This test mainly evaluates the data transmission delay and the packet loss effect on the proposed value prioritized global controller. Fig. 4.7a depicts the changes in the energy deviation between the sub-areas under study with respect to time delays. It can be seen that the performance of fully trained RL controller is worsen with time delays increment. However, for delays more than 200ms, the value priority algorithm directs the control action to local controllers. Further analysis is done on data discontinuation or missing packet (Fig. 4.7b). As it is shown, packet loss for 120ms (12 time steps) and above leads to deteriorated performance of the global controller, which is recognized by value prioritization algorithm. The method then shifts the control signal giving priority to local controllers and keeping the overall system resilient and the design robust and



(a) With transmission delay (b) With data packet loss

Figure 4.7: Case B.1. inter-area energy deviation between sub-area 1 and 2.

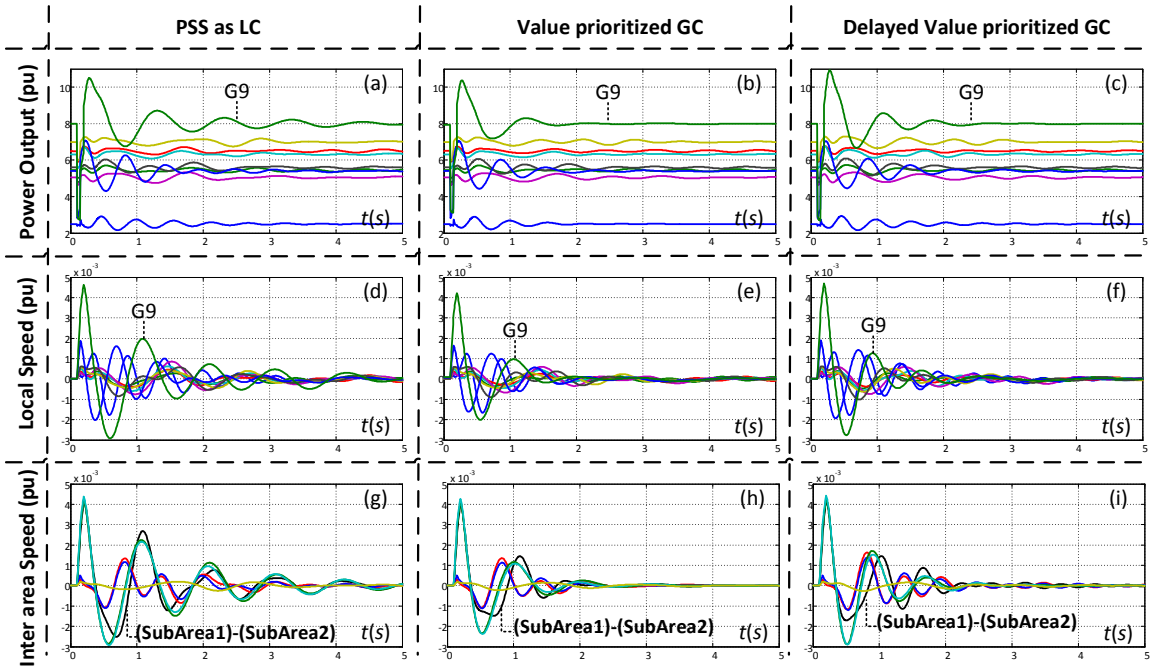


Figure 4.8: Case B.1. time domain simulation for the 68-bus system: the sub-figures present power output of the generators, their frequency, and inter-area frequency oscillations for local control, proposed global control both without delay and in presence of 100 ms delay. A three phase fault at 0.1 s in the middle of tie-line between sub-area 1 and 2 is occurred, which clears at 0.2 s.

reliable.

Furthermore, time-domain performance of the proposed controller with fully accessible WAM and 100 ms delayed data is provided in Fig. 4.8: see Figs. 4.8a and 4.8b for comparisons on the generators power output, Figs. 4.8d and 4.8e for local speed oscillations, and Figs. 4.8g and 4.8h for speed oscillations between sub-areas. As it can be concluded from these figures the proposed controller has higher damping capabilities especially on G9 leading to mitigating inter-area oscillations. Yet, Figs. 4.8c, 4.8f, and 4.8i shows the effect of transmission delay of 50 ms in both sending and receiving data to WAC from WAM. It has been demonstrated in previous tests that the proposed global controller outperforms the local controllers when an inter-area disturbance occurs. The same conclusions also hold in the presence of this severe delay as the global controller can predict the dynamics of the system in a forward

view and compensate for the delay. It can be seen that with the deterioration of the quality of the data transmission, oscillation overshoot increases and the damping is decreased. This test validates the value priority capability in recognizing the negative delay impact, and mitigating it through the WANNID adaptation as described before.

4.4.4 Case B.2. 68-bus system: Comparison with conventional controllers

The purpose of this test is to compare the effectiveness of the proposed value prioritized global controller with the conventional local controller PSS and conventional global power system stabilizer (GPSS) [4, 116] with the transfer function of:

$$u_{GPSS} = 1\frac{10s}{1+10s}(\frac{1+0.6280s}{1+0.1025s})^3\Delta\dot{\delta}_{jk} \tag{4.28}$$

GPSS designed as a wide-area controller, is based on global loop which involves a single differential frequency signal between two suitably selected areas of j and k. The parameters are tuned based on small-signal analysis.

This case as a transient test study affects the oscillatory mode between area 1 and 2. Three tie-lines connecting areas 1 and 2 are disconnected sequentially, which finally generates two separate areas. First lines 1-27 and 8-9 are disconnected, allowing the proposed global controller with initial random weights to gain enough knowledge about the topology and optimal policy. The events are 40 sec long, allowing the system to converge to post-fault equilibrium points. Furthermore, line 1-2 is disconnected due to a self-clearing 100 ms 3-phase fault, which leads to severe mismatch between the energies and complete separation of these two areas. An eigenvalue analysis of the open-loop and closed-loop system with different control schemes was carried out for the events to examine the performance of the designed controller in terms of improving the frequency and damping ratios of the corresponding inter-area mode. Parameters of the oscillations is derived by means of prony analysis, which is summarized in Table 4.2. It can be seen that the damping ratios of oscillations fol-

Table 4.2: Case B.2. oscillations parameters for sequence of tie-line outages.

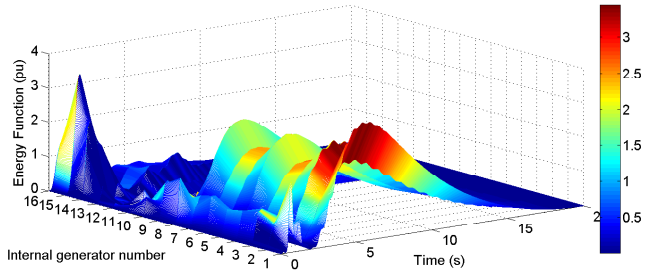
Event	No control		LC		GPSS		Value prioritized GC	
	f(Hz)	ζ(%)	f(Hz)	ζ(%)	f(Hz)	ζ(%)	f(Hz)	ζ(%)
Event 1 (8-9)	0.5899	0.0321	0.5967	0.1324	0.5764	0.1543	0.5961	**0.1324**
Event 2 (1-27)	1.1494	0.0177	1.1998	0.0344	1.2011	0.0358	1.2023	**0.0382**
Event 3 (1-2)	1.9014	0.0164	1.8905	0.0324	1.8558	0.0389	1.8502	**0.0511**

lowing the sequence of tie-lines outages has decreased and the frequency of inter-area oscillation has increased. At the first event, GPSS's damping (0.1543) is better than that of value prioritized GC (0.1324). However, in the presence of the value prioritized global controller with initialized random weights these characteristics improves during the course of events. This sheds light on the RL advantage, as it can *reinforce* the performance and allows the supervised RL to gain more knowledge about the optimal policy of action and energy during the course of events. Moreover, it is apparent that direct and predictive nature of the proposed global controller over the static linear design of GPSS can significantly enhance damping support in transient cases. This study validates the reliability and speed of convergence of the proposed architecture.

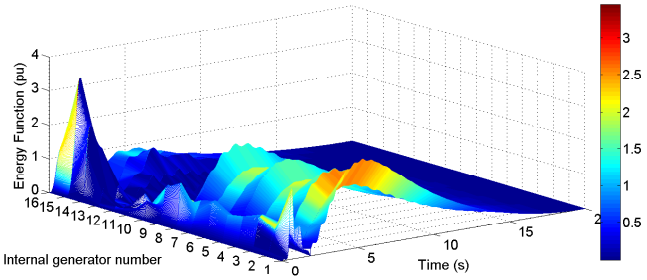
Time-domain performance of the proposed value prioritized global controller for the last outage, fault on line 1-2, is compared to local PSS and conventional GPSS. Figs. 4.9a, 4.9b, and 4.9c depicts the energy function changes of each generator with respect to time, in the presence of different control architecture. This figure indicates that considerable energy damping has been gained by the proposed architecture.

4.4.5 Case B.3. 68-bus system: Real-time Simulation

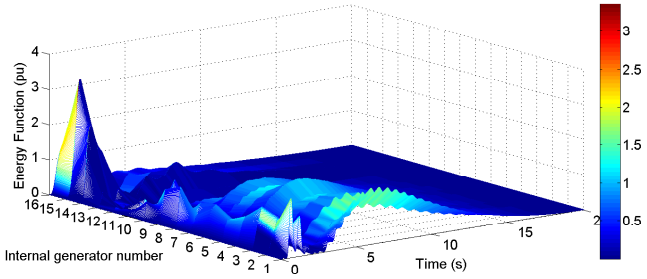
In order to investigate the performance of the proposed global controller for real-time transient stability improvement, a real-time power system simulation using the Hypersim simulator has been analysed [117]. The New England part of the test



(a) Local controller (PSS)



(b) Conventional global controller (GPSS)

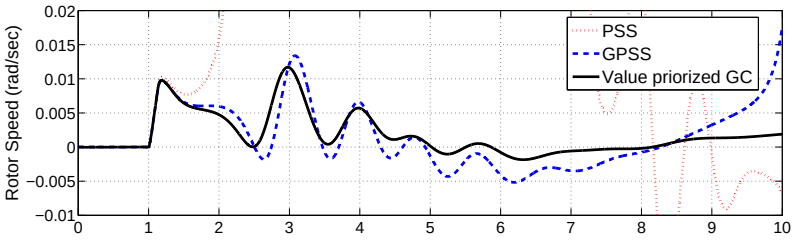


(c) Value priorized global controller

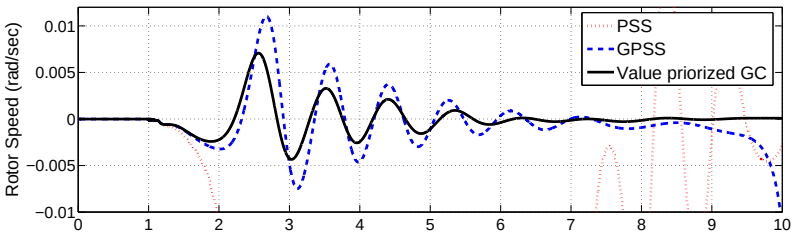
Figure 4.9: Case B.2. energy function of IEEE 68-bus system in the case of separation of area 1 and 2 in event 3 of Table 4.2 with different controllers

system, so-called IEEE 39-bus system, has been developed using the software library. The modeled power systems generators is then connected with exciter and PSS acting as a local controllers. These controllers run at the speed of real-time simulation, which is 50 us. The model is then imported from Matlab-Simulink through the C-code conversion and deployed in the Hypersim environment. The input of the exciter is then augmented by the global control actions running at the time step of 10 ms. Also, a conventional GPSS is simulated for comparison purpose.

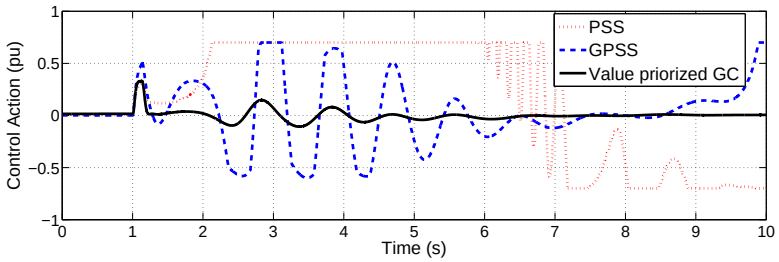
Using the described real-time platform, the proposed controller performance in a



(a) Local rotor speed deviation of G9



(b) Inter-area speed oscillation between sub-area 2 and 3



(c) Control actions

Figure 4.10: Case B.3. real-time response of the system and controllers

non-linear time domain case is simulated. For this purpose, self-clearing three-phase short circuit faults at bus 17 with duration of 100 ms, 120 ms, and 130 ms are initiated. These faults affect the inter-area modes of the system, specifically, sub-area 2 and 3. The fault with the duration of 130 ms can cause the system with the local and global PSS have unstable behavior. Fig. 4.10 shows the performance of the controllers as the local speed deviation at G9, inter-area speed between sub-area 2 and 3, and control action chosen. In this case, the value prioritized global controller has maintained the stability of the system and damped the oscillation considerably. It shows that the value priority not only has chosen the closest to optimal solution, also, has maintained stability of the system.

4.5 Learning-based Adaptive Controller Topology

Furthermore, we expand the proposed hybrid design to a learning-based adaptive control system as shown in Fig. 4.11. As demonstrated in the figure, the control architecture develops a signal u (a combination of an adaptive controller signal u^{ad} and a neuro-controller u^r) that can control the system in presence of parametric changes and functional nonlinearities and uncertainties. Both controllers are updated based on a specific adjustment mechanism developed using a reference model. For the adaptive controller, reference model generates a stable reference output to be tracked by the system. The adaptation rule is on the basis of a Lyapunov function, a linear model of the system, and the closed-loop error. As a measurement-based tool, neuro-controller is updated based on RL algorithm and ACD technique. In this method, using the Identifier NN, the Critic NN for current and future time steps, and controller NN, the optimal cost function in approximated. The combined identified and reference model consist of utility and cost functions that are optimized in a predictive manner. The adaptive controller, here MRAC, forces the system output to track the reference model during parametric changes. Theoretically this task bounds the deviation in the system function. Thus the NN RL learns the system functional changes in this bounded region. The main advantage of this approach is that NNs can perform the functional adaptation without excessive tuning and saturation.

The value priority block acts as a manger to generate a performance trajectory, which tracks the stability and optimality of the closed-loop system. Receiving the data needed, RL performance is first compared to adaptive controller with respect to optimal action backward-in-time. Further, identification and stability indexes are calculated in Lyupunov function and input-output stability framework. The approach thus guarantees quadratic stability of the system ($\dot{J}_k^{ad} < 0$), while improving optimal performance ($\frac{\partial J_k^r}{\partial u_k^r} = 0$). Overall, the technical issues addressed by the hybrid architecture are: allocating different networks to learn different tasks based on value

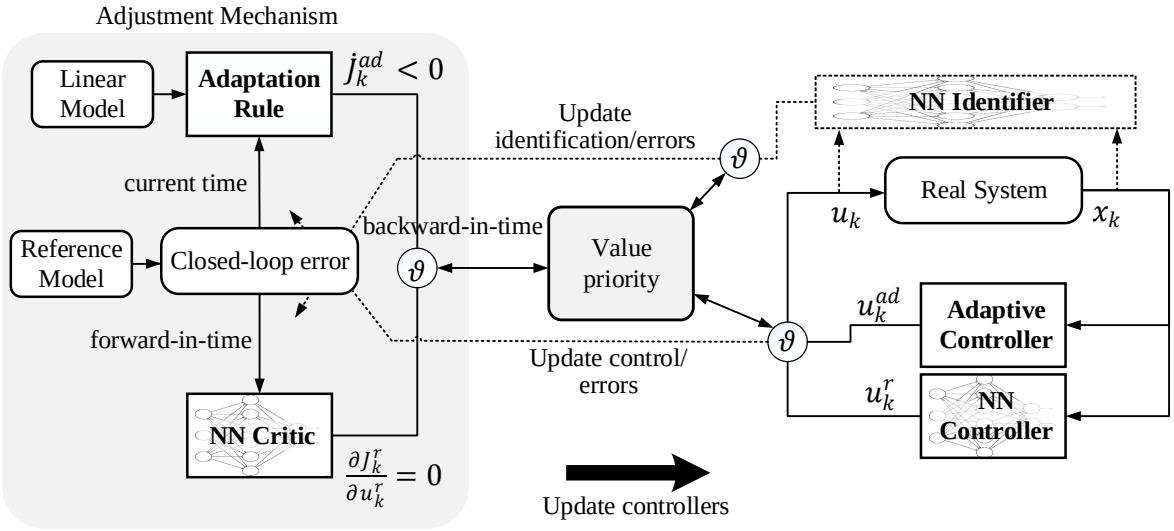


Figure 4.11: Learning-based adaptive controller topology

priority probabilistic rules, detecting the RL training patterns and operating regions, and shifting the control output toward RL as it has gained knowledge from the system and the linear controller.

4.6 Mathematical preliminaries

4.6.1 System Identification, Energy Function, and Control

Let $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ denote the system states and control inputs vectors, respectively. The discrete system dynamics are separated into two components of a linear parametric and an unknown nonlinear functional adaptation. Consider a general class of nonlinear system as

$$x_{k+1} = \underbrace{Ax_k + Bu_k}_{\text{linear model}} + \underbrace{f(x_k) + g(x_k)u_k}_{\text{nonlinear model}} + d_k \tag{4.29}$$

where, k refers to the kth time step, $f(.) \in \mathbb{R}^n$ and $g(.) \in \mathbb{R}^{n \times m}$ are unknown bounded nonlinear functions and $f(0) = 0$ and $g(0) = 0$, $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are the unknown parametric change, and d is unknown and bounded uncertain disturbance.

The goal of this work is to find an estimate of A, B, $f(.)$, and $g(.)$ and an associated

confidence interval of this estimation from measurement data. This information is furthermore used to design a hybrid adaptive nonlinear controller that stabilizes the system despite the uncertainty in the estimates. In general, the control task is to determine optimal control law for the system state vector x to converge the system from a known initial state x_0 to a new stable equilibrium x^*. To accomplish this task, we use a linear and nonlinear direct energy function methods, which focuses on the transient dynamics of the system. In other words, the goal of the chapter is to develop computationally tractable optimal control and transient stability cost function policy guaranteeing system convergence to the optimal equilibrium. The energy function of the overall system (4.29) can be defined as a superposition of linear and nonlinear Lyapunov functions proposed on tracking error $\Delta x = x - x^*$ as,

$$J(x_k) = \Delta x_k^T P \Delta x_k + J^{nl}(\Delta x_k) \tag{4.30}$$

The first element in the above equation refers to energy of the linear part. The second element, J^{nl}, is a function mapping $\mathbb{R}^n$ to $\mathbb{R}$ for nonlinear Lyapunov function [22,118]. It can be proved that the system is asymptotically stable in the sense of Lyapunov stability criteria such that:

$$J(x_k) \geq 0, \quad \Delta J(x_k) \equiv J(x_k) - J(x_{k-1}) \leq 0 \tag{4.31}$$

For this, the weighting matrix P is of appropriate dimension and required to be positive-definite ($\lambda_i(P) > 0$), and,

$$AP + PA^T + Q_0 = 0 \tag{4.32}$$

where $Q_0 > 0$. Also, $f(x)$ in (4.29) should satisfy $f(x)^T N x > 0$ for some constant real matrix N, $J^{nl}(x) \geq 0$, and $\nabla J^{nl}(x) = Q^T f(x)$ for some constant real matrix

Q [118]. Then,

$$\Delta J(x_k) = \Delta x_k^T Q_0 \Delta x_k + u_k^T R u_k + \Delta x_k^T P f(x_k) + Q^T f(x_k)(x_k - x_{k-1}) \tag{4.33}$$

Overall, as long as the effect of nonlinearity is relatively small and bounded, smaller cone of Lyapunov function can be defined to clarify the stability of x^* and estimate the domain of attraction [22].

To this end it is desired to find the control policy $\pi = \{u_0, u_1, ...\}$ that satisfies the aforementioned Lyapunov stability criteria. Moreover, an optimal control policy u^* can be found that minimizes the infinite-horizon cost function $J_\pi(x_0)$. Considering that dynamics of the system can be identified, controller action can be presented as linear and non-linear parts. Denoting u^{ad} as initial controller around the plant equilibrium point which is adaptive in nature, a nonlinear-based controller can be used to approximate the system functional dynamics. The RL-based controller can be activated to learn the nonlinear subsystem dynamics of $f(.)$ and $g(.)$ for which no prior knowledge is available.

4.6.2 MRAC Adaptation

In general, MRAC works so that the state vector x tracks a desired state vector as specified by the reference input under the condition of the state vector remaining bounded. Let the tracking reference model be represented as

$$x_{k+1}^{ref} = -A_m x_k^{ref} + B_m r_k \tag{4.34}$$

where, r is the input command, and A_m and B_m are reference coefficient matrices. In this chapter, direct adaptive approach is employed, where plant parameters are considered as known and the control parameters are adjusted directly to minimize the closed-loop error between system and reference model states [119]. Based on

(4.29), the first-order linear time-invariant system can be described by

$$x_{k+1}^{ad} = Ax_k + Bu_k^{ad} \tag{4.35}$$

With the knowledge of system dynamics, an adaptive linearizable control feedback can be estimated as

$$u_k^{ad} = \theta_x x_k + \theta_r r \tag{4.36}$$

The closed-loop error equations for $\Delta x = x - x^{ref}$ can then be derived in terms of adaptation parametric set of $\theta = [\theta_x, \theta_r]^T$ as,

$$\Delta x_{k+1} = -A_m \Delta x_k + B\Delta\theta_x x_k + B\Delta\theta_r r \tag{4.37}$$

where, $\Delta\theta = \theta - \theta^*$ are controller parameter error. An adaptive law can be generated based on Lyapunov function candidate to converge these estimates to the actual parametric values denoted with the superscript $*$. At every time step, the aim is to determine adaptive law for θ updating continuously so that the overall system has bounded solution and the state error Δx tends to zero asymptotically. This adaptive law can be derived as,

$$\Delta\theta_x = -sgn(B)\Delta x.x, \quad \Delta\theta_r = -sgn(B)\Delta x.r \tag{4.38}$$

where $\Delta\theta_k = \theta_k - \theta_{k-1}$ is the parameter update at time step k. From the definition of the closed-loop error and adaptive laws it can be shown that a Lyapunov function exists as linear parts of (4.30). This proves the stability of the controller. It should be noted that an additional augmentation of growth rate of parameters can also be used to ensure the boundedness of the signals [119].

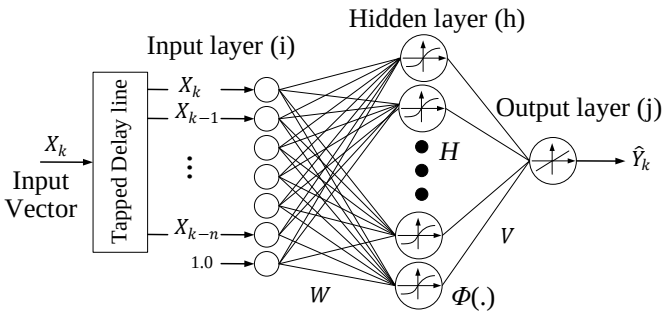


Figure 4.12: Feed forward neural network structure

4.6.3 ACD Adaptation as Reinforcement Learning

In general, ACDs can perform optimal control policy for a given system based on RL by means of ADP principles. Essence of ACD is approximation of the Hamilton-Jacobi-Bellman (HJB) equation associated with optimal control theory and meeting a certain objective function, here, Lyapunov stability function [7]. The HJB equation is developed using $\Delta x = x - x^{ref}$ around the reference point for *present* time step and *future* window. System nonlinear identification is carried out by means of adapting FFNN presented in Fig. 4.12, represented as

$$\hat{x}^r_{k+1} = \hat{W}_I^T \Phi_I(V_I^T[x_k, u_k]) + \epsilon_I(x_k, u_k) \tag{4.39}$$

where, $\Phi \in \Re^h$ is the corresponding nonlinear mapping function of the states which is usually considered as sigmoidal function, $W \in \Re^{h,j}$ and $V \in \Re^{i,h}$ are the parameter vector of weights of the FFNN where $i, h, j \in \mathbb{N}$ are dimensional space of the feature vector representing the input, hidden, and output layers of the NN, and ϵ is the NN functional approximation error. If we fix the first-layer weights, then NN has only one layer of tunable weights. Thus the identification error can be defined as,

$$e_k^I = x_k - \hat{W}_I^T \Phi_I(V_I^T[x_k, u_k]) \tag{4.40}$$

The design of an optimal controller is based on the fact that the cost function

or policy approximation is converged to a best optimal value, J^*, with an optimal control action, u^*. In the ACD architecture this task is performed by,

$$\hat{u}_k^r = \hat{W}_A^T \Phi_A(V_A^T x_k) + \epsilon_A(x_k) \tag{4.41}$$

$$\hat{J}^r(x_k) = \hat{W}_C^T \Phi_C(V_C^T x_k) + \epsilon_C(x_k) \tag{4.42}$$

where J^r represents the nonlinear mapping in (4.29). It is required to $J^r(x_0) = 0$ as a boundary condition to serve as a Lyapunov function [49].

4.7 Hybrid Global Control Structure

In this section we expand the value priority scheme developed in previous sections for local and RL-based WAC toward global adaptive and RL-based WAC. This will allow us to build a hybrid controller that provides quadratic stability, while improving the performance of the overall closed loop system. The RL, the expert and the value priority networks form a hybrid scheme that provides a composite control action to the system as,

$$u_k = \vartheta_k^{ad} u_k^{ad} + \vartheta_k^r u_k^r \tag{4.43}$$

where, ϑ^r and $\vartheta^{ad} \in [0, 1]$ are value priority trajectory of controllers. The RL network competes with the expert to learn the training patterns and the value priority network mediates this competition. The learning policies used to train the RL and the value priority networks are based on the minimization of HJB equation, respectively, forward and backward in time. For this, intuition is expressed mathematically as follows.

4.7.1 Value Priority Network Update

As discussed before, the value priority computes the action, u, as a weighted sum of the actions received by both component policies, u^r and u^{ad} (see Fig. 4.13). The

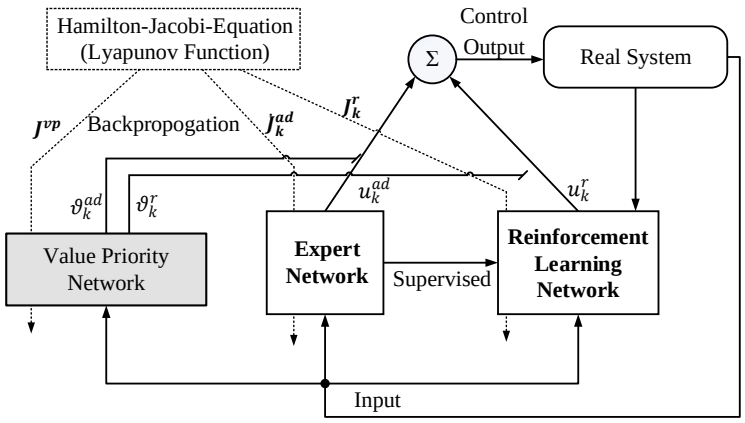


Figure 4.13: Hybrid control structure with value priority network

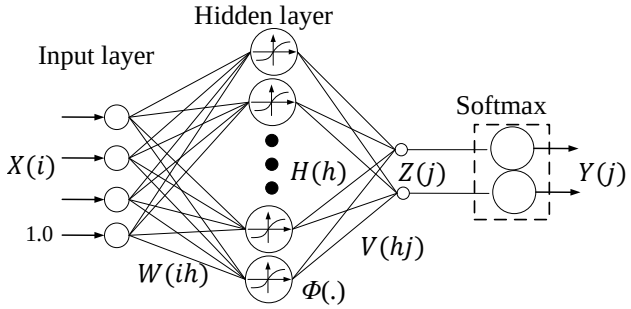


Figure 4.14: Value priority network structure

activation of these output units must be non-negative and sum to one. To meet these constraints, we use Softmax activation function [114] at the output layer of the FFNN as depicted in Fig. 4.14 and represented as,

$$\vartheta^r = e^{z^r}/(e^{z^r} + e^{z^{ad}}), \quad \vartheta^{ad} = 1 - \vartheta^r \tag{4.44}$$

where z is the total weighted inputs received by the corresponding output as,

$$z = \hat{W}_{vp}^T \Phi_{vp}(V_{vp}^T x) \tag{4.45}$$

The training value priority error function for the controllers, J^{vp} is derived from the Lyapunov stability function of the system, (4.29). Evaluating the total energy, which

is the cumulative effect of changes in the energy ΔJ from previous states and the two control actions, the sensitivity of each controller with respect to the total energy can be extracted. Hence, the priority of each controller that reduces this energy function can be learned. The proposed function is combination of the identification and controllers performance as,

$$J_k^{vp} = \sum\nolimits_{n=0}^{k} \lambda^{k-n}(e_n^I Q_0 e_n^I + \tilde{u}_n R \tilde{u}_n) \tag{4.46}$$

The first part is a measure of the identification error of the controllers over a finite backward window $n \in (0, k]$ in order to distinguish the operating region. The significance of this element as a part of value priority function is that, as the error of identification for RL gets smaller, it can provide closer to optimal estimate of the cost and control policies. On the other hand, if the system is known to MRAC, it can take control of the system.

The second part is the performance function developed on $\tilde{u}_n = u_n - u_n^*$ as the estimated control error at time step n. In order to derive this element, quality of the controllers is iteratively evaluated with respect to u^*. Towards this, considering the current time as the terminal state, calculations are done backward for J and u as,

$$J_i(x_n) = U(x_n, u_{ni}) + J_i(x_{n+1}) \tag{4.47}$$

$$u_{ni} = \arg\min_u(J_i(x_n)) \tag{4.48}$$

where, i is iteration number indicating the number of times the window has been swiped, and u_n is estimated as,

$$\begin{aligned} u_n &\approx R^{-1}(\frac{\partial x_{n+1}}{\partial u_n})^T \frac{\partial U(x_{n+1}, u_{n+1})}{\partial x_{n+1}} \\ &\approx R^{-1}(\frac{\partial x_{n+1}}{\partial u_n})^T (Q_0 x_{n+1} + R u_{n+1} \frac{\partial x_{n+1}}{\partial u_n}) \end{aligned} \tag{4.49}$$

The state estimation is carried on in forward manner, to estimate the new trajectory and sensitivity by

$$x_{n+1} = \vartheta^{ad}_{k-1}(Ax_n + Bu_{ni}) + \vartheta^{r}_{k-1}(f(x_n) + g(x_n)u_{ni}) \tag{4.50}$$

$$\frac{\partial x_{n+1}}{\partial u_n} = \vartheta^{ad}_{k-1}B + \vartheta^{r}_{k-1}g(x_n) \tag{4.51}$$

It is assumed that the estimation of the model is accurate enough for the visited states based on identification value priority criteria, and the HJB equation converges to a solution in this window. If we differentiate the value priority function with respect to the outputs of the value priority network we get a signal for training the network. Overall, the problem formulation requires that whenever there is a system trajectory movement towards RL, then ϑ^r is raised, meaning that the RL has learned the optimal cost function and generates less than the average squared error of the controllers. The rate of this learning is

$$\frac{\partial J^{vp}}{\partial z^r} = \vartheta^r(J^{vp|r} - J^{vp|ad}) \tag{4.52}$$

Furthermore, back-propagation of this error function yields update to the weights, which generate the value priority trajectory. The hybrid shifting rule is initialized by $\vartheta^r_0 = 0$ as representation of adaptive control activating the system.

4.7.2 Operating Region and Identification Update

The goal of the hybrid control problem considered in this work is to stabilize the system around a fixed desired operating point through the energy function minimization. Due to the fact that *a priori* operating points are given by the adaptive controller and associated steady-state inputs are known, we can iteratively update the operating point for the unknown dynamics of the ACD controller as $\Delta x^r_k = x_k - x^{ref}_k$. Hence, for the ACD combination, the reference model output will act as a predictive

stable reference output. The neuro-identifier predicts plant output that needs to be tracked at time $k+1$, and the action NN will generate control signal accordingly in a predictive manner.

Furthermore, the goal of RL is to eventually evolve to learn both linear and non-linear regions. This yields to operating region of $\hat{x}_{k+1} = \vartheta^{ad}\hat{x}^{ad}_{k+1} + \vartheta^{r}\hat{x}^{r}_{k+1}$. The aforementioned rule thus makes the action NN independent of the adaptive controller in the case of learned scenarios.

4.7.3 Supervised Reinforcement Learning Update

In this part we present an iterative supervised ADP algorithm, where the cost function and the control law are updated recursively. In the proposed method the cost-to-go function is given based on the Lyapunov stability function as follows:

$$J^{r}(x_k) = \sum\nolimits_{n=k}^{\infty} \gamma^{n-k} U(x_n, u_n) \tag{4.53}$$

where, U is the utility function and γ is the discount factor [7]. An explicit utility function is used for reward/punishment for RL, or as incremental cost function in Lyapunov stability concept based on,

$$U(x_t, u_t) = \Delta x_t^T Q_0 \Delta x_t + u_t^T R u_t + Q^T f(x_k)(x_k - x_{k-1}) \tag{4.54}$$

which is equal to ΔJ_k in (4.33) with neglection of linear and nonlinear interaction element.

The process of training a NN, as the implementation benchmark of this method, requires computing an error value that describes how the NNs output varies from the target value. In the proposed method, the value priority parameter, ϑ, not only provides u, but also plays an important role to modify the RL's policy as well. After each state transition, the parameters of RL are updated according to the update rule

derived as

$$\Delta W_C = \alpha[\vartheta^r e^C + \vartheta^{ad}(J^{ad} - J^r)]\nabla_{W_C} J^r \quad (4.55)$$

$$\Delta W_A = \alpha[\vartheta^r e^A + \vartheta^{ad}(u^{ad} - u^r)]\nabla_{W_A} u^r \quad (4.56)$$

where, α is small step size learning parameter, J^{ad} and u^{ad} are adaptive references for Critic and Action NNs, e^C and e^A are critic and action error in RL algorithm. This feature leads the RL controller to merge faster towards supervised controller especially early in the learning process when the critic has a poor estimate of the optimal cost function based on the value priority criteria. Furthermore, it is proven that the update method for the proposed structure of supervised RL yields to optimal value and is converging to a compact optimal set.

4.8 Case A. Implementation as a Power System Stabilizer

The states chosen for monitoring the rotor angle stability to provide damping and synchronizing forces for oscillations as mentioned before are rotor angle and speed of G3 (δ_3, $\dot{\delta}_3$) and the control action is excitation voltage reference (V_{e3}). The output of the controller is augmented to the excitation system, mitigating the kinetic energy of the synchronous generator through the field energy control. The input to the adaptive controller is the reference command based on linear characteristic equation of the system, the reference model and the plant output. The reference model ensures that the generator is stable at all times and incorporates the dynamic preferences such as rise time and damping. Since the degree of the system is $n = 2$ the error polynomial can be represented as $s^2 + z_1 s + z_0 = 0$. Setting $z_0 = 3.56$ and $z_1 = 2.67$, the roots can be placed at left-half plane in the frequency domain. The reference model was chosen to have a settling time of 3 seconds and a damping ratio of 0.707, which is an industry accepted standard. Choosing $Q_0 = diag(1, 1)$, and then by solving (4.33) we get $P = [0.6149, -0.5000; -0.5000, 0.8539]$. Control coefficient is considered as

$R = 0.1$ and nonlinear coefficient $Q = [1; 0]$.

Specification of the ACD is as follows. The Action NN architecture consists of a three layer FFNN with two inputs, and a single hidden layer with 15 neurons and one output. The inputs are the generator angle and speed. These inputs are time delayed by 20 ms and, together with two previously delayed values, comprise the three inputs. The number of neurons in the hidden layer is identified in a heuristic manner. Critic NN has the specification with the output of one cost function. As discussed in previous sections, Identifier NN is used to learn the plant linear and nonlinear dynamics. The input to this NN is the states and the hybrid control. These inputs are time delayed by 20 ms and together with two previously delayed values, are used for learning.

For NN offline training, first, a set of random control inputs and respective outputs of the power system model for 100 sec (5000 time steps) is fed into the identifier, and the outputs (the next time step's states) are extracted. The offline training as a batch learning algorithm is stopped when the Root Mean Square Error (RMSE) between the actual plant states and the identified ones, represented as e^I, reaches a user defined small number ($\epsilon = 0.01$). The learning rate of the NNs is $\alpha = 0.01$. During online training, incremental learning is performed, in which, the identifier weights are adjusted based on the plant output and the reference model at each iteration. The same methodology is used for training the Action NN and Critic NN with their respective targets. Action training is started with the target output of MRAC. Furthermore, it is trained alongside the Critic NN, to perform optimization of the cost function. Utility function is developed based on Q_0, R, and Q to monitor the energy function deviation and the power system stabilizers is updated to mitigate the oscillations. Once the critic weights have converged, the Action NN is connected to the generator's exciter to augment the MRAC. These controller's outputs are then fed to the plant through value priority NN block.

The effectiveness of the proposed hybrid controller is evaluated on G3 of the 5-machine 8-bus power system. For this, dynamic equivalent frequency dependent models of the generators is developed using electromagnetic transient simulation (EMTP) software $PSCAD^{TM}$.

4.8.1 Case A.1. Short Circuit

In this case study, a 100ms three-phase short circuit at the middle of one of the transmission lines between area 1 and area 2 is simulated. In order to assess the performance of the controllers, the transmission line is disconnected by means of breakers at each ends of the line. This line is then re-closed after 7s. This case study as a transient test effects all oscillatory modes of the system and the system topology. During this time, the system moves away from the pre-fault operating points $\delta_2^0 = 0.257$ and $\delta_3^0 = 0.1962$ and experiences a nonlinear post-fault transition dynamics after the re-closer action. This action causes a mismatch between generator reference and electrical powers, which leads to frequency increment and change in the tie-line reactance. This yields to new operating points of $\delta_2^* = 0.3683$ and $\delta_3^* = 0.3111$. From Fig. 4.15a it can be observed that the initial and final operating points satisfy the nonlinear Lyapunov function constraints, and the energy is strictly decreasing in the region and converges to the post-fault optimal points. The reference value is obtained by means of the reference model response defined in previous section. Fig. 4.15b shows the G3 rotor angle response of the system with the MRAC and the proposed adaptive learning-based controller. The NN-based RL controller is able to track the nonlinear dynamics and energy function, and enhance the performance of the controllers with respect to damping capability of $\dot{\delta}_3$ oscillation, as shown in Fig. 4.15c. As it can be seen from the figure, a notable damping improvement is gained, when the proposed controller with trained RL is installed on G3.

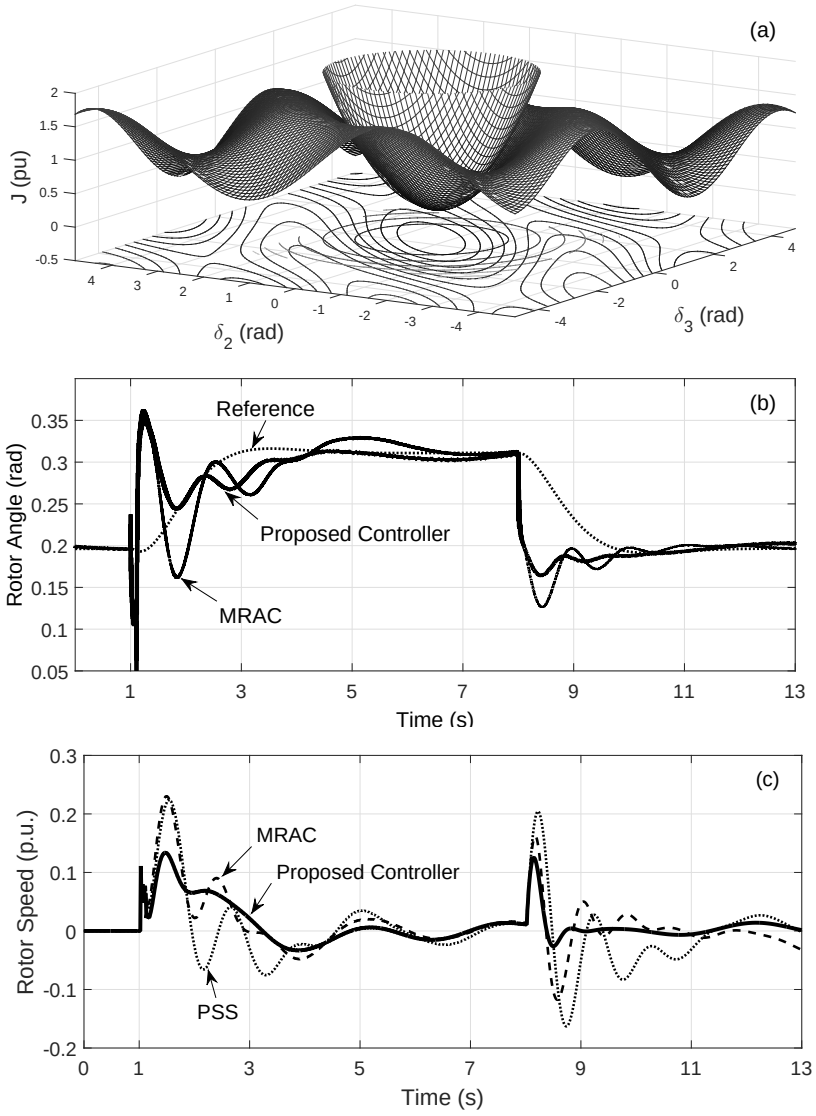


Figure 4.15: Case A.1. (a) linear and nonlinear energy function landscape as a function of δ_2 and δ_3, (b) G3 rotor angle deviation, δ_3, (c) G3 rotor speed deviation, $\dot{\delta}_3$, with different controllers.

4.8.2 Case A.2. Voltage Reference Change

In this case, a 0.05 p.u. step increase in exciter reference voltage of G3 is applied at $t = 1s$ and removed after 7s. The main purpose of this study is to show the capability of the value priority network in recognition of identification and performance criteria. In this case study, NNs with initial random weights is utilized for the RL-based controller. The terminal voltage response of G3 in the case of different controllers is

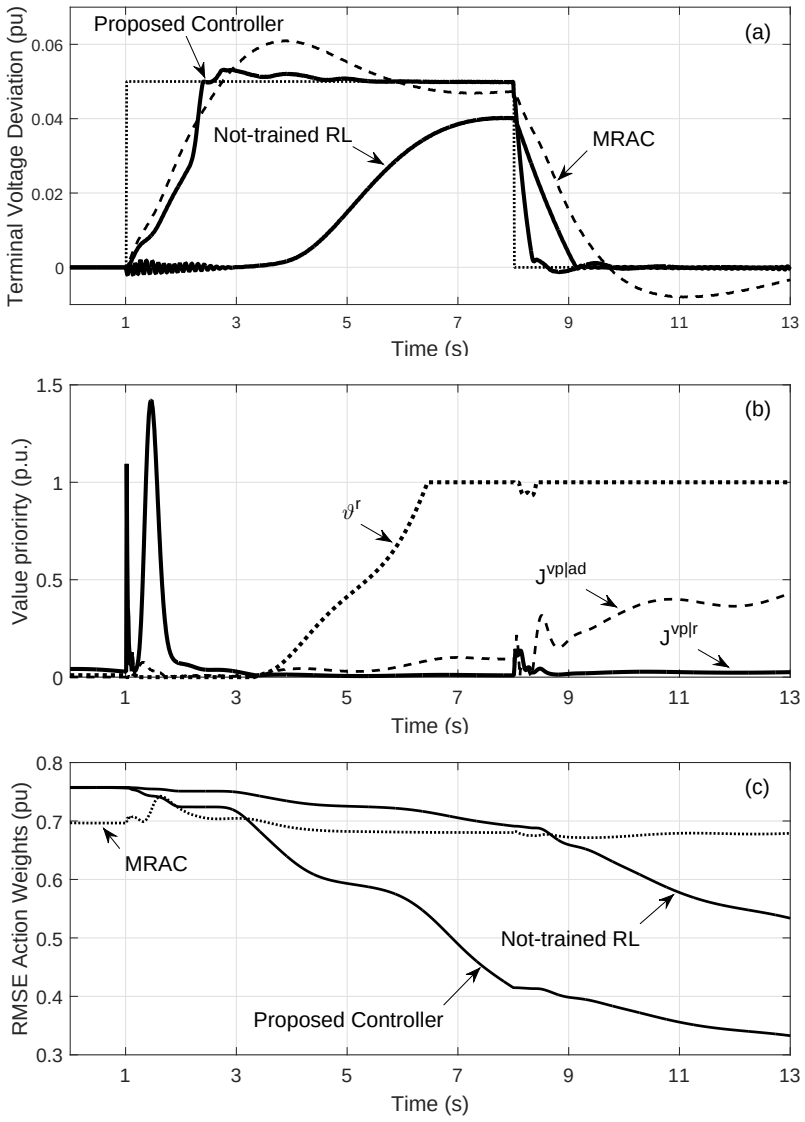


Figure 4.16: Case A.2. (a) dynamics of the G3 as terminal voltage deviation with different controllers, (b) value priority trajectory and elements, (c) RMSE of the control weights from the optimal weights.

shown in Fig. 4.16a.

It can be seen that the proposed control scheme provides a quick, well-damped response in voltage. It can be noted that, the gain scheduler made a gradual shift from full supervision to full autonomy as the NNs acquired enough knowledge to reach the optimal performance target. Fig. 4.16b shows the corresponding proposed value priority functions for adaptive and RL controllers ($J^{vp|ad}, J^{vp|r}$) and RL value

priority trajectory (ϑ^r). As it can be noted, initially $J^{vp|r}$ is larger than $J^{vp|ad}$ due to identification error. However after learning it will lean towards RL controller. It can also be seen that this method allows the RL to reach the optimal weights faster with the help of supervisor, leading to lower RMSE of Action NN weights (Fig. 4.16c). The optimal weights are calculated based on 100 epochs of batch learning for this specific scenario.

4.9 Case B. Implementation as a Wide-Area Controller

To illustrate the performance of the proposed approach as a WAC, tests are carried out on the 68-bus 16-machine IEEE test power system simulated in PST toolbox. To reduce the size of the data set for training, generator coherent groups are developed based on section 4.3.1.3. The tie-line 1-27 is disconnected due to a self-clearing 100 ms three-phase short circuit fault, which leads to severe mismatch between the energies of the two systems and angle deviation of 3.12°. Proposed controller and MRAC are designed based on global loop, which involves a multiple differential frequency signals of the selected areas with the uncertainty of the transmission delay. In this case study, it is assumed that the wide-area system consists of 100 ms transport delay. In order to predict the plant state vector at time step $k+10$, the values of state at time steps k, $k-10$, and $k-20$ are used. Fig. 4.17 depicts the time-domain performance of the controllers with 100 ms delayed data set. The local PSS acts on the local modes of oscillation with the frequency of $0.9824Hz$, while the proposed WAC tracks the inter-area oscillation with the frequency of $0.6492Hz$. Further, Fig. 4.18 depicts the Action and Critic NN performance in tracking their respective linear references and learning evoluation that minimize the actual Lyapunov function.

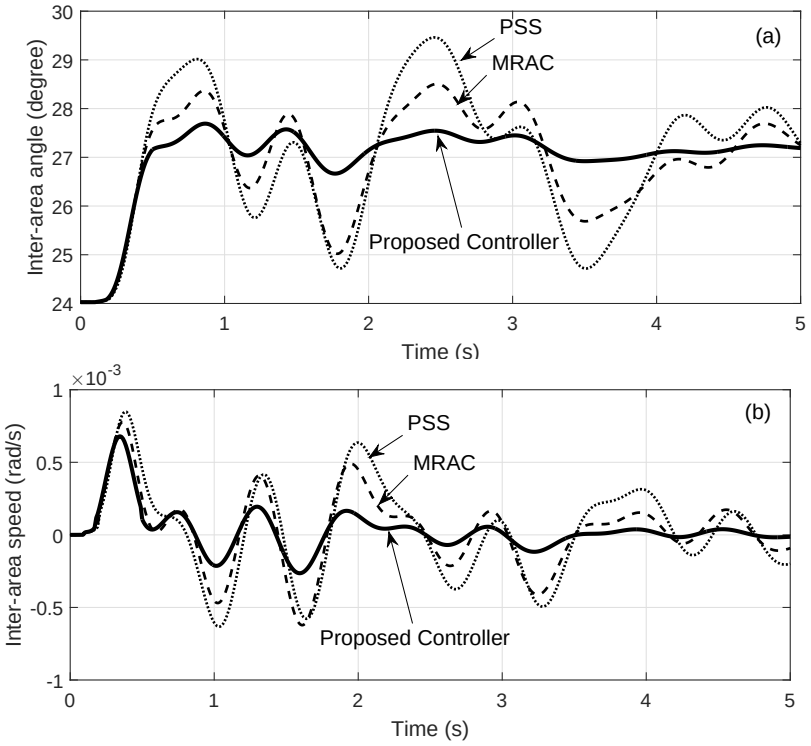


Figure 4.17: Case B. dynamics of the IEEE 68-bus system in event 2, short circuit in tie-line 1-27, as inter-area 1 2 angle and speed oscillations.

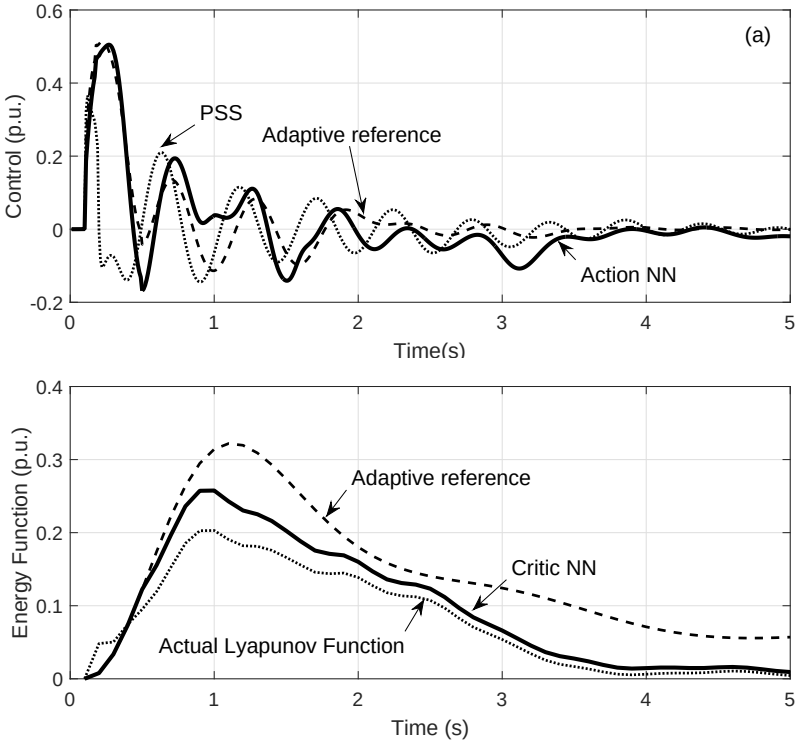


Figure 4.18: Case B. action and critic NN performance.

4.10 Conclusion

In this chapter, a new value prioritized global damping controller based on supervised reinforcement learning for transient stability improvement has been proposed. The proposed architecture shows better damping and transient response than the traditional local and global controllers. The design has been implemented in two scenarios. In the first scenario, the nonlinear learning-based controller has been implemented as a WAC design to augment the conventional local controllers. In the second scenario, the learning-based controller has been augmented to an adaptive controller which could model linear domain. The main advantage of this algorithm is that it is precise, feasible, and more reliable than other nonlinear adaptive controllers. Simulation results are presented for various contingency scenarios of stable and unstable cases including WAM constraints to show the effect of the proposed architecture in augmenting the local PSS installed on the generators on a 8-bus system and 68-bus IEEE system to damp the oscillations.

CHAPTER 5: TRANSIENT STABILITY ENHANCEMENT CONTROL

5.1 Introduction

As discussed in previous chapters, most of the works in the area of wide-area rotor-angle stability have utilized classical frequency domain based techniques in the form of Global PSS (GPSS) by incorporating additional remote signals to the local controllers [4]. Such methods rely on modal controllability and observability residues with a-priori passivity information for sensor and actuator locations [120]. However, the complexity of the large scale power system makes such analytical control techniques impractical. In addition, majority of these works are designed based on the model of the system linearized around a particular operating point. This arises the issue of generality, i.e. the proposed techniques are suitable only for a specific application. Additionally, these model-based techniques are not able to damp the oscillations in transient domain when power system is subjected to severe disturbances. Application of WAM in real-time disturbance analysis and electro-mechanical mode estimation has been studied in literature [121, 122]. The temporal information available from WAM can be used in Wide Area Damping Control (WADC) designs to perform transient stability enhancement, which can improve the power transfer capability of a transmission system and prevent the system from generation or load disconnection, or catastrophic failure following a sequence of disturbances in the system [72].

Traditionally, numerical methods and energy function based direct methods have been the most important conceptual frameworks for real-time transient stability assessment. However, numerical methods have shown considerably weak performance in real-time implementation, as they require accurate information of the power network topology; while, direct methods may lead to excessive simplifications. Such

problem formulation provides only the sufficient conditions for assessing the stability using direct calculation of the transient energy function [29]. Although some control techniques based on energy methods have been proposed, such as the Boundary of stability region based Controlling Unstable equilibrium point (BCU) [20], still the computational modeling complexity is considered as a challenge. With the advent of technology new techniques and approaches based on combination of these methods have been developed. These approaches mainly rely on equivalent modeling which can actually be integrated much faster than real-time. Further, direct energy methods can be utilized to predict the transient stability status of the system, as well as, the stability margins [30]. Another hybrid method to tackle the problem is early-termination criteria, allowing simulation of stable cases to be aborted as soon as possible [123]. This criteria for numerical simulations can be defined on the basis of coherency, transient energy conversion between kinetic energy and potential energy, and the product of system variables [124].

On the other hand, application of intelligent-based techniques has been successfully investigated in recent literature for post-fault transient stability assessment and power system control [28, 124–126]. Decision Tree algorithms, Fuzzy Logic techniques, Neural Networks (NNs), and Support Vector Machines (SVMs) are among these methods that effectively learn and map the system dynamics from the relationship between specified inputs and outputs without any prior knowledge of the system. Most of the works in this area has been designed for classification and remedial action schemes using supervised learning algorithm in order to predict the post-fault stability status.

This chapter aims at designing an optimal WADC to enhance the transient stability of the system in online/real-time application. There are mainly three issues critical to such a measurement-based control designs.

- First, the design should have the capability of dealing with nonlinear and non-stationary power system dynamics in the presence of uncertainities.

- Second, the method should satisfy closed loop system stability and global optimality conditions.

- Finally, the it should be able to project actual stability condition of the system, guaranteeing the stability of the power system augmented with the online control schemes; i.e. ensuring the gradient of the cost-to-go function is decreasing over the time.

In order to address the first issue, Reinforcement Learning (RL) technique is used to train the artificial intelligence. This approach, as opposed to supervised method can make use of the new knowledge to inturn improve learning in real-time during the course of events and actions. This technique can find optimal solution to the cost function by means of Adaptive Dynamic Programming (ADP) forward-in-time and provide an effective benchmark to construct an optimal controller by exploiting function approximators, e.g. NNs [127, 128], and has been implemented as power system controller in several research works presented in chapter 2. In chapter 3, RL has been employed as a WADC on real-time benchmark with wider horizon of prediction to take account for WAM delays. Major disadvantages noted in such designs are strictness of global optimal solution, the objective function definition, and the need for extensive offline training requirements.

Recently it has been shown that RL could be used as an optimal controller guaranteeing global optimal conditions for a non-convex functions [50]. Taking advantage of such a design, in this chapter, we show a method that mitigates the last aforementioned issue. In the proposed hybrid approach, the cost function in RL problem is defined based on the energy function damping and tuned online based on Prony method to ensure that the most suitable energy function is estimated. The controller is designed in order to ensure that the gradient of the cost function is strictly negative and minimized over this time period. Consideration of RL based method is due to the fact that for a given time duration it can reach to the global optimal solutions

in a nonlinear and uncertain environment [50]. Since the cost function is derived based on energy functions, unlike in conventional methods where it is updated by the Euclidean norm of the desired states, this method guarantees the system convergence to post-fault equilibrium. This is due to the fact that the proposed energy function is utilized to screen the level of stability of the system.

Fig. 5.1 present the overall scheme of the proposed WAC connection to optimal control theory and direct energy function method. In summary, the construction of WADC is initialized based on offline data derived from direct energy method and supervised learning. This design is further adapted to RL for online implementation and policy iteration with only partial knowledge of system dynamics. One of the main advantages of the proposed method, also shown in [28], is that the NNs would be able to capture the underlying relationship with smaller sized training data set and with higher accuracy when meaningful energy features are used as inputs. Moreover, through deployment of RL and by monitoring the system modes online the control scheme can be performed in real-time as well. This method is evolved from our preliminary work [129] for real-time wide-area monitoring and control of large power system.

The remainder of the chapter is organized as follows. The second section presents system modeling and direct energy function development and Section 5.3 discusses the proposed optimal WADC design. Section 5.4 and 5.5 present the power and control systems implementation test bed and simulation results followed by conclusions in section 5.6.

5.2 System Modeling and Energy Function Development

Considering x^* is the reference equilibrium point of the dynamic system expressed in the form of nonlinear continuous-time equations

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) \tag{5.1}$$

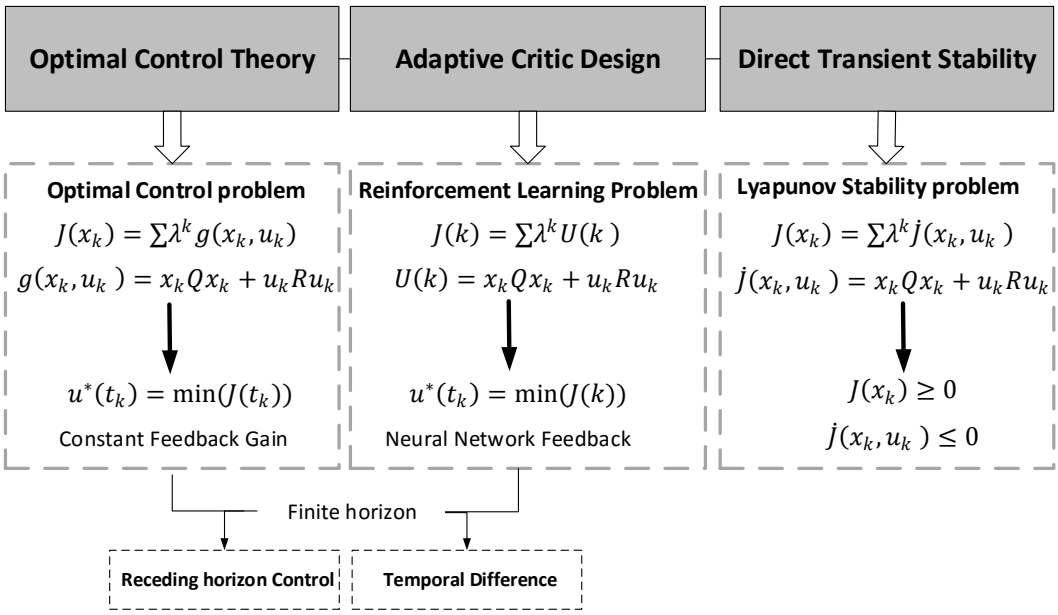


Figure 5.1: Integration of adaptive critic design to optimal control theory and direct transient stability

then $f(x^*(t)) = 0$. $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$ are the state vector and the control action vector at time t, respectively, and f and g are nonlinear functions. Let $J(x(t))$ be an energy function of trajectory x at a particular time t, the system (5.1) is asymptotically stable in the sense of Lyapunov stability criteria when

$$J(x(t)) \geq 0, \quad \dot{J}(x(t), u(t)) \leq 0 \tag{5.2}$$

Based on this, the sufficient conditions for stability assuming this energy function model can be evaluated and the candidates Lyapunov function can be developed to mathematically prove convergence of the system to the reference operating points x^*. This dynamic and energy function is developed next for synchronous generators in transient disturbances. Furthermore, $\dot{J}$ is studied to investigate the damping performance of the oscillations in time.

5.2.1 Area Oscillation Modeling

In the context of rotor angle stability, the dynamics of each synchronous generator bus can be can be characterized by the complex terminal voltage $V_t\angle\delta$, where δ is the rotor angle with respect to synchronously rotating reference frame. The rotor speed is given by $\dot{\delta}$. A weakly coupled power network does not display any coherent oscillation behavior, whereas a strongly coupled network with sufficiently homogeneous natural frequencies is amenable to synchronization in the form of coherent areas. Hence, slow coherency technique based on Center Of Inertia (COI) can be employed to cluster the generators in one area [107]. Let a state space vector $x = [\delta_1, ..., \delta_{\mathcal{N}}, \dot{\delta}_1, ..., \dot{\delta}_{\mathcal{N}}]^T$ be composed of the vector of coherent generators rotor angle and speed, with $\mathcal{N}$ being the total number of areas. The dynamics of the equivalent generators can be described through the expansion of the electro-mechanical single generator model [130]. Additionally, in order to incorporate the impacts of damping controller the third order dynamics of the generator is also included in this chapter as:

$$m_j\ddot{\delta}_j + d_j\dot{\delta}_j = P_j - \sum\nolimits_{k\in N_j} B_{jk}E'_{qj}E'_{qk}\sin(\delta_{jk}) \tag{5.3}$$

$$T'_{doj}\dot{E}'_{qj} = \Delta E_{fj} - \Delta E_{qj} \tag{5.4}$$

with

$$\delta_j = (1/m_j)\sum\nolimits_{i=1}^{n_{g_j}} m_i\delta_i, \quad \dot{\delta}_j = (1/m_j)\sum\nolimits_{i=1}^{n_{g_j}} m_i\dot{\delta}_i \tag{5.5}$$

$$m_j = \sum\nolimits_{i}^{n_{g_j}} m_i, \quad P_j = \sum\nolimits_{i}^{n_{g_j}} P_i \tag{5.6}$$

where,

δ_j Area j rotor angle in COI frame;

$\dot{\delta}_j$ Area j rotor (electrical) speed in COI frame;

i, j	Generator and area index;
n_{g_j}	Number of generators in the area j;
N_j	Set of neighboring buses of the j^{th} bus
m, d	Inertia and damping parameters;
P_j	The power injection from area j;
B_{jk}	(j,k)th entry of the reduced lossless admittance matrix;
E_q, E'_q	The q-axis internal voltage and transient emf;
T'_{do}	The open-circuit transient time constant;
E_f	Excitation voltage;
Δ	Deviation of the variables from the reference.

The transient instability, in general rotor angle oscillations, is caused by a mismatch between injected power P_j of each unit and the power flows along the transmission lines $\sum_{k\in N_j} B_{kj}E'_{qj}E'_{qk}\sin(\delta_{jk})$. This mismatch is usually due to a fault on power system lines or other components, which is followed by the action of protection system. These actions may change or restore the topology of the system, which leads to a transition from pre-fault to post-fault equilibrium points. The transition can be viewed from energy landscape to investigate the convergence as well as damping performance.

5.2.2 Energy Function

Power system rotor angle stability can be assessed by the ability of the synchronous machines in absorbing or releasing the energy accumulated during a disturbance to reach a post-fault equilibrium points. An energy-type Lyapunov function for such a system model comprises of the sum of the system kinetic energy and potential energy

with respect to the relevant equilibrium points δ_j^* and $\dot{\delta}_j^*$. Such a function can be represented as,

$$J(\delta,\dot{\delta}) = \sum_j \underbrace{\frac{1}{2} m_j \Delta\dot{\delta}_j^2}_{J_{Kj}} \underbrace{- \int_{\delta_j^*}^{\delta_j} P_j d\delta_j + \int_{\delta_j^*}^{\delta_j} \sum_{k\in N_j} B_{jk} E'_{qj} E'_{qk} \sin\delta_{jk} d\delta_j}_{J_{Pj}} \tag{5.7}$$

where, $\Delta\delta_j = \delta_j - \delta_j^*$ denotes the deviation from post-fault equilibrium point.

It should be noted that, decreasing nature of the energy function provides the main certificate of local stability; however, it is not the only function that can be linked to decrease in the energy dynamics around a given equilibrium point $\delta^* = [\delta_1^*, ..., \delta_{\mathcal{N}}^*]$. Thus the key is to find a Lyapunov function and verify that it has the required stability properties regarding the domain of state trajectory. It could be understood from (5.7) that the stability depends on the system pre- and post-fault conditions. It can be shown that these equilibrium points satisfy the net power in each area as $P_j = \sum_{k\in N_j} B_{jk} E'_{qj} E'_{qk} \sin\delta_{jk}^*$. If the changes of the states are in small-signal domain the nonlinear part of potential energy could be linearized around the operating point as $\Delta\delta_j^2 = (\delta_j - \delta_j^*)^2$; whereas, in the case of severe transient changes the post-fault points falls into a nonlinear region. The transient problem involves finding the stability boundary which is highly nonlinear. This feature shown in Fig. 5.2, is important for convergence and stability certificate as it can be proved that there exist a boundary that nonlinear energy function has positive and strictly decaying behavior in the region, meaning here the δ_j converges to the stable global optimal point δ_j^*.

Overall, nonlinear energy function can be described as a function of angular differ-

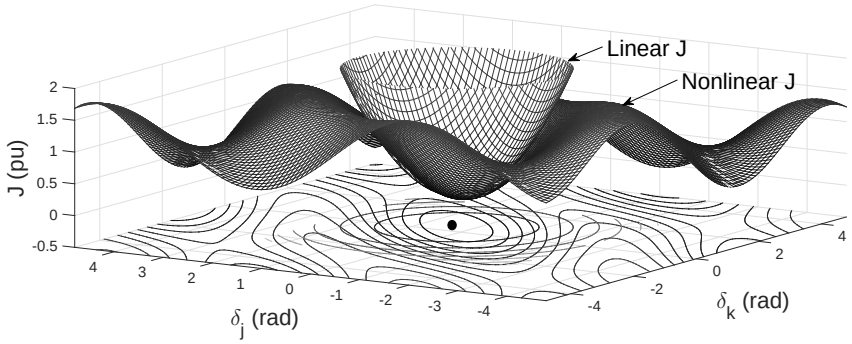


Figure 5.2: Linear and nonlinear energy function representation

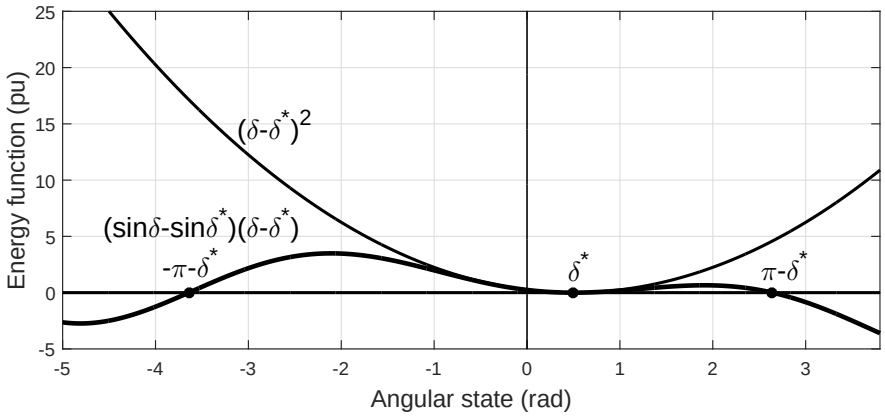


Figure 5.3: Boundedness of nonlinear energy function.

ences as,

$$J^{nl}(\delta_i) = \int_0^{\Delta\delta_i} f(\delta)d\delta \tag{5.8}$$
$$= -\sum_j K_{ij}[\cos(\delta_{ij}) - \cos(\delta_{ij}^*) - (\delta_{ij} - \delta_{ij}^*)\sin(\delta_{ij}^*)]$$

This definition yields to $\Delta J^{nl}(\tilde{\delta}_i) = f(\tilde{\delta}_i)\tilde{\delta}_i$, which satisfies the constraints of nonlinear energy boundedness $0 \leq f(\tilde{\delta}_i)\tilde{\delta}_i \leq \tilde{\delta}_i^2$ in the region of attraction $(\delta_i + \delta_i^*) \leq \pi$ as shown in Fig. 5.3. This feature, also discussed in [22] is important for convergence and stability criteria as it can be proved that the overall energy function of (5.7) has positive and strictly decaying behavior in this region, meaning the x converges to the x^*.

5.2.3 Energy Function Damping

In general, rotor oscillations of increasing amplitude due to lack of sufficient damping torque can lead to instability [131]. In order to incorporate the impacts of excitation control another term is added to energy function so-called field or control energy [41] based on internal voltage:

$$J_{Fj} = \frac{1}{2}\beta_j(\Delta E'_{qj})^2 \tag{5.9}$$

to capture the damping of the energy with respect to time, the derivative of the aggregated kinetic, potential, and field energy functions is derived and can be simplified as,

$$\dot{J}(\dot{\delta}, E_q, E_f) = -\sum\nolimits_j d_j \Delta\dot{\delta}_j^2 + \frac{1}{T'_{d0j}\Delta X_{dj}}(\Delta E_{qj}^2 + \Delta E_{qj}\Delta E_{fj}) \tag{5.10}$$

with the assumption on higher orders of synchronous generator as $\Delta E_f = K\Delta E_q$, $\Delta E_q = X_{ad}\Delta i_f$, and $\Delta i_f = K'\Delta V_t$ with the gains K and $K' > 0$, Δi_f as a field current deviation, and ΔV_t as the terminal voltage deviation or the control input to the field component, (5.10) yields

$$\dot{J}(\dot{\delta}, V_t) = -\sum\nolimits_j d_j \Delta\dot{\delta}_j^2 + r_j \Delta V_{tj}^2. \tag{5.11}$$

with $r_j \equiv \frac{1+K_j}{T'_{d0j}\Delta X_{dj}} X_{adj}^2 K'_j$, which can be derived based on aggregated model of synchronous generators [130].

This equation reveals:

- The states needed to monitor the transient energy deviation of each area is the $\dot{\tilde{\delta}}$. The impact of each area's state on the inter-area oscillation is projected by the damping coefficients d_j. Overall, damping of inter-area oscillations depends on the strength of the transmission system, generator control systems and dynamics

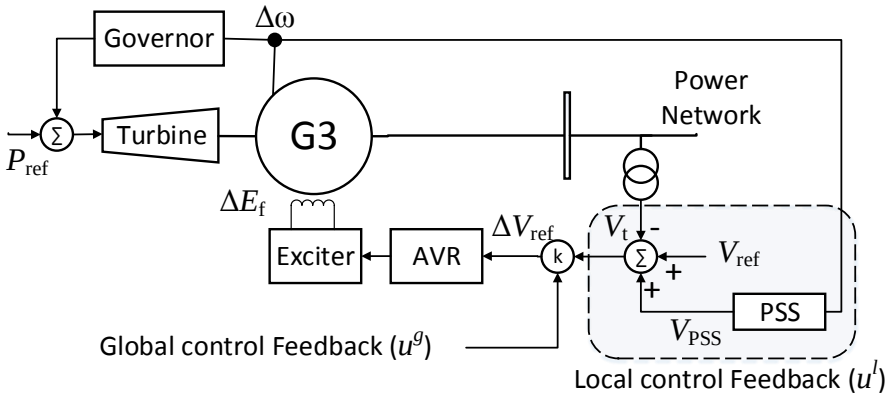


Figure 5.4: Schematic of synchronous generator local control

of inertia-less generators. Even though this parameter is fixed to individual generator, for the system as a whole the damping coefficients may be changing not be known precisely at given time of interest and is needed to be calculated. Identification of energy function considering the actual oscillation frequency and damping parameters can further improve the performance of the controller.

- Another observation in the above equation is the impact of ΔV_t as the input to the synchronous generator excitation field. This input without controllers is the error between the reference and the terminal voltages of the generator ($u = \Delta V_t$). However, this signal can be augmented by local and wide-area damping control actions to further improve the energy function damping.

5.3 Intelligent WADC Construction

In this section, the goal is to develop an optimal intelligent WADC to stabilize and damp the inter-area oscillations occurring in power system.

As mentioned in previous section and shown in Fig. 5.4, the input to the synchronous generator excitation field could be augmented with u_{loc} as local damping signal derived from PSS and enhanced by u_{wac} as a wide-area level damping controller. Hence, the control action can be presented as 2-level combination of local and

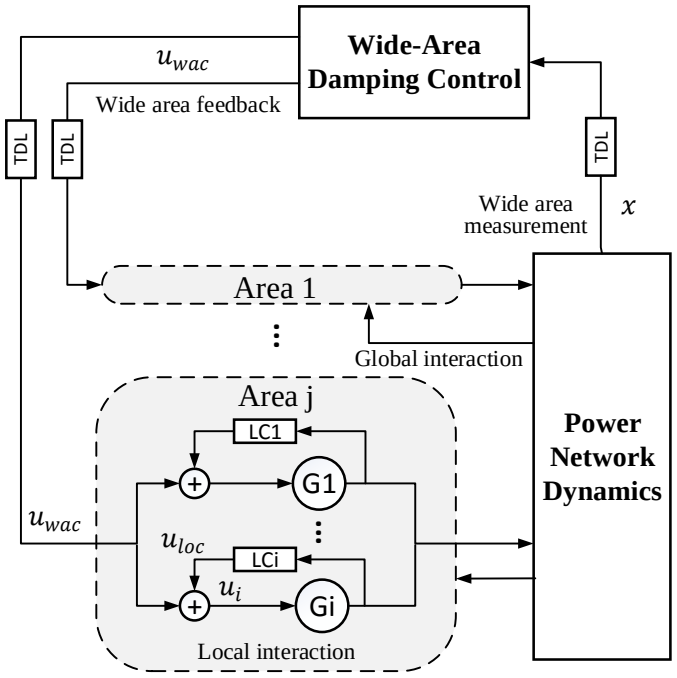


Figure 5.5: Structure of the 2-level control design combining local and wide-area damping control (TDL: time delay line, LC: local control)

wide-area parts as

$$u(t) = u_{loc}(t) + u_{wac}(t) \tag{5.12}$$

as shown in Fig. 5.5. The PSS monitoring the local states is able to damp local oscillations; while, the WADC monitoring rotor angle and speed of all generators using wide-area measurements in COI frame can form the system's energy function (5.7) and its damping (5.11) as in direct methods and enhance the damping performance by u_{wac}.

5.3.1 Optimal Control Design

Considering the transient energy function developed in previous section as the cost function, optimal controller can be designed to minimize this energy function forward in time. Overall, the cost function can be implicitly specified in discrete time domain

with the time step of Δt by

$$J(x(k)) = \sum_{\tau=k}^{\infty} \dot{J}(x(\tau), u(\tau)) \tag{5.13}$$

where, k is the time step index and $\dot{J}$ is the short-time cost function associated with transient state errors and control effort equivalent to derivative function in Lyapunov stability concept. Hence, it can be written as,

$$\dot{J}(x(k), u(k)) = -\Delta x(k)^T Q \Delta x(k) - u(k)^T R u(k) \tag{5.14}$$

where,

$$x = [\delta_1, ..., \delta_{\mathcal{N}}, \dot{\delta}_1, ..., \dot{\delta}_{\mathcal{N}}]^T \tag{5.15}$$

$$u = [\Delta V_{t1}, ..., \Delta V_{t\mathcal{N}}]^T \tag{5.16}$$

$$Q = \text{diag}(0, ..., 0, d_1, ..., d_{\mathcal{N}}) \in \mathbb{R}^{2\mathcal{N} \times 2\mathcal{N}} \tag{5.17}$$

$$R = \text{diag}(r_1, ..., r_{\mathcal{N}}) \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}} \tag{5.18}$$

which are based on actual values of coefficient of (5.11) and are positive definite. One can simply use these functions and assess transient stability based on direct methods and design decentralized controller such as in [41]. However, such designs are not global optimal. The optimal solution of the problem of (5.13) subjected to the dynamic system (5.1) is given by minimization of HJB equation as,

$$u^*(k) = \arg\min_{u} \{\dot{J}(x(k), u(k)) + J(x(k+1))\} \tag{5.19}$$

Applying the first order optimality conditions with the dynamic programming algorithm leads to,

$$u^*(k) = -1/2R^{-1}\frac{\partial x(k+1)}{\partial u(k)}\frac{\partial J(x(k+1))}{\partial x(k+1)} \tag{5.20}$$

It can be seen from equation above that the optimal control action $u^*(k)$ is defined on the future dynamics of the system $x(k+1)$ and the energy function $J(x(k+1))$, which implies the necessity of nonlinear identification. Additionally, the energy function itself is highly dependent on the parameters of the system such as m, B, or β, even though simplified model of the system without structure preserving is considered in the design.

In order to overcome the aforementioned issues and cover nonlinear regions, artificial intelligence is employed for mapping from a parameter space into the space of functions they aim to represent. A common approach is to deploy NN to map the nonlinearities of the system identification, control and the cost function. In this scheme, three networks called *identifier*, *critic*, and *action* NNs are trained to approximate the dynamics of the system, the control action, and the cost function, respectively. In this work, using Feed Forward Neural Network (FFNN) the outputs are computed as,

$$\hat{x}(k) = \hat{W}_I(k)^T\Phi_I(x(k-1), u(k-1)) + \epsilon_I(k) \tag{5.21}$$

$$\hat{u}(k) = \hat{W}_A(k)^T\Phi_A(x(k)) + \epsilon_A(k) \tag{5.22}$$

$$\hat{J}(x(k)) = \hat{W}_C(k)^T\Phi_C(x(k)) + \epsilon_C(k) \tag{5.23}$$

where, sub-scripts I, A, and C denotes identifier, action, and critic networks, respectively, "$\hat{}$" denotes estimated values, $\Phi(.) \in \Re^h$ is the corresponding nonlinear mapping function of the states, $W(k) \in \Re^h$ is the parameter vector of approximated weights of the FFNN at time step k, with $h \in \mathbb{N}$ dimensionality of the feature vector

representing each state, and ϵ is the approximation error. Based on the universal approximation property of NN it is assumed that ϵ tends to zero [132, p. 52].

NNs should be trained to provide accurate approximation of their respective outputs and back-propagation estimation and cover whole region of operations. In general, training the NNs means adjusting the parameters or weights in an iterative process to reduce the error between the target outputs and the actual outputs. Intelligent WADC construction could be implemented in two phases:

5.3.2 Offline Supervised Learning Construction

This method is referred to problems involving static input/output mappings and minimization of a vector error signal, with no explicit dependence on how training examples are gathered. It is assumed that full knowledge of the problem context is available. In offline phase the parameters and operating point are known; hence, the target output for training the NNs could be calculated. The supervised training for WADC are based on the following errors:

$$e_I = \sum\nolimits_k x(k) - \hat{x}(k) \tag{5.24}$$

$$e_C = \sum\nolimits_k J(x(k)) - \hat{J}(x(k)) \tag{5.25}$$

$$e_A = \sum\nolimits_k u^*(k) - \hat{u}(k) \tag{5.26}$$

where, "ˆ" denotes estimated values based on (5.21)-(5.23).

The training process of the critic NN is performed with the target of conventional energy function (5.7). The post-fault equilibrium point can be calculated simply based on the post-fault load flow to derive this function. Further, the training of the action NN is based on optimal control derived in (5.20). It should be noted that, the two derivative elements in (5.20) are simply calculated by propagating the respective output back through the identifier and critic NNs.

Change of the NNs weights at each iteration i can be derived from the deviation of

NN's output to its optimal value e^i. In offline training, batch learning is performed in which adjustment of the weights are accumulated over all training data to give an aggregated error as,

$$\hat{W}^{i+1} - \hat{W}^i = \alpha e^i \Phi(.) \tag{5.27}$$

where, α is a small learning rate. This optimization iteration is performed by means of training the NNs through gradient descent via back-propagation. From the viewpoint of optimal control theory, this task is the same as the first-order calculus of variation to find the continuous-time equations derivations.

5.3.3 Online Reinforcement Learning Construction

In this chapter, RL technique is employed for adaptation of artificial intelligence-based WADC in online application. By means of this method, the NNs parameters and weights are updated based on measurements instead of conservative offline assumptions. RL is often applied to problems involving sequential dynamics and optimization of a scalar performance objective, with online exploration of the effects of actions as it can adapt itself to fit the changing environment. The identifier NN is updated online with the monitored states similar to supervised learning; however, different procedure is used for critic and action NNs training. In the RL method cost function at each iteration i is approximated by adaptation of Approximate Dynamic Programming (ADP) as,

$$e_C^i(k) = \hat{J}^i(x(k)) - \hat{J}^i(x(k+1)) - \dot{J}(x(k), u^i(k)) \tag{5.28}$$

and, training of the action NN at each iteration is based on

$$e_A^i(k) = 2Ru^i(k) + \frac{\partial \hat{x}(k+1)}{\partial \hat{u}^i(k)} \frac{\partial \hat{J}^i(x(k+1))}{\partial \hat{x}(k+1)} \tag{5.29}$$

and, weight adjustment is done incrementally at each iteration and time step as,

$$\hat{W}^{i+1}(k) - \hat{W}^{i}(k) = \alpha e^{i}(k)\Phi(.) \tag{5.30}$$

The stability properties of the ADP is discussed in [50].

5.3.4 Online Energy Function Construction

As presented in previous subsection the incremental critic NN training is based on $\hat{J}(x(k+1))$ and $\dot{J}(x(k), u(k))$ at each time step. The first element is derived from the critic NN with the input of $\hat{x}(k+1)$ which is accessible from the identifier NN. The later is derivative of energy function defined in (5.14) in quadratic form representing the damping of energy function. This function is dependent on d which is characteristic of damping of oscillations of each area's states and needed to be identified online.

The Prony method, the best known parameter identification method in the power system community [133], is used to determine the unknown eigenvalues of the system and extract the damping coefficient features. This method is based on measured global states, $x(k)$, being expressed in z-transform domain as a sum of n damped complex sinusoids,

$$x(k) = \sum_{l=1}^{n} \bar{R}_l z_l^k \tag{5.31}$$

with characteristic equation of $z^n + a_1 z^{n-1} + ... + a_n = 0$, where R_l is an output residue corresponding to the mode λ_l. The vector $A' = [a_1, .., a_n]$ leads to the eigenvalues, z_ls, of the system which are the roots of the system characteristic equation. Once the roots of the system characteristic equation are obtained, the eigenvalues with high frequencies are neglected. This analysis leads to obtaining the eigenvalues of the system which is then used to adapt the weighing matrix of the states accordingly. For this purpose, an auto regressive (AR) model estimation method has been employed.

One can find a values by solving a least squares problem defined on the 2-norm of a vector with an adaptive algorithm that recursively optimizes the criterion.

$$\min_{a} \| \begin{bmatrix} x(n) \\ x(n+1) \\ \vdots \\ x(n+l) \end{bmatrix} - \begin{bmatrix} x(n-1) & \dots & x(0) \\ x(n) & \dots & x(1) \\ \vdots & \ddots & \vdots \\ x(n+l-1) & \dots & x(l) \end{bmatrix} \begin{bmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_n \end{bmatrix} \| \tag{5.32}$$

After computing a at time step k, roots of the characteristic polynomial is derived to generate eigenvalues $\lambda_j = -\zeta_j\omega_j \pm \omega_j\sqrt{\zeta_j^2 - 1}$ for global oscillations with ω natural frequency and ζ damping factor. Further, damping coefficient of area j can be derived as

$$d_j = \omega_l \zeta_l m_j \tag{5.33}$$

Subsequently, Q and $\dot{J}$ can be updated based on $d_j(k)$ at each iteration i as,

$$Q^{i+1}_{n+j,n+j}(k) = Q^{i}_{n+j,n+j}(k) + \gamma(d^{i+1}_j(k) - d^{i}_j(k)) \tag{5.34}$$

$$\dot{J}^{i+1}(x(k), u(k)) = \tilde{x}(k)^T Q^{i+1} \tilde{x}(k) + u(k)^T R u(k) \tag{5.35}$$

where, γ is the scaling factor.

Remark: It is assumed that the power system is operating at the same operating point for a certain time that enables the estimated d to converge. It is noted that this is not a constraint in practice, since the estimated model parameters converge to their new values fast enough compared to the dynamics [19].

5.3.5 Comparison of the Techniques

In this chapter RL method with online tuning has been employed to design the WADC. The advantages of the proposed method over conventional methods are as follows:

1. Both methods of direct and supervised learning are developed directly on the energy function J, which is defined on the post-fault equilibrium points of δ^* and $\dot{\delta}^*$. This post-fault δ^* maybe different than pre-fault δ^0 due to network re-configuration after a fault and protection control actions. However, RL method is updated based on $\dot{J}$ which is only defined on $\dot{\delta}^*$. This value is considered 1 pu in both pre- and post-fault conditions.

2. Direct and supervised learning methods rely on the parameters of the system such as m and B, which may have uncertainty or be changing in time. RL method on the other hand is capable of updating it's policies over time and under conditions of noise and uncertainty through state-action interaction [128].

3. Since actions should be taken at each time step and their effect is not known until the end of the sequence, it is not possible to design an optimal controller in online application using the traditional SL. RL method allows the WADC to account for the present control actions $u(k)$ and *future* consequences on the system in $J(x(k+1))$ term used in (5.28), and *present* consequences in the form of short-time cost function in $\dot{J}(x(k), u(k))$ term.

4. The term $\dot{J}$ in RL is used for updating the cost function and is called performance function. Monitoring this function directly instead of J is well-suited for the purpose of WADC design as it shows how is the *performance* of the controller in terms of transient energy function damping.

5.4 Implementation

5.4.1 Power System Test Case

In order to assess the capability of the proposed method, IEEE 68-bus 16-machine power system has been modeled using Power System Toolbox (PST). The model is a reduced order equivalent of the inter-connected New England and New York power

Table 5.1: PSS parameters for power system test case

Generator	K_s	T_s	$T1$	T_2	T_3	T_4	T_5	T_6
G_1:G_8,G_{10}:G_{12}	20	15	0.15	0.04	0.15	0.04	0.15	0.04
G_9	12	10	0.09	0.02	0.09	0.02	1	1

systems with nine areas as shown in Fig. 5.6. The New England and New York power grids are represented by group of generators, whereas the power import from each of the three other neighboring areas are approximated by equivalent generator models. Grouping consideration is based on the oscillation modes, dynamic data of the generators, and the parameters of the transmission system [134]. Dominant modes of inter-area oscillations can be derived by small-signal stability analysis, which is presented in Fig.5.6. In this simulation, the slow-dynamics of the governors are ignored. Two types of excitation systems for the generators is considered: IEEE standard DC exciter (DC4B) and the standard static exciter (ST1A). All generators are equipped with PSS using the parameters in Table 5.1. The local damping controller derives $u_i^l = \Delta V_{refi}$ to reduce the local energy function by tracking the state $\dot{\delta}_i$ as

$$u_{loci}(s) = K_{si}\frac{sT_{si}}{1+sT_{si}}\frac{1+sT_{1i}}{1+sT_{2i}}\frac{1+sT_{3i}}{1+sT_{4i}}\frac{1+sT_{5i}}{1+sT_{6i}}\Delta\dot{\delta}_i(s) \tag{5.36}$$

where, $u(s)$ is the local control in frequency domain.

First, small signal stability analysis has been performed to derive the damping ratio, frequency, and participation factors of the generators in the dominant oscillatory modes (Table. 5.2). This modes has been derived without the presence of any PSS in the system, and it is also provided in [134]. It can be seen that, all the inter-area modes have high participation from machines G13 to G16, and the local modes have high participation from the corresponding local machines.

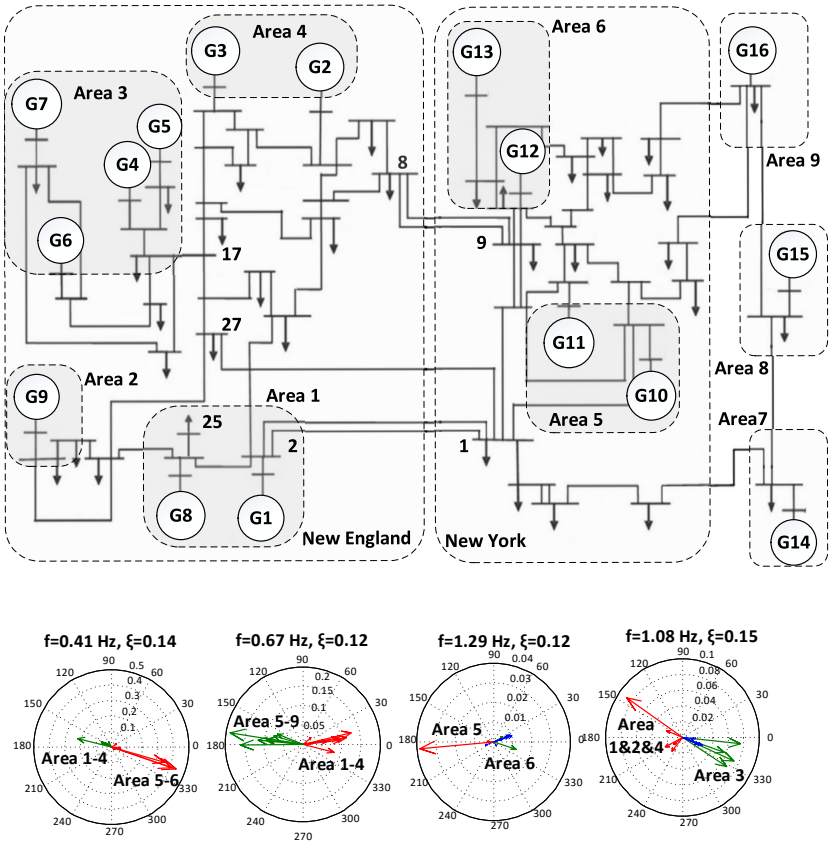


Figure 5.6: Single line diagram and polar dominant modes of the 68-bus 16-machine New England New York power system.

5.4.2 Proposed WADC Architecture

Input vector of the WADC is the states derived from WAM $(\delta(k), \dot{\delta}(k))$ in order to estimate the energy function $(J(x(k)))$, and the output is the optimal damping control $(\Delta V_t(k))$ at time step k. Here, we assume that the communication system can transmit the signals to the control center at the speed of 100 frames/sec; hence, WADC communication time step is considered as $\Delta t = 10ms$. It is also assumed that each generating unit bus has a sensor that transmits signals to WADC system. For synchronous generators buses signals are aggregated in WADC center in COI frame of (5.5) to generate area's state. Fig. 5.7 presents the overall proposed WADC architecture for offline and online energy function estimation, the link to intelligent

Table 5.2: Electro-mechanical Modes and participating generators of the IEEE 68-bus system without PSS

Damping ratio (%)	frequency (Hz)	Gen/pf	Gen/pf	Gen/pf
-0.438	0.404	G13/1	G15/0.556	G14/0.524
0.937	0.526	G14/1	G16/0.738	G13/0.114
-3.855	0.61	G13/1	G12/0.137	G6/0.136
3.321	0.779	G15/1	G14/0.305	x
0.256	0.998	G2/1	G3/0.913	x
3.032	1.073	G12/1	G13/0.179	x
-1.803	1.093	G9/1	G1/0.337	x
3.716	1.158	G5/1	G6/0.959	x
3.588	1.185	G2/1	G3/0.928	x
0.762	1.217	G10/1	G9/0.426	x
1.347	1.26	G1/1	G10/0.756	x
6.487	1.471	G8/1	G1/0.435	x
7.033	1.487	G4/1	G5/0.483	x
6.799	1.503	G7/1	G6/0.557	x
3.904	1.753	G11/1	x	x

control, and the RL and supervised training process of NN blocks.

The construction of intelligent system starts with offline supervised learning in order to set the initial weights for online implementation. The training is performed based on the I/O signals for each of the NNs derived using Matlab NN toolbox. These weights and parameters are extracted in data base for online implementation. Table 5.3 lists the NNs parameters used for the study. It should be noted that variables are time delayed by one time step to capture the dynamics of the system. For identifier NN, pseudo random inputs and related outputs of the power system model in COI frame is captured and fed into the NN data base. Then, gradient descent batch learning algorithm is performed with a learning rate of about 0.001. It has been seen that e_I reaches a small number in 100 epochs. Furthermore, the critic NN is trained based on (5.7) for different fault scenarios and operating points in multiple time-domain simulations to learn the relative energy functions. Parameters of energy function including m, B, β can be derived from [134]. It should be noted that B is the Kron-reduced susceptance matrix with the loads removed from consideration. Next,

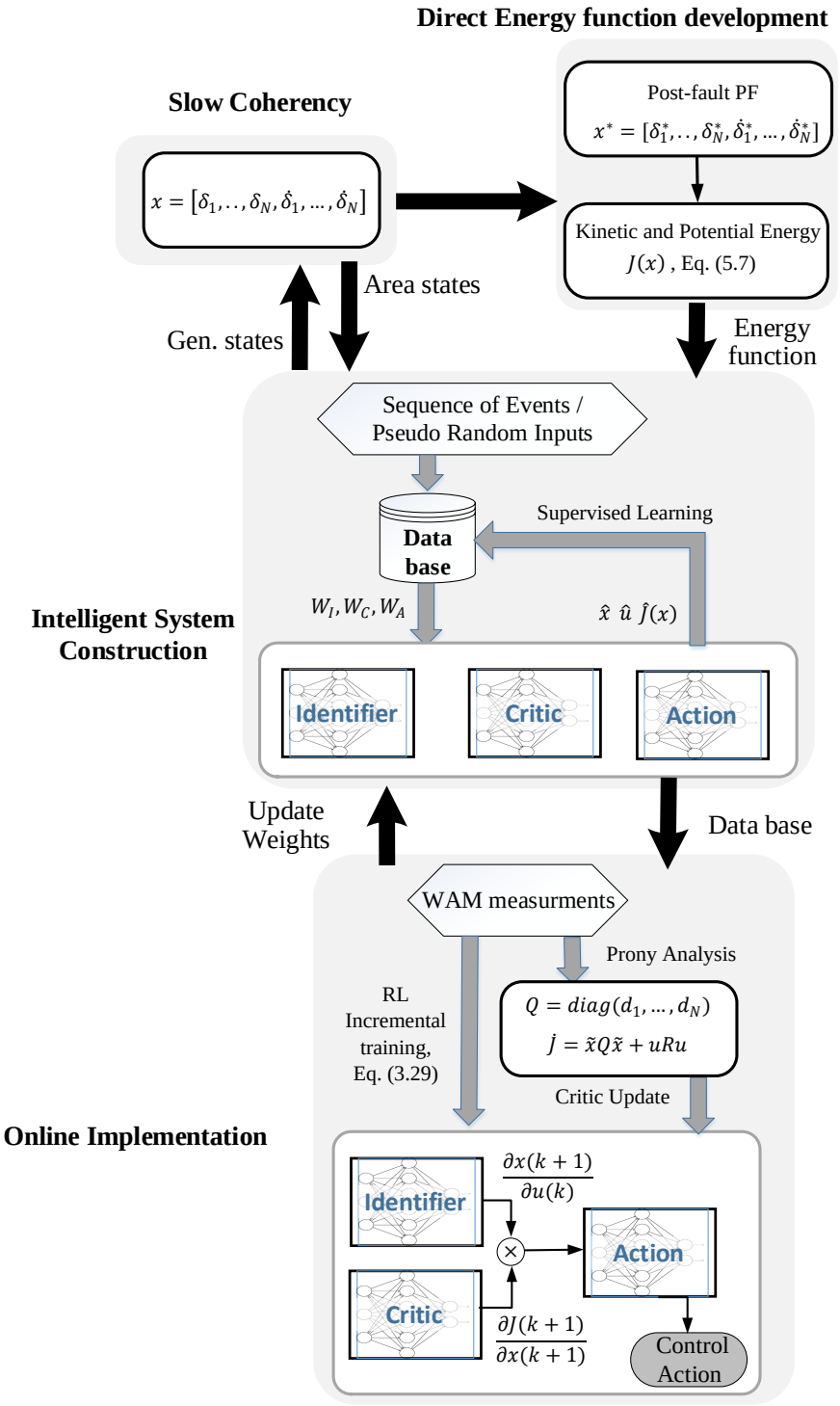


Figure 5.7: Structure of the wide-area damping control center design

action NN is trained based on optimal control theory (5.20). Once the NNs weights have converged ($e_C \& e_A < 0.01$), then the action NN is connected to the generator's

Table 5.3: Configuration of Neural Networks

NN	Input (nodes)	Output (nodes)	Hidden nodes
Identifier	δ, $\dot{\delta}$, ΔV_t (54)	δ, $\dot{\delta}$ (18)	70
Action	δ, $\dot{\delta}$ (36)	ΔV_t (9)	50
Critic	δ, $\dot{\delta}$ (36)	J (9)	50

exciter to augment the PSS.

The process of online training starts by monitoring the states to incrementally train the identifier NN. This is followed by critic and action update iterations. In the online process, the critic NN is updated by derivative of energy function. Here, for sake of simplicity r is considered as 0.1. Prony algorithm is applied to identify the modes and assess the damping ratios of individual area angular oscillations. For this purpose, the values of δ_j is captured at time steps of $5.\Delta t$ and used for tuning the derivative energy function. Since there are 9 sixth-order areas, the algorithm should ideally solve 54th order polynomial. However, our simulation showed that choosing order of 15 yields a satisfactory estimates of the inter-area modes.

5.5 Simulation Results

5.5.1 Case A. Proposed Method vs Conventional Artificial Intelligence Methods

In this case study the damping performance of the proposed hybrid RL and energy-based method WADC design has been investigated. The load connected to bus 25 (224 MW) has been disconnected for 200 ms due to a short circuit fault, creating a local mode of oscillation in area 1. During the fault the area 1 moves away from the pre-fault equilibrium point ($\delta^0_{A1} = 0.22$). After the re-closing action ($t = 0.3s$), the system configuration is the same as pre-fault ($\delta^*_{A1} = 0.22$) and the system experiences the post-fault transient dynamics of Fig. 5.8a. This figure presents during- and post-fault trajectory of the $\dot{\delta}_{A1}$ in the case of different control scenarios. As shown, the proposed WADC compared to supervised learning algorithm has provided better performance with respect to overshoots and damping. The transient energy function dynamics

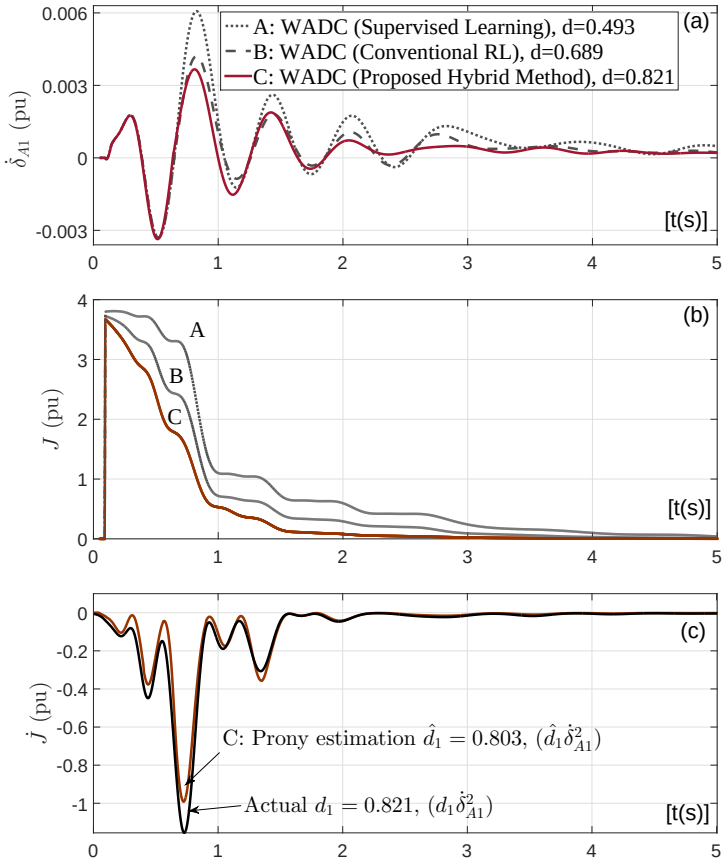


Figure 5.8: Case A. Comparison of the proposed hybrid method with conventional supervised and RL: (a) Area 1 speed oscillation, (b) Energy function area 1, (c) Derivative of energy function area 1.

has been demonstrated in Fig. 5.8b. As it can be seen the proposed WADC is able to provide more damping than supervised learning due to online optimal control adaptation and exploration.

Additionally, the proposed method has been compared to conventional RL-based WADC. The RL method being implemented in several works as a WADC such as in [10, 13, 15], where the cost function is defined in the form of quadratic function of states with heuristic coefficients based on linearized model of the system. This definition, however, is not representative of energy function damping as opposed to the proposed hybrid method. The proposed method with the help of Prony analysis can estimate the actual damping coefficient ($d_1 \approx \hat{d}_1 = 0.821$) and subsequently energy function damping performance as shown in Fig. 5.8c.

Table 5.4: Case B. performance of NNs in trained dataset

Event	Identifier		Action		Critic	
	RMSE	Max Iter.	RMSE	Max Iter.	RMSE	Max Iter.
1	0.010	100	0.021	100	0.016	100
2	0.010	71	0.010	83	0.010	84
3	**0.010**	43	**0.010**	54	**0.010**	55

5.5.2 Case B. Robustness to Parameters

This test presents a case study to demonstrate one of the motivations for choosing the RL algorithm as the benchmark to perform the online optimal function approximation. Initially, a scenario of cascading failures due to faults on inter-area tie-lines has been performed. In this test, three tie-lines connecting New England and New York power systems are disconnected sequentially, separating the two grids at the end. Under nominal condition 1170 MW is transferred to New England system through these tie-lines. First, line 1-27 is disconnected (event 1) followed by line 8-9 (event 2) allowing the RL to gain enough knowledge about system dynamics and optimal policy. Finally, line 1-2 is disconnected due to a self-clearing 100 ms three-phase short circuit fault (event 3), which leads to complete separation of these two systems and huge energy mismatch. The WADC is independently validated for performance and accuracy using two data sets of trained and non-trained NNs. Tables 5.4 and 5.5 reveal the advantage of the RL method by allowing the NNs to gain more knowledge about the optimal policy of action and energy during the course of events. In the case of trained NN set, identifier, action, and critic NNs have been trained offline and then used in the WADC. In the other data set, the NNs weights are assigned randomly. In this table, the Root Mean Square Error (RMSE) of the latter is higher; however, it improves with each course of event because of online RL learning. Also it can be seen that the number of iterations during last fault occurrence is less than the first event implying the faster convergence of NNs. For this test, maximum number of iteration at each time step is 100 and minimum error assigned is 0.01.

Table 5.5: Case B. performance of NNs in not trained dataset

Event	Identifier		Action		Critic	
	RMSE	Max Iter.	RMSE	Max Iter.	RMSE	Max Iter.
1	0.530	100	0.476	100	0.489	100
2	0.110	100	0.153	100	0.112	100
3	**0.021**	100	**0.064**	100	**0.053**	100

5.5.3 Case C. Proposed WADC vs Linear Controllers

Further, performance of the proposed WADC for the last event in case B, i.e. fault on line 1-2, is compared to PSS and conventional GPSS with the parameters of $K_s = 1$, $T_s = 10$, $T_1 = T_3 = 0.6280$, and $T_2 = T_4 = 0.1025$ [4,116] (see Fig. 5.9). For this test, GPSS uses a single differential frequency signal between two selected areas of 1 and 5 ($\dot{\delta}_{A1} - \dot{\delta}_{A5}$). The parameters are tuned based on small-signal analysis. Fig. 5.9a demonstrates the modes of the inter-area oscillation with frequency of 0.830 Hz and damping factor of 0.605%. As it can be seen, modes of area 5 has been shifted closer to the area 1 with the proposed WADC. This result is validated for inter-area oscillation as well as local speed deviation as they have been better damped in presence of the proposed WADC. Fig. 5.9d shows the derivative of energy function with respect to time and in terms of elements of control, local and area states.

5.5.4 Case D. Robustness to Delays

It is worth noting that transfer of WAM measurements to the control center may incur certain time-delays. As a measurement-based control design, neglecting this property of WAM may degrade the performance or even destabilize the control system. Usually, the delay of the WAM signals in a high-bandwidth communication system is expected to be small for the purpose of the WADC design. In [70] the WAM infrastructure and various possible communication delays have been covered. In this case study the robustness of the proposed WADC to possible communication delays has been investigated. All the PMU signals are delayed by 100 ms, which is larger than the expected delay in the realistic system. Fig. 5.10 shows the system

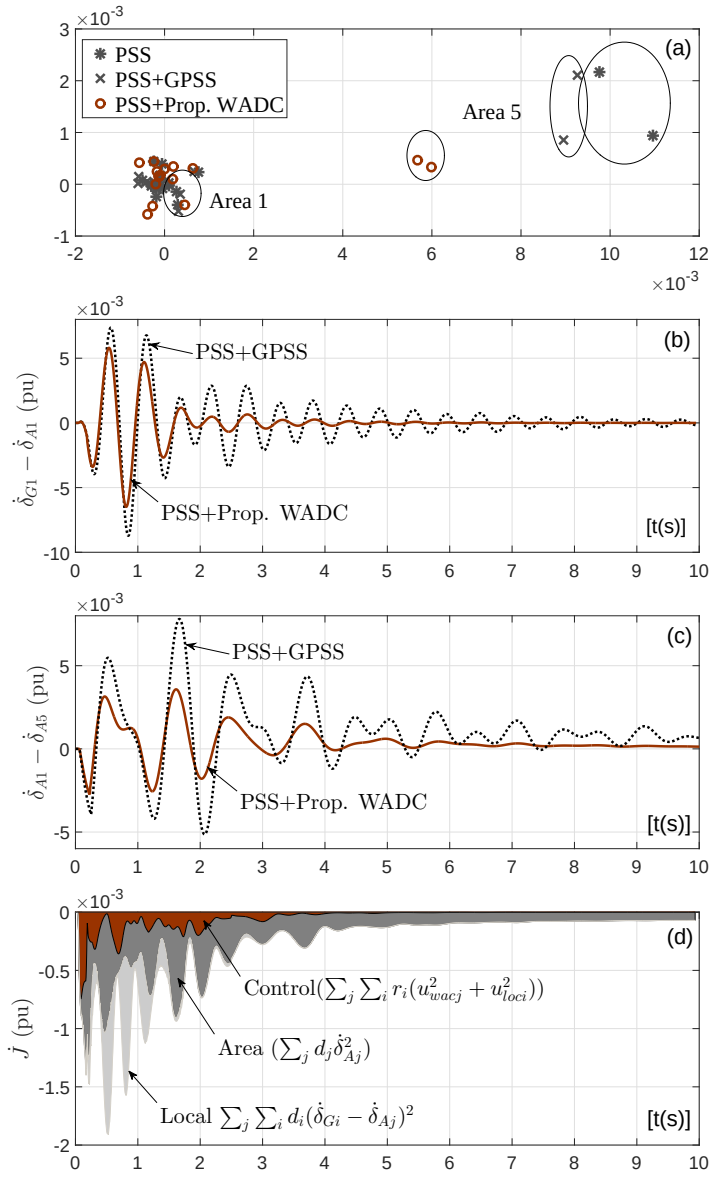


Figure 5.9: Case C. comparison of proposed WADC with conventional PSS and GPSS, (a) area oscillation modes (b) local oscillation of G1 ($\zeta = 0.7805\%, f = 1.8957Hz$, (c) inter-Area oscillation of A1 and A5 ($\zeta = 0.6054\%, f = 0.83Hz$, (d) derivative of energy function.

dynamic response of a the system (local and inter-area oscillations) with the same fault scenario as Case C. It can be seen that delays has deteriorated the WADC performance, however, the system oscillations still damp faster than the case without WADC. Moreover, the performance of proposed WADC in the most extreme scenario is evaluated by increasing the delay time in PMU signals. Further simulations showed

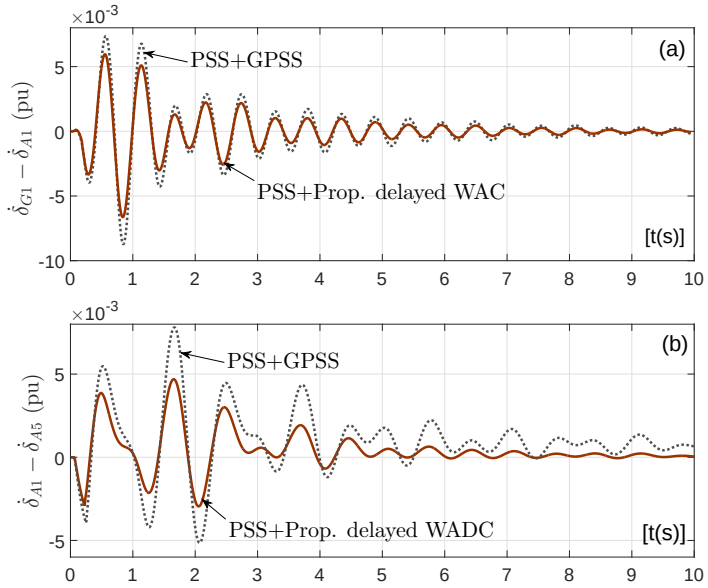


Figure 5.10: Case D. comparison of proposed 100 ms delayed WADC with conventional PSS and GPSS, (a) local oscillation of G1, (b) inter-area oscillation of A1 and A5.

that the WADC is robust to delays of 350 ms in communication network.

5.5.5 Case E. Real-Time Simulation

In order to investigate the performance of the proposed WAC for real-time transient stability improvement, the New England part of the test system, so-called IEEE 39-bus system has been modeled using real-time simulator Hypersim [117]. The simulation is based on EMTP nodal method capable of running the simulations by parallel computation. The modeled power system generators are connected to exciter and PSS acting as local controllers. These controllers run at the speed of 50 us. For real-time development of the WAC, the offline pre-training stage of NNs is obtained from the power system modeled in PST. The WAC model is then imported from Matlab-Simulink through the C-code conversion and deployed in the Hypersim environment. The input of exciter is augmented by the WAC actions running at the simulation time step of 10 ms.

In order to analyze the performance of the proposed architecture, a self-clearing 125ms three-phase short circuit at bus 17 is simulated. This scenario affects all the

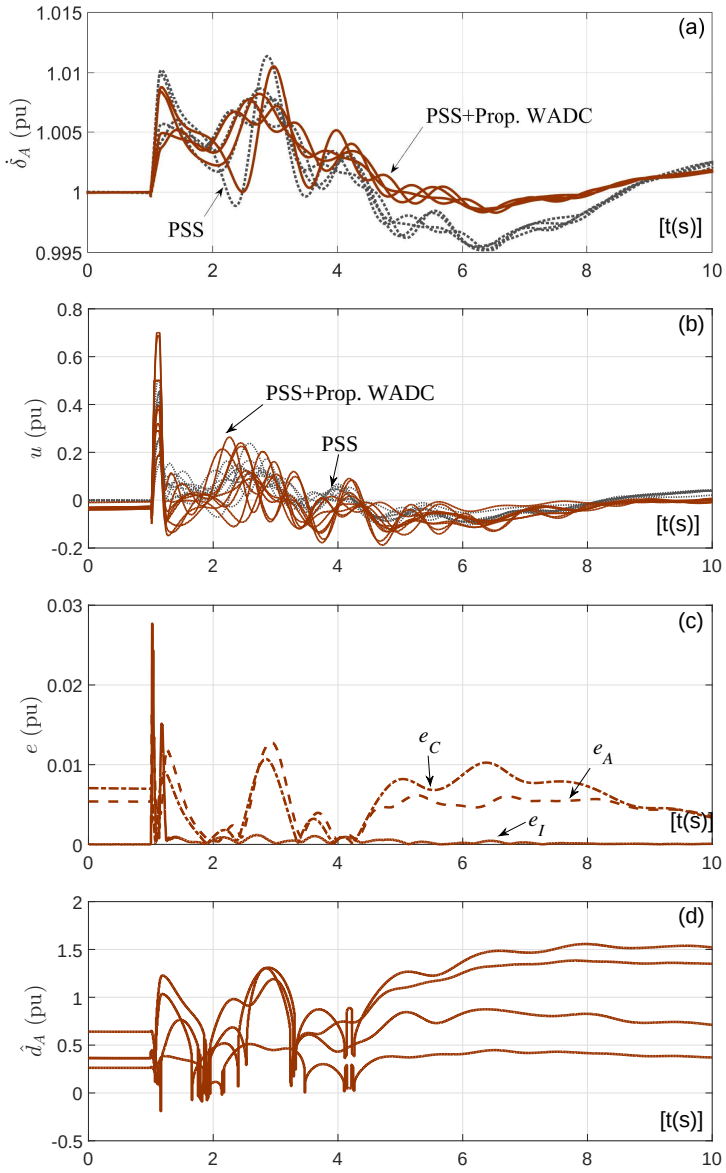


Figure 5.11: Case E. real-time dynamic performance of the power and control system, (a) Area Oscillations, (b) control signals, (c) NNs errors, (d) damping coefficient approximation with Prony analysis.

oscillatory modes of the devices and the network topology. Area speed oscillations of the 4 areas of the test system are presented in Fig. 5.11a. As shown, a considerable damping improvement is gained when the proposed architecture is adopted. In addition, overshoots and undershoots of oscillation are reduced as well. The local control action of PSS and the WADC are presented in Fig. 5.11b. Further, Fig. 5.11c presents

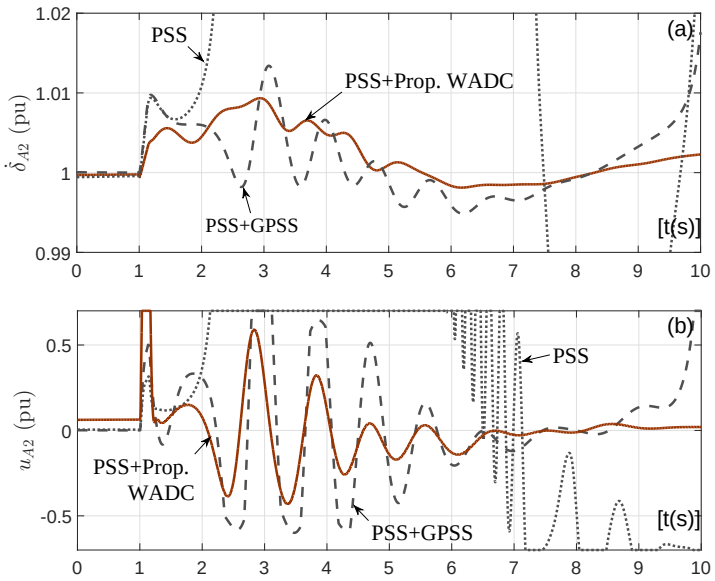


Figure 5.12: Case F. real-time dynamic performance of the power and control system (a) area 2 Oscillation, (b) control signals of Area 2.

the WADC NNs error. As it can be seen the identifier, action, and critic NNs have performed satisfactory in real-time with the maximum error limited to 0.025 pu. It is worth noting that as mentioned before, the critic tuning is performed in real-time using the proposed Prony analysis. The method is capable of identifying the oscillation parameters, the damping coefficient, as shown in Fig. 5.11d and developing the tuned energy function. The initial values of Q is extracted from mode 1 and participation factors.

5.5.6 Case F. Unstable Case

Further, in order to show the efficiency of the proposed WADC during unstable disturbances, the fault duration of previous case study in real-time simulator has been increased to 135 ms. It can be seen that this condition has led to system instability with the presence of the local PSS acting alone or with the GPSS re-tuned based on inter-area modes of oscillation frequencies as shown in Fig. 5.12a. However, the proposed WADC has maintained the system stability. The respective controller contributions are shown in Fig. 5.12b. This case study demonstrates the effectiveness

of the proposed technique in enhancement of the transient stability as the CCT or duration of the fault has been increased.

5.6 Conclusion

In this chapter a new intelligent energy-based wide-area damping control design for improved power system transient stability is presented. The algorithm adapts the reinforcement learning to optimally solve for a control action policy through approximate dynamic programming. The cost function is defined to improve the inter-area oscillations and define transient energy function candidate. Furthermore, an offline and online tuning method is designed based on extracted eigenvalues using small signal stability analysis and Prony analysis and is linked to performance index generation. More in depth information about the proposed methods in this chapter could be found in our recently published journal papers [70, 135]. Results on IEEE 68-bus 16-machine system showed that the proposed method is able to guarantee the generalized energy function in real-time and converge it to optimal operating point with higher damping.

CHAPTER 6: DFIG IMPACT ON TRANSIENT STABILITY ENHANCEMENT

6.1 Introduction

It has been shown in previous chapters that Wide-Area Control (WAC) designs, thanks to recent advances in the Wide-Area Monitoring (WAM) technologies such as Phasor Measurement Units (PMUs), can effectively mitigate these oscillations and improve wide-area system stability. Further, Optimal enhancement of transient stability using WAC was capable of improving the power transfer capability of a transmission grid and preventing the system from generation or load disconnection, or catastrophic failure following a sequence of disturbances. This chapter seek to investigate the impact of renewable energy resources in transient stability enhancement.

Among the various renewable energy resources,DFIG-based wind power is assumed to have one of the most favorable technical and economical prospects [73]. Considerable research effort have been devoted to investigate the performance of these resources in the presence of disturbances [73, 77, 79, 136]. These efforts have defined the problem from the renewable energy resource view. In [79], a local control scheme has been developed for grid connected wind turbines aiming at transient stability restoration. Also several research works has been reported for the design of local or WAC for the DFIG-integrated power system, in order to improve the performance of the overall system [78, 137]. Generally, the wind farms connected to the grid do not actively take part in damping of power system oscillations. However, it has been illustrated that as the *penetration level* of these wind generators increases, the overall inertia 'seen' by the system decreases. Also when the tie-line power flow changes, damping of power system oscillations are adversely effected [77]. Ref. [78] shows that when active power flow changes due to a disturbance in the system, the way the wind

turbine provides active or reactive power is critical in maintaining rotor-angle stability of synchronous generators. Conventional control techniques using small-signal stability methods are widely used for the power management of renewable resources, mimicking the excitation control of a synchronous generator design such as in [138]. However, these models are based on large number of parameters of power system network, synchronous and wind generators which are not available in practice. Additionally, these techniques depends on linear models of the system that cannot be used in transient stability domain. Appendix B provides more discussion and literature review in this area.

Generally, transient stability analysis involves numerically solving large number of nonlinear differential equations [21]. The most straightforward approach in this area is based on direct or time-domain simulations of transient dynamics following a disturbance. Direct methods are based on energy or Lyapunov functions, which guarantees the system convergence to stable equilibrium points without the knowledge of post-fault trajectory. These techniques has been successfully used in conventional power systems [21,22,28]. In these works, the stability is assessed based on the potential and kinetic energies required at a specific time. Energy function has also been considered in renewable energy resource controller designs. In [139], active and reactive power support capability of DFIG is exploited using Lyapunov function based control laws in order to increase the damping of the oscillation modes of power systems. Here, the dynamics of DFIGs are simplified to load models and a control scheme is developed based on negative derivative of energy function. Excessive simplifications are also performed in these methods to make the approach computationally efficient [28]. However, these non-optimal methods only provide the sufficient conditions for assessing the stability and are mainly used as preventive or emergency control schemes.

An alternative set of approaches deemed promising to address the above issues are machine learning, generally termed as measurement-based techniques. Neural

Networks (NNs) has been successfully investigated in literature for post-fault transient stability assessment and control [140], conventional power system WAC design [10], wind farm local control [89], or dynamic energy management system for micro grids [141]. The supervised learning methods are very fast but again their accuracy highly depends on offline training. Reinforcement Learning (RL) as one of the algorithms in this area possesses the ability to learn from the interaction with system and update its policy during state-action interactions [7]. Additionally, they can find the optimal solution to the approximated cost function using *ADP*, thus dealing with non-stationary power system dynamics.

In this chapter a novel method based on combination of energy functions with machine learning techniques has been exploited to design the WAC for wind integrated power grid. In order to address the transient stability, we propose a method to link the energy function estimates to the cost function of the RL approach. The controller is designed in order to ensure that the gradient of the cost function is strictly negative and *minimized* forward-in-time. The advantage of this method is that, it not only provides transient stability index which guarantees the system convergence to post-fault equilibrium, but also can enhance the transient performance of the system by increasing the damping of the oscillation, transfer capacity, and critical clearing time (CCT). Another consideration is the transient dynamics of DFIG-based wind generators. In this work, it is shown that wind farms are capable of maintaining their local stability while system is exposed to transient inter-area oscillations, if these oscillations are not severe enough to activate the protection devices in wind farms. Further, this chapter proposes exploitation of DFIG in the form of *potential*, *reactive*, and *inertial* energies to consume/support some energy from/to the system. The system dynamics is represented in coupled oscillatory platform with synchronous generators as coherent group of generators containing *potential*, *kinetic*, and *field* energies. Through adaptation of the proposed multi-agent WAC, wind farms power references along with synchronous

generators excitation voltage references are controlled to optimize the overall energy exchange in the system.

Overall, the proposed strategy:

- is based on the Lyapunov stability theory deriving an energy function for the DFIG which can be used as a cost function in the WAC,
- is independent on the network topology and the post-fault equilibrium points,
- is optimal,
- is robust in handling WAM uncertainty such as delays and system parameters,
- can be easily incorporated to the local wind farm controller as an added control function.

The rest of the chapter is organized as follows. The second section provides the theoretical framework for the system modeling and energy function development for synchronous generators and wind farms. In section 6.3, the control hierarchy and the proposed WAC design is illustrated. Section 6.4 presents the power system test case and WAC architecture implementations details. Then, the simulation results is provided in Section 6.5, followed by conclusion in section 6.6.

6.2 Theoretical Framework

Transient rotor angle stability, focuses on the angular oscillations of coherent machines and their related transient energies during disturbances. In this chapter, generator buses with a strong coupled network which display coherent behavior are aggregated as one equivalent generator [130]. Consider $\mathcal{N} \equiv \{1, 2, ..., |\mathcal{N}|\}$ as the set of reduced order buses of the power system network, $\mathcal{A} \equiv \{1, 2, ..., |\mathcal{A}|\}$ as the sets of areas aggregated buses and $\mathcal{W} \equiv \{|\mathcal{A}| + 1, ..., |\mathcal{N}|\}$ as wind farm buses. The dynamics of each bus $j \in \mathcal{N}$ can be characterized by the complex voltage $V_j \angle \delta_j$, where δ_j

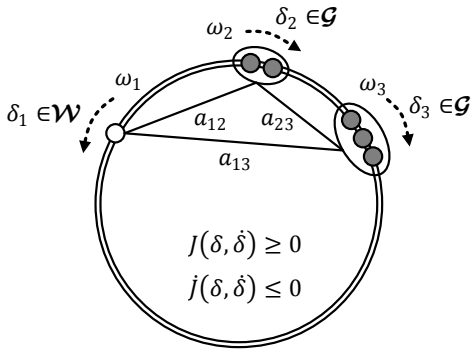


Figure 6.1: Coupled oscillatory system

is the phase angle with respect to synchronously rotating reference frame. The bus frequency is given by $\omega_j = \dot{\delta}_j$.

Mathematically, the inter-area oscillations can be described by a coupled oscillator network of Fig. 6.1, where each oscillatory element has a specific natural frequency and synchronization-enforcing coupling [142]. Considering the generating units ($j \in \mathcal{N}$) as the oscillatory elements the transient instability of the system is caused by a mismatch between injected power of each unit and the power flows along the transmission lines. This mismatch is usually due to a fault on power system lines or other components, which is followed by the action of protection system. These actions may change or restore the topology of the system, which leads to a transition from pre-fault to post-fault equilibrium points. The transition can be viewed from energy landscape or transient energy function. Considering δ and ω as the states of the oscillators and J as the energy function in the form of combination of Lyapunov function candidate of all oscillatory elements, the system is asymptotically stable when [21],

$$J(\delta, \omega) \geq 0, \quad \dot{J}(\delta, \omega) \leq 0 \tag{6.1}$$

Based on this, the sufficient conditions for stability assuming this energy function model can be evaluated and the candidates for proposed Lyapunov function can be

developed. This work aims to develop an energy function to capture dynamics of synchronous generators and wind farms as the oscillators and minimize it by means of WAC in order to damp systems oscillations.

6.2.1 Synchronous Generators Areas

In the context of transient stability assessment due to slow variation of voltage magnitude (V) in comparison to the angle, for synchronous generators, it is assumed that $E = V$ with E being the internal voltage magnitude. The third order model of synchronous generator is utilized [41] such that the effects of generator control and excitation system in damping the oscillations of coherent generators can be studied. This is given by:

$$m_j\dot{\omega}_j + d_j\omega_j + \sum\nolimits_{k\in\mathcal{N}_j} B_{jk}E'_{qj}E'_{qk}\sin(\delta_{jk}) = P_j, j \in \mathcal{A} \tag{6.2}$$

$$T'_{doj}\dot{E}'_{qj} = \Delta E_{fj} - \Delta E_{qj}, \quad j \in \mathcal{A} \tag{6.3}$$

where, $\delta_{jk} = \delta_j - \delta_k$ is the inter-area angle, δ_j and ω_j are bus angle and angular frequency in Center Of Inertia (COI) frame [143] as

$$\delta_j = \tfrac{1}{m_j}\sum\nolimits_i^N m_i\delta_i, \quad \omega_j = \tfrac{1}{m_j}\sum\nolimits_i^N m_i\omega_i, \quad m_j = \sum\nolimits_i^N m_i \tag{6.4}$$

where, i is the generator index in the area j, m_j and d_j are the equivalent inertia constant and damping coefficient of the area, $P_j = \sum_i^N P_i$ is the power injection, E'_q the q-axis component of transient internal voltage, B_{jk} is the (j,k)th entry of the reduced lossless admittance matrix; $\sum_{k\in N_j} B_{kj}E'_{qj}E'_{qk}\sin(\delta_{jk})$ is the power flow along the transmission lines connected to bus j, N_j is the set of neighboring buses of the j^{th} bus, E_q is the internal voltage in q-axis, E_f the excitation voltage, and T'_{do} the equivalent transient time constant. In these equations, operator Δ denotes the deviation of the variables from the reference equilibrium point.

An energy function for such an aggregated synchronous generator model comprises of the sum of the kinetic energy, potential energy, and field or control energy with respect to the relevant equilibrium points [41]. Such a function can be represented for each area as,

$$J_j(\delta,\omega,E_q) = \underbrace{-\int_{\delta_j^*}^{\delta_j} P_j d\delta_j + \int_{\delta_j^*}^{\delta_j} \sum\nolimits_{k\in N_j} B_{jk}E'_{qj}E'_{qk}\sin\delta_{jk}d\delta_j}_{PE_j}$$
$$+ \underbrace{0.5m_j\Delta\omega_j^2}_{KE_j} + \underbrace{0.5\beta_j(\Delta E'_{qj})^2}_{FE_j}, \quad j\in\mathcal{A} \tag{6.5}$$

where, $\Delta\omega_j = \omega_j - \omega_j^*$, β is a parametric coefficient based on equivalent X_d and X'_d as synchronous and transient reactances. Next, the derivative of J with respect to time along the trajectory of the system (6.2) and (6.3) is evaluated [41]. The derivative of such a function after simplifications can be represented as,

$$\dot{J}_j(\omega,E_q,E_f) = -d_j\Delta\omega_j^2 - \tfrac{1}{T'_{d0j}\Delta X_{dj}}(\Delta E_{qj}^2 + \Delta E_{qj}\Delta E_{fj}) \tag{6.6}$$

with the assumptions on higher orders as $\Delta E_f = K\Delta E_q$, $\Delta E_q = X_{ad}\Delta i_f$, and $\Delta i_f = K'\Delta V_{ref}$ with the gains K and $K' > 0$, X_{ad} the d-axis armature reaction reactance, Δi_f as a field current deviation, and $\Delta V_{ref} = V_{ref} - V_t$ as the terminal voltage deviation or the control input to the field component, (6.6) yields

$$\dot{J}_j(\omega,V_{ref}) = -d_j\Delta\omega_j^2 - r_j\Delta V_{refj}^2, \quad j\in\mathcal{A} \tag{6.7}$$

with $r_j \equiv \frac{1+K_j}{T'_{d0j}\Delta X_{dj}}X_{adj}^2K'_j$. Equation (6.7) reveals that the states and actions affecting the transient energy deviation of each area in time are $\Delta\omega_j$ and ΔV_{refj}, and the respective parameters d_j and r_j.

6.2.2 DFIG-based Wind Farms

It is known that wind generators are not much affected by the low frequency oscillations due to fast control capability of power electronic-based controllers [144]. However, as components in the coupled oscillator model they can have an impact on the oscillations and transient energy of synchronous generators. Considering the wind farm buses as the PQ buses, the concept of controllable loads can be applied to the wind farms. Let P_j and Q_j denote the total real and reactive power leaving the jth bus via transmission lines, and P_j^{ss} and Q_j^{ss} as steady state power outputs and ΔP_j and ΔQ_j as changes in power due to control references, then $P_j = P_j^{ss} + \Delta P_j$ and $Q_j = Q_j^{ss} + \Delta Q_j$. The power equations on the wind farm buses could be written as

$$\sum\nolimits_{k\in\mathcal{N}_j} B_{jk}V_jV_k\sin(\delta_{jk}) = P_j \quad , \quad j \in \mathcal{W} \tag{6.8}$$

$$-\sum\nolimits_{k\in\mathcal{N}_j} B_{jk}V_jV_k\cos(\delta_{jk}) = Q_j \quad , \quad j \in \mathcal{W} \tag{6.9}$$

The output power from internal point of view in DFIG is generated from stator and rotor as $P = P_s + P_r$ or $P = \omega_r P_s$ at steady state operating point, where P_s and P_r are stator and rotor powers and ω_r is the rotor speed. This power in dynamic and transient domain considering the impacts of ω_r state can be derived as follows:

$$m_j\dot{\omega}_{rj}\omega_{rj} = P_{mj} - Pj \quad , \quad j \in \mathcal{W} \tag{6.10}$$

where, m_j is the rotor inertia and P_{mj} is the mechanical reference based on the wind speed, wind turbine speed and pitch angle.

An energy-type Lyapunov function for such a wind farm includes the potential

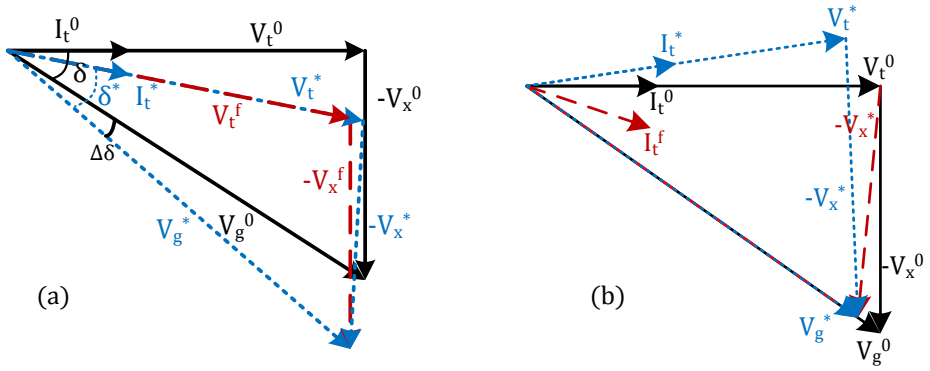


Figure 6.2: Vector diagram of (a) angle deviation (b) voltage deviation. V_g: grid voltage, V_t: DFIG stator terminal voltage, V_x: line voltage drop, I_t: current flowing through DFIG, δ angle between terminal and grid voltage, superscript: 0: initial steady state vectors, f: transient vectors (after the fault), $*$: final steady state vectors

energy, reactive energy, and kinetic energy based on δ, V, and ω_r as,

$$
\begin{aligned}
J_j(\delta, V, \omega_r) = & \underbrace{\int_{\delta_j^*}^{\delta_j} \sum\nolimits_{k\in\mathcal{N}_j} B_{jk}V_jV_k\sin(\delta_{jk})d\delta_j - \int_{\delta_j^*}^{\delta_j} P_j^{ss}d\delta_j}_{PE_j} \\
& + \underbrace{\int_{V_j^*}^{V_j} \sum\nolimits_{k\in\mathcal{N}_j} B_{jk}V_jV_k\cos(\delta_{jk})dV_j - \int_{V_j^*}^{V_j} Q_j^{ss}dV_j}_{QE_j} \\
& + \underbrace{0.5m_j\Delta\omega_{rj}^2}_{KE_j}, \quad j \in \mathcal{W}
\end{aligned} \tag{6.11}
$$

where, $\Delta\omega_r^2$ is the deviation of rotor speed from its MPPT reference. The first two elements, PE and QE, are based on power outputs and the respective states which are terminal voltage angle and magnitude.

Due to asynchronous nature of the DFIG, the wind generator voltage angle and grid angle are decoupled. Fig. 6.2a shows the operation of DFIG in PQ control mode during a sudden change in the grid voltage angle (due to a transient fault). It can be observed from the figure that the DFIG control can adjust its terminal voltage such that the active and reactive power can be maintained at the same level before the grid disturbance. Fig. 6.2b shows the operation of DFIG during a grid voltage deviation.

It can be seen that even after the voltage sag in the grid, DFIG tries to maintain the reference active and reactive power, by changing its terminal voltage and current based on changes in the grid. In the above case, it can be observed that after the grid voltage changed from V_g^0 to V_g^*, the DFIG terminal voltage changes from V_t^0 to V_t^*, and current changes from I_t^0 to I_t^* such that the initial value of the active and reactive power is maintained. It should be noted that for the above phasor diagram, the severity of the changes in grid voltages is not large enough to cause DFIG wind generator to be disconnected from the grid or any of its maximum allowable operating limit to be violated. Hence, the energy function defined above yields:

$$\begin{aligned} J_j(\delta, V, \omega_r) = &- \sum\nolimits_{k\in\mathcal{N}_j} B_{jk}V_jV_k(\cos(\delta_{jk}) - \cos(\delta_{jk}^*)) + P_j^{ss}\Delta\delta_j \\ &+ \sum\nolimits_{k\in\mathcal{N}_j} 0.5B_{jk}V_k\cos(\delta_{jk})(V_j^2 - V_j^{*2}) + Q_j^{ss}\Delta V_j \\ &+ 0.5m_j\Delta\omega_{rj}^2, \quad j \in \mathcal{W} \end{aligned} \tag{6.12}$$

The third element in this equation, KE, is released and extracted automatically from the rotating mass for a change in power reference and consequently the rotor speed. The proposed method exploits the kinetic energy stored in the rotating mass of wind turbines, such that the additional amount of power supplied by the wind generator to the grid is proportional to $\Delta\omega_r^2$.

From the network point of view, DFIG buses behave like negative active and reactive power loads, hence, active and reactive powers are considered as control variables. With the assumption of energy stored in the outer and inner control loops as negligible $\Delta P = \Delta P_{ref}$ and $\Delta Q = \Delta Q_{ref}$, then the derivative of wind farm energy function

could be written as a function of these control actions as,

$$\dot{J}_j(\delta, V, \omega_r, P_{ref}, Q_{ref}) = \Delta P_{refj}\Delta\dot{\delta}_j + \Delta Q_{refj}\Delta\dot{V}_j + [\frac{\partial P_{mj}}{\partial \omega_{rj}} - \frac{\partial P_j}{\partial \omega_{rj}}]\Delta\omega_{rj}^2, \quad j \in \mathcal{W} \tag{6.13}$$

Further, characteristics of the output power can be linearized around the ω_{rj}^*, as the initial rotor speed set point as

$$P_j = P_{mj}\omega_{rj}/\omega_{rj}^* + \Delta P_{sj}\omega_{rj} \quad , \quad j \in \mathcal{W} \tag{6.14}$$

where $\Delta P_{sj} = \Delta P_{refj}$. Without loss of generality, considering the Maximum Power Point Tracking (MPPT) operation of DFIG, the sensitivity of the mechanical power to rotor speed could be assumed as zero as long as rotor speed deviation is limited. However exact value could be extracted through the MPPT curve and estimated by initial set point. Hence,

$$\dot{J}_j(\delta, V, \omega_r, P_{ref}, Q_{ref}) = \Delta P_{refj}\Delta\dot{\delta}_j + \Delta Q_{refj}\Delta\dot{V}_j + \Delta P_{ref}\Delta\omega_{rj}^2, \quad j \in \mathcal{W} \tag{6.15}$$

This equation reveals that, whenever ΔP is injected from wind farm along with increase in the angular speed, the wind farm deviates from the equilibrium point in the form of increase in the energy or positive $\dot{J}$. Similar effect happens to the voltage and ΔQ. With respect to kinetic energy, any deviation from ω_r^* introduces an energy, which has increasing (ascending) or decreasing (descending) behavior based on the sign of ΔP and the electrical and mechanical powers difference.

6.3 Proposed Hierarchical Controller Design

The controller action of overall system including the synchronous and wind generators can be presented as 3-level combination of local, wide-area, and steady state

Table 6.1: WAC's state, control, and energy signals

Signals	Area ($A = 1, ..., \|\mathcal{A}\|$)	Wind Farm ($W = 1 + \|\mathcal{A}\|, ..., \|\mathcal{N}\|$)
$x = [x_A, x_W]$	$x_A = [\delta_A, \omega_A]$	$x_W = [\delta_W, V_W, \omega_{rW}]$
$J = \sum J_A + J_W$	$J_A = PE_A + KE_A + FE_A$	$J_W = \quad PE_W + QE_W + KE_W$
$u^{wac} = [u_A, u_W]$	$u_A = \Delta V_{refA}^{wac}$	$u_W = [\Delta P_{refW}^{wac}, \Delta Q_{refW}^{wac}]$

counterparts as

$$u(t) = u^{loc}(t) + u^{wac}(t) + u^{ss}(t) \tag{6.16}$$

The WAC tracks the generators and wind farms states $x = [x_A, x_W]$ through wide-area measurements, derives wide-area feedback $u^{wac} = [u_A, u_W]$, to minimize the energy function $J = J_A + J_W$ and enhance the global system behavior. Table 6.1 presents the states, control action, and energy functions definition for areas and wind farms.

6.3.1 Overall Architecture

Local control is the first control level in control hierarchy that has the fastest response. This control responds to local system dynamics and ensures that the variables track their reference values. In the case of synchronous generator, input to the excitation field without damping controllers is the error between steady state reference V_{ref}^{ss} and the terminal voltage of generators as shown in Fig. 6.3. This voltage error could be augmented with ΔV_{ref}^{loc} as local damping signals derived from PSS and enhanced by ΔV_{ref}^{wac} as a wide-area level controller. In the case of DFIG, conventional vector control as shown in Fig. 6.3 is designed as local controller, where the q-axis of the rotor current controls the stator active power P_s and d-axis of the rotor current controls the stator reactive power Q_s. The controller consists of the inner loop current control and outer loop power control. The active power reference P_{ref}^{ss} comes from the MPPT, which takes rotor speed ω_r as its input. The reactive power reference Q_{ref}^{ss} is supplied by the grid operator based on the reactive power needs of the grid. In addition, a frequency based droop control has been implemented such that the terminal

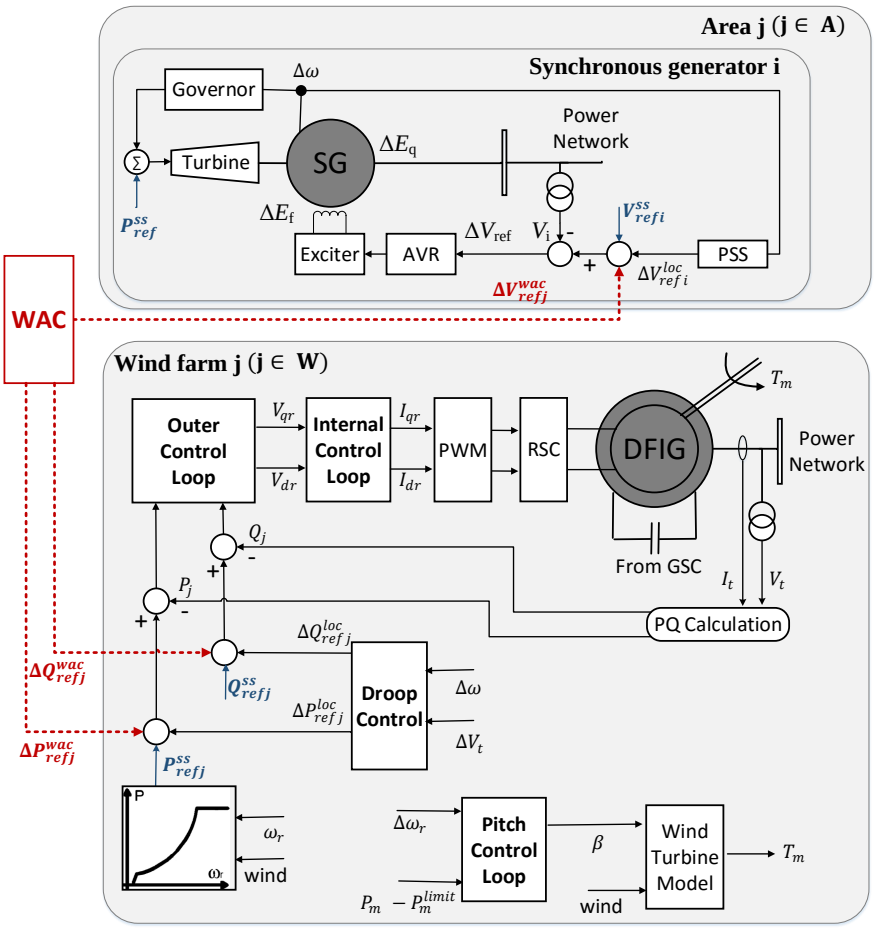


Figure 6.3: Structure of the proposed integrated WAC hierarchy considering synchronous generator and DFIG: WAC signals for wind farms and areas $u^{wac} = [u_W^{wac}, u_A^{wac}]$ where $u_W^{wac} = [\Delta P_{ref}^{wac}, \Delta Q_{ref}^{wac}]$, $u_A^{wac} = \Delta V_{ref}^{wac}$, synchronous generator local control: $u^{loc} = \Delta V_{ref}^{loc}$ and steady state reference:$u^{ss} = \Delta V_{ref}^{ss}$; wind generator primary control: $u^{loc} = [\Delta P_{ref}^{loc}, \Delta Q_{ref}^{loc}]$, and steady state reference: $u^{ss} = [\Delta P_{ref}^{ss}, \Delta Q_{ref}^{ss}]$.

voltage frequency can be maintained around the nominal value and a voltage based droop controller has been used in the reactive power control loop such that it helps faster recovery of the stator terminal voltage following a disturbance. On top of the existing power control loops, a supplementary damping WAC has been used in this work. The detailed control loops implemented in DFIG is shown in Fig. 6.3.

6.3.2 WAC Design

Decreasing nature of the energy function provides the main certificate of local stability; If the energy function is less than the critical energy, the system is stable

and the system might be unstable otherwise. The direct method of Lyapunov stability based on derived J and $\dot{J}$ in section II provides only the sufficient conditions for assessing the stability. Additionally, these energies are nonlinear functions of multiple parameters of the generating units, post-fault operating points, and network topology. To overcome these issues machine learning techniques could be utilized in conjunction with the energy functions in WAC design to cover all possible operating conditions and topology changes by learning system parameters.

A common approach is to deploy NNs to map the nonlinearities and uncertainties of the system identification, control and the cost function. Here, Feed Forward Neural Network (FFNN) is used as

$$Y(k) = W(k)^T \Phi(X(k)) + \epsilon(k) \tag{6.17}$$

where, $X(k)$ and $Y(k)$ are the vector input and output of NN, $\Phi(.) \in \Re^h$ is the corresponding nonlinear mapping function of states, $W(k) \in \Re^h$ is the parameter vector of weights at time step k, with $h \in \mathbb{N}$ dimensionality of the feature vector representing each state, and ϵ is the approximation error.

In general, the task of the controller is to generate a control action policy u that transfers the states from known initial values x^0 to specified final equilibrium points x^* with the minimum cost-to-go or energy as objective. Using the described NN structure, the *critic*, *action*, and *identifier* NN blocks are constructed to approximate the cost function $J(k)$, the control action $u(k)$, and the dynamics of the system $x(k)$, respectively, at each time step k. The architecture employed in this chapter to construct the WAC as the RL approach. It should be noted that simple supervised learning algorithm could be used to train the critic NN with the target of the energy function presented in Table 6.1. Even though the parameters of the generating units and network are known for the system as a whole, some parameters may be changing

with time. Thus, the trained critic NN could show short-comings in online application and does not project precise energy functions of the system. On the other hand, RL is capable of dealing with the noise and uncertainty through state-action interactions [7]. Besides, RL can learn the Hamilton-Jacobi-Bellman equation associated with optimal control theory through a critic NN, and find the optimal control through action NN. Training of action NN is based on selecting sequence of actions that minimize the estimated cost function (J). Using RL approach, the energy function at time step k is given as

$$J(k) = \sum_{n=0}^{\infty} \dot{J}(k+n) \tag{6.18}$$

where, $\dot{J}$ is the utility function used for reward/punishment in terms of RL concept, or incremental cost function in Lyapunov stability concept. This function based on (6.7) and (6.15) can be represented as,

$$\begin{aligned} \dot{J}(k) &= -x_A(k)^T Q x_A(k) - u_A(k)^T R u_A(k) \\ &+ u_W^T(k) R'(x_W(k+1) - x_W(k)) + x_W(k) Q' x_W(k) R^{"} u_W^T(k) \end{aligned} \tag{6.19}$$

where, coefficient matrices are in the form diagonal matrix with the elements of

$$\begin{aligned} Q_{jj} &= [0_{|\mathcal{A}|\times 1}, d_1, ..., d_{|\mathcal{A}|}] \\ R_{jj} &= [r_1, ..., r_{|\mathcal{A}|}] \\ R'_{jj} &= [1_{2|\mathcal{W}|\times 1}, 0_{|\mathcal{W}|\times 1}] \\ Q^{'}_{jj} &= [0_{2|\mathcal{W}|\times 1}, 1_{|\mathcal{W}|\times 1}] \\ R^{"}_{jj} &= [1_{|\mathcal{W}|\times 1}, 0_{|\mathcal{W}|\times 1}] \end{aligned}$$

Another advantage of using RL over supervised learning in online process is that $\dot{J}$ is not dependent on δ^* which could be different than δ^0 and not known in practice.

In this chapter, an iterative learning scheme is used for finding the optimal control solution u^* corresponding to the optimal cost function J^*. The algorithm starts with J^0 and u^0 at iteration $i = 0$ based on offline tuning, and then solves for u and J as follows:

$$\frac{\partial \dot{J}(k)}{\partial u^{i+1}(k)} + \frac{\partial x(k+1)}{\partial u^i(k)} . \frac{\partial J^i(x(k+1))}{\partial x(k+1)} = 0 \tag{6.20}$$

$$J^{i+1}(x(k)) = \dot{J}(x(k), u^i(k), x^i(k+1)) + J^i(x(k+1)) \tag{6.21}$$

where,

$$\frac{\partial \dot{J}(k)}{\partial u(k)} = 2Ru_A(k) + R'(x_W(k+1) - x_W(k)) + x_W(k)Q'x_W(k)R" \tag{6.22}$$

This optimization iteration is performed by means of training the NNs through gradient descent via back-propagation which is discussed in detail in previous chapters.

6.4 Implementation and WAC Construction

In order to assess the capability of the proposed method, a modified IEEE 68-bus 16-machine power system has been modeled using PSCAD and is presented in Fig. 6.3. The original IEEE test system is a reduced order equivalent of the interconnected New England and New York power systems [134]. The New England and New York power grids are represented by group of generators. Grouping consideration is based on the dynamic data of the generators and the parameters of the transmission system [134]. Preliminary results on the real-time simulation of 2-area model in OPAL-RT and the proposed method in this chapter have recently been reported in our conference paper [145].

6.4.1 Wind Integrated Power System Test-Bed

In the modified system, W4 and W5 are replaced with the same size conventional generators in original system accounting for 9.75% penetration in area 1 and 15.53%

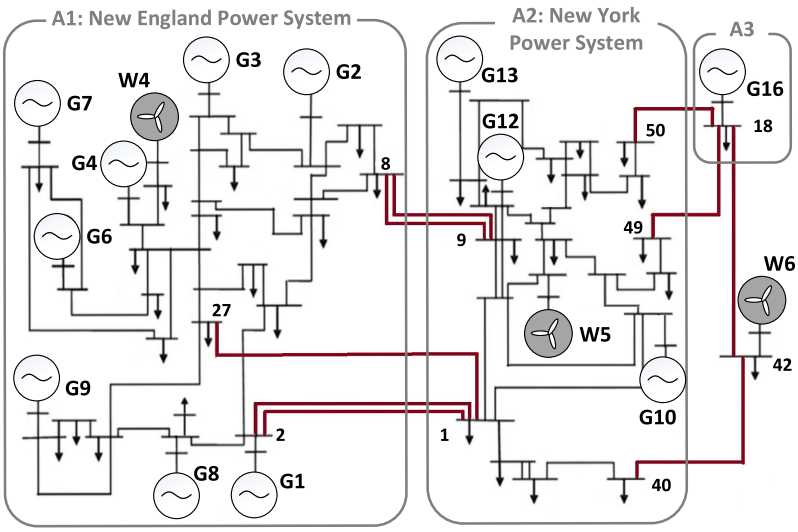


Figure 6.4: Modified IEEE 68-bus New England New York power system test network including three areas ($|\mathcal{A}| = 3$, $j = 1:3$) and three wind farms ($|\mathcal{W}| = 3$,$j = 4:6$).

penetration in area 2. The standard test network also contains three large generators representing external networks which import approximately 2.1 GW into area 2 under nominal loading. In the modified network, two of these generators (G14 and G15) are replaced with a wind farm (W6). This unit generates 2.8 GW, supporting 800 MW to the neighboring areas. The impact of increased DFIG penetration in the form of inertia change is investigated in the small signal stability domain performed using Power System Toolbox and is presented in Fig. 6.5. It has been observed that the system modes of oscillation shows considerable change with increased penetration of wind farms especially with respect to the damping capabilities and frequencies. For instance, damping of inter-area oscillation between area 1 and 2, has been decreased from 0.142% to 0.104%. The main purpose of this thesis is to capture these oscillations in the form of transient energy and enhance the damping with WAC.

For EMTP simulations the time step is considered as 50 us. All synchronous generators are represented by detailed full sixth-order model equipped with governors, exciters, and PSS. Two types of excitation systems for the synchronous generators is considered: IEEE standard DC exciter (DC4B) and the standard static exciter

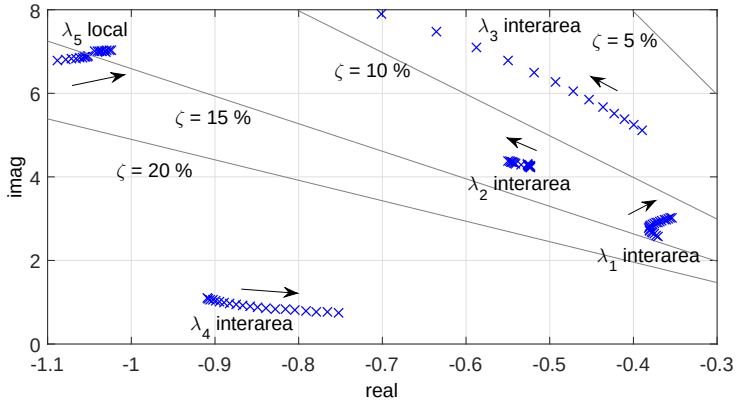


Figure 6.5: Electro-mechanical modes of the original 68-bus system as a function of reduced inertia of G5, G11, G14, G15.

(ST1A). The governor model for all generators is the standard TGOV1. The generator local damping control architecture is conventional PSS design with the consideration and procedure of washout filters and lead/lag elements. Full system details and generator and controllers parameters can be found in [134, 146].

The wind farms are modeled as an equivalent single machine. It has been shown in [147] that the equivalent aggregated machine representation of DFIG will accurately depict the dynamic performance of wind farms. In this work, a detailed model of DFIG with the power capacity of 500 MW, 1000 MW, and 2800 MW at maximum rated speed are used including Grid Side Converter (GSC), Rotor Side Converter (RSC), dynamic model of the Wound Rotor Induction Generator (WRIG), two mass model for the wind turbine, and aerodynamic model of the turbine to get the mechanical torque from wind speed, as well as the associated controls for each of the components. The converters used in this work are three phase two level converters. The parameters of the DFIG used are obtained from NREL's report [148] and are scaled accordingly to the power requirements. The controllers are optimized based on the p.u. model of the 1.5 MW turbine used in the report. Q_{ref}^{ss} for wind generators are set based on load flow analysis which are: 116 MVar, 122 MVar, and 780 MVar. The frequency droop control is neglected; however, the voltage droop controller has

been used with the gain of 1 such that it helps in faster voltage establishment during the start-up of the simulation as well as faster recovery of the terminal voltage following disturbances. The outer loop PI controllers have a proportional gain of 0.01, 0.04, 0.04, and integral time constant of 0.01, 0.08, 0.012 for W4, W5, and W6, respectively. Additionally, pitch control loop time step has been considered as 1 s. The rest of the DFIG parameter can be found in [148]. It should be noted that, all the operating limits of the DFIG are addressed in this work so that realistic operations can be captured. The rotor and stator currents are limited to 1.2 p.u. for dynamic overloading of the machine. The maximum allowed rotor speed is 1.2 p.u. and the minimum allowed rotor speed is 0.7 p.u. Also, it has been assumed that the DFIG turbine shaft can bear the short term stress applied to its ends because of the deviation between mechanical power and electrical power. Another consideration in this thesis is on the wind speed, which is assumed to be constant during transient events.

6.4.2 Proposed WAC Architecture

Input vector of the WAC is the states derived from WAM in order to estimate the energy function, and the output is the optimal damping control presented in Fig.6.6. Here, we assume that the communication system can transmit the signals to WAC at the speed of 100 frames/sec; hence, WAC communication time step is considered as $\Delta t = 10$ ms. It is also assumed that each generating unit bus has a sensor that transmits signals to WAC system. For the synchronous generators buses, signals are aggregated at WAC center based on COI frame (see (6.4)) to generate states of each system area. It is considered that WAM is able to transmit the ω_r as the aggregated rotor speed of wind generators in a wind farm.

The construction of the WAC starts with offline supervised learning algorithm in order to set the initial weights. The training is performed based on the I/O signals for each of the NNs derived using Matlab NN toolbox. Table 6.2 lists the NNs signals used for the study. To capture the dynamics of the system, input signals are accompanied

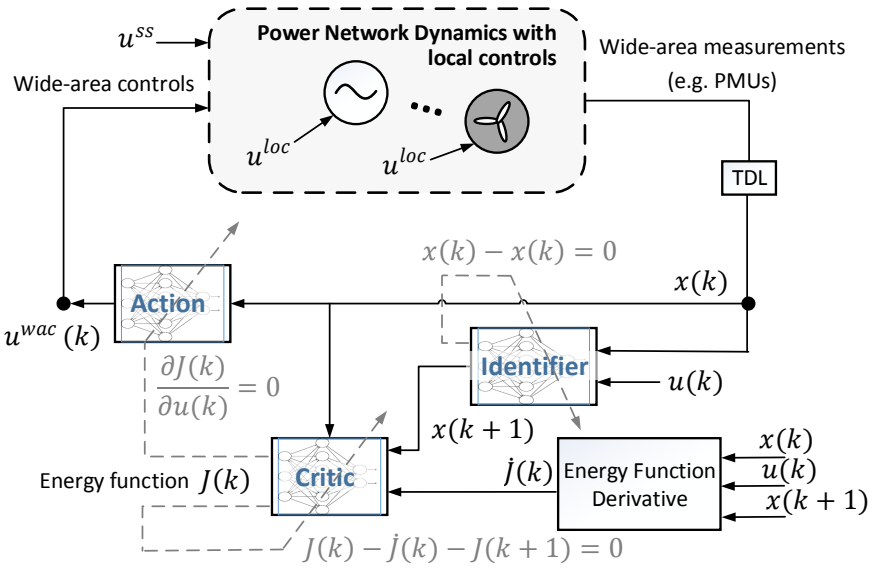


Figure 6.6: Proposed online WAC design.

with 10 time step delayed signals. For identifier NN, pseudo random inputs and related outputs of the power system model is captured and fed to the NN data base. Then, gradient descent batch learning algorithm is performed with a learning rate of 0.001. It has been observed that e_{Id} reaches its threshold in 100 epochs. Initial training of the action NN is based on local controllers references described in the previous subsection. Critic NN is trained based on (6.5) and (6.11) for different fault scenarios and operating points to learn the relative energy functions. Parameters of energy function including m, B, β are given in [134]. It should be noted that B is the Kron-reduced susceptance matrix with the loads removed from consideration. Once the NNs weights have converged ($e < 0.01$), then the action NN is connected to the generator's exciter and wind farm's power loops to augment the local controllers. The

Table 6.2: Configuration of Neural Networks

NN	Input (nodes)	Output (nodes)	Hidden nodes
Identifier	$[x(k), x(k-10), u(k), u(k-10), 1]$ (49)	$x(k+1)$ (15)	70
Action	$[x(k), x(k-10), 1]$ (31)	$u(k)$ (9)	50
Critic	$[x(k), x(k-10), 1]$ (31)	$J(k)$ (1)	50

process of online training starts by monitoring the states to incrementally train the identifier. This is followed by critic and action update iterations. The weights of critic NN in each iteration is updated by means of derivative of energy function $\dot{J}(k)$ and estimated future energy function $J(k+1)$. Without lose of generality, parameters of $\dot{J}(k)$ are assumed as $d_j = 0.1$ and $r_j = 0.01$. Furthermore, action NN is tuned online to minimize the estimated energy function.

6.5 Simulation and Test Results

6.5.1 Case A: Wind Farm Contribution

In this case study, an investigation on the impacts of wind farm controls for transient stability enhancement is performed. Therefore, WAC for synchronous generators are disabled and only the WAC for wind generators are enabled. The effectiveness of augmenting the local control of wind farms using WAC is evaluated by applying a three phase fault at tie-line 18-49 at near end of the bus 18 and comparing the stability characteristics of the system with and without WAC. The fault is cleared after 300 ms by opening the tie-line breakers. Initially, 450 MW is transferred from area 3 to area 2 through this tie-line. After the fault this power is transferred through line 18-50 and Area 1. In order to visualize the performance of the proposed WAC, the dynamic response of the system with and without WAC is presented in Fig. 6.7. It can be seen that with WAC on the wind farms, not only the angular oscillations of the areas ($j = 1:3$) damps faster, but the angular deviation and magnitude of the terminal voltage of wind farms ($j = 4:6$) are improved.

Fig. 6.8 shows the active and reactive power contribution of the wind farms along with the WAC references during the transient period such that the overall system moves to the new equilibrium point faster. Also it should be noted that no operating limits of the wind farms are violated during the transient period. With respect to active power, the WAC tries to minimize the power difference of the two areas in the form of potential energy function. In addition, through reactive power control,

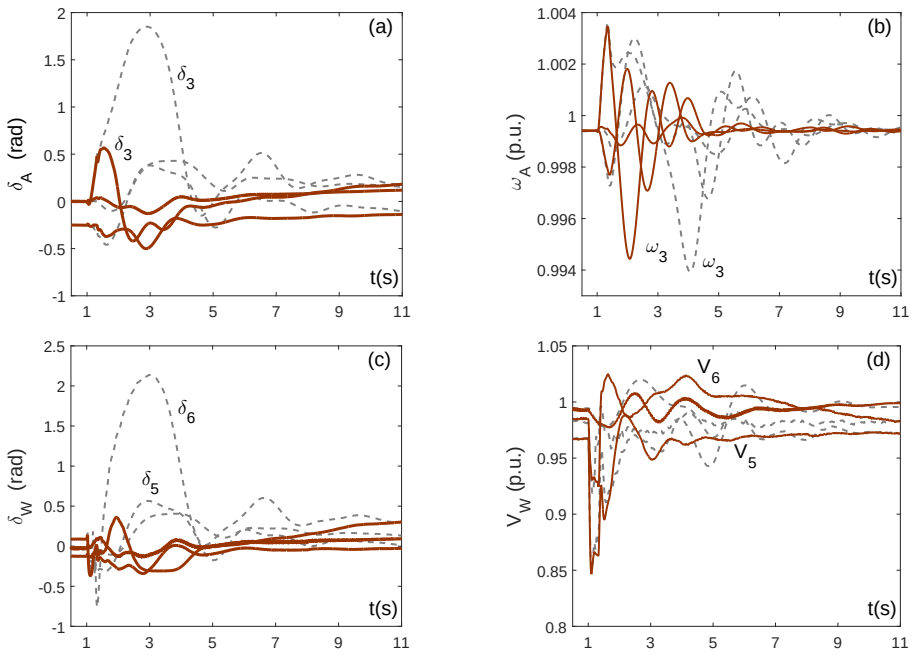


Figure 6.7: Case A. rotor and voltage oscillations of the synchronous generator areas and wind farms for Case A with WAC (solid) and without WAC (dashed)

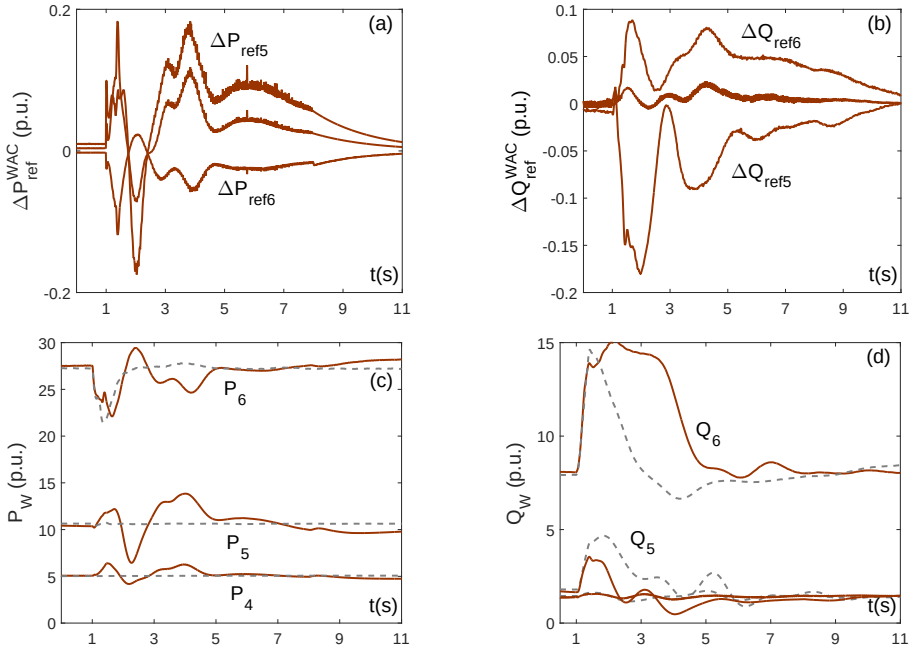


Figure 6.8: Case A. wind farms WAC signals and power outputs: with WAC (solid line) and without WAC (dashed line)

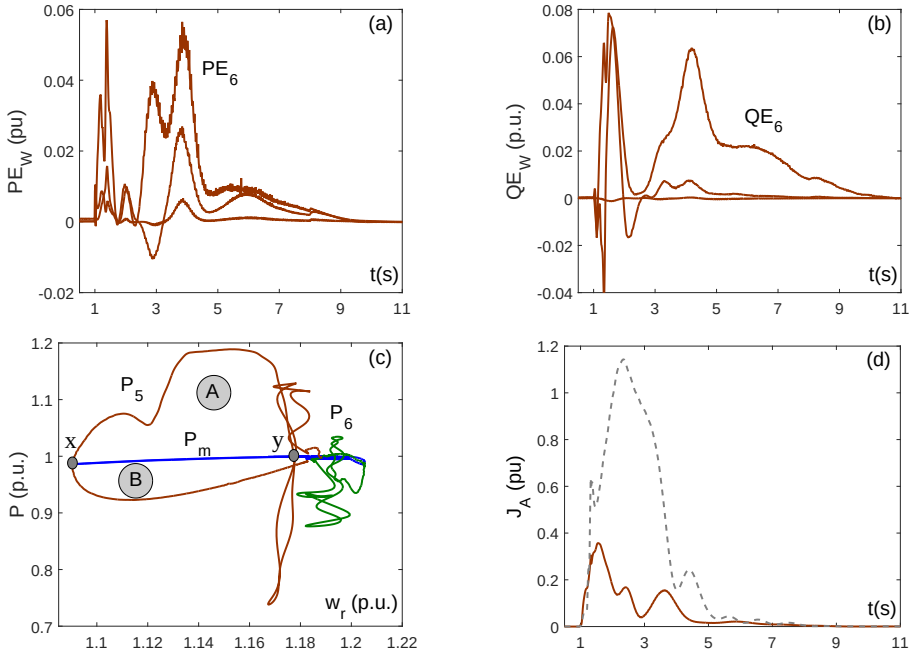


Figure 6.9: Case A. transient energy elements of wind farms and comparisons of transient energy interactions between areas (solid:with WAC, dashed:without WAC).

the architecture provides additional voltage and reactive support is provided for the system.

It can be seen that without the WAC, the local controllers of wind farms are able to maintain the stability of the wind farms. This could be especially seen in active power output of the W4 and W5 and reactive power output for W4 (P_4,P_5,Q_4). Reactive power control is augmented by voltage droop control to maintain the voltage in addition to Q^{ss}, leading to reactive power changes in the case of local controllers. In a separate simulation a test was carried out with only Q control which resulted in constant Q output from wind, however, the settling time for wind voltages was increased by 6 s. Further, Fig. 6.9 presents the changes in the system energies. Fig. 6.9c shows the deviation of the operating point of the wind farms because of the implementation of WAC. It can be observed that as the power from the W5 (P_5) increases, the wind farm rotor decelerates and also the mechanical power decreases (at a slower rate); hence, the wind farm can reach a new equilibrium at point "x",

but as that point is not the steady state operating point, the machine slowly moves towards point "y" as the WAC signal is taken off from wind farm (region "B"). As for W6 (P_6), the power reference is decreased; however, pitch control acts and resists ω_r deviation. In summary it is noticed that the proposed WAC enhances the transient performance of the system as shown in Fig. 6.9d in the form of transient energy reduction for synchronous generators.

The CCT is a widely-used benchmark for the transient stability assessment, which is obtained by multiple time-domain simulations for the aforementioned fault on the system. The comparison of the CCT with and without WAC implemented on wind farms is presented in Table 6.3. This results shows considerable improvement in system stability using the WAC. The effectiveness of the WAC could be dependent on the wind power and initial states of the system. To investigate this criteria, CCT is evaluated for three wind farm conditions: sub-synchronous, synchronous, and super-synchronous with $w_r = 0.8pu$, $w_r = 1pu$, and $w_r = 1.2pu$, respectively, based on the reference wind speed. It should be noted that wind speed v is considered constant at the period of transition. The result shows that the CCT of the system is improved with WAC for all the conditions. However, for different wind speeds and operating points the contribution of wind farms is changed as well. Transient stability is mainly dependent on the pre-fault operating points and post-fault conditions. As for the system with $w_r = 1$ it is observed that the power transfer through tie-line 18-49 is reduced to 40 MW from 450 MW, which leads to higher CCT. However, as the power contribution of the wind generators is reduced the WAC contribution is also reduced. Other factors of wind generators affecting the transient stability enhancements are the maximum and minimum w_r limits and MPPT curve at the operating point.

6.5.2 Case B: Global WAC Performance

This test presents a transient case study to evaluate the effectiveness of the proposed WAC applied to synchronous generators and wind farms for damping the inter-area

Table 6.3: Case A. comparison of transient stability (CCT)

Scenarios	CCT without WAC	CCT with WAC
$w_r = 1.2pu\ v = 11m/s\ P^{ss} = 100\%$	320 ms	410 ms
$w_r = 1.0pu\ v = 9.2m/s\ P^{ss} = 58\%$	340 ms	390 ms
$w_r = 0.8pu\ v = 7.1m/s\ P^{ss} = 27\%$	260 ms	320 ms

oscillations. A 300 ms short circuit fault has been applied to the tie-line 1-2 connecting areas 1 and 2, which is followed by opening the line. Initially, considering the line 8-9 out of service, 1050 MW and 200 MW are transferred from area 1 to 2 through tie-lines 1-2 and 1-27, respectively. The loss of the low impedance line of 1-2 causes a transient oscillation as shown in Fig. 6.10a. One of the main reasons for choosing the RL as the algorithm for implementation of WAC is its capability to perform optimal control action. This multi-agent technique can be easily employed as a hierarchical controller. Figs. 6.10b, 6.10c, and 6.10d demonstrate the WAC and PSS signals for the aggregated generators and the WAC, primary power sharing control, and local feedback for W5. Here droop control is enabled and the gains of droop control are set as 1 for both voltage and frequency control. The error in WAC with respect to time has been demonstrated in Fig. 6.10e. It can be seen that, by using the proposed WAC the inter-area oscillations has been damped faster.

The stability enhancement is shown as the change in energy function Fig. 6.11. After the fault, G1, G10, W5, and G15 are affected more than other generating units. However, as wind farm supports optimal excitation control of generators, better transient performance has been gained as the reduction in energy functions. To show the effectiveness on stability margin, the system has been tested under higher loading conditions. The load demand of area 2 has been increased to gain 35% increase in tie-line 1-2 power transfer. This reduces the stability margin. Table 6.4 provides the comparison of CCT of the system with the described fault on line 1-2. To show the effectiveness as the enhanced transient stability controller, the loading of area 2 has been increased further. It has been observed that the line loading should be equal

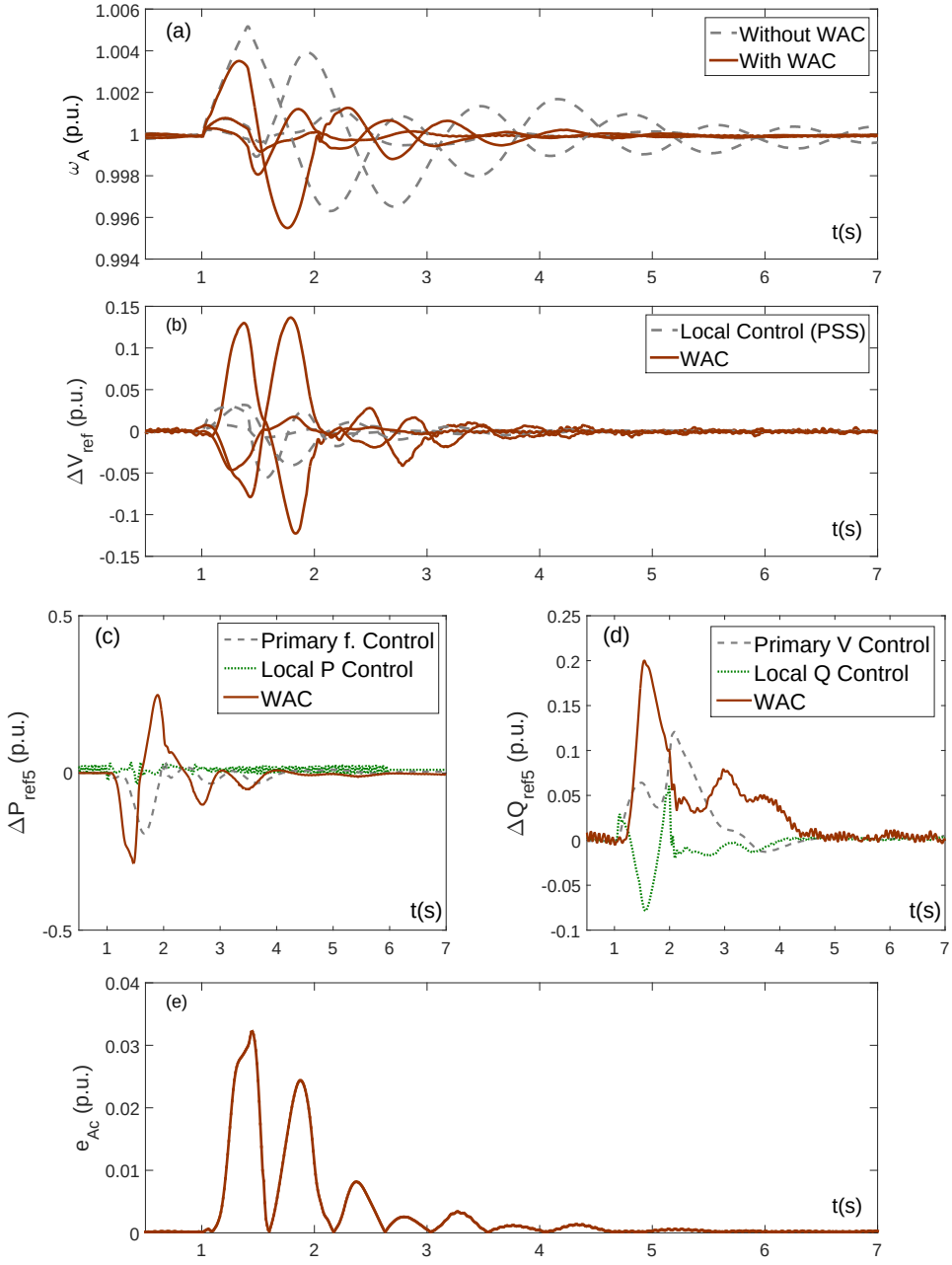


Figure 6.10: Case B. system, local control, and WAC response

to 161.2% to get the CCT of 385 ms (CCT without WAC and 135% loading). This shows extensive enhancement in active power margin.

6.5.3 Case C: Robustness

One of the main concerns in implementation of WACs is the WAM infrastructure and uncertainty in signals and parameters. The WAC depends on the timely mea-

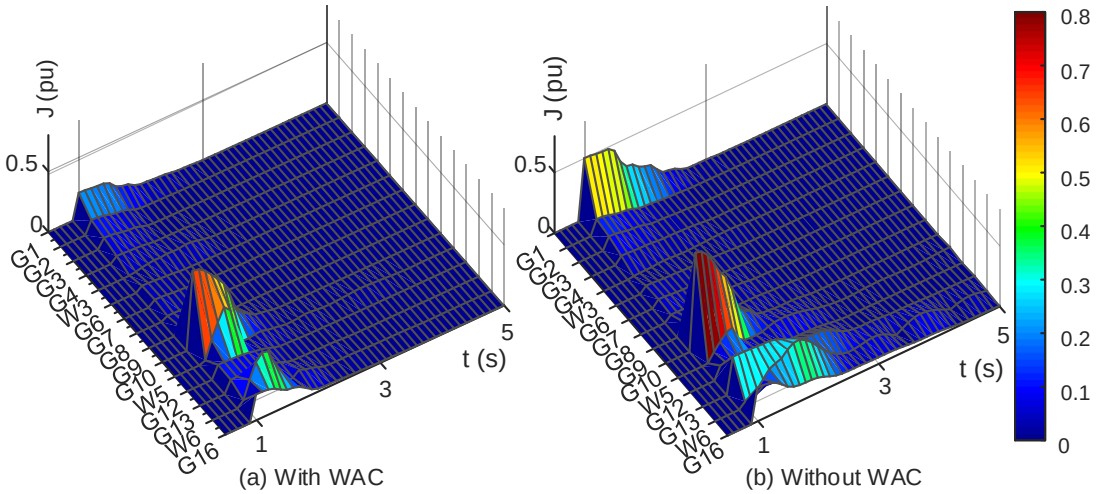


Figure 6.11: Case B. synchronous generators and wind farms energy function.

Table 6.4: Case B. comparison of transient stability (CCT)

CCT without WAC	CCT with WAC
385 ms	502 ms

surement and control signals to efficiently improve the system stability. However, the delay of the signals in the communication network can have an adverse effect on the performance of the WAC which can be minimized with efficient and robust WAM infrastructure. Usually, the delay of the PMU signals in a high-bandwidth communication system is expected to be small for WAC. In [70] the WAM infrastructure and various possible communication delays have been discussed.

In this case study the robustness of the proposed RL-based WAC to possible communication delays has been investigated. All the PMU signals are delayed by 100 ms and 200 ms, which is larger than the expected delay in the realistic system. Fig.6.12a shows the system dynamic response of a 350 ms three phase fault at tie-line 8-9 in different control scenarios. It can be seen that delays has deteriorated the WAC performance, however, the system oscillations still damps faster when compared to the system without WAC. Fig. 6.12b demonstrates the identifier NN performance as the error in the identification in these scenarios. These delays in reality is not expected to

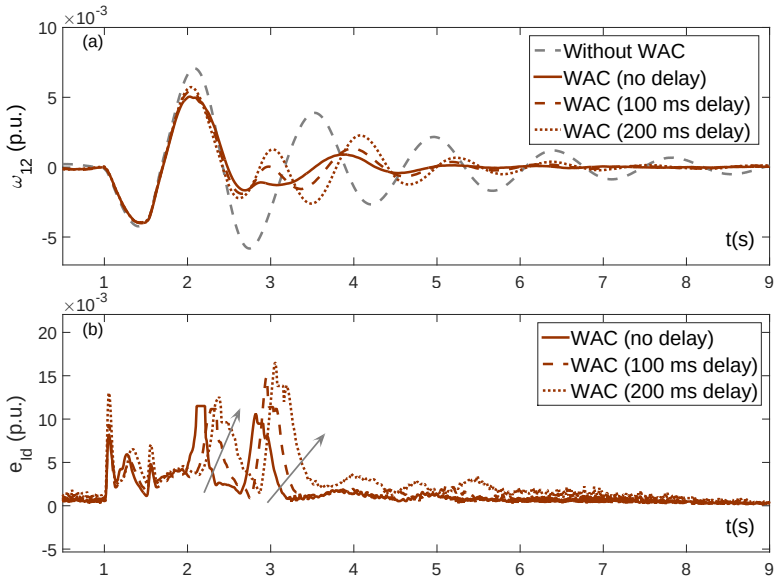


Figure 6.12: Case C. system response and identifier error-function of delay

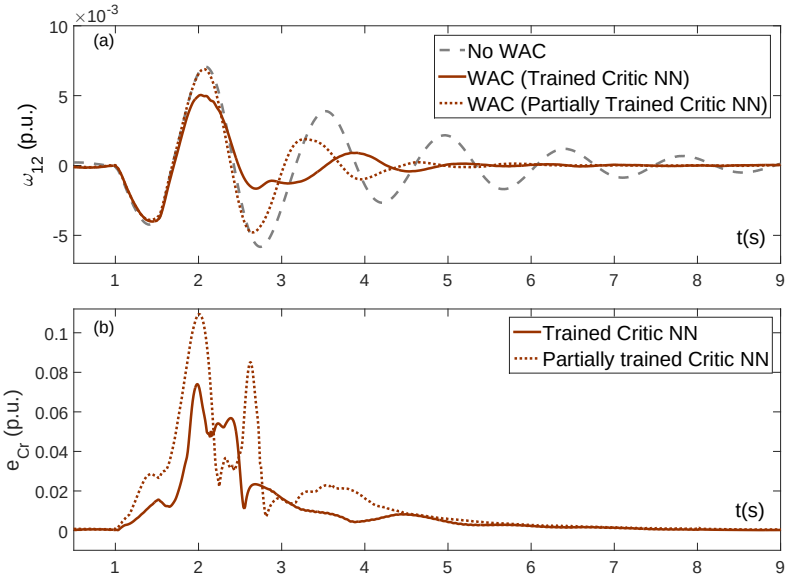


Figure 6.13: Case C. system response and critic error - function of weights

be larger than the simulated case; however, the performance of proposed WAC in the most extreme scenario is evaluated by increasing the delay time in PMU signals. The simulations showed similar results to the case without WAC, when delay in WAMs is 380 ms.

The parametric uncertainty is another challenge in WAC design. As a large scale dynamic power system controller, considerable number of parameters are involved in

the design. Knowing exact values of all these parameters in practice is not feasible. However, RL technique through the concept of multi-agent based modeling and employment of NN can estimate these parameters in online process, and overcome the necessity of needing exact values of the model. This technique has advantages over the supervised learning. For instance, as mentioned before, the critic NN is trained offline based on the energy function, which depends on several parameters which are projected in the weights. The weights of the critic NN have been multiplied by 0.5 to make a partially trained Critic NN. Fig. 6.13 demonstrates the performance of the WAC in these cases, as the RL can learn the energy function and yet provide satisfactory results in comparison to the system without the proposed WAC.

6.6 Conclusion

In this chapter an energy-based WAC is proposed on a wind integrated power grid. RL method has been employed as an optimal control for transient stability improvement, augmenting the excitation system of the synchronous generator and local active and reactive power control of DFIG. Transient stability has been quantified by energy elements derived for synchronous generators and wind farms. Simulations on the modified IEEE 68-bus system showed improved system responses with the proposed RL-based WAC. It has been observed that the WAC is able to damp the area oscillations faster, increase the CCT, and stability margins of tie-lines.

CHAPTER 7: CONCLUSIONS AND FUTURE WORKS

In this dissertation, novel methods for wide-area monitoring and control of large scale power system including renewable energy resources have been designed. A big step forward to improve power system wide area stability is wide area power system monitoring to communicate the specific-node information to a remote station providing higher observability and controllability. In the area of large-scale power systems, inter-area response are more effectively damped through the use of wide-area measurements. The complexity, nonlinearity, and uncertainty of power system in rotor angle stability problem led us to artificial Neural Networks (NNs) methods.

We first adapted NN based control design on WAC system. The challenges included optimality, robustness in presence of wide-area measurements uncertainty and delay, and generality. In order to address the first issue we used Reinforcement Learning (RL) algorithm to handle the NNs and perform optimal control theory. The WAC by means of this method could coordinate the actions of several agents, here generating units, toward a unified global cost function. Next, temporal difference was used in a finite moving average optimization horizon to improve the learning and prediction of NNs. The WAC actions is augmented to local excitation system of synchronous generators, which could raise several issues such as contradicting controls. In addtion, the NNs themselves have generality issue; i.e. performing unsatisfactory if the new operating point observed in real-time or online process may not cover the offline training domain. These problems led us to the "value priority" algorithm, in which the performance and identification of controllers were quantified based on Lyapunov function. The algorithm was performed for two sets of hybrid control design (local conventional PSS and global adaptive controller along with learning-based control

system). In both cases, the dynamic performance of the system has been considerably improved when the proposed WAC design has been used in terms of transient stability and oscillation damping.

In summary such a RL method has enabled us to construct a measurement-based WAC. However, the WAC design needed to be defined based on transient stability problem as the objective of this dissertation. Hence, we used energy functions of synchronous generator in the form of kinetic, potential, and field energies and linked them to the cost function of RL problem. We aimed at minimization of such energy function by means of WAC actions, which could further damp this energy. Finally this energy function was observed to be affected by other agents such as renewable energy resources as well. The renewable energy resources, in this thesis DFIG, could absorb or release some energy in the form of potential, reactive, and inertial energies and further damp the synchronous generators energy function. This transient stability enhancement led to faster convergence of the system toward the post-fault operating point which could be quantified by reduced critical clearing time or increased power transfer capability.

Practical Planning Aspects: As mentioned before, wind farms mainly affect the power system stability through change of system?s inertia and power flow through the tie-lines. For many cases the systems are not able to cope with these changes as the location of wind farms are mainly decided based on topological factors rather than power system operational and stability criteria. Considering these impacts and the considerable penetration of wind energy to the power system (16 GW in Ercot as for 2016), detailed studies is needed to see the impacts of such units on the power system stability. Each new wind project to be added to the current system should pass stability test as well as the system steady state operating test. The stability test includes Low Voltage Ride Through (LVRT) and transient stability criteria for several events such as line outage, generator outage, multiple faults, and so on.

The mitigation plans are needed to meet the standards if there are violations in case studies. These mitigations include reactive power compensator, line addition, or decreasing the wind dispatch. Such mitigation plans are further studied from economic view point. Usually, the wind farms are dispatched, at their max operating limits (MPPT), due to power generation costs. This brings up the wide-area control problem as more critical issue in such systems. The system is planned for the worst case scenario, which happens rarely. This issue along with the variability of the wind makes such static plans non cost effective. On the other hand, the dynamic WAC can mitigate such power flows in the case of fault eliminating the necessity of new facility installments. Such a WAC can support the system in such scenarios and, furthermore, it can provide more capacity for new wind power penetration installments.

The work proposed in this manuscript leads to a number of future research plans for developing appropriate analytical tools for in-depth information extraction from power system networks and, thereafter, to build control methods for ensuring system-wide stability and performance.

- Considering the recent improvement in terms of NNs learning algorithm the control design could be enhanced in terms of scalability. This thesis tackled this issue with slow coherency technique. This offline method is based on fixed number of areas or fixed electrical distance between them. Either way, the WAC is not able to monitor the states inside the area and perform optimization. On the other hand, if a dynamic or online coherency technique has been used the observability could be increased in any layer. This is achievable by modern NNs method such as unsupervised learning methods.

- Other than scalability, a big challenge in transient stability of large scale power system is the level of simplification and consideration of power system components. In general, when an energy-based methods are used like in this work, a trade off is performed for the system modeling and speed of calculation. How-

ever, as shown in this dissertation, some components such as DFIGs could considerably improve the performance of transient stability when included in WAC design. This consideration could be extended to other components of the power system, such as controllable loads, PV units in distribution system, and in general any power system component that can have flexible active and reactive powers.

REFERENCES

[1] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal. Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions. *Power Systems, IEEE Transactions on*, 19(3):1387–1401, 2004.

[2] Jeffrey Taft and Paul De Martini. Ultra Large Scale Power System Control Architecture. (October):1–25, 2012.

[3] R. Giovanini, K. Hopkinson, D.V. Coury, and J.S. Thorp. A primary and backup cooperative protection system based on wide area agents. *Power Delivery, IEEE Transactions on*, 21(3):1222–1230, 2006.

[4] I. Kamwa, R. Grondin, and Y. Hebert. Wide-area measurement based stabilizing control of large power systems-a decentralized/hierarchical approach. *Power Systems, IEEE Transactions on*, 16(1):136–153, 2001.

[5] S. Ray, G.K. Venayagamoorthy, B. Chaudhuri, and R. Majumder. Comparison of adaptive critic-based and classical wide-area controllers for power systems. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 38(4), 2008.

[6] A Ghosh, G Ledwich, OP Malik, and GS Hope. Power system stabilizer based on adaptive control techniques. *Power Apparatus and Systems, IEEE Transactions on*, (8):1983–1989, 1984.

[7] A.; Powell W.; Wunsch D. Si, J.; Barto. *Handbook of Learning and Approximate Dynamic Programming*. Wiley-IEEE Press, 2004.

[8] R.S. Sutton and A.G. Barto. Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks*, 9(5):1054–1054, September 1998.

[9] D. K. Chaturvedi and O. P. Malik. Generalized neuron-based adaptive pss for multimachine environment. *Power Systems, IEEE Transactions on*, 20(1):358–366, 2005.

[10] S. Mohagheghi, G.K. Venayagamoorthy, and R.G. Harley. Optimal wide area controller and state predictor for a power system. *Power Systems, IEEE Transactions on*, 22(2):693–705, 2007.

[11] S. Ray, G.K. Venayagamoorthy, and E.H. Watanabe. A computational approach to optimal damping controller design for a gcsc. *Power Delivery, IEEE Transactions on*, 23(3):1673–1681, 2008.

[12] S. Ray and G.K. Venayagamoorthy. Real-time implementation of a measurement-based adaptive wide-area control system considering communication delays. *Generation, Transmission Distribution, IET*, 2(1):62–70, 2008.

[13] D. Molina, Jiaqi Liang, R.G. Harley, and G.K. Venayagamoorthy. Virtual generators: Simplified online power system representations for wide-area damping control. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8, 2012.

[14] G. Venayagamoorthy and S. Jetti. A dual function neuron based external controller for a static var compensator. In *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, pages 1–1, 2008.

[15] R. Hadidi and B. Jeyasurya. Reinforcement learning based real-time wide-area stabilizing control agents to enhance power system stability. *Smart Grid, IEEE Transactions on*, 4(1):489–497, March 2013.

[16] Kenneth O. Stanley and Risto Miikkulainen. Efficient reinforcement learning through evolving neural network topologies. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2002)*, page 9, San Francisco, 2002. Morgan Kaufmann.

[17] D V Prokhorov and D C Wunsch. Adaptive critic designs. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 8(5):997–1007, January 1997.

[18] C.W. Taylor, Dennis C. Erickson, K.E. Martin, R.E. Wilson, and V. Venkatasubramanian. Wacs-wide-area stability and voltage control system: R d and online demonstration. *Proceedings of the IEEE*, 93(5):892–906, May 2005.

[19] S. Nabavi, Jianhua Zhang, and A. Chakrabortty. Distributed optimization algorithms for wide-area oscillation monitoring in power systems using interregional pmu-pdc architectures. *IEEE Trans. Smart Grid*, 6(5):2529–2538, Sept 2015.

[20] Hsiao-Dong Chiang, Felix F Wu, and Pravin P Varaiya. A bcu method for direct analysis of power system transient stability. *IEEE Transactions on Power Systems*, 9(3):1194–1208, 1994.

[21] Hsiao-Dong Chiang. *Direct methods for stability analysis of electric power systems: theoretical foundation, BCU methodologies, and applications*. John Wiley & Sons, 2011.

[22] T.L. Vu and K. Turitsyn. Lyapunov functions family approach to transient stability assessment. *IEEE Transactions on Power Systems*, PP(99):1–9, 2015.

[23] N AMJADI and M Ehsan. Transient stability assessment of power systems by a new estimating neural network. *CANADIAN JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING*, 1997.

[24] G. Garcia, J. Bernussou, and M. Berbiche. Pattern recognition applied to transient stability analysis of power systems with modelling including voltage and

speed regulation. *Electric Power Applications, IEE Proceedings B*, 139(4):321–335, Jul 1992.

[25] MA Pai. *Energy function analysis for power system stability*. Springer, 1989.

[26] G.A. Jones. Transient stability of a synchronous generator under conditions of bang-bang excitation scheduling. *Power Apparatus and Systems, IEEE Transactions on*, 84(2):114–121, Feb 1965.

[27] H.E. Brown, H.H. Happ, C.E. Person, and C.C. Young. Transient stability solution by an impedance matrix method. *Power Apparatus and Systems, IEEE Transactions on*, 84(12):1204–1214, Dec 1965.

[28] J. Geeganage, U.D. Annakkage, T. Weekes, and B.A. Archer. Application of energy-based power system features for dynamic security assessment. *IEEE Transactions on Power Systems*, 30(4):1957–1965, 2015.

[29] J. Hazra, R.K. Reddi, K. Das, D.P. Seetharam, and A.K. Sinha. Power grid transient stability prediction using wide area synchrophasor measurements. In *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, pages 1–8, Oct 2012.

[30] Daniel Ruiz-Vega and Mania Pavella. A comprehensive approach to transient stability control. i. near optimal preventive control. *IEEE Transactions on Power Systems*, 18(4):1446–1453, 2003.

[31] F. DorIer, M.R. Jovanovic, M. Chertkov, and F. Bullo. Sparsity-promoting optimal wide-area control of power networks. *IEEE Transactions on Power Systems*, 29(5):2281–2291, Sept 2014.

[32] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi, C. A. Canizares, R. Iravani, M. Kazerani, A. H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke, G. A. Jimenez-Estevez, and N. D. Hatziargyriou. Trends in microgrid control. *IEEE Trans. on Smart Grid*, 5(4):1905–1919, July 2014.

[33] S.D.J. McArthur, E.M. Davidson, V.M. Catterson, A.L. Dimeas, N.D. Hatziargyriou, F. Ponci, and T. Funabashi. Multi-agent systems for power engineering applications—part i: Concepts, approaches, and technical challenges. *Power Systems, IEEE Transactions on*, 22(4):1743–1752, 2007.

[34] Jian Xiao, Fushuan Wen, C.Y. Chung, and K.P. Wong. Wide-area protection and its applications - a bibliographical survey. In *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, pages 1388–1397, Oct 2006.

[35] M. Amin. Special issue on energy infrastructure defense systems. *Proceedings of the IEEE*, 93(5):855–860, May 2005.

[36] Qian Liu, V. Vittal, and N. Elia. Lpv supplementary damping controller design for a thyristor controlled series capacitor (tcsc) device. *Power Systems, IEEE Transactions on*, 21(3):1242–1249, Aug 2006.

[37] K. Tomsovic, D.E. Bakken, V. Venkatasubramanian, and A. Bose. Designing the next generation of real-time control, communication, and computations for large power systems. *Proceedings of the IEEE*, 93(5):965–979, May 2005.

[38] A. Chakrabortty. Wide-area damping control of power systems using dynamic clustering and tcsc-based redesigns. *Smart Grid, IEEE Transactions on*, 3(3):1503–1514, Sept 2012.

[39] Hui Ni, G.T. Heydt, and L. Mili. Power system stability agents using robust wide area control. *Power Systems, IEEE Transactions on*, 17(4):1123–1131, 2002.

[40] G.C. Zweigle and V. Venkatasubramanian. Wide-area optimal control of electric power systems with application to transient stability for higher order contingencies. *Power Systems, IEEE Transactions on*, 28(3):2313–2320, Aug 2013.

[41] J. Machowski, S. Robak, J.W. Bialek, J.R. Bumby, and N. Abi-Samra. Decentralized stability-enhancing control of synchronous generator. *Power Systems, IEEE Transactions on*, 15(4):1336–1344, Nov 2000.

[42] B.G. Horne. Progress in supervised neural networks. *Signal Processing Magazine, IEEE*, 10(1):8–39, Jan 1993.

[43] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

[44] Michael T Rosenstein and Andrew G Barto. Supervised Actor-Critic Reinforcement Learning. *In Practice*.

[45] Andrew G Barto and Thomas G Dietterich. Reinforcement Learning and its Relationship to Supervised Learning.

[46] B. Widrow, Narendra K. Gupta, and Sidhartha Maitra. Punish reward learning with a critic in adaptive threshold systems. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-3(5):455–465, 1973.

[47] Fei-yue Wang, Huaguang Zhang, and Derong Liu. Programming : An Introduction. (May):39–47, 2009.

[48] Thaddeus T Shannon and George G Lendaris. A new hybrid critic-training method for approximate dynamic programming. pages 1–11.

[49] Asma Al-Tamimi, Frank L Lewis, and Murad Abu-Khalaf. Discrete-time nonlinear hjb solution using approximate dynamic programming: convergence proof. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 38(4):943–949, 2008.

[50] A. Heydari and S. N. Balakrishnan. Global optimality of approximate dynamic programming and its use in non-convex function minimization. *Applied Soft Computing*, 24(1):291–303, 2014.

[51] Dejan J Sobajic, Yoh-han Pao, Critical Clearing Time, Machine Learning Systems, Adaptive Pattern Recognition, and Rumelhart Neural-net. ARTIFICIAL NEURAL-NET BASED DYNAMIC SECURITY. 4(1):220–228, 1989.

[52] Kumpati S Narendra and Kannan Parthasarathy. Identification and control of dynamical systems using neural networks. *Neural Networks, IEEE Transactions on*, 1(1):4–27, 1990.

[53] Wenxin Liu, Ganesh K Venayagamoorthy, and Donald C Wunsch. Design of an adaptive neural network based power system stabilizer. *Neural networks*, 16(5):891–898, 2003.

[54] DK Chaturvedi and OP Malik. Generalized neuron-based adaptive pss for multimachine environment. *Power Systems, IEEE Transactions on*, 20(1):358–366, 2005.

[55] Jian He and OP Malik. An adaptive power system stabilizer based on recurrent neural networks. *Energy Conversion, IEEE Transactions on*, 12(4):413–418, 1997.

[56] Ravi Segal, ML Kothari, and Shekhar Madnani. Radial basis function (rbf) network adaptive power system stabilizer. *Power Systems, IEEE Transactions on*, 15(2):722–727, 2000.

[57] M.G. Adamiak, A.P. Apostolov, M.M. Begovic, C.F. Henville, K.E. Martin, G.L. Michel, A.G. Phadke, and J.S. Thorp. Wide area protection technology and infrastructures. *Power Delivery, IEEE Transactions on*, 21(2):601–609, 2006.

[58] Yi Deng, Hua Lin, A.G. Phadke, S. Shukla, J.S. Thorp, and L. Mili. Communication network modeling and simulation for wide area measurement applications. In *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*, pages 1–6, 2012.

[59] Kun Zhu, M. Chenine, and L. Nordstrom. Ict architecture impact on wide area monitoring and control systems' reliability. *Power Delivery, IEEE Transactions on*, 26(4):2801–2808, 2011.

[60] M. Davoudi, J. Sadeh, and E. Kamyab. Parameter-free fault location for transmission lines based on optimisation. *IET Generation, Transmission Distribution*, 9(11):1061–1068, 2015.

[61] I. N. Moghaddam, Z. Salami, and S. Mohajeryami. Generator excitation systems sensitivity analysis and their model parameter's reduction. In *Power Systems Conference (PSC), 2014 Clemson University*, pages 1–6, March 2014.

[62] Hongxia Wu, K.S. Tsakalis, and G.T. Heydt. Evaluation of time delay effects to wide-area power system stabilizer design. *Power Systems, IEEE Transactions on*, 19(4):1935–1941, 2004.

[63] S. Ray and G.K. Venayagamoorthy. Wide area signal based optimal neurocontroller for a upfc. In *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, pages 1–1, 2008.

[64] Salatas John. Implementation of elman recurrent neural network in weka.

[65] B. Luitel, G.K. Venayagamoorthy, and C.E. Johnson. Enhanced wide area monitoring system. In *Innovative Smart Grid Technologies (ISGT), 2010*, pages 1–7, 2010.

[66] K. Balasubramaniam, B. Luitel, and G.K. Venayagamoorthy. A scalable wide area monitoring system using cellular neural networks. In *Neural Networks (IJCNN), The 2012 International Joint Conference on*, pages 1–8, 2012.

[67] I. Kamwa, S.R. Samantaray, and G. Joos. Development of rule-based classifiers for rapid stability assessment of wide-area post-disturbance records. *Power Systems, IEEE Transactions on*, 24(1):258–270, Feb 2009.

[68] R. Yousefian and S. Kamalasadan. Design and real-time implementation of optimal power system wide-area system-centric controller based on temporal difference learning. *IEEE Transactions on Industry Applications*, 52(1):395–406, Jan 2016.

[69] J.W. Stahlhut, T.J. Browne, G.T. Heydt, and V. Vittal. Latency viewed as a stochastic process and its impact on wide area power system control signals. *Power Systems, IEEE Transactions on*, 23(1):84–91, 2008.

[70] R. Yousefian and S. KAMALASADAN. Energy function inspired value priority based global wide-area control of power grid. *IEEE Transactions on Smart Grid*, PP(99):1–1, 2016.

[71] G K Venayagamoorthy, R G Harley, and D C Wunsch. Implementation of adaptive critic-based neurocontrollers for turbogenerators in a multimachine power system. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 14(5):1047–64, January 2003.

[72] Sina Zarrabian, Rabie Belkacemi, and Adeniyi A Babalola. Reinforcement learning approach for congestion management and cascading failure prevention with experimental application. *Electr. Pow. Syst. Res.*, 141:179–190, 2016.

[73] D. Gautam, V. Vittal, and T. Harbour. Impact of increased penetration of dfig-based wind turbine generators on transient and small signal stability of power systems. *Power Systems, IEEE Transactions on*, 24(3):1426–1434, Aug 2009.

[74] S. Eftekharnejad, V. Vittal, G.T. Heydt, B. Keel, and J. Loehr. Impact of increased penetration of photovoltaic generation on power systems. *Power Systems, IEEE Transactions on*, 28(2):893–901, May 2013.

[75] A. Mendonca and J.A. Peas Lopes. Impact of large scale wind power integration on small signal stability. In *Future Power Systems, 2005 International Conference on*, pages 5 pp.–5, Nov 2005.

[76] Wenjuan Du, Haifeng Wang, and L-Y. Xiao. Power system small-signal stability as affected by grid-connected photovoltaic generation. *European Transactions on Electrical Power*, 22(5):688–703, 2012.

[77] JG Slootweg and WL Kling. The impact of large scale wind power generation on power system oscillations. *Electric Power Systems Research*, 67(1):9–20, 2003.

[78] Eknath Vittal, Mark O'Malley, and Andrew Keane. Rotor angle stability with high penetrations of wind generation. *Power Systems, IEEE Transactions on*, 27(1):353–362, 2012.

[79] Tao Sun, Zhe Chen, and Frede Blaabjerg. Transient stability of dfig wind turbines at an external short-circuit fault. *Wind energy*, 8(3):345–360, 2005.

[80] Yufei Tang, Haibo He, Jinyu Wen, and Ju Liu. Power system stability control for a wind farm based on adaptive dynamic programming. *Smart Grid, IEEE Transactions on*, 6(1):166–177, Jan 2015.

[81] J. Rocabert, A. Luna, F. Blaabjerg, and P. Rodriguez. Control of power converters in ac microgrids. *Power Electronics, IEEE Transactions on*, 27(11):4734–4749, Nov 2012.

[82] J.C. Vasquez, J.M. Guerrero, J. Miret, M. Castilla, and L.G. de Vicuna. Hierarchical control of intelligent microgrids. *Industrial Electronics Magazine, IEEE*, 4(4):23–29, Dec 2010.

[83] Trishan Esram, Patrick L Chapman, et al. Comparison of photovoltaic array maximum power point tracking techniques. *IEEE Transactions on Energy Conversion EC*, 22(2):439, 2007.

[84] E. Vittal and A. Keane. Rotor angle stability with high penetrations of wind generation. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–1, July 2012.

[85] Shuhui Li, Donald C Wunsch, Edgar A O'Hair, and Michael G Giesselmann. Using neural networks to estimate wind turbine power generation. *Energy conversion, ieee transactions on*, 16(3):276–282, 2001.

[86] Soteris A Kalogirou. Artificial neural networks in renewable energy systems applications: a review. *Renewable and sustainable energy reviews*, 5(4):373–401, 2001.

[87] H.M. Jabr, Dongyun Lu, and N.C. Kar. Design and implementation of neuro-fuzzy vector control for wind-driven doubly-fed induction generator. *Sustainable Energy, IEEE Transactions on*, 2(4):404–413, Oct 2011.

[88] V. Galdi, A. Piccolo, and P. Siano. Designing an adaptive fuzzy controller for maximum wind energy extraction. *Energy Conversion, IEEE Transactions on*, 23(2):559–569, June 2008.

[89] Orlando Soares, Henrique Goncalves, Antonio Martins, and Adriano Carvalho. Nonlinear control of the doubly-fed induction generator in wind power systems. *Renewable Energy*, 35(8):1662–1670, 2010.

[90] Shuhui Li, M. Fairbank, C. Johnson, D.C. Wunsch, E. Alonso, and J.L. Proao. Artificial neural networks for control of a grid-connected rectifier/inverter under disturbance, dynamic and power converter switching conditions. *Neural Networks and Learning Systems, IEEE Transactions on*, 25(4):738–750, April 2014.

[91] J. Han, S. Khushalani-Solanki, J. Solanki, and J. Liang. Adaptive critic design-based dynamic stochastic optimal control design for a microgrid with multiple renewable resources. *Smart Grid, IEEE Transactions on*, PP(99):1–1, 2015.

[92] Wei Qiao, G.K. Venayagamoorthy, and R.G. Harley. Dhp-based wide-area coordinating control of a power system with a large wind farm and multiple facts devices. In *Neural Networks, 2007. IJCNN 2007. International Joint Conference on*, pages 2093–2098, Aug 2007.

[93] P. Zhang, D.Y. Yang, K.W. Chan, and G.W. Cai. Adaptive wide-area damping control scheme with stochastic subspace identification and signal time delay compensation. *Generation, Transmission Distribution, IET*, 6(9):844–852, 2012.

[94] R. Yousefian and S. Kamalasadan. System-centric control architecture for wide area monitoring and control of power system. In *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, pages 1–7, Feb 2013.

[95] G.K. Venayagamoorthy and S. Ray. A neural network based optimal wide area control scheme for a power system. In *Industry Applications Conference, 2005. Fourtieth IAS Annual Meeting. Conference Record of the 2005*, volume 1, pages 700–706 Vol. 1, 2005.

[96] Richard S. Sutton, Hamid Reza Maei, Doina Precup, Shalabh Bhatnagar, David Silver, Csaba SzepesvÃ¡ri, Eric Wiewiora, and Maei Recently Introduced. Fast gradient-descent methods for temporal-difference learning with linear function approximation. In *In Danyluk et*, pages 993–1000, 2009.

[97] J.N. Tsitsiklis and B. Van Roy. An analysis of temporal-difference learning with function approximation. *Automatic Control, IEEE Transactions on*, 42(5):674–690, May 1997.

[98] Roland E. Suri and Wolfram Schultz. Temporal difference model reproduces anticipatory neural activity, 2000.

[99] Richard S. Sutton. Learning to predict by the methods of temporal differences. In *MACHINE LEARNING*, pages 9–44. Kluwer Academic Publishers, 1988.

[100] A. G. Barto, R. S. Sutton, and C. J.C.H. Watkins. Learning and sequential decision making. Technical report, Amherst, MA, USA, 1989.

[101] Michael J. Kearns and Satinder P. Singh. Bias-variance error bounds for temporal difference updates. In *Proceedings of the Thirteenth Annual Conference on Computational Learning Theory*, COLT '00, pages 142–147, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.

[102] H. van Seijen and R. Sutton. True online td(λ). In *Proceedings of the 31st International Conference on Machine Learning*, page 9, Beijing, China, 2014.

[103] S. Teleke, M.E. Baran, S. Bhattacharya, and A.Q. Huang. Optimal control of battery energy storage for wind farm dispatching. *Energy Conversion, IEEE Transactions on*, 25(3):787–794, 2010.

[104] P.M. Anderson, A.A.A. Fouad, Institute of Electrical, and Electronics Engineers. *Power system control and stability*. IEEE Press power engineering series.

[105] S. Nabavi and A Chakrabortty. Topology identification for dynamic equivalent models of large power system networks. In *American Control Conference (ACC), 2013*, pages 1138–1143, June 2013.

[106] IEEE 39 bus system benchmark model.

[107] J. Chow and P. Kokotovic. Time scale modeling of sparse dynamic networks. *IEEE Trans. Autom. Control*, 30(8):714–722, Aug 1985.

[108] Innocent Kamwa, Robert Grondin, and Yves Hébert. Wide-area measurement based stabilizing control of large power systems-a decentralized/hierarchical approach. 16(1):136–153, 2001.

[109] Junbo Zhang, C.Y. Chung, and Yingduo Han. A novel modal decomposition control and its application to pss design for damping interarea oscillations in power systems. *Power Systems, IEEE Transactions on*, 27(4):2015–2025, Nov 2012.

[110] S. Kamalasadan, G.D. Swann, and R. Yousefian. A novel system-centric intelligent adaptive control architecture for power system stabilizer based on adaptive neural networks, 2013.

[111] J.W. Chapman, M.D. Ilic, C. A King, L. Eng, and H. Kaufman. Stabilizing a multimachine power system via decentralized feedback linearizing excitation control. *Power Systems, IEEE Transactions on*, 8(3):830–839, Aug 1993.

[112] Qiang Lu and Yuan-Zhang Sun. Nonlinear stabilizing control of multimachine systems. *Power Systems, IEEE Transactions on*, 4(1):236–241, Feb 1989.

[113] Mohamed Bahita and Khaled Belarbi. Neural stable adaptive control for a class of nonlinear systems without use of a supervisory term in the control law. *Journal of Engineering Science and Technology*, 7(1):97–118, 2012.

[114] Robert A Jacobs, Michael I Jordan, Steven J Nowlan, and Geoffrey E Hinton. Adaptive mixtures of local experts. *Neural computation*, 3(1):79–87, 1991.

[115] M. Chenine and L. Nordstrom. Modeling and simulation of wide-area communication for centralized pmu-based applications. *IEEE Trans. Power Del.*, 26(3):1372–1380, July 2011.

[116] Wei Yao, Lin Jiang, Jinyu Wen, Qinghua Wu, and Shijie Cheng. Wide-area damping controller for power system interarea oscillations: A networked predictive control approach. *IEEE J CST*, 23(1):27–36, 2015.

[117] Hypersim.

[118] N. Kakimoto, Y. Ohsawa, and M. Hayashi. Transient stability analysis of multimachine power system with field flux decays via lyapunov's direct method. *Power Apparatus and Systems, IEEE Transactions on*, PAS-99(5):1819–1827, 1980.

[119] Manuel A Duarte and Kumpati S Narendra. Combined direct and indirect approach to adaptive control. *Automatic Control, IEEE Transactions on*, 34(10):1071–1075, 1989.

[120] Saeid Khosravani, Iman Naziri Moghaddam, Ahmad Afshar, and Mahdi Karrari. Wide-area measurement-based fault tolerant control of power system during sensor failure. *Electr. Pow. Syst. Res.*, 137:66–75, 2016.

[121] Atena Darvishi and Ian Dobson. Threshold-based monitoring of multiple outages with pmu measurements of area angle. *IEEE Trans. on Power Syst.*, 31(3):2116–2124, 2016.

[122] Christian Dufour and Jean Bélanger. On the use of real-time simulation technology in smart grid research and development. *IEEE Trans. Ind. Appl*, 50(6):3963–3970, 2014.

[123] M. La Scala, G. Lorusso, R. Sbrizzai, and M. Trovato. A qualitative approach to the transient stability analysis [of power systems]. *Power Systems, IEEE Transactions on*, 11(4):1996–2002, Nov 1996.

[124] I. Kamwa, R. Grondin, and L. Loud. Time-varying contingency screening for dynamic security assessment using intelligent-systems techniques. *Power Systems, IEEE Transactions on*, 16(3):526–536, Aug 2001.

[125] N. Amjady and S.F. Majedi. Transient stability prediction by a hybrid intelligent system. *Power Systems, IEEE Transactions on*, 22(3):1275–1283, Aug 2007.

[126] Seung-Mook Baek, Jung-Wook Park, and Ganesh Kumar Venayagamoorthy. Power system control with an embedded neural network in hybrid system modeling. *IEEE Trans. Ind. Appl*, 44(5):1458–1465, 2008.

[127] D.P. Bertsekas and J.N. Tsitsiklis. Neuro-dynamic programming: an overview. In *Proc. of the 34th IEEE Conf. on Decision and Control*, volume 1, pages 560–564 vol.1, Dec 1995.

[128] Draguna Vrabie, O Pastravanu, Murad Abu-Khalaf, and Frank L Lewis. Adaptive optimal control for continuous-time linear systems based on policy iteration. *Automatica*, 45(2):477–484, 2009.

[129] R. Yousefian, A. Sahami, and S. Kamalasadan. Hybrid energy function based real-time optimal wide-area transient stability controller for power system stability. In *Industry Applications Society Annual Meeting, 2015 IEEE*, pages 1–8, Oct 2015.

[130] M Larbi Ourari, Louis-A Dessaint, and Van-Que Do. Dynamic equivalent modeling of large power systems using structure preservation technique. *IEEE Transactions on Power Systems*, 21(3):1284–1295, 2006.

[131] Michael J Basler and Richard C Schaefer. Understanding power-system stability. *IEEE Trans. Ind. Appl*, 44(2):463–474, 2008.

[132] Simon Fabri and Visakan Kadirkamanathan. *Functional adaptive control: an intelligent systems approach*. Springer Science & Business Media, 2012.

[133] D. Ruiz-Vega, A.R. Messina, and M. Pavella. Online assessment and control of transient oscillations damping. *IEEE Transactions on Power Systems*, 19(2):1038–1047, May 2004.

[134] IEEE 68 bus system benchmark model.

[135] R Yousefian and S Kamalasadan. A lyapunov function based optimal hybrid power system controller for improved transient stability. *Electric Power Systems Research*, 137:6–15, 2016.

[136] Mohsen Rahimi and Mostafa Parniani. Transient performance improvement of wind turbines with doubly fed induction generators using nonlinear control strategy. *IEEE J EC*, 25(2):514–525, 2010.

[137] Durga Gautam et. al. Control strategy to mitigate the impact of reduced inertia due to doubly fed induction generators on large power systems. *Power Systems, IEEE Transactions on*, 26(1):214–224, 2011.

[138] Zhixin Miao, Lingling Fan, Dale Osborn, and Subbaraya Yuvarajan. Control of dfig-based wind generation to improve interarea oscillation damping. 24(2):415–422, 2009.

[139] RD Fernandez, PE Battaiotto, and RJ Mantz. Wind farm non-linear control for damping electromechanical oscillations of power systems. *Renewable Energy*, 33(10):2258–2265, 2008.

[140] Fahd Hashiesh et. al. An intelligent wide area synchrophasor based system for predicting and mitigating transient instabilities. *IEEE Trans. Smart Grid*, 3(2):645–652, 2012.

[141] G. K. Venayagamoorthy, R. K. Sharma, P. K. Gautam, and A. Ahmadi. Dynamic energy management system for a smart microgrid. *IEEE Trans. Neural Netw. Learn. Syst.*, 27(8):1643–1656, 2016.

[142] T. L. Vu and K. Turitsyn. A framework for robust assessment of power grid stability and resiliency. *IEEE J AC*, PP(99):1–1, 2016.

[143] A Michel, A Fouad, and Vijay Vittal. Power system transient stability using individual machine energy functions. *IEEE Transactions on Circuits and Systems*, 30(5):266–276, 1983.

[144] F.M. Hughes, O. Anaya-Lara, N. Jenkins, and G. Strbac. Control of dfig-based wind generation for power network support. *Power Systems, IEEE Transactions on*, 20(4):1958–1966, Nov 2005.

[145] R. Yousefian, R. Bhattarai, and S. Kamalasadan. Direct intelligent wide-area damping controller for wind integrated power system. In *2016 IEEE Power Energy Society General Meeting*, July 2016.

[146] Graham Rogers. *Power system oscillations*. Springer Science & Business Media, 2012.

[147] Vijay Vittal and Raja Ayyanar. *Dynamic Models for Wind Generators*, pages 99–113. Springer New York, New York, NY, 2013.

[148] Mohit Singh and Surya Santoso. *Dynamic models for wind turbines and wind power plants*. National Renewable Energy Laboratory, 2011.

[149] Rui Zhang, Yan Xu, Zhao Yang Dong, and Kit Po Wong. Post-disturbance transient stability assessment of power systems by a self-adaptive intelligent system. *Generation, Transmission Distribution, IET*, 9(3):296–305, 2015.

[150] Louis Wehenkel, Daniel Ruiz-Vega, Damien Ernst, and Mania Pavella. Preventive and emergency control of power systems. *Real Time Stability in Power Systems*, pages 199–232, 2005.

[151] Chih-Wen Liu, Mu-Chun Su, Shuenn-Shing Tsay, and Yi-Jen Wang. Application of a novel fuzzy neural network to real-time transient stability swings prediction based on synchronized phasor measurements. *Power Systems, IEEE Transactions on*, 14(2):685–692, May 1999.

[152] S. Rovnyak, S. Kretsinger, J. Thorp, and D. Brown. Decision trees for real-time transient stability prediction. *Power Systems, IEEE Transactions on*, 9(3):1417–1426, Aug 1994.

[153] A. Rajapakse, F. Gomez, K. Nanayakkara, P. Crossley, and V. Terzija. Rotor angle instability prediction using post-disturbance voltage trajectories. In *Power and Energy Society General Meeting, 2010 IEEE*, pages 1–1, July 2010.

[154] I. Kamwa, S.R. Samantaray, and G. Joos. Catastrophe predictors from ensemble decision-tree learning of wide-area severity indices. *Smart Grid, IEEE Transactions on*, 1(2):144–158, Sept 2010.

[155] Ž Eleschová, M Smitková, and A Belán. Evaluation of Power System Transient Stability and Definition of the Basic Criterion. 4(1):9–16, 2010.

[156] Y. Guo, D.J. Hill, and Youyi Wang. Global transient stability and voltage regulation for power systems. *Power Systems, IEEE Transactions on*, 16(4):678–688, 2001.

[157] D. Aia and G.T. Heydt. On-line transient stability evaluation by system decomposition-aggregation and high order derivatives. *Power Apparatus and Systems, IEEE Transactions on*, PAS-102(7):2038–2046, 1983.

[158] K. Meng, Z.Y. Dong, K.P. Wong, Y. Xu, and F.J. Luo. Speed-up the computing efficiency of power system simulator for engineering-based power system transient stability simulations. *Generation, Transmission Distribution, IET*, 4(5):652–661, May 2010.

[159] Xiaochen Wu, Jinquan Zhao, Aidong Xu, Hui Deng, and Peng Xu. Review on transient stability prediction methods based on real time wide-area phasor measurements. In *Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2011 4th International Conference on*, pages 320–326, July 2011.

[160] Chih-Wen Liu and J.S. Thorp. New methods for computing power system dynamic response for real-time transient stability prediction. *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, 47(3):324–337, Mar 2000.

[161] M.H. Haque and A.H.M.A. Rahim. Determination of first swing stability limit of multimachine power systems through taylor series expansions. *Generation, Transmission and Distribution, IEE Proceedings C*, 136(6):373–380, Nov 1989.

[162] F.F. Song, T.S. Bi, H.Y. Li, Q.X. Yang, and Y. Liu. The perturbed trajectories prediction method based on wide-area measurement system. In *Transmission and Distribution Conference and Exhibition, 2005/2006 IEEE PES*, pages 1467–1471, May 2006.

[163] Qiang Guo, Xiao-peng Liu, Shi-rong Lü, and Dao-zhi XIA. Application of gps synchronized clock to power system transient stability predict and control. *Automation of Electric Power Systems*, 22(6):11–13, 1998.

[164] Zhaoyan Liu, Quanyuan Jiang, and Yijia Cao. Fast learning algorithm for transient stability prediction based on wide-area measurement system. *Automation of Electric Power Systems*, 31(21):1–4, 2007.

[165] Mingsong Liu, Hongbin Sun, Boming Zhang, Liangzhong Yao, Min Han, and Wenchuan Wu. Pmu measurements and ems models based transient stability on-line forecasting. In *Power Energy Society General Meeting, 2009. PES '09. IEEE*, pages 1–8, July 2009.

[166] K. Yamashita and H. Kameda. Out-of-step prediction logic for wide-area protection based on an autoregressive model. In *Power Systems Conference and Exposition, 2004. IEEE PES*, pages 307–312 vol.1, Oct 2004.

[167] Hussain Hassan and Al Marhoon. A Practical Method for Power Systems Transient Stability and Security. 2011.

[168] M. Sherwood, Dongchen Hu, and V.M. Venkatasubramanian. Real-time detection of angle instability using synchrophasors and action principle. In *Bulk Power System Dynamics and Control - VII. Revitalizing Operational Reliability, 2007 iREP Symposium*, pages 1–11, Aug 2007.

[169] C. Jing, J.D. McCalley, and M. Kommareddy. An energy approach to analysis of interarea oscillations in power systems. *Power Systems, IEEE Transactions on*, 11(2):734–740, May 1996.

[170] Lei Chen, Yong Min, and Wei Hu. An energy-based method for location of power system oscillation source. *Power Systems, IEEE Transactions on*, 28(2):828–836, May 2013.

[171] D. Ruiz-Vega and M. Pavella. A comprehensive approach to transient stability control. ii. open loop emergency control. *Power Systems, IEEE Transactions on*, 18(4):1454–1460, Nov 2003.

[172] D. Ernst, D. Ruiz-Vega, M. Pavella, P.M. Hirsch, and D. Sobajic. A unified approach to transient stability contingency filtering, ranking and assessment. *Power Systems, IEEE Transactions on*, 16(3):435–443, Aug 2001.

[173] GC Ejebe, C Jing, JG Waight, V Vittal, G Pieper, F Jamshidian, D Sobajic, and P Hirsch. On-line dynamic security assessment: Transient energy based

screening and monitoring for stability limits. In *1997 IEEE Summer Meeting, Panel Session on "Techniques for Stability Limit Search*, 1997.

[174] Rui Zhang, Yan Xu, Zhao-Yang Dong, Ke Meng, and Zhao Xu. Intelligent systems for power system dynamic security assessment: Review and classification. In *Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2011 4th International Conference on*, pages 134–139, July 2011.

[175] T.E. Dy-Liacco. Enhancing power system security control. *Computer Applications in Power, IEEE*, 10(3):38–41, Jul 1997.

[176] A.D. Rajapakse, F. Gomez, K. Nanayakkara, P.A. Crossley, and V.V. Terzija. Rotor angle instability prediction using post-disturbance voltage trajectories. *Power Systems, IEEE Transactions on*, 25(2):947–956, May 2010.

[177] Xu Tao, He Renmu, Wang Peng, and Xu Dongjie. Applications of data mining technique for power system transient stability prediction. In *Electric Utility Deregulation, Restructuring and Power Technologies, 2004. (DRPT 2004). Proceedings of the 2004 IEEE International Conference on*, volume 1, pages 389–392 Vol.1, April 2004.

[178] Fahd Hashiesh, Hossam E. Mostafa, A.-R. Khatib, I. Helal, and Mohamed M. Mansour. An intelligent wide area synchrophasor based system for predicting and mitigating transient instabilities. *Smart Grid, IEEE Transactions on*, 3(2):645–652, June 2012.

[179] J.C. Cepeda, J.L. Rueda, D.G. Colome, and D.E. Echeverria. Real-time transient stability assessment based on centre-of-inertia estimation from phasor measurement unit records. *Generation, Transmission Distribution, IET*, 8(8):1363–1376, August 2014.

[180] E.M. Voumvoulakis and N.D. Hatziargyriou. A particle swarm optimization method for power system dynamic security control. *Power Systems, IEEE Transactions on*, 25(2):1032–1041, May 2010.

[181] M. Yamamoto and O. Motoyoshi. Active and reactive power control for doubly-fed wound rotor induction generator. *Power Electronics, IEEE Transactions on*, 6(4):624–629, Oct 1991.

[182] G. Tsourakis and C. Vournas. A controller for wind generators to increase damping of power oscillations. In *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pages 2195–2198, May 2010.

[183] Zhixin Miao, Lingling Fan, D. Osborn, and S. Yuvarajan. Control of dfig based wind generation to improve inter-area oscillation damping. In *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, pages 1–7, July 2008.

[184] José Luis Domínguez-García, Oriol Gomis-Bellmunt, Fernando D Bianchi, and Andreas Sumper. Power oscillation damping supported by wind power: a review. *Renewable and Sustainable Energy Reviews*, 16(7):4994–5006, 2012.

[185] Huazhang Huang and C.Y. Chung. Coordinated damping control design for dfig-based wind generation considering power output variation. *Power Systems, IEEE Transactions on*, 27(4):1916–1925, Nov 2012.

[186] F. Wu, X.-P. Zhang, K. Godfrey, and P. Ju. Small signal stability analysis and optimal control of a wind turbine with doubly fed induction generator. *Generation, Transmission Distribution, IET*, 1(5):751–760, September 2007.

[187] Y. Mishra, S. Mishra, Fangxing Li, Zhao Yang Dong, and R.C. Bansal. Small-signal stability analysis of a dfig-based wind power system under different modes of operation. *Energy Conversion, IEEE Transactions on*, 24(4):972–982, Dec 2009.

[188] N. Kshatriya, U.D. Annakkage, F.M. Hughes, and A.M. Gole. Optimized partial eigenstructure assignment-based design of a combined pss and active damping controller for a dfig. *Power Systems, IEEE Transactions on*, 25(2):866–876, May 2010.

[189] R. Shah, N. Mithulananthan, and K.Y. Lee. Large-scale pv plant with a robust controller considering power oscillation damping. *Energy Conversion, IEEE Transactions on*, 28(1):106–116, March 2013.

[190] Rakibuzzaman Shah, Nadarajah Mithulananthan, Arthit Sode-Yome, and Kwang Y Lee. Impact of large-scale pv penetration on power system oscillatory stability. In *Power and Energy Society General Meeting, 2010 IEEE*, pages 1–7. IEEE, 2010.

[191] Rakibuzzaman Shah, N Mithulananthan, and RC Bansal. Oscillatory stability analysis with high penetrations of large-scale photovoltaic generation. *Energy Conversion and Management*, 65:420–429, 2013.

APPENDIX A: TRANSIENT STABILITY ASSESSMENT AND CONTROL

Synchronous generators are in essential nonlinear, non-stationary, fast acting, Multi-Input-Multi-Output (MIMO) devices, especially, in interconnected and global aspect [7]. Fully reliable transient stability analysis and control, considering these devises is, still, a challenging problem. In general, transient stability problem, with respect to time of the calculation, outputs, and speed can be divided into four main categories:

A.1 Transient Stability Assessment Problem

Transient Stability Assessment (TSA) is performed offline at pre-fault stage before the disturbance happen actually. The literature in this category is respectfully rich, with various methods employed, [23–27]. TSA focuses on the critical clearing time of the faults in power system. A comprehensive simulation of faults provides useful information regarding the vulnerable points of the system and makes sure of safe stability margin. Besides, these evaluations can provide helpful database for real-time analysis.

A.1.1 Prediction problem and Preventive Control

Conventionally, transient instability events are prevented by means of event based actions. Set of unstable contingencies and respective actions are designed offline. Once, one of the known contingencies is monitored, the predefined control action will be triggered. Although being relatively fast, their implementation is not comprehensive. Additionally, the problem of transient stability prediction based on response has been proposed, which is simpler and more accurate than the traditional event-based methods [149]. In this method the progress of the power system transient and oscillations are monitored, and prediction of stability of swings is carried on for classification. Nominal output of the controller represents a discrete status. This prediction results in preventive control actions such as generation rescheduling, reactive compensation,

or load shedding [150]. Various methods, such as numerical routines, pattern recognition, direct methods, and intelligent methods have been widely proposed and used in this area, which will be further discussed [125, 151, 152].

A.1.2 Emergency Control

Emergency or remedial control is used in the post-disturbance conditions. Several utilities have developed protection schemes known as Special Protection Systems (SPSs) or Remedial Action Schemes (RASs) [153]. This control scheme predicts the system stability status during the disturbance, and the assessment results are used to trigger emergency control actions such as generator tripping or load shedding to terminate the propagation of the instability. Usually, practical remedial control are event-based, and response-based designs are not yet adequately fast enough [154]. These systems directly detect selected faults that would lead to instability using binary signals and take predefined corrective control actions. The main disadvantage is the practicality issue in number of outages and initial states investigated [40].

A.1.3 Mitigating and damping Control

It is known that transient instability may happen in further swings due to lack of damping. The control actions can also be used as the output, to make the problem dynamic mitigating and damping control. This type of control schemes are designed not only to provide a stable final state but also minimize state excursions along the trajectory and increase the power system stability margins [40].

A.2 Transient Stability Assessment Methods

Various approaches to real time transient stability assessment and control have been proposed in the power systems literature. Following detection of a fault, real time TSA will run and prediction of stability of the system is performed. Numerical Integration and direct methods were utilized in this category in early years, each having their own advantages and disadvantages. Further, with the advent of technology and complexity

of the real-time wide area transient stability problem new techniques and approaches based on combination of these methods and integration of intelligent methods have evolved.

A.2.1 Numerical Integration

In this method, numerical integration algorithms are utilized to solve the system dynamics represented by huge number of differential equations. The most common methods in numerical integration are Euler method, Runge-Kutta method, and Implicit Integration method [155]. Given the complexity of power system models in wide-area domain, the most effective analysis tool for predicting the states trajectory and transient stability is full time domain simulations [156]. However, with real-time operation limitations, in such simulations each fault must be treated separately, and stability margins must be calculated by separate trials. In addition, time-step intervals are needed to be effectively managed to ensure accuracy and numerical stability [156]. Considering all these issues, the overall process is considered very time consuming. Various methods have been proposed to speed up the integration time, such as decomposition-aggregation integration [157], or parallel computing architecture [158]. One of the main break through in this area has been prediction of trajectories through observed states. The real-time wide area state trajectory prediction are performed via three methods [159]:

- Network dynamic equivalence [160]
- Curve fitting based trajectory extrapolation
  - Polynomial model based prediction method [161]
  - Trigonometric function model based prediction method [162]
  - Auto regression prediction method [163]
- The angular velocity prediction and integration method [164]

In reduced order methods, following fault detection, the algorithm starts and a post-fault reduced order system model is constructed with initial variables calculated. Then, the time domain simulation can be performed on the model in real-time speed, and the transient stability of the power system can be predicted online [165]. However, such methods show considerable drawbacks in real-time computation in the case of wide area use, where, it requires accurate information of the power network topology [29]. In another methods, the stability of the power system is detected by using the characteristic roots the pulse transfer function of the auto regressive model of the system [166].

A.2.2 Direct Methods

Direct methods analyze the transient stability using direct calculation of the Transient Energy Function (TEF). The Energy Functions are fully described in references such as [21]. In these methods, it is not needed to explicitly calculate the system solution by differential equation of the electric power system over time [155]. Direct stability analysis technique has been proposed based on second method of Lyapunov's stability theory. The objective is to define an energy or Lyapunov function with certain characteristics to obtain a direct measure of the stability region [118]. At the time of the change in the power system, the energy is captured and compared to pre-determined critical value. The criteria of Lyapunov stability is considered sufficient, and if the states are not in the boundary stability of the system cannot be inquired. It should be noted that, integration of the system equations during fault is still needed to obtain the post-fault initial conditions and the critical clearing time. In general, direct methods are more preferable than time domain simulations for online operation in dynamic stability assessment since it only requires simple mathematical operations [167]. Literature in this category is respectfully rich. In [168], the potential and kinetic energy functions are approximated real-time in heuristic manner to define the total energy of each generator in the system. Further, in [169], the interchange of

energy between areas is used to analyze inter-area oscillation. In [170], this method is employed to find the location of oscillation sources in power systems.

A.2.3 Hybrid Numerical and Direct Method

considering time domain and direct methods separately, both methods are not fully reliable and efficient enough to be applied in real-time TSA. In order to overcome the pitfalls of these methods, mentioned briefly before, hybrid methods based on the combination of these methods have been proposed.

In order to reduce the work load of time domain methods, and eliminating the dependency on the parameters of the system, trajectory prediction is performed real time based on the observed WAM data. This approach is based on the fact that a equivalent model can actually be integrated much faster in real-time. Based on Single Machine Equivalent (SIME) method, the reduced-order model is derived by observing the post-fault swing curve data and clustering the machines into coherent groups of critical and noncritical generators by means of SMIB can be performed. Further, Direct energy methods, Equal-Area Criterion, can be utilized to predict the transient stability status of the system, as well as, the stability margins [29, 30, 171, 172].

Another hybrid method to tackle the problem is early-termination criteria, allowing simulation of stable cases to be aborted as soon as possible [123, 173]. This criteria for time domain simulations is based on of coherency, transient energy conversion of kinetic and potential energies, and the product of system variables [123, 124, 173].

A.2.4 Intelligent methods

Intelligent methods have shown encouraging application potential due to their speed in TSA [174]. Various advanced artificial intelligent techniques as well as machine learning approaches have been proposed to develop TSA and promising results have been gained. These methods effectively learn and map the process behavior from relationship between specified inputs and outputs, without any prior knowledge

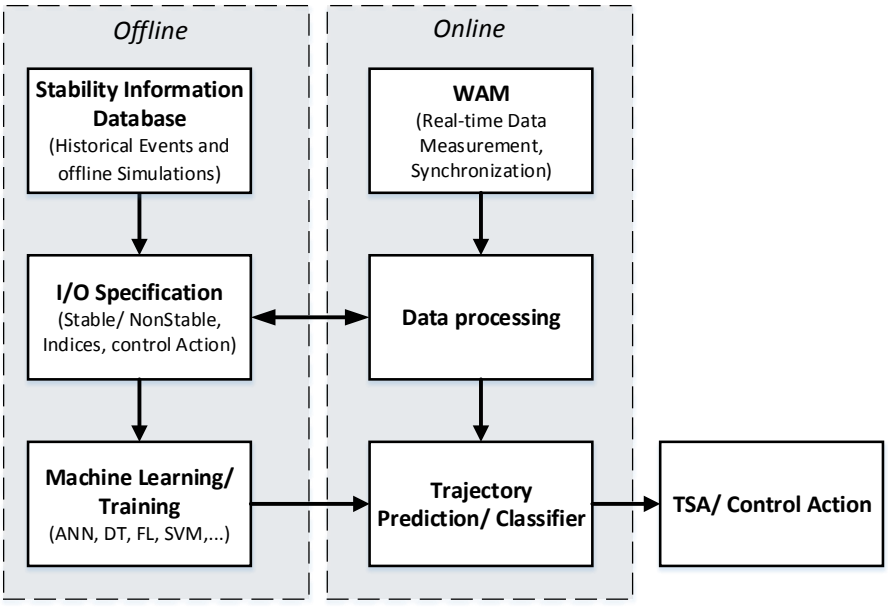


Figure A.1: Intelligent transient stability assessment/control scheme

of the process behavior. The initial idea was introduced as a pattern recognition in 1968 [175]. Decision Tree (DT) algorithm as a common method in this area, recursively splits the training data in terms of class membership [125]. It clusters states into groups with similar values for the response variable [154]. DT for transient stability prediction and classification involves considerable amount parameters and complex mapping functions, which usually requires many processing, terminal nodes, and training samples. moreover, Fuzzy Logic (FL) techniques have high applicability joint with DT as a classifier [67, 124, 176].

Another method that has been proposed for this purpose is Neural Networks (NNs). NNs also, lack the capability to learn all input output mapping relationships, causing less consideration in the industrial applications. Further, several methods such as parallel NN and Support Vector Machines (SVMs), has been proposed to overcome this problem [125, 177].

Fig. A.1 depicts a unified scheme of these TSA methods. As a machine learning methods, offline training should be performed before taking any actions in real-time.

Table A.1: Intelligent transient stability methods

Input	Design Approach	Ref.
δ^{COI}, ω^{COI}, V, f-domain criteria	DT, FL	[67]
δ^{COI}, ω^{COI} , f-domain indices	DT, Random Forest learning	[154]
δ^{COI}, ω^{COI} , f-domain indices	T-D Simulation, FL and NN	[124]
δ, ω	ANN	[125]
δ^{COI}	ANN	[178]
δ^{COI}, velocity and acceleration	Fuzzy hyper rectangular NN	[151]
V Trajectory	Fuzzy C-mean Clustering, SVM	[176]
COI δ and ω	SVM	[179]
$\Delta\delta$ security Criterion	NN, Particle swarm optimization	[180]
State (6th order)	Model-based finite Window predictive control	[40]

The source of the training so-called stability information database is constructed based on historical data and excessive offline simulations to cover a wide range of operating points. Next, Input and Output (I/O) specification based on the proposed method is set. These input specification can be PMU based states, trajectories, and indices. The outputs are classification, preventive, remedial, or damping control actions. Further, offline training is performed based on the I/O. In this stage, the learning algorithms (ANN, SVM, DT, FL) are used to extract the knowledge on the mapping relationship between input and output, and formulate the knowledge as a classifier, predictor, or control. These methods are very fast but again their accuracy highly depends on offline learning [29]. Table 8.1, provides a brief list of real-time intelligent TSA methods.

APPENDIX B: RENEWABLE ENERGY RESOURCES CONTROL

B.1 Local Control of renewable resources

Local control is the first control level in the control hierarchy of renewable resources that has the fastest response due to its distance to measurements. This control monitors the local system dynamics and ensures that the variables which are frequency and voltage track their set points. The majority of these units are connected to the grid by Voltage-Sourced Converter (VSC) units. Voltage control of a VSC in the dq-frame can be achieved based on inner current control and outer voltage control loops. The controller of the inner loop maintains the converter current, and controller of the outer loop regulates the output voltage [80–82].

Like synchronous generators power sharing is used in the renewable resources by including virtual inertia and droop method. Each inverter has an external power loop to share active and reactive power among the units and improve the system performance and stability, adjusting at the the frequency and voltage. The droop control scheme can be defined as [82]

$$\omega = \omega^* - m(P - P^*) \tag{B.1}$$

$$E = E^* - n(Q - Q^*) \tag{B.2}$$

where, ω^* and E^* are the frequency and the amplitude of the terminal voltage, respectively, m and n coefficients define the corresponding slopes, and P^* and Q^* are the active and reactive power references based on local feedback. Fig. B.1 depicts the overall scheme of the control hierarchy for the renewable energy sources. Moreover, the virtual impedance loop can be used to fix the output impedance of the inverter by subtracting a processed portion of the output current to the voltage reference of the inverter. The droop control is highly related to R/X ratio of the line, which is deteriorated with a small mismatch in the grid impedance estimation [81]. In the

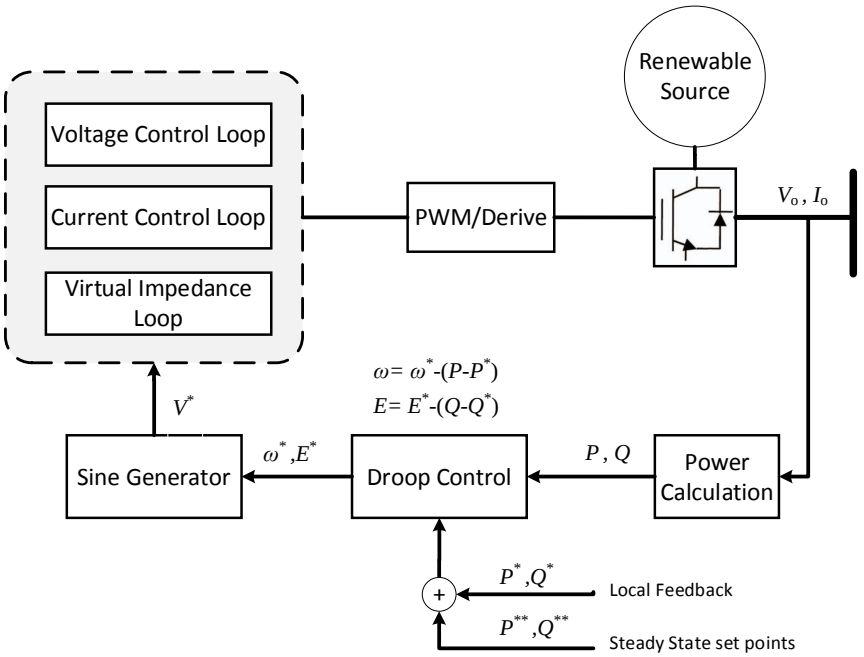


Figure B.1: Structure of the adaptive critic design

case of DFIG, rotor side control is responsible for maintaining the power.

In order to maintain the stability of the renewables for longer time, the next level controller is employed to determine the references for the power sharing control. The secondary control works as a centralized AGC and compensates the steady-state errors of the states. This controller using WAMs, coordinate the action of all the units within a specified area. The time response of this control level is in the range of minutes, having a slow dynamic compared to local controllers. Finally, the tertiary control level could be used for optimizing the operation of the system.

B.1.1 Wide-area Control in Presence of Renewable Sources

Renewables are usually far from the synchronous generators where the oscillations happen, therefore, WAMs are used to capture the oscillations and coordinate the control actions of these resources with the rest of network in the form of WAC. Rotor angles and speed of synchronous generators are directly affected by the type of renewable's power control [80, 84]. Reference [84] shows that when active power flow changes, the the way that the wind farm provides reactive power to the system

is critical in maintaining rotor angle stability of synchronous generators. This also applies to active power control, due to the fact that the oscillations are produced by active power differences between generation and load. Therefore, the implementation of appropriate control strategies in renewable sources can mitigate large rotor angle swings.

B.1.1.1 Conventional Techniques

Conventional control techniques such as PI, are widely used for the active and reactive power of renewable sources [181]. Reference [144] has used a strategy of controlling the magnitude and position of the rotor flux vector, mimicking the excitation control of a synchronous generator design. Reference [144] has uses local signals to design a PSS for system damping. A global PSS has also been proposed in this area, such as [39], that exchanges information with local PSSs to improve power systems oscillatory stability. A reduced version of the conventional PSS has been proposed in [182], comprising a filter and a proportional controller with the frequency as input and the power reference as output, exhibiting an effective damping of oscillations. Root locus methods has also been used to design the controllers by adding pole-zero pairs to attract the root locus towards left of s-plane [183]. These techniques are required to identify the most adequate pair of input-output signals based on small-signal stability analysis or controllability-observability analysis. These techniques needs accurate models of the renewable and power system, with large number of parameters to be tuned.

In order to perform the tuning of such controllers, optimal control theory could be used to minimize stability criteria [184]. For instance, particle swarm optimization method has been used in [185, 186]. Among these methods genetic algorithm is used in [187] to set the parameter control of DFIG. Furthermore, Reference [188] has proposed a mixed control of eigenvalue structure and a nonlinear optimization method for the conventional PSS of the DFIG to damp the system oscillations. Robust control

techniques has also been used for this application. For instance, [189] has employed a minimax linear quadratic gaussian-based power oscillation damper for PV plants. The proposed controller works adaptively under different operating conditions.

B.1.1.2 Lyapunov-based methods

Another control method in this field is based on the Lyapunov stability function. This method is more effective in transient stability area. In order to apply the direct method of TSA, the Lyapunov function for a renewable unit can be proposed from the droop controller variables, which are frequency and voltage. A control strategy in [139] is proposed based on overall Lyapunov function of the system including the DIFGs. The advantage of such method is that the strategy is independent of the network topology and is based on that the derivative of the energy function of the system being negative.

B.1.1.3 Intelligent methods

Intelligent control strategies, such as NNs and fuzzy logic, has been successfully applied to control of modern power system in different applications [87,88]. In Chapter 2 we have covered such designs based on supervised and reinforcement learning.

B.2 Stability Analysis and Control for Renewable Connected Power Grid

It is known that, with the increased penetration of renewables in transmission system, the effective inertia of the system is reduced and system stability following a disturbances could be significantly affected [73]. In the case of a DFIG-based wind turbines the inertia of the turbine is decoupled from the system as shown in chapter 6. The power electronic converter of the DFIG controls the performance and acts as an interface between the machine and the power grid. Transient disturbances in the system change the speed and the angle of the rotor. However, due to the asynchronous operation, the position of the rotor flux vector does not rely on the physical position of the rotor [144]. In the case of PV units, another impact is

lack of reactive power support [74]. The PV units are mainly utilized as sources of active power and no reactive power is generated by these resources. In general, the introduction of renewable energy resources has the potential to change the damping performance of the system through displacing synchronous machines; impacting line power flows; displacing synchronous machines that have PSSs.

Recently several works have been done to address the impacts of renewable resources on power system stability [73, 75–77]. Generally, it is believed that the renewable *type* does not significantly take part in power system oscillations; rather, the *penetration level* has a damping effect due to reduction in the size of conventional generators [77]. In addition, [75] shows that increased penetration of renewable resources is accompanied by congestion at weak interconnection lines, which leads to reduced damping. Comprehensive study on the impacts of high PV penetrations on local oscillation mode has been provided in [76]. The article shows that there is a threshold of PV penetration level beyond that PV tends to decrease the damping of local oscillations mode. Ref. [190] also shows how increased PV penetrations could influence the local mode of the system, and advocates that the large-scale PV tends to increase the damping of these modes. Further, [191] has shown that the distributed PV is better than the concentrated penetration of these units for inter-area mode damping.